Reply to Reviewer #1

Interactive comments on "PeRL: A Circum-Arctic Permafrost Region Pond and Lake Database" by Sina Muster et al.

We sincerely thank the reviewer for his/her positive and constructive comments on our manuscript.

Our responses to the reviewer's comments are highlighted in bold.

Changes done in the manuscript are marked in italic.

Page and line numbers refer to the marked-up manuscript version which is attached to this review as well as the new supplement.

Anonymous Referee #1

This manuscript describes a database (PeRL) that maps ponds and lakes for the Circum-Arctic Permafrost region. The authors provide detailed information on the generation of the database, how images were processed and classified. Additionally, they discuss classification accuracy and uncertainty as well as the potential use of the database.

This database is a very valuable contribution to the scientific community as well as potential stakeholders from outside science. As the authors describe, classifying the area of waterbodies in the Arctic is crucial, as it impacts permafrost degradation and carbon fluxes of permafrost lowlands. The introduction is very well written and delivers the right message on why this database is useful. The second chapter defines ponds and lakes, which I thought was really useful, since I usually don't think about lakes and ponds, and chapter three explained the study areas. Up to this point I could easily follow the manuscript and I don't have any major comments. Chapter four deals with the generation of the PeRL database and because image processing or classification are not my expertise I cannot comment much on the correctness of the methods. However, I would like to comment on the understandability of this section of the manuscript. I got completely lost in all the technical details and had a hard time staying focused. I understand that technical details have to be provided but I would suggest to refine the language to the extent that a non-expert can understand it. Ultimately, you want this database to be useful to a large audience and the user of this database may not be the expert for all the technical details.

We agree with the reviewer that there is a lot of technical detail in the manuscript that makes it at times hard to follow. We have therefore restructured the manuscript by moving some of the more technical descriptions and information into a supplement. We have also revised the overall language to address some technical issues in simpler language or added explanations for more clarity.

Parts of the manuscript that were moved to the supplement include

- **Processing of TSX imagery**
- **Selection of box size for subgrid sampling**
- **Description of original permafrost landscape maps**
- **Metadata on image processing and classification**

In addition, we have also adapted the tables with metadata information. Some information is immediately relevant to all users such as accuracy, available classes (open water, overgrown water), map extent, acquisition date and whether the data includes only waterbodies or also streams and rivers. We have therefore rearranged this information in a new table which is shown in the main part of the manuscript (Table 1). We find the paper is now easier to read for the average user; however all the supplement provides detailed technical information is available for the interested expert and ensures a full characterization of the image and data processing.

Besides having a hard time with the technical part of this manuscript, I could not follow what the benefits of the statistical analyses for all waterbodies are or even why they were performed. Was that explained somewhere and I missed it?

We have revised section 4 to better explain the motive for the statistical analysis [page 7, lines 1-17]

Statistics such as areal fraction of water or average waterbody surface area are meaningful measures to compare waterbody distributions between individual study areas and permafrost landscapes. **Statistics were calculated for all waterbodies, as well as separately for ponds and lakes. We calculated areal fraction, i.e., the area fraction of water relative to land (the total mapped area), waterbody density, i.e. the number of waterbodies per km^2, and mean, median, and standard deviation of surface area for each site using the software package R version 3.3.1.** *However, statistics are subject to the size of the study area. Very small study areas may not capture larger waterbodies which may nonetheless be characteristic of the larger landscape. Very large study areas, on the other hand, may show more spatial variation in waterbody distribution than smaller study areas. In order to make statistics comparable between study areas, we subdivided larger study areas into boxes of 10x10 km.*

Then, section 4.6.1-4.6.3 explains permafrost landscape maps for Alaska, Canada, and Russia (Table 2) but I couldn't find details for Europe. Are there none available? There are a few study areas in Scandinavia (Figure. 1) and so I would have assumed to read something about Europe as well. This is probably an easy fix but it should be addressed.

We have added an explanation [page 9, lines 23-24]

Extrapolations were done in Alaska, Canada and Russia for waterbody maps with a (combined) extent of 100 km^2 or larger, but not for Europe where available waterbody maps were too small.

One questions that I had that I couldn't quite figure out from the manuscript is whether wetlands can falsely be classified as ponds or lakes, depending on precipitation and other factors, some areas might be really wet in some years and could get misinterpreted as ponds although they

aren't real ponds. There is some text about that on page 18, line 13ff but it would be good to address this in the manuscript in more detail.

