A New Land Cover Map of Two Watersheds under Long-Term Environmental Monitoring in the Swedish Arctic Using Random Forest Classification of Sentinel-2 Data

Yves Auda *, Erik J. Lundin, Jonas Gustafsson, Oleg S. Pokrovsky, Simon Cazaurang, Laurent Orgogozo

Posted Date: 12 July 2023

doi: 10.20944/preprints202307.0841.v1

Keywords: land cover; sentinel-2 images; random forest; boreal forest; alpine tundra

Article

A New Land Cover Map of Two Watersheds under Long-Term Environmental Monitoring in the Swedish Arctic Using Random Forest Classification of Sentinel-2 Data

Yves Auda [1,*], Erik J. Lundin [2], Jonas Gustafsson [2], Oleg S. Pokrovsky [1,3], Simon Cazaurang [4] and Laurent Orgogozo [2]

[1] GET (Géosciences Environnement Toulouse), UMR 5563 CNRS/UR 234 IRD/UPS, Observatoire Midi Pyrénées, Université de Toulouse, Toulouse, France; oleg.pokrovski@get.omp.eu
[2] Swedish Polar Research Secretariat, Abisko Scientific Research Station, Sweden; erik.lundin@polar.se (E.J.L.); volehunter@gmail.com (J.G.); laurent.orgogozo@get.omp.eu (L.O.)
[3] BIO-GEO-CLIM Laboratory, Tomsk State University, Tomsk, Russian Federation
[4] Toulouse Institute of Fluid Mechanics (IMFT), National Polytechnic Institute of Toulouse, Toulouse, F-31400, France; simon.cazaurang@imft.fr
* Correspondence: yves.auda@get.omp.eu; Tel.: +33(0)6 02 24 51 13

Abstract: A land cover map of two arctic catchments, nearby the Abisko Scientific Research Station, was obtained from a classification of a Sentinel-2 satellite image and a ground survey performed in July 2022. The two contiguous catchments, Miellajokka and Stordalen, are covered by various ecotypes, from boreal forest to alpine tundra and peatland. The random forests algorithm correctly identified 88% of polygon pixels reserved for testing. The developed workflow relied solely on open source software and acquired ground observations. Space organization was directed by the altitude as demonstrated by the intersection of the land cover with the topography. Comparison between this new land cover map and previous ones based on data acquired between 2008 and 2011 shows some trends of vegetation cover evolution in response to climate change in the considered area.

Keywords: : land cover; sentinel-2 images; random forests; boreal forest; alpine tundra

1. Introduction

The nature of the land cover, including vegetation covers, bare rock outcrops and surface water bodies are of major importance to understand hydrological and biogeochemical fluxes on continental surfaces [1,2,3]. It is especially true in the Arctic, where permafrost conditions may strongly control the present ecotypes and their distributions [4,5,6,7], while vegetation cover variability may in turn significantly impact thermo-hydrological conditions [8,9]. In permafrost-affected soils, strong coupling between water and heat transfers occurs, and thus the thermal buffering of the vegetation cover is a key determinants of permafrost dynamics [10,11,12,13,14]. Evapotranspiration fluxes may also be a dominant term of the water budget in permafrost regions [15,16]. For all these reasons, permafrost modeling requires detailed knowledge of up to date land cover distribution.

The vast extension and the remoteness of the Arctic regions make the establishment of field survey – based land cover maps difficult. Moreover, fine resolutions and open data maps are needed for many applications [17], including permafrost modelling. Thus there is a growing interest in airborne [18,19] and remote sensing [20,21,22] capable of producing fine resolution vegetation maps in the Arctics. These regions are experiencing intensive climate change [23]. Permafrost thawing results in methane and carbon dioxide emissions [24] which contributes to the greenhouse effect. These modifications induce changes in ecotypes [25] that are visible at the landscape level. Thus, there is a need for not only high spatial resolution maps, but also for high temporal resolution survey.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2023 doi:10.20944/preprints202307.0841.v1

In order to produce regularly updated land cover map for large areas, the use of remote sensing data from long term satellite missions combined with in-situ information is required [21].

Here we present a workflow for creating high resolution vegetation map using only open data and open source software along with dedicated field data. The workflow is applied to two watersheds in the Swedish Arctic, for parts of which previous vegetation maps at lower resolutions and/or in past climatic conditions where already available [18,26,27]. The obtained map is used for investigating a link between topography and vegetation distribution, and assessing the temporal evolution of the vegetation cover during a 14 years period (2008-2022). This information is crucially important forpermafrost modeling works of the studied sites within a permaFoam approach [16,28], and it will also provide new insights on contemporary landscape evolution in this type of environment.

2. Materials and Methods

2.1. General geographic information about the study area

We selected two watersheds close to Abisko Scientific Research Station (INTERACT Network)(Figure 1). The first one, from West to East, is Miellajokka, a sub-alpine catchment which includes the iconic mounts of Tjuonavagge. This 51.5 km² catchment presents altitudes ranging from 383 to 1731 m above sea level [29]. The most eastern watershed is Stordalen, a 16 km² catchment with a lake-rich, peat-rich Northern part, and a sub-alpine Southern part, with elevation between 350 and 770 m above sea level [30,31,32,33]. In Stordalen vegetation maps of the Northern, low elevation part has been already produced based on airborne data of 2000 [26]. Later on, another vegetation map for the whole watershed has been produced using airborne data of 2008 [27]. Both Stordalen and Miellajokka are encompassed in the area studied by Reese et al. [18], with a vegetation map established on the basis of 2010 satellite images, using also data acquired by a lidar survey.

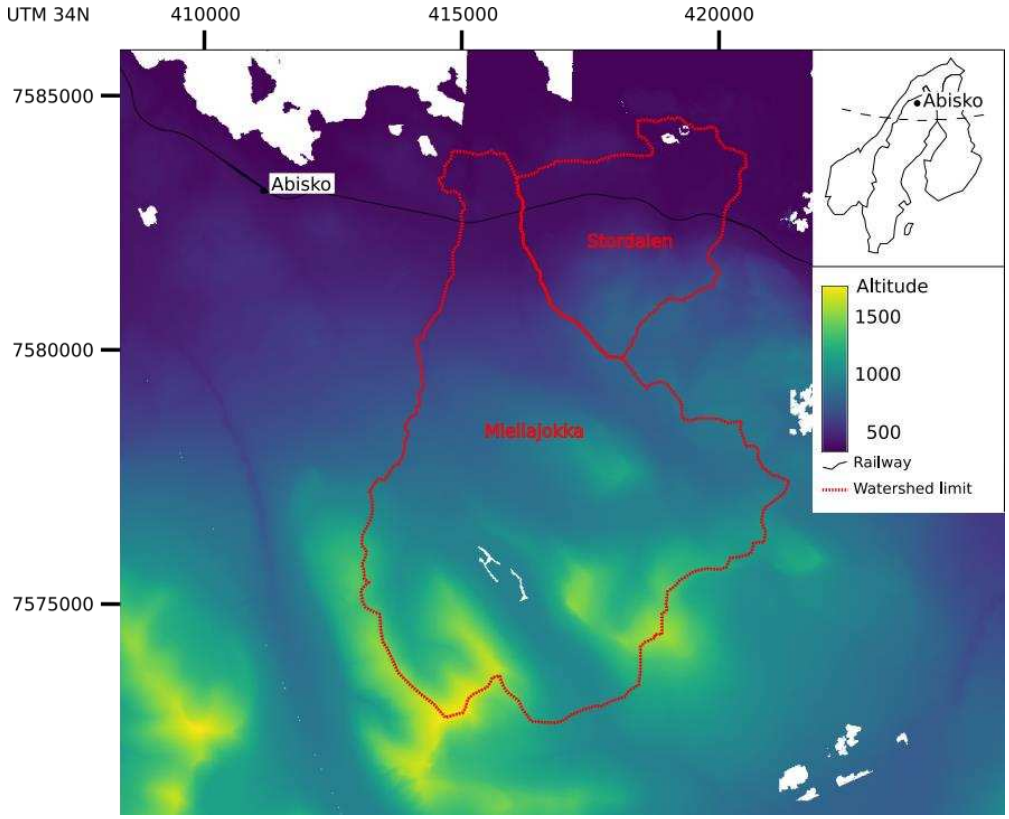

Figure 1. Location of the study zone on the articDEM map.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2023 doi:10.20944/preprints202307.0841.v1