This is an important question that we now discuss in more detail [page 21, lines 16-24]

Seasonal processes, such as snowmelt, progressing thaw depth, evaporation, and precipitation do affect the extent of surface water. Waterbody maps therefore reflect the local water balance at the time of image acquisition. Seasonal reduction of surface water extent, however, is largest in the first 2 weeks following snowmelt (Bowling et al. 2003). All PeRL maps date from the late summer season so that snowmelt and the early summer season are excluded. Changes of water extent in late summer are primarily due to evaporation and precipitation. In a study area on the Barrow Peninsula, Alaska, we find that the open water extent varies between 6 % and 8 % between the beginning and end of August of different years. However, the effect is hard to quantify as other factors such as spectral properties and resolution also impact classifications of different times at the same site. Seasonal variations may be larger in case of heavy rain events right before image acquisition but ultimately depend on local conditions which control surface and subsurface runoff.

Some small comments: Table 1: why do you abbreviate Permafrost depth as PT? That doesn't make any sense and is certainly not intuitive.

The abbreviation was corrected to "PD".

Duplicate paragraph: 4.5 Accuracy assessment of classification. This paragraph is a duplication of what is written a page earlier, p. 8, lines19-26.

The duplicated paragraphs were deleted and their information instead incorporated into Table 1 and section *"5.1 Site-level waterbody maps"* [page 13, lines 16-20] and into section *"6.1 Classification accuracy and variability"* [page 21,22]

PeRL: A Circum-Arctic Permafrost Region Pond and Lake Database

Sina Muster[1], Kurt Roth[2], Moritz Langer[3], Stephan Lange[1], Fabio Cresto Aleina[4], Annett Bartsch[5], Anne Morgenstern[1], Guido Grosse[1], Benjamin Jones[6], A. Britta K. Sannel[7], Ylva Sjöberg[7], Frank Günther[1], Christian Andresen[8], Alexandra Veremeeva[9], Prajna R. Lindgren[10], Frédéric Bouchard[11,13], Mark J. Lara[12], Daniel Fortier[13], Simon Charbonneau[13], Tarmo A. Virtanen[14], Gustaf Hugelius[7], Juri Palmtag[7], Matthias B. Siewert[7], William J. Riley[15], Charles D. Koven[15], and Julia Boike[1]

[1] Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Telegrafenberg A43, 14473 Potsdam, Germany
[2] Institute for Environmental Physics, Heidelberg University , Germany
[3] Humboldt University, Berlin, Germany
[4] Max Planck Institute for Meteorology, Hamburg, Germany
[5] Zentralanstalt für Meteorologie and Geodynamik, Vienna , Austria
[6] U.S. Geological Survey - Alaska Science Center, Anchorage, AK 99508, USA
[7] Stockholm University, Department of Physical Geography and the Bolin Centre for Climate Research, 10691 Stockholm, Sweden
[8] Los Alamos National Laboratory, Los Alamos, NM, USA
[9] Institute of Physicochemical and Biological Problems in Soil Science, Russian Academy of Sciences, Pushchino, Russia
[10] Geophysical Institute, University of Alaska Fairbanks, Fairbanks, AK, USA
[11] Institut national de la recherche scientifique (INRS), Centre Eau Terre Environnement (ETE), Québec QC, G1K 9A9, Canada
[12] Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[13] Geography Department, University of Montréal, Montréal QC, H3C 3J7, Canada
[14] Department of Environmental Sciences, University of Helsinki, Finland
[15] Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, USA

Correspondence to: Sina Muster (sina.muster@awi.de)

Abstract. Ponds and lakes are abundant in Arctic permafrost lowlands. They play an important role in Arctic wetland ecosystems by regulating carbon, water, and energy fluxes and providing freshwater habitats. However, ponds, i.e. waterbodies with surface areas smaller than 1.0E+04 m², have not been inventoried at global and regional scales. The Permafrost Region Pond and Lake Database (PeRL) presents the results of a circum-arctic effort to map ponds and lakes from modern (2002-2013) high-resolution aerial and satellite imagery with a resolution of 5 m or better. that resolve waterbodies with a surface area between 1.0E+02 m² and 1.0E+06 m². The database also includes historical imagery from 1948 to 1965 with a resolution of 6 m or better. PeRL includes 69 maps covering a wide range of environmental conditions from tundra to boreal regions and from continuous to discontinuous permafrost zones. Waterbody maps are linked to regional permafrost landscape maps which provide information on permafrost extent, ground ice volume, geology and lithology. This paper describes waterbody classification and accuracy, and presents statistics of waterbody distribution for each site. Maps of permafrost landscapes in Alaska, Canada and Russia are used to extrapolate waterbody statistics from the site level to regional landscape units. PeRL presents pond and lake estimates for a total area of 1.4E+06 km² across the Arctic, about 17 % of the Arctic lowland (<300 m a.s.l.) land surface area. PeRL waterbodies with