2.2. Satellite image and digital elevation model

Obtaining images in the Arctic zone to study vegetation cover is difficult. These geographical areas are covered with snow for a large part of the year, which prevents any satellite study of the vegetation cover. In addition, frequent clouds hinder the acquisition of optical images. A single Sentinel-2 image acquired on 25 August 2022 was downloaded from https://peps.cnes.fr. Ten bands were selected for land cover classification (B02 – blue, B03 – green, B04 – red, B05 – red-edge 1, B06 – red-edge 2, B07 - red-edge 3, B08 – NIR, B08A – narrow NIR, B11 – SWIR 1, B12 – SWIR 2). The image is not corrected for atmospheric effects (Level-1C). The images are stored in the UTM34N reference coordinate system and all calculations are performed in this system to avoid altering the radiometry by re-projection.

On the basis of four (B03, B04, B08, B11) out of the 10 acquired channels, four derived indices were calculated: Bright, NDVI, NDWI, NDII (Table 1). The bright index is very sensitive to albedo. It distinguishes between light and dark soils. The NDWI (Normalized Difference Water Index) was used to detect water areas. The NDVI expressed the photosynthesis of the vegetation cover. The use of NDII [34,35] did not improve the results and was not retained for the final classification.

Table 1. Vegetation indices. Band notation correspond to MSI sensor of sentinel-2 satellite.

Index	Formula
Bright	$\sqrt{(B04 * B04)/(B08 * B08)}$
NDVI	$(B08 - B04)/(B08 + B04)$
NDWI	$(B03 - B08)/(B03 + B08)$
NDII	$(B08 - B11)/(B08 + B11)$

Since vegetation in mountainous areas is related to altitude, the digital terrain model is a very useful data source. ArcticDEM is an NGA-NSF public-private initiative to automatically produce a high-resolution digital surface model of the Arctic using optical stereo imagery. The majority of ArcticDEM data was generated from the panchromatic bands of the WorldView-1, WorldView-2, and WorldView-3 satellites and, for a small percentage of data, from the GeoEye-1 satellite. For this study, ArticDEM Release 7 "mosaic" format files with a spatial resolution of 2 m were downloaded at https://data.pgc.umn.edu/elev/dem/setsm/ArcticDEM/mosaic/v3.0/.

2.3. Field survey

The ground-truth survey took place from 21 July 2022 to 24 July 2022 in the Miellajokka and Stordalen watersheds, northern Sweden. We geolocalized areas of the different land cover types in field using a GPS, GLONASS, Beidou and Galileo navigation systems supported Samsung Galaxy Tab S6 Lite tablet. The Qfield software was used for data entry in the field. Its compatibility with QGIS simplifies data collection and subsequent analysis [28].

A field survey database including a color composite of Sentinel-2 image channels B08/B04/B03, the Open Street Map data and an empty vector layer was prepared in QGIS then transferred to Qfield.

Areas of observed and land cover types were highlighted as polygons on Figure 2. Each surveyed polygon serves a ground truth location. As a complement, a photo of the most characteristic observations was taken with the tablet camera.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2023 doi:10.20944/preprints202307.0841.v1

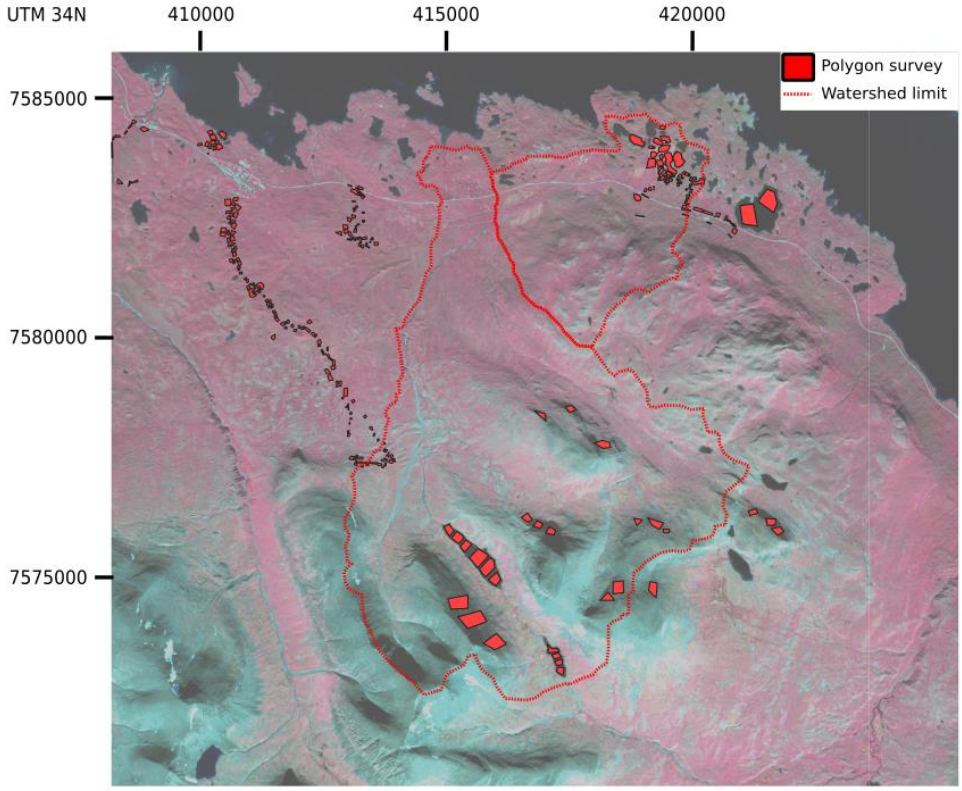

Figure 2. Field surveys drawn on color composite of Sentinel-2 images channels B08,B04,B03.

The 270 observations conducted during the field survey only identified seven out of the 12 classes by Reese et al. [18]. "Alpine meadow" was not encountered enough to constitute an individual class. Likewise, the "Mountain birch - meadow" class was only observed in six ground truth polygons and was grouped with the "Mountain birch - moss" class to form a single "Mountain birch" class. Snowbeds were poorly represented and are not included. "Grass heaths" were not encountered. Further, "Rock" class mainly represents bed rock outcrops but may also include thin organic soil and sediment. We also added "Human infrastructure" as a new class, mainly representing the road and the railway passing through the mapped area. Shadows in the steep areas to the south of the study area hinder recognition of the landscape they cover. To avoid confusion, especially with water, a "Shadow" class was created, summarizing the total number of classes to nine (Table 2).

Table 2. Land cover classes.

Class	Number of polygons	Number of pixels
Rock	25	361
Dry heath	35	889
Mesic heath	21	801
Wetland	29	1614
Alpine willow	19	402
Mountain birch	105	4587
Water	30	8568
Human infrastructure	13	312
Shadow	18	7026

As far as possible, we balanced the number of survey polygon between the representative classes. Poor accessibility due to difficult terrain limited the choices of locations (Figure 2). Thus, a

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2023 doi:10.20944/preprints202307.0841.v1

randomized field survey design was not possible in this natural environment, the transition from one land-use class to another is sometimes gradual, making it difficult to assign an area to a specific class. For this reason, surveys are only carried out in areas that are homogeneous in terms of the characteristics of the land cover classes.

2.4. Classification

The random forests method is widely used in image classification studies [37,38]. It uses decision trees and random draws without replacement of samples and variables to classify the Sentinel-2 image. Within each class, 30% of the surveyed polygons are randomly drawn and reserved for classification quality assessment. The classification is trained with the remaining 70%. GRASS software is used for the calculations [39]. The extension r.learn.ml2 interfaces with the Scikit-learn library written in python to perform random forests classification. Two parameters were selected in order to produce the forest trees. The number of trees was set to 100. The number of variables to be selected and tested for the best split when growing the trees was set to the square root of the number of input variables.

3. Results

3.1. Vegetation map in current climatic conditions

The classified image (Figure 3) shows patterns consistent with the knowledge of the terrain. The transport infrastructures are described with precision in their continuity. The lake Torneträsk at the North is homogeneously identified.

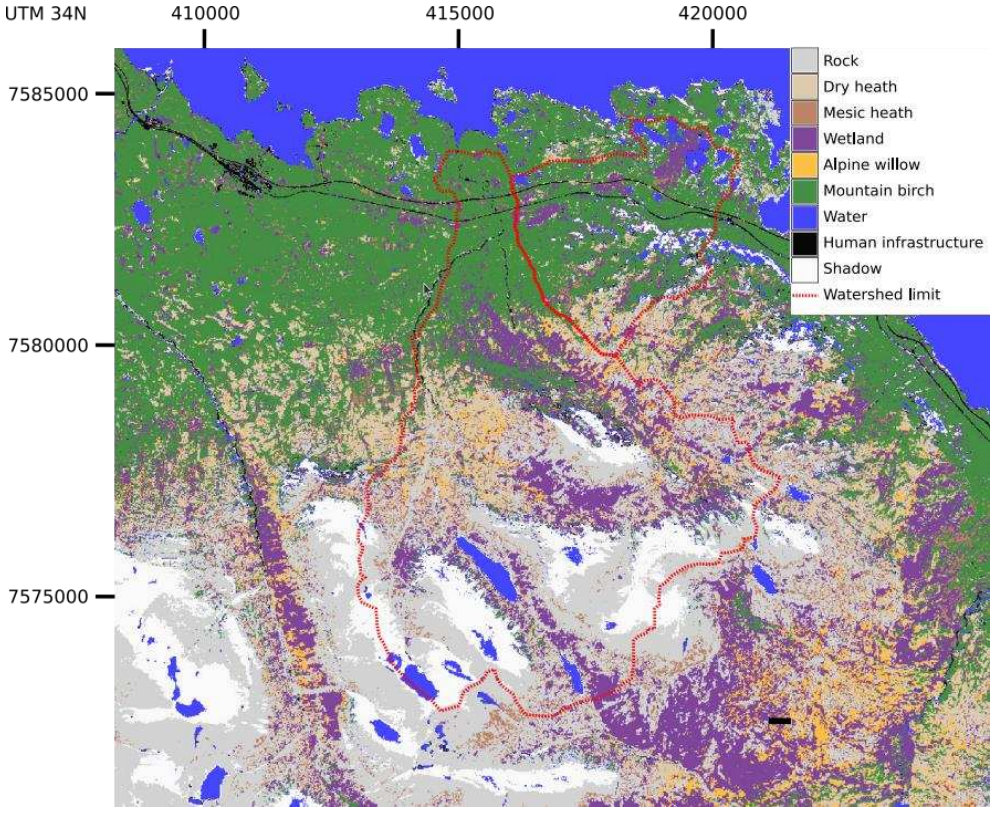

Figure 3. Image classified from the July 2022 field survey.

The statistics computed from surveyed polygons reserved for classification quality assessment confirm the quality of the classification. The percentage of pixels correctly classified by random forests is 88%. The confusion matrix (Table 3) provides an analysis of the accuracy of the classification used for building our map at the class level. The largely dominant weight of the diagonal terms in this confusion matrix demonstrates the quality of the classification. The confusion between "Dry

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2023 doi:10.20944/preprints202307.0841.v1

heath" and "Mesic heath" is understandable because these two formations are differentiated primarily by canopy height, a feature not accessible from the images used in our study. Likewise, the confusion between "Alpine willow" and "Wetland" is due to the difficulty of recognizing spaces occupied by a few willow plants. Besides, with such a pixel classification approach, places that are temporarily flooded at the moment of the satellite image acquisition are difficult to distinguish from true wetlands, i.e. places that are under water almost all along the active season. This could overestimate the wetland area, including the places with other vegetation types such as meadow which are temporarily flooded by ground water discharge or snowmelt water. Another important point is the detection of temporary high elevation open water bodies in several places around the Tjuonavagge lake, according to both this classification and the two indices NDII, NDVI values of the pixels. These ones may be generated by late snowmelt in the highest places of the landscape. Finally, the confusion between "Dry heath", "Mesic heath" and "Mountain birch" are classes that can be contiguous and even associated in some places. It describes mixed spaces where several classes coexist, i.e. ecotone between these classes.

Table 3. Confusion matrix of the random forests classification. The asteriks highlights in confusion matrix indicate the most significant confusions. The columns show the field surveys and the rows show the classification results. For instance, the number 117 corresponds at the cross of the "Mountain birch" line and the "Mesic heath" column means that 117 pixels that have been classified as "Mountain birch" belong to "Mesic heath" according to the field survey.

	Rock	**Dry heath**	**Mesic heath**	**Wetland**	**Alpine willow**	**Mountain birch**	**Water**	**Human infrastructure**	**Shadow**
Rock	84	2	15	0	0	2	0	0	0
Dry heath	0	168	59*	11	18	2	0	0	0
Mesic heath	17	18	49	5	6	8	0	0	0
Wetland	0	0	0	484	51*	6	4	1	0
Alpine willow	0	2	4	0	40	1	0	0	0
Mountain birch	0	98*	117*	102*	38*	1312	8	1	0
Water	2	0	0	1	0	0	2913	12	309
Human infrastructure	1	0	0	0	0	0	8	47	0
Shadow	5	0	0	0	0	0	44	10	2051

Classification accuracy was accessed by class using producer's accuracy, user's accuracy, class area estimates and standard error of class estimate in table 4 [40,41]. Concerning the producer's accuracy criterion, only two classes "Mesic heath" and "Alpine willow" are poorly assessed. These classes represent only 4% of the whole image (table 5). "Alpine willow" is confused with "Wetland", which is understandable given that this plant grows in very wet areas. "Mesic heath" is confused with "Dry heath". These two classes are similar and are differentiated mainly by the height of the vegetation from a remote sensing point of view. The "Dry heath" producer's accuracy criterion is poorly assessed because it is confused with "Mesic heath". This confusion is explained for the reason given above.

All the pixels classified as "Alpine willow" belong to this class. An area classified as "Alpine willow" therefore corresponds to this class. However, an area of land covered by this plant is not always assigned to this class. Thus, the area classified to "Alpine willow" is underestimated. Assuming that the distribution of surface areas follows a normal distribution, the approximate 95% confidence interval is obtained by the formula

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2023 doi:10.20944/preprints202307.0841.v1

$\hat{A}_k \mp 1.96\ S(\hat{A}_k)$ (1)

$\hat{A}_k$ class area estimate

$S(\hat{A}_k)$ standard error of class area estimate

All the class areas are estimated with low standard error. The standard error values are particularly good for classes with large areas: 17 pixels for 1331 pixels for "Mountain birch", 19 pixels for 2977 pixels for "Water".