sizes of 1.0E+06 m² down to 1.0E+02 m² contributed up to 21% to the total water fraction. Waterbody density ranged from 1.0E+00 per km² to 9.4E+01 per km². Ponds are the dominant waterbody type by number in all landscapes with 45 % to 99 % of the total waterbody number. The implementation of PeRL size distributions into land surface models will greatly improve the investigation and projection of surface inundation and carbon fluxes in permafrost lowlands. Waterbody maps, study area boundaries and maps of regional permafrost landscapes including ~~link to~~ detailed metadata are available at https://doi.pangaea.de/10.1594/PANGAEA.868349.

1 Introduction

Globally, Arctic lowlands underlain by permafrost have both the highest number and area fraction of waterbodies (Lehner and Döll, 2004;Grosse et al., 2013;Verpoorter et al., 2014). ~~They~~ These landscapes play a key role~~s~~ as a freshwater resource, as habitat for wildlife, and as part of the water, carbon, and energy cycles (Rautio et al., 2011;CAFF, 2013).

The rapid warming of the Arctic affects the distribution of surface and subsurface water due to permafrost degradation and increased evapotranspiration (Hinzman et al., 2013). Remote sensing studies have found both increasing and decreasing trends in surface water extent for ~~lakes~~ waterbodies in permafrost regions across broad spatial and temporal scales (eg Carroll et al. (2011);Watts et al. (2012);Boike et al. (2016);Kravtsova and Rodionova (2016)). These studies, however, are limited in their assessment of changes in surface inundation since they only include lakes, *i.e.*, waterbodies with a surface area of 1.0E+04 m² or larger. Ponds with a surface area smaller than 1.0E+04 m², on the other hand, have not yet been inventoried on the global scale. Yet ponds dominate the total number of waterbodies in Arctic lowlands, accounting for up to 95% of individual waterbodies, and may contribute up to 30% ~~of~~ to the total water surface area (Muster et al., 2012;Muster, 2013). Arctic ponds are characterized by intense biogeophysical and biogeochemical processes. They have been identified as a large source of carbon fluxes compared to the surrounding terrestrial environment (Rautio et al., 2011;Laurion et al., 2010;Abnizova et al., 2012;Langer et al., 2015;Wik et al., 2016;Bouchard et al., 2015). Due to their small surface areas and shallow depths, ponds are especially prone to change; various studies reported ponds to dry out or to increase in abundance due to new thermokarst or drainage of large lakes (Jones et al., 2011;Andresen and Lougheed, 2015;Liljedahl et al., 2016). Such changes in surface inundation may significantly alter regional water, energy, and carbon fluxes (Watts et al., 2014;Lara et al., 2015). Both monitoring and modeling of pond and lake development are therefore crucial to better understand the trajectories of Arctic land cover dynamics in relation to climate and environmental change. Currently, however, the direction and magnitude of these changes remain uncertain, mainly due to the limited extent of high-resolution studies and the lack of pond representation in global databases. Although recent efforts have produced global land cover maps with resolutions of 30 m (Liao et al., 2014;Verpoorter et al., 2014;Feng et al., 2015;Paltan et al., 2015), these data sets only include lakes.

To complement previous approaches, we present the Permafrost Region Pond and Lake Database (PeRL), a circum-arctic effort t~~hat compiles o~~ 69 map~~s~~ of ponds and lakes from remote sensing data with high spatial resolution (of ≤ 6 m) (Fig. 1). This database fills the gap in available global databases that have cutoffs in waterbody surface area at 1.0E+04 m² or above. In addition, we link PeRL waterbody maps with existing maps

of permafrost landscapes to extrapolate waterbody distributions from the individual study areas to larger landscapes units. Permafrost landscapes are terrain units characterized by distinct properties such as climate, surficial geology, parent material, permafrost extent, ground ice content, and topography. These properties have been identified as major factors in the evolution and distribution of northern waterbodies (Smith et al., 2007;Grosse et al., 2013;Veremeeva and Glushkova, 2016).

The core objectives of the PeRL database are to (i) archive and disseminate fine-resolution geospatial data of northern high latitude waterbodies ~~(≤ 6 m spatial resolution)~~, (ii) quantify the intra- and interregional variability in waterbody size distributions, and (iii) provide regional key statistics, including the uncertainty in waterbody distributions, that can be used to benchmark site-, regional-, and global-scale land models.