Table 4. Accuracy estimates computed from polygons reserved for classification quality.

	Rock	Dry heath	Mesic heath	Wetland	Alpine willow	Mountain birch	Water	Human infrastructure	Shadow
Producer's accuracy (%)	77	58	20	80	26	99	98	66	87
User's accuracy (%)	82	65	48	89	85	78	90	84	97
Class area estimate (pixel)	109	288	244	603	153	1331	2977	71	2360
Standard error of class area estimate (pixel)	6	13	15	13	11	17	19	6	19

The "Mountain birch" class is the most represented (Table 5). It occupies more than a quarter of the space. The "Rock" class and the wetlands ("Wetland" and "Alpine willow") share 40% of the spaces. The other classes are much smaller.

Table 5. Distribution of land cover classes of our classification. The shadow class is not taken in account.

Class	Percentage
Rock	21
Dry heath	9
Mesic heath	4
Wetland	16
Alpine willow	4
Mountain birch	29
Water	16
Human infrastructure	1

3.2. Influence of altitude on land cover

Land cover appears to be strongly conditioned by altitude. Altitudinal zonation of land cover is encountered in various arctic contexts [42,43,44], including in the Abisko region [45,46]. In Miellajokka catchment, the seasonal variability of the hydrogeochemistry of the stream indicates a strong altitudinal control on hydrological processes, especially during the spring freshet [47], while hydrological conditions strongly interact with vegetation and carbon dioxyde fluxes [48]. Table 6, constructed by cross-referencing the land cover map with the ArticDEM, illustrates this phenomenon in the studied watersheds.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2023 doi:10.20944/preprints202307.0841.v1

Table 6. Percentage of land cover classes according to altitudinal levels. The "Water", "Human infrastructure" and "Shadow" classes are not taken in account.

Land cover	Level class (m)			
	Subalpine < 600	**Low alpine [600,800[**	**High alpine [800,1100[**	**Nival > 1100**
Rock	2	4	33	88
Dry heath	3	19	16	0
Mesic heath	3	7	5	6
Wetland	11	24	30	6
Alpine willow	1	6	9	0
Mountain birch	80	40	7	0

The subalpine level at altitudes below 600 m is mainly occupied by the "Mountain birch" class. The Alpine stage includes the "Dry heath", "Mesic heath", "Wetland" and "Alpine willow" formations. It is divided into two sub-stages: Between 600 and 800 m of altitude, the "Mountain birch" class is still very present. Above 800 m, this class gives way to the rock. The nival stage is composed only of rock, probably because it depends on harsher life conditions and more intense erosive processes at higher elevations.

3.3. Comparison between past and present vegetation maps

Three maps of parts of our study area have been produced by different authors. A map was constructed from aerial images of 8 August 1970 and 29 July 2000 [26], but the area covered is too small to allow comparison with our data. Another map of the Stordalen watershed by Lundin et *al.* [31] was obtained from images obtained from a helicopter flight on 1 August 2008. The most recent map of the Miellajokka watershed was produced by Reese et *al.* [18,49] from SPOT5 images of 28 July 2011 and laser data acquired under leaf-on conditions from two scanning dates (20 August 2010 and 9 September 2010). As the semantics of these maps are not identical to ours, an analysis of the variable typologies is necessary prior to the study of the landscape evolution. In order to overlay the maps and then calculate statistics, the Lundin and Reese maps extracted from the publications were georeferenced from control points identified in the landscape.

Comparison with the map of Reese et *al.* (based on data acquired in 2010)

Reese et *al.* [18,49] produced a land cover map with a larger number of classes. In order to compare the Reese map with our data, some classes of the Reese map are merged. "Snow ice" and "Snow bed" classes are grouped together. "Dry heath", "Extremely dry heath", "Grass heath" are also grouped together. The "Human infrastructure" class is not considered because it does not exist in Reese's work. "Alpine meadow" and "Tall Alpine meadow" were not confirmed by ground observations during our field trip.

The spatial distribution of the classes is slightly different (Table 7). The "Rock" class accounts for 36% for our classification and only 14% for the Reese map.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2023 doi:10.20944/preprints202307.0841.v1

Table 7. Percentage of land cover classes of Reese [18] and our classification in the Miellajokka watershed. The asteriks indicate classes which were not observed during our field trip.

Class	Reese (2010)	Our classification
Rock	14	36
Dry heath / Extremely dry heath / Grass heath	26	11
Mesic heath	4	5
Wetland	4	22
Alpine willow	19	4
Mountain birch	13	18
Water	4	4
Snow Ice, Snow bed	5	*
Alpine meadow / Tall alpine meadow	11	*

The change matrix shown in Table 8 encompasses the differences in semantics of classes and the changes in the landscape between the dates of the two maps (i.e.: 2014 and 2022), and may be also discrepancies due to the use of different methodologies. Table 8 and Figure 3 show that some "Rock" areas in the middle our map are covered by "Grass heath", "Dry heath" and "Alpine willow" on the Reese map. The forest is also growing slightly to the south in sparse patches.

Table 8. Change matrix comparing our classification to the map of Reese [18]. Each column corresponds to the percentage of pixels of a class obtained by our classification in function of the Reese map classes. The "Shadow" class is not taken in account.

		Our classification						
		Rock	Dry heath	Mesic heath	Wetland	Alpine willow	Mountain birch	Water
Reese map (2010)	Rock	32	3	10	3	2	1	12
	Dry heath / Extremely dry heath / Grass heath	32	36	41	24	33	8	9
	Mesic heath	1	4	3	5	6	10	1
	Wetland	2	6	3	7	9	3	1
	Alpine willow	15	29	22	24	31	14	10
	Mountain birch	0	4	8	12	7	53	1
	Water	3	2	3	3	2	4	38
	Snow Ice, Snow bed	8	2	2	2	1	1	24
	Alpine meadow / Tall alpine meadow	7	14	8	20	9	6	4

On the other hand, "Alpine willow" is also more represented on Reese map without any conclusion being drawn because this class is misclassified by Reese [18]: the confusion matrix indicates 20 of 44 pixels misclassified. In addition, "Alpine meadow" which we have not taken into account, is identified as "Wetland" on our map (Figure 4), maybe because most of "Alpine meadow" places were temporarily flooded at the time of observation(see also section 3.1.). This class "Wetland" exists on the map published by Borgelt [49] but it is not present in the confusion matrix published by Reese [18]. The other classes do not show significant difference in in their proportion and spatial distribution.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2023 doi:10.20944/preprints202307.0841.v1

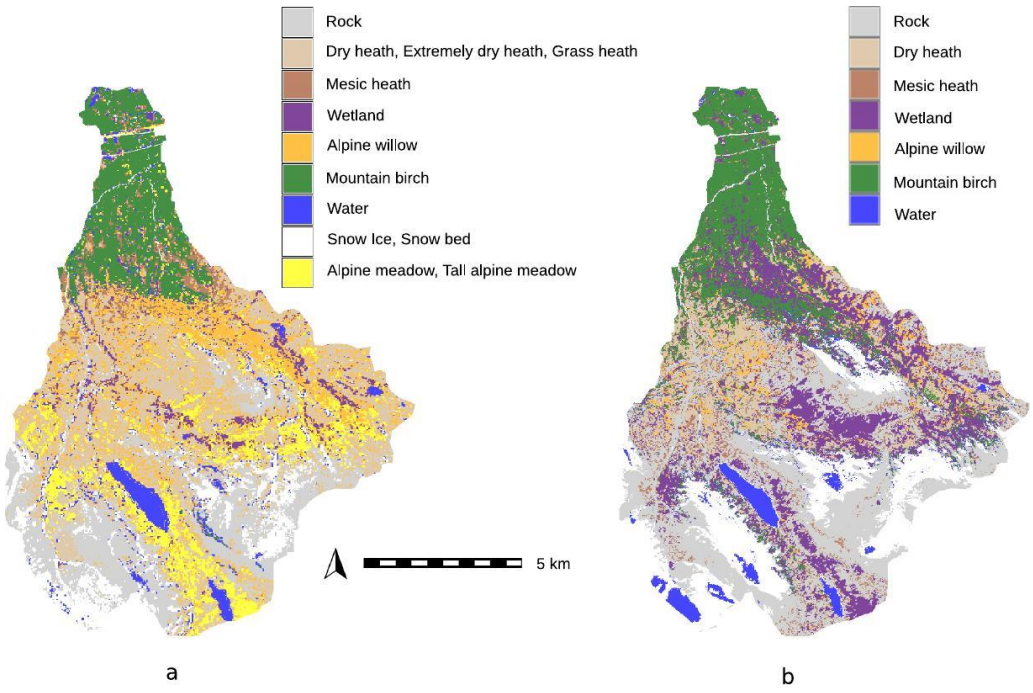

Figure 4. Land cover map by Reese (2010) [18] (a) and our reclassified map (2022) (b) in the Miellajokka watershed.

Despite the differences between the two maps, figure 3 shows a great consistency in the distribution of space. Three elements can explain these differences. 1) The landscape is natural. There are no parcels to structure it. Between areas occupied by two vegetation classes there is often a transition zone which is difficult to assign to a class. 2) The class definition of Reese [18] takes into account the height of the stratum using metrics derived from laser acquisitions, a technology we did not employ. 3) The landscape has evolved between 2010 and 2022. For example, some areas of rock have become vegetated.

Comparison with the map of Lundin et *al.* (based on data acquired in 2008)

The definition of classes in Lundin et *al.* [31] is significantly different from the one presented in this work, with a smaller number of classes in the map of Lundin. So the classes we used have to be modified to achieve a consistency between two maps. Grouping "Dry heath", "Mesic heath", "Alpine willow" classes of our classification allows them to be compared to the "Alpine tundra" class of the map of Lundin. Similarly, our "Wetland" class is compared to the "Peatland" class of the map of Lundin. The classes "Human infrastructure" and "Non vegetated" correspond to road, railway and building. The latter class is much more represented on the map of Lundin (Table 9). The differences are related to a larger road and railway footprint, which does not affect the landscape dynamics. This observation shows the satisfactory overlay of the two maps.

Table 9. Percentage of land cover classes of the map of Lundin [31] in the Stordalen watershed. In our classification, the "Alpine tundra" class corresponds to the grouping of classes "Dry heath", "Mesic heath", "Alpine willow". The class "Peatland" corresponds to "Wetland". The class "Non vegetated" corresponds to "Human infrastructure".

Class	Lundin (2008)	Our classification (2022)
Rock	9	5
Alpine tundra	13	11
Peatland	11	19
Forest	51	56
Water	7	7
Non vegetated	9	2

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 July 2023 doi:10.20944/preprints202307.0841.v1

The change matrix (Table 10) shows that the "Rock" class is also more represented on the map of Lundin, which covers twice as much space. The areas of this class that are not classified as “Rock” for our map are located in the south of the map (Figure 5). They are contiguous to the “Rock” areas of our map. It is also possible that there has been a forest expansion between 2008 and 2002 in this area. Indeed, the forest has grown between 2008 and 2022. It has gained some space in all land cover categories. Furthermore, some areas of “Peatland” appear to be transformed into “Wetland” but Lundin et *al.* [31] indicate "Peatlands were subdivided into wet areas (fen) and dry areas (bog) proportionally to what was found by Malmer et *al.* (2005)". It is therefore not possible to draw a conclusion from this observation. In summary, it is possible to compare our map with the map of Lundin after a semantic analysis of the categories. The main differences between the two maps concerns the south of the Stordalen watershed where Alpine toundra and forest are intermixed. The forest seems to have taken over areas previously occupied by tundra.

Table 10. Change matrix comparing our classification (2022) to the map of Lundin (2008).[31]. Each column corresponds to the percentage of pixels of a class obtained by our classification in function of the classes of the Lundin map. The shadow class is not taken in account.

		Our classification (2022)					
Lundin map (2008)		Rock	Dry heath Mesic heath Alpine willow	Wetland	Mountain birch	Water	Human infrastructure
	Rock	40	27	8	4	4	5
	Alpine tundra	27	26	12	11	4	8
	Peatland	3	9	30	7	8	7
	Forest	16	26	37	66	34	28
	Water	4	3	6	4	42	6
	Non vegetated	10	9	7	8	8	46

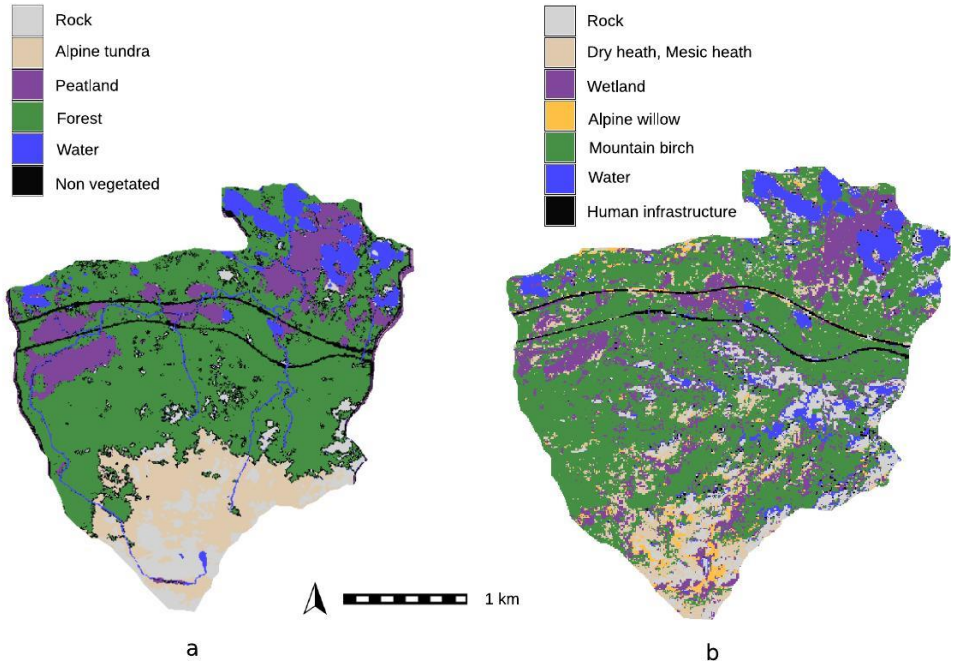

Figure 5. Land cover map by Lundin (2008) [31] (a) and our reclassified map (2022) (b) in the Stordalen watershed.

 doi:10.20944/preprints202307.0841.v1

4. Discussion

The methodology implemented in this study requires little sampling effort to quickly obtain a land cover map. This operation is feasible every year to monitor the evolution of land cover. These results are particularly important in a region subject to climate change which is currently undergoing major upheaval.