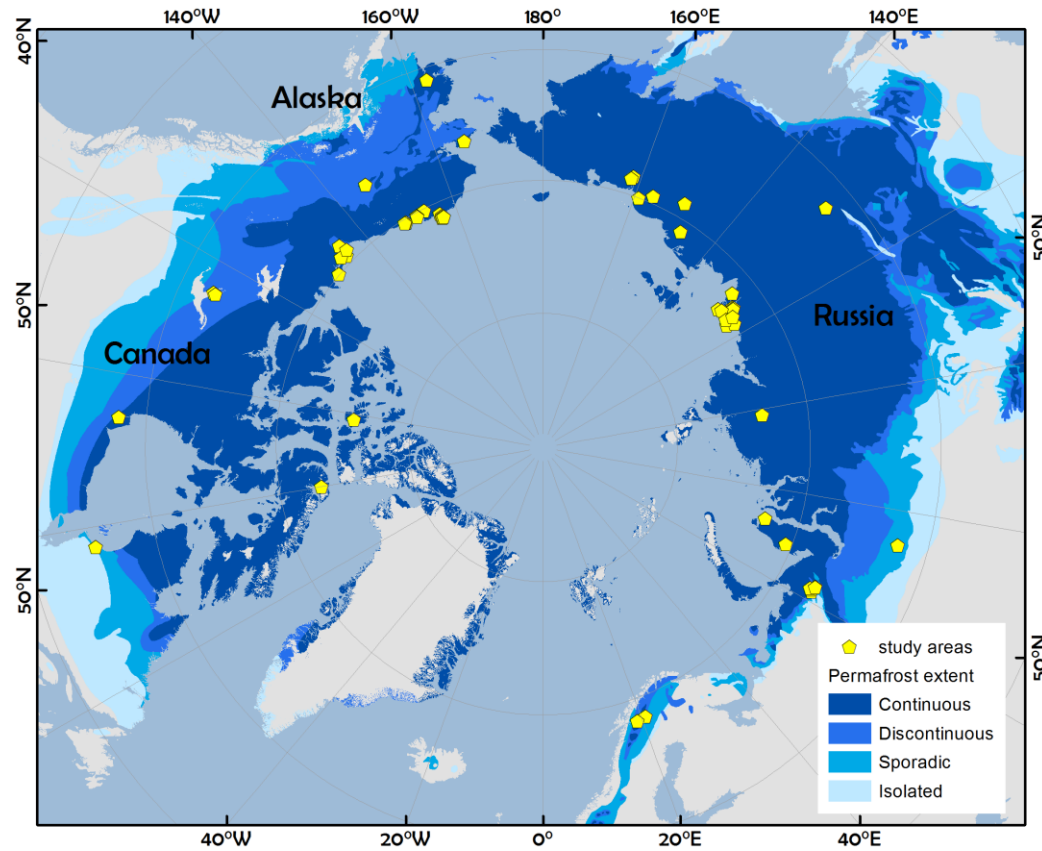



Fig. 1: Distribution of PeRL study areas. Permafrost extent according to Brown et al. (1998).

2 Definition of ponds and lakes

The definition of ponds and lakes varies in the literature and depends on the chosen scale and goal of studies when characterizing waterbodies. The Ramsar classification scheme defines ponds as permanently inundated basins smaller than 8.0E+04 m² in surface area (Ramsar Convention Secretariat, 2010). Studies have also used surface areas smaller than 5.0E+04 m² (Labrecque et al., 2009) or 1.0E+04 m² (Rautio et al., 2011) to distinguish ponds from lakes.

In remote sensing studies, surface area is the most reliably inferred parameter related to waterbody properties. Physical and biogeochemical processes of waterbodies, however, also ~~depend~~ strongly depend on waterbody depth. Differences in thermodynamics are associated with water depth, where deeper lakes may develop a stratified water column while shallow ponds remain well mixed. In high latitudes, waterbodies with depths greater than 2 m are likely to remain unfrozen at the bottom throughout winter, with ~~implications for~~providing overwintering habitat ~~availability~~ for fish and other aquatic species. In permafrost regions, a continuously unfrozen layer (talik) may develop underneath such deeper ~~ponds~~waterbodies, which strongly affects carbon cycling in these sediments (Schuur et al., 2008). Several studies have shown a positive correlation between waterbody surface area and depth (Langer et al., 2015;Wik et al., 2016). However, there is large variability in the area-depth relationship, i.e., there are large but shallow lakes that freeze to the bottom and small but deep ponds that develop a talik, and these characteristics may also change over time with changes in water level and basin~~s~~ morphology. In this study we distinguish ponds and lakes based on their surface area. We adopt the distinction of Rautio et al. (2011) and define ponds as bodies of largely standing water with surface area smaller than 1.0E+04 m² and lakes as waterbodies with surface area of 1.0E+04 m² or larger.

~~**3 Study areas**~~

~~PeRL study areas are widely distributed throughout Arctic lowlands in Alaska, Canada, Europe, and Russia and cover a latitudinal geographic gradient of 20° (55.3°N to 75.7°N), including tundra to boreal regions, continuous to discontinuous permafrost zones (Fig. 1). Mean annual temperature ranges from 0 °C to -20 °C and average annual precipitation ranges from 97 mm to 650 mm (Table 1).~~

4̶3 PeRL Database Generation

4̶3.1 Data sources and processing

PeRL's goal is to make high-resolution waterbody maps available to a large science community. PeRL compiles both previously published and unpublished fine-scale waterbody maps. Maps were included if they met the resolution criteria of 5m or less for modern imagery and 6 m for historical imagery. Historical imagery was included to enable high-resolution change detection of ponds and lakes.

~~PeRL compiles both previously published and unpublished fine-scale waterbody maps.~~ Twenty-nine maps were specifically produced for PeRL to complement the published maps in order to represent a broad range of landscape types with regard to permafrost extent, ground ice content, geology and ecozone. All waterbody maps were derived from optical or radar airborne or satellite imagery ~~with spatial resolutions of ≤ 6m~~ that were acquired between mid and late summer (July−September), thereby excluding the snow melt and early summer season. Modern imagery dates from 2002−2013 and historical imagery dates from 1948−1965. ~~Historical imagery allows high-resolution change detection of ponds and lakes.~~

Previously published maps ~~were~~ are the product of many independent studies ~~originally created for a variety of purposes and from various experts~~ which leads to a broad range of image types and classification methods used. For details on image processing and classification procedures for already published maps (n=31) we refer to the supplement (Supplementary ~~T~~table S~~B~~1 ~~and B2~~) and the respective publications. ~~Twenty-eight maps were specifically produced for~~ The processing of new PeRL maps ~~PeRL for which we provide the methods in the following~~ is described in section ~~3~~4.2. ~~Sections 4.3 to 4.6 then describe methods applicable to all waterbody maps.~~

~~3~~4.2 PeRL image processing and classification

~~4~~3.2.1 ~~TerraSAR-X imagery~~Image processing

Available high-resolution imagery used for PeRL map production included optical aerial and satellite imagery (GeoEye, QuickBird, WorldView-1 and -2, and KOMPSAT-2) and radar (TerraSAR-X) imagery.

Most optical imagery provided a near-infrared band that was used for classification with the exception of WorldView 1 that only has a panchromatic band. Preprocessing of the optical imagery involved georeferencing or orthorectification depending on the availability of high resolution digital elevation models (Supplementary Table S1).

~~TerraSAR-X imagery was available for sites in the Mackenzie Delta, Tuktoyaktuk Coastlands, and Tanzin Uplands in Canada; the Beaufort Coastal Plain and the Yukon Delta in Alaska; and the West Siberian Lowlands and Yamal Peninsula in Russia.~~ TerraSAR-X (TSX) imagery was acquired in Stripmap Mode with an HH polarization as the geocoded Enhanced Ellipsoid Corrected (EEC) product or as the Single Look Slant Range Complex (SSC) product which was then processed to EEC (Supplementary text S1). TSX images were filtered in ENVI v4.7 (ITTVIS) in order to reduce image noise: a lee filter with a 3x3 pixel window followed by a gamma filter with a 7x7 or 11x11 window depending on the image quality (Klonus and Ehlers, 2008).

~~Geocoded EEC products obtained from DLR are delivered in radar brightness. They are projected to the best available DEM, *i.e.*, SRTM X band DEMs (30 m resolution) and SRTM C band DEMs (90 m resolution). For the remaining areas, the 1 km resolution GLOBE30 DEM is used. The EEC is a detected multi-look product with reduced speckle and approximately square cells on ground. The slant range resolution of the image is 1.2 m, which corresponds to 3.3–3.5 m projected on the ground for incidence angles between 45° and 20° and an azimuth resolution of 3.3 m (Eineder et al., 2008). SSC were geocoded to the DUE Permafrost DEM (Santoro and Cartus, 2010) and no multi-looking was applied. TSX data were despeckled in ENVI 4.7 using a lee filter with a 3x3 pixel window followed by a gamma filter with a 7x7 or 11x11 window depending on the image quality (Klonus et al., 2008).~~