The results of the Sentinel-2 image classification show a structuring of the landscape as a function of elevation, with different vegetation level along with altitude [50]. In the Miellajokka and Stordalen catchments, three levels are present. The subalpine level (< 600 m) is mainly occupied by birch forest. The alpine level [600 m , 1100 m] is characterized by heath and willow. This level could be split into two sub-levels according to the respective abundance of birch forest at lower altitude and outcrops at upper altitude. The upper level, the nival level (> 1100 m), is only composed of rock. Some temporary open water bodies were localised at high altitude, which is a surprising feature, may be linked to late melt of high elevation snow bodies.

Comparison between past and present vegetation maps is not straightforward due to a lack of a common typology field survey protocols. Nevertheless, it was possible to identify a change in the landscape between 2008 and 2022. Analysis comparison of the maps of Lundin [31] and Reese [18] with the one presented in this work demonstrated an extension of the forest on the tundra towards the south (i.e., toward higher elevations) during the 2008-2022 period. This finding is in agreement with Rundqvist's work [51] which shows an upward movement of species observed over a study period between 1976 and 2010. Nevertheless one should be careful with this possible interpretation because of the statistical uncertainty of the different classifications. This trend could be a consequence of the on-going climate warming, demonstrated across the Arctic [52,53].

5. Conclusion

In this work we provided a new land cover map for two watersheds located nearby the Abisko Scientific Research Station, to be used in future permafrost modelling work in the framework of the HiPerBorea project (hiperborea.omp.eu) based on permaFoam approach [16,28]. This new map also provides some insights on recent land cover changes in the studied area, by comparison with previous maps based on data acquired in 2008 [31] and 2010 [18]. High elevation temporary water bodies have been detected, which requires further investigation in the future.

The proximity of the Abisko observation station makes the Miellajokka and Stordalen watersheds a privileged study area for the evolution of landscapes in the Arctic zone, in particular, where thawing of the permafrost at high altitudes is attested. This monitoring requires annual surveys according to a unified protocol in terms of sampling, definition of classes and method of recording in order to monitor the evolution of this region under on-going climate change. The present study presents a protocol that would be suitable for such purpose.

At the same time, the study of land cover in the Arctic zone poses a number of difficulties. The areas are covered by snow for a large part of the year, which limits observations to a few months of summer. Acquisition by passive optical satellites can only be made during the six months of perpetual daylight, from 21 June to 21 December. Fortunately, this period includes the summer months when snow cover is at its minimum extension. Cloud cover frequently hampers optical acquisitions. For the year 2022, only one Sentinel-2 image could be used, which illustrates the problems to rely on optical images only in such environments. Future work could involve radar images whose acquisition is not affected by clouds and polar night.

Author Contributions: The following statements should be used "Conceptualization, Y. Auda and L. Orgogozo; methodology, Y. Auda and L. Orgogozo; software, Y. Auda; validation, Y. Auda, E. Lundin, J. Gustafsson, O.S. Pokrovsky, S. Cazaurang and L. Orgogozo; writing—original draft preparation, Y. Auda and L. Orgogozo; investigation, Y. Auda, L. Orgogozo and S. Cazaurang; writing—review and editing, Y. Auda, E. Lundin, J. Gustafsson, O.S. Pokrovsky, S. Cazaurang and L. Orgogozo; project administration, L. Orgogozo; funding acquisition, L. Orgogozo. All authors have read and agreed to the published version of the manuscript." Please turn to the CRediT taxonomy for the term explanation. Authorship must be limited to those who have contributed substantially to the work reported.

Acknowledgments: This work has been funded by the French Research Agency ANR (grant n ◦ ANR-19 CE46-0003-01), O.S. Pokrovsky was partially supported by the TSU Development Programme "Priority-2030". The authors would like to thank Reiner Giesler and Emily Pickering Pedersen for their help for preparing the field work.

References

1. Walvoord: M.; Kurylyk B. Hydrologic Impacts of Thawing Permafrost—A Review. Vadose Zone J. 2016, 15,6 ; DOI:10.2136 /vzj2016.01.0010
2. Meyer, G.; Humphreys, E.R.; Melton, J.R.; Cannon, A.J.; Lafleur, P.M. Simulating shrubs and their energy and carbon dioxide fluxes in Canada's Low Arctic with the Canadian Land Surface Scheme Including Biogeochemical Cycles (CLASSIC). Biogeosciences 2021, 18, 3263–3283; DOI:10.5194/bg-18-3263-2021
3. Jones, B.J.; Grosse, G.; Farquharson, L.M.; Roy-Léveillée, P.; Veremeeva, A.; Kanevskiy, M.Z.; Gaglioti, B.V.; Breen, A.L; Parsekian, A.D.; Ulrich, M.; Hinkel, K.M. Lake and drained lake basin systems in lowland permafrost regions. Nat. Rev. Earth Environ. 2022, 3, 85–98; DOI:10.1038/s43017-021-00238-9
4. Cable, W.L.; Romanovsky, V.E.; Jorgenson, M.T. Scaling-up permafrost thermal measurements in western Alaska using an ecotype approach. The Cryosphere, 2016; DOI:10, 2517–2532; DOI:10.5194/tc-10-2517-2016
5. Douglas, T.A.; Zhang, C., Machine learning analyses of remote sensing measurements establish strong relationships between vegetation and snow depth in the boreal forest of Interior Alaska. Environ. Res. Lett. 2021, 16 065014; DOI:10.1088/1748-9326/ac04d8
6. Zhang, Y.; Lantz, T.C.; Touzi, R.; Feng, W.; Kokelj, S.V. Landscape-scale variations in near-surface soil temperature and active-layer thickness: Implications for high-resolution permafrost mapping. Permafrost and Periglac Process. 2021, 32, 4, 627–640; DOI:10.1002/ppp.2104
7. Heijmans, M.M.P.D.; Magnússon, R.Í.; Lara, M.J.; Frost, G.V.; Myers-Smith, I.; van Huissteden, J.; Jorgenson, Fedorov, N.A.; Epstein, H.E.; Lawrence, D.M.; Limpens J. Tundra vegetation change and impacts on permafrost. Nat. Rev. Earth Environ. 2022, 3, 68–84; DOI:10.1038/s43017-021-00233-0
8. Bagard, M.-L.; Schmitt, A.-D.; Chabaux, F.; Pokrovsky, O.S.; Viers, J.; Stille, P.; Labolle, F.; Prokushkin, A. Biogeochemistry of stable Ca and radiogenic Sr isotopes in a larch-covered permafrost-dominated watershed of Central Siberia. Geochimica et Cosmochimica Acta 2013, 114, 169–187; DOI:10.1016/j.gca.2013.03.038
9. 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.; Boike, J.; Miller, N.B.; Beringer, J.; López-Blanco, E.; Stoy, P.C.; Sullivan, R.C.; Kejna, M.; Parmentier, F.-J.W.; Gamon, J.A.; Mastepanov, M.; Wille, C.; Jackowicz-Korczynski, M.; Karger, D.N.; Quinton, W.L.; Putkonen, J.; van As, D.; Christensen, T.R.; Hakuba, M.Z.; Stone, R.S.; Metzger, S.; Vandecrux, B.; Frost, G.V.; Wild, M.; Hansen, B.; Meloni, D.; Domine, F.; te Beest, M.; Sachs, T.; Kalhori, A.; Rocha, A.V.; Williamson, S.N.; Morris, S.; Atchley, A.L.; Essery, R.; Runkle, B.R.K.; Holl, D.; Riihimaki, L.D.; Iwata, H.; Schuur, E.A.G.; Cox, C.J.; Grachev, A.A.; McFadden, J.P.; Fausto, R.S.; Göckede, M.; Ueyama, M.; Pirk, N.; de Boer, G.; Bret-Harte, M.S.; Leppäranta, M.; Steffen, K.; Friborg, T.; Ohmura, A.; Edgar, C.W.; Olofsson, J.; Chambers, S.D. Vegetation type is an important predictor of the arctic summer land surface energy budget. *Nat. Commun. 2022, 13, 6379*; DOI:10.1038/s41467-022-34049-3
10. Lenoir, J.; Graae, B.J.; Aarrestad, P.A.; Alsos, I.G.; Armbruster, W.S.; Austrheim, G.; Bergendorff, C.; Birks, H.J.B.; Bråthen, K.A.; Brunet, J.; Bruun, H.H.; Dahlberg, C.J.; Decocq, G.; Diekmann, M.; Dynesius, M.; Ejrnæs, R.; Grytnes, J.-A., Hylander, K.; Klanderud, K.; Luoto, M.; Milbau, A.; Moora, M.; Nygaard, B.; Odland, A.; Ravolainen, V.T.; Reinhardt, S.; Sandvik, S.M.; Schei, F.H.; Speed, J.D.M.; Tveraabak, L.U.; Vandvik, V.; Velle, L.G; Virtanen, R.; Zobel, M.; Svenning, J.-C. Local temperatures inferred from plant communities suggest strong spatial buffering of climate warming across Northern Europe. Glob. Change Biol. 2013, 19, 1470-1481; DOI:10.1111/gcb.12129