~~4.2.2 Optical imagery~~

~~Optical imagery used for PeRL map production included aerial imagery, GeoEye, QuickBird, WorldView 1 and -2, and KOMPSAT 2 imagery. Pixel resolutions ranged from less than 1 m to 4 m. Most imagery provided a near-infrared band that was used for classification with the exception of WorldView 1 that only has a panchromatic band. Preprocessing of the imagery involved georeferencing or orthorectification depending on availability of high resolution digital elevation models (Table B2).~~

~~4~~3.2.~~3~~2 Classification of open water

Imagery was ~~Both TerraSAR-X and optical imagery were~~ classified using either a density slice or an unsupervised k-means classification in ENVI v4.8 (ITTVIS). The panchromatic, the near-infrared and the X-Band (HH-polarization) show a sharp

contrast between open water and surrounding vegetation. Visual inspection of the imagery could therefore be used to determine individual thresholds values (in case of density slice) or to assign classes (in case of k-means unsupervised classification) for the extraction of open water surfaces. Threshold values and class boundaries varied between images and sites due to differences int illumination, and acquisition geometry and different sensor spectroradiometryradiometric properties of images. Detailed information on remote sensing imagery and classification procedure for each site is listed in the supplement (Supplementary Ttable BS1 and B2).

4.3 Post-processing of waterbody maps

The classification procedure in ENVI produces raster images that were converted to ESRI vector files Classification results from raster image processing were converted to ESRI vector files in ENVI v4.8 (ITTVIS) so that each waterbody is represented as a single polygon. Vector files were then manually processed in ArcGIS v10.2 (ESRI) to fill gaps in waterbody surfaces, remove streams, rivers, shadows due to clouds or topography and partial lakes along the study area boundaries. Minimum waterbody size was set to at least 4 pixels. This equals less than 4 m² for the highest resolutions of less than 1 m and 64 m² for the lowest resolution of 4 m for modern imagery (1.4E+02 m² for historical imagery with resolutions of 6 m). All classified objects smaller than the minimum size were removed. Partial lakes along the study area boundaries, segments of streams and rivers, and shadows due to clouds or topography were manually removed.

Commission errors occurred where high reflectance from lake bottoms, turbid waters, shadows or waves on the water surface resulted in higher digital number (DN) values and therefore a misclassification of water as land. Many of these commission errors result in gaps or holes in the lake surface which were removed with a gap-filling procedure in ArcGIS. Partial lakes within the study area were edited manually. Partial lakes along the study area boundaries, segments of streams and rivers, and shadows due to clouds or topography were removed. Minimum waterbody size was set to at least 4 pixels and all classified objects smaller than the minimum size were removed.

Three study areas in the Hudson Bay Lowlands (Canada), Lapland (Sweden), and the Usa River basin (Russia), feature multi-temporal but very small waterbody maps with extents less than 4 km² and less than 300 waterbodies. They also contain partial waterbodies along the study area boundaries. Partial lakes were not removed in order to retain the maximum information about waterbody boundaries for change detection analysis. Very large waterbody maps in the Lena Delta and Yakutsk region include river and stream segments that were not manually corrected. Compared to manually refined classifications in the same area, these RapidEye classifications showed a higher estimation of pond number by 35% and of pond surface area by 21%. Information about whether a map includes partial lakes or misclassification areas is given in the metadata (Table B1).

4.43.3   Study area boundaries

Each waterbody map is associated with a vector file that delineates the study area's boundary. Boundaries were calculated for each map – whether new or previously published - in ArcGIS by first producing a positive buffer of 1 km to 3 km around each waterbody in the map and merging the individual buffers into one single polygon. From that single buffer polygon we then subtracted again the same distance which rendered the study area boundary. The area of the boundary is referred to as the total mapped area extent of that site (Table 1). For sites with multi-temporal data, the total mapped area extent of the oldest classification was chosen as a reference in order to calculate changes in pond and lake statistics over time.

4.6 PeRL statistical analysis

Statistics such as areal fraction of water or average waterbody surface area are meaningful measures to compare waterbody distributions between individual study areas and permafrost landscapes. Statistics were calculated for all waterbodies, as well as separately for ponds and lakes. We calculated areal fraction, i.e., the area fraction of water relative to land (the total mapped area), and waterbody density, i.e. the number of ~~individual~~ waterbodies per km² ~~, and mean, median, and standard deviation of surface area~~ for each site using the software package R version 3.3.1. However, statistics are subject to the size of the study area. Very small study areas may not capture larger waterbodies which may nonetheless be characteristic of the larger landscape. Very large study areas, on the other hand, may show more spatial variation in waterbody distribution than smaller study areas. In order to make statistics comparable between study areas, we subdivided larger study areas into boxes of 10x10 km. Box size was chosen as a function of the standard error (Supplementary text S2). We calculated the statistics for each box and then averaged the statistics across all boxes within each study area. This subgrid analysis was conducted for all study areas larger than 300 km² for which at least 4 boxes could be sampled.

Statistics are also subject to image resolution which defines the minimum object size that can be confidently mapped. For all modern imagery, the minimum waterbody size included in the calculation of statistics was therefore set to 1.0E+02 m² (1.4E+02m² for historical imagery). Very large lakes are not representative of all study areas and may only be partially mapped within a 10x10 km box. We therefore chose a maximum waterbody size of 1.0E+06 m² to be included in the calculation of statistics.

~~The representativity of these parameters is subject to the study area size, its location within a landscape, and the image resolution. Image resolution defines the minimum object size that can be confidently mapped, whereas study area size determines if very large lakes can be representatively captured. For statistical comparison, we therefore included only waterbodies with sizes of 1.0E+06 down to 1.0E+02 m². Additionally, imagery may capture a certain variation of the landscape's waterbody distribution which may not be representative of the larger landscape. In order to determine a representative total mapping area we compared the variability of probability density functions (PDF) and distribution parameters within three study areas in Russia, Canada, and Alaska. In each study area, we performed a subgrid analysis, i.e., we selected waterbodies from a minimum of 5 and up to 50 randomly distributed boxes with varying sizes of 5x5 km, 10x10 km, and 20x20 km. We calculated the standard error of the mean of all statistical parameters across all boxes of the same size. Relative error of density (waterbody number per km²) and waterbody mean surface area was lowest for 10x10 km boxes. Relative error increased with 20x20 km boxes which is probably due to the significantly lower number of boxes that can be sampled from the study areas. Only 12 PeRL sites have a study area larger than 1000 km² that would allow us to sample a minimum of 5 boxes of 20x20 km in size. A box size of 10x10 km allows the subsampling of 26 sites with a minimum of 5 boxes. Taking into account the overall variability of distributions and the possible number of subgrid samples, a box size of 10x10 km was chosen for subgrid analysis. Subgrid analysis was conducted for study areas larger than 300 km² and the mean of each statistical parameter along with the relative error across the 10x10 km subsets is reported.~~

4 Extrapolation of site-level waterbody statistics to permafrost landscapes

~~4.6 Regional maps of pond and lake distributions~~

~~Regional maps of permafrost landscapes were used to extrapolate waterbody maps for lowlands with elevations lower than~~

~~300 m a.s.l..~~

4.1 Regional maps of permafrost landscapes

Waterbody statistics of each site were extrapolated to permafrost landscapes based on the assumption that distributions of ponds and lakes are similar for similar permafrost landscapes, i.e. areas with similar properties regarding ~~We define permafrost landscapes as a unique combination of~~ climate, geology, lithology (soil texture), permafrost extent and ground ice volume. Vector maps of ~~these landscape properties~~ permafrost landscapes (PLM) are available on the regional level: the Alaskan map of permafrost characteristics (AK2008) (Jorgenson et al., 2008), the National Ecological Framework for Canada (NEF) (Marshall et al., 1999), and the Land Resources of Russia (LRR) (Stolbovoi and McCallum, 2002). Despite differences in mapping approaches and terminology, the~~se~~ databases report similar landscape characteristics at comparable scales. All regional maps were available as vector files which were converted to a common North Pole Lambert Azimuthal Equal-Area (NPLAEA) projection. All ~~PL maps~~PLM were clipped in ArcGIS v10.4 with a lowland mask including only areas with elevations of 300 m or lower. The lowland mask was derived for the whole Arctic using the digital elevation model GTOPO30 (USGS). ~~Vector files of permafrost landscapes are available for Alaska, Canada and Russia.~~ Details on the properties of each PLM are provided in the supplement (Supplementary Text S3, S4, and S5 and supplementary tables S2, S3, and S4). The original PLM were merged in ArcGIS to produce a unified circum-arctic vector file and map representation. Landscape attributes that were retained from the original PLM were ecozone, permafrost extent, ground ice volume, surficial geology, and lithology. Variable names were consolidated using uniform variable names (Supplementary table S5).