11. Aalto, J.; Scherrer, D.; Lenoir, J.; Guisan, A.; Luoto, M. Biogeophysical contr ols on soil-atmosphere thermal differences: implications on warming Arctic ecosystems. Environ. Res. Lett. 2018, 13, 7; DOI:10.1088/1748-9326/aac83e
12. De Frenne, P.; Zellweger, F.; Rodríguez-Sánchez, F.; Scheffers, B.R.; Hylander, K.; Luoto, M.; Vellend, M.; Verheyen, K.; Lenoir, J. Global buffering of temperatures under forest canopies. Nat. Ecol. Evol. 2019, 3, 744-749; DOI:10.1038/s41559-019-0842-1
13. Stuenzi, S.M.; Boike, J.; Gädeke, A.; Herzschuh, U.; Krise, S.; Pestryakova, A.; Westermann, S.; Langer, M. Sensitivity of ecosystem-protected permafrost under changing boreal forest structures. Environ. Res. Lett. 2021, 16,8; DOI:10.1088/1748-9326/ac153d
14. Stuenzi, S.M.; Boike, J., Cable, W.; Herzschuh, U.; Kruse, S.; Pestryakova, L.A.; Schneider von Deimling, T.; Westermann, S.; Zakharov, E.S.; Langer, M. Variability of the surface energy balance in permafrost-underlain boreal forest, Biogeosciences. 2021, 18,2, 343–365; DOI:10.5194/bg-18-343-2021
15. Park, H.; Yamazaki , T.; Yamamoto K., Ohta, T. Tempo-spatial characteristics of energy budget and evapotranspiration in the Eastern Siberia. Agricultura and Forest Meteorology, 148, 12, 1990-2005; DOI:10.1016/j.agrformet.2008.06.018
16. Orgogozo, L.; Prokushkin, A.S.; Pokrovsky, O.S.; Grenier, C.; Quintard, M.; Viers, J.; Audry, S. Water and energy transfer modeling in a permafrost-dominated, forested catchment of Central Siberia: The key rôle of rooting depth. Permafr. Periglac. Process. 2019, 30 75–89; DOI:10.1002/ppp.1995
17. Bartsch, A.; Höfler, A.; Kroisleitner, C.; Trofaier, A.M. Land Cover Mapping in Northern High Latitude Permafrost Regions with Satellite Data: Achievements and Remaining Challenges. Remote Sens. 2016, 8, 979; DOI:10.3390/rs8120979
18. Reese, H.; Nyström, M.; Nordkvist, K.; Olsson, H. Combining airborne laser scanning data and optical satellite data for classification of alpine vegetation. International Journal of Applied Earth Observation and Geoinformation 2014, 27, 81–90; DOI:10.1016/j.jag.2013.05.003
19. Greaves, H.E.; Eitel, J.U.H; Vierling, L.A.; Boelman, N.T.; Griffin, K.L.; Magney T.S.; Prager C.M. 20 cm resolution mapping of tundra vegetation communities provides an ecological baseline for important research areas in a changing Arctic environment. Environ. Res. Commun. 2019, 1; DOI:10, 10.1088/2515-7620/ab4a85
20. Langford, Z.L.; Kumar, J.; Hoffman, F.M.; Breen, A.L.; Iversen, C.M. Arctic Vegetation Mapping Using Unsupervised Training Datasets and Convolutional Neural Networks. Remote Sens. 2019, 11, 69; DOI:10.3390/rs11010069
21. Beamish, A.; Raynolds, M.K.; Epstein, H.; Frost, G.V.; Macander, M.J.; Bergstedt, H.; Bartsch, A.; Kruse, S.; Miles, V.; Tanis, C.M.; Heim, B.; Fuchs, M.; Chabrillat, S.; Shevtsova, I.; Verdonen, M.; Wagner ,J. Recent trends and remaining challenges for optical remote sensing of Arctic tundra vegetation: A review and outlook. Remote Sens. Environ. 2020, 246, 111872; DOI:10.1016/j.rse.2020.111872
22. Rudd, D.A.; Karami, M.; Fensholt, R. Towards High-Resolution Land-Cover Classification of Greenland: A Case Study Covering Kobbefjord, Disko and Zackenberg. Remote Sens. 2021, 13, 3559; DOI:10.3390/rs13183559
23. IPCC, 2019. Summary for Policymakers. In IPCC Special Report on the Ocean and Cryosphere in a Changing Climate. H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, M. Nicolai, A. Okem, J. Petzold, B. Rama, N. Weyer (eds.), https://www.ipcc.ch/site/assets/uploads/sites/3/2019/12/SROCC_FullReport_FINAL.pdf (accessed on 12 April 2023)
24. Patzner, M.S.; Logan, M.; McKenna, A.M.; Young R.B.; Zhou, Z.; Joss, H.; Mueller, C.W.; Carmen, H.; Scholten, T.; Straub, D; Kleindienst, S.; Borch, T.; Kappler A.; Bryce C. Microbial iron cycling during palsa hillslope collapse promotes greenhouse gas emissions before complete permafrost thaw. Commun. Earth Environ. 2022, 3, 76; DOI:10.1038/s43247-022-00407-8
25. Loranty, M.L.; Abbott, B.W.; Blok, D.; Douglas, T.A.; Epstein, H.E.; Forbes, B.C.; Jones, B.M.; Kholodov, A.L.; Kropp, H.; Malhotra, A.; Mamet, S.D.; Myers-Smith, I.H.; Natali, S.M.; O'Donnell, J.A.; Phoenix, G.K.; Rocha, A.V.; Sonnentag, O.; Tape, K.D.; Walker, D.A. Reviews and syntheses: Changing ecosystem influences on soil thermal regimes in northern high-latitude permafrost regions. Biogeosciences 2018, 15, 5287–5313; DOI:10.5194/bg-15-5287-2018
26. Malmer, N.; Johansson, T.; Olsrud, M.; Christensen, T.R. Vegetation, climatic changes and net carbon sequestration in a North-Scandinavian subarctic mire over 30 years. Glob. Change Biol. 2005, 11, 1895–1909; DOI:10.1111/j.1365-2486.2005.01042.x
27. Tang, J.; Miller, P.A.; Persson, A.; Olefeldt, D.; Pilesjö, P.; Heliasz, M.; Jackowicz-Korczynski, M.; Yang, Z.; Smith, B.; Callaghan, T.V.; Christensen, T.R., 2015. Carbon budget estimation of a subarctic catchment using a dynamic ecosystem model at high spatial resolution. Biogeosciences 2015, 12,9, 2791-2808; DOI:10.5194/bg-12-2791-2015