~~Processing of each map is described in detail in the sections 4.6.1, 4.6.2 and 4.6.3. Attributes of all regional and circum-arctic vector files of permafrost landscapes are described in Tables C1, C2, C3, and C4.~~

~~4.6.1 Alaskan permafrost landscape maps~~

~~The PL map of Alaska reports surficial geology, MAAT, primary soil texture, permafrost extent, ground ice volume, and primary thermokarst landforms (Jorgensen et al., 2008). A rule-based model was used to incorporate MAAT and surficial geology. Permafrost characteristics were assigned to each surficial deposit under varying temperatures using terrain-permafrost relationships and expert knowledge (Jorgensen et al., 2008).~~

~~4.6.2 Canadian permafrost landscape maps~~

~~Permafrost landscapes of Canada are described in the National Ecological Framework (NEF). The NEF distinguishes four levels of generalization nested within each other. Ecozones represent the largest and most generalized units followed by ecoprovinces, ecoregions, and ecodistricts. Ecodistricts were delineated based mainly on differences in parent material, topography, landform and soil development derived from the Soil Landscapes of Canada (Soil Landscapes of Canada Working Group, 2010) at a map scale of 1:3,000,000 to 1:1,000,000 (Ecological Stratification Working Group, 1995; Marshall, 1999) whereas ecoregions and ecoprovinces are generalized based mainly on climate, physiography, and vegetation. Ecodistricts were therefore chosen as most appropriate to delineate permafrost landscapes. NEF reports the areal fraction of the underlying soil landscape units and attributes nested within each ecodistrict. The dominant fraction of surficial geology, lithology, permafrost extent and ground ice volume was chosen to describe each ecodistrict. Ecodistricts with the same permafrost landscape type within the same ecozone were then merged to PL units.~~

4.6.3 Russian permafrost landscape characterization

In Russia, information about permafrost extent, ground ice content, generalized geology and lithology was available only as individual vector maps. The individual maps were combined in ArcGIS 10.4 to delineate Russian PL units similar to the Canadian and Alaskan databases. Russian ecozones were mapped using the global scale map by Olson et al. (2001) which conforms to the Alaskan and Canadian ecozones. The geometric union of ecozone, ground ice content, and permafrost extent was calculated in ArcGIS 10.1 via the tool "intersect". Each unique combination of these three variables was assigned the dominant fraction of geology and lithology type.

4.6.4 Integration of regional permafrost landscapes

Table 2: Terminology for permafrost properties in the regional permafrost databases of Alaska (AK2008), Canada (NEF), and Russia (LRR).

Description	PeRL	AK2008	NEF	LRR
ecozone	ECOZONE	NA	ecozone	NA
surficial geology	GEN_GEOL	AGGRDEPOS	UNIT	PARROCK
lithology	LITHOLOGY	TEXTURE	TEXTURE	TEXTUR
permafrost extent	PF_EXTENT	PF_EXTENT	PERMAFROST	EXTENT_OF_
ground ice	GROUND_ICE	ICECLOWASS	PERMAFROST	MIN_MAX

4.6.54.2 Extrapolation of waterbody statistics to permafrost landscapes

Waterbody maps were spatially linked with their associated permafrost landscape. Maps within the same landscape were combined whereas maps spanning two or more landscapes were divided by selecting all waterbodies that intersected with the respective permafrost landscape. Generally, waterbody statistics were extrapolated for study areas with single or combined total mapped areas of 1.0E+02 km² or larger. If Sseveral maps were present within one permafrost landscape unit they were combined and average statistics were calculated across all maps in that permafrost landscape unit. Historical maps and RapidEye unedited classifications affected by large scale inclusion of rivers and streams (see section 4.5) were not included used in the extrapolation.

Extrapolations were done in Alaska, Canada and Russia for waterbody maps with a (combined) extent of 100 km² or larger, but not for Europe where available waterbody maps were too small. Maps in the Canadian High Arctic were smaller than 1.0E+02 km² but represent typical wetlands in that region and were therefore included in the extrapolation. Figures A1, A2, A3, and A4 in the appendix show the location of waterbody maps within their associated permafrost landscape.

Extrapolated values statistics were assigned two confidence classes: (1) high (1) and (2) low (2) confidence. Permafrost landscapes were assigned a high confidence if a map was present in the permafrost landscape of that ecozone. If in the A low confidence indicates that statistics were derived from the same permafrost landscape but in a different ecozone., the

same waterbody statistics were assigned with a low confidence. Due to differences in the mapping and generalization of landscape properties methodology of mapping and extrapolating permafrost and landscape characteristics of the regional PLM, the extrapolation was conducted only within each region. Figures D1, D2, D3, and D4 show the location of waterbody maps within their associated