28. Orgogozo, L.,; Xavier, T.; Oulbani, H.; Grenier, C. Permafrost modelling with OpenFOAM®: New advancements of the permaFoam solver. Comput. Phys. Commun. 2023, 282, 108541; DOI:10.1016/j.cpc.2022.108541
29. Giesler, R.; Lyon, S.W.; Mörth, C.-M.; Karlsson, J.; Karlsson, E.M.; Jantze, E.J.; Destouni, G.; Humborg, C.; 2014. Catchment-scale dissolved carbon concentrations and export estimates across six subarctic streams in northern Sweden. Biogeosciences 2014, 11, 2, 525–537; DOI:10.5194/bg-11-525-2014
30. Lundin, E.J.; Giesler, R.; Persson, A.; Thompson, M.S.; Karlsson, J. Integrating carbon emissions from lakes and streams in a subarctic catchment. JGR Biogeosciences 2013, 118, 1–8; DOI:10.1002/jgrg.20092
31. Lundin, E.J.; Klaminder, J.; Giesler, R.; Persson, A.; Olefeldt, D.; Heliasz, M.; Christensen, T.R.; Karlsson, J. Is the subarctic landscape still a carbon sink? Evidence from a detailed catchment balance. Geophys. Res. Lett. 2016 , 43, 1988–1995; DOI:10.1002/2015GL066970
32. Mzobe, P.; Berggren, M.; Pilesjö, P.; Lundin, E.; Olefeldt, D.; Roulet, N.T.; Persson, A. Dissolved organic carbon in streams within a subarctic catchment analysed using a GIS/remote sensing approach. PLOS ONE 2018, 13, 7, e0199608; DOI:10.1371/journal.pone.0199608
33. Mzobe, P.; Yan, Y.; Berggren, M.; Roulet, N.T.; Persson, A. Morphometric Control on Dissolved Organic Carbon in Subarctic Streams. Journal of Geophysical Research: Biogeosciences 2020, 125, e2019JG005348; DOI:10.1029/2019JG005348
34. Hardisky, M.A.; Kiemas, V.; Daiber, F.C. Remote sensing salt marsh biomass and stress detection. Adv. Space Res. 1983, 2, 8, 219-229; DOI:10.1016/0273-1177(82)90243-5
35. Gao, B.C. NDWI - a normalized difference water index for remote sensing of vegetationfor liquid water from space. Remote Sens. Environ. 1996, 58, 257-266; DOI:10.1016/S0034-4257(96)00067-3
36. Ostadabbas, H.; Weippert, H.; & Behr, F. J. Using the synergy of qfield for collecting data on-site and qgis for interactive map creation by alkis® data extraction and implementation in postgresql for urban planning processes. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. 2020, 679-683; DOI:10.5194/isprs-archives-XLIII-B4-2020-679-2020
37. Breiman, L. Statistical Modeling: The Two Cultures, Statistical Science, 2001, 16, 3, 199-231; DOI: 10.1214/ss/1009213726
38. Belgiu, M.; Dragut, L. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 2016, 114, 24-31; DOI:10.1016/j.isprsjprs.2016.01.011
39. Neteler, M.; Mitasova, H. 2008. Open Source GIS: A GRASS GIS Approach, 3rd ed. Springer, New York.
40. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock C.E.; Wulder M.A. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 2014, 148, 42-57; DOI:10.1016/j.rse.2014.02.015
41. Nguyen, H.T.T.; Doan; T.M.; Tomppo, E.; McRoberts R.E. Land Use/Land Cover Mapping Using Multitemporal Sentinel-2 Imagery and Four Classification Methods. A case Study from Dak Nong, Vietnam. Remote Sens. 2020, 12, 9, 1367; DOI:10.3390/rs12091367
42. Ermakov, N.; Shaulo, D.; Maltseva T. The class Mulgedio-Aconitetea in Siberia. Phytocoenologia 2000, 30, 2 145-192; DOI:10.1127/phyto/30/2000/145
43. Hjort, J.; Luoto M. 2009. Interaction of geomorphic and ecologic features across altitudinal zones in a subarctic landscape. Geomorphology 2009, 112, 324-333; DOI:10.1016/j.geomorph.2009.06.019
44. Sieg, B.; Drees B.; Hasse, T. High-altitude vegetation of continental West Greenland. Phytocoenologia 2009, 39,1, 27–50; DOI:10.1127/0340-269X/2009/0039-0027
45. Sundqvist, M.; Giesler, R.; Graae, B.J.; Wallander, H.; Fogelberg, E.; Wardle, D.A. Interactive effects of vegetation type and elevation on aboveground and belowground properties in a subarctic tundra. Oikos 2011, 120, 128–142; DOI:10.1111/j.1600-0706.2010.18811.x
46. Sundqvist, M.; Liu, Z.; Giesler, R.; Wardle, D.A. Plant and microbial responses to nitrogen and phosphorus addition across an elevational gradient in subarctic tundra. Ecology 2014, 95, 7 1819–1835; DOI:10.1890/13-0869.1
47. Lyon, S.W.; Ploum, S.W.; van der Velde, Y.; Rocher-Ros, G.; Mörth, C.-M.; Giesler, R. Lessons learned from monitoring the stable water isotopic variability in precipitation and streamflow across a snow-dominated subarctic catchment. Arctic, Antarctic, and Alpine Research 2018, 50, 1, e1454778; DOI:10.1080/15230430.2018.1454778
48. Rocher-Ros, G.; Sponseller, R.A.; Lidberg, W.; Mörth, C.-M.; Giesler, R. Landscape process domains drive patterns of CO2 evasion from river networks . Limnol Oceanogr Lett, 2019, 4, 87-95; DOI:10.1002/lol2.10108
49. Borgelt, J. Terrestrial respiration across tundra vegetation types - Implications for arctic carbon modelling. Master Degree report of Umeå University. 2017.
50. Troll, C. High mountain belts between the polar caps and the equator: their definition and lower limit. Arctic and Alpine Research 1973, 5, 3, A19-A27, https://www.jstor.org/stable/1550149 (accessed 7 July 2023)
51. Rundqvist, S.; Hedenås, H.; Sandström, A.; *Emanuelsson, U.; Eriksson, H.;* Jonasson, C.; Callaghan, T.V. Tree and Shrub Expansion Over the Past 34 Years at the Tree-Line Near Abisko, Sweden. *AMBIO* 2011, 40, 683–692; DOI:10.1007/s13280-011-0174-0

52. Rees, W.G.; Hofgaard, A.; Boudreau, S.; Cairns, D.M.; Harper, K.; Mamet, S.; Mathisen, I.; Swirad, Z.; Tutubalina, O. Is subarctic for-est advance able to keep pace with climate change? Glob. Change Biol. 2020, 26, 3965–3977; DOI:10.1111/gcb.15113
53. Berner, L.T.; Goetz, S.J. Satellite observations document trends consistent with a boreal forest biome shift. Glob. Change Biol. 2022, 28, 3275-3292. DOI: 10.1111/gcb.16121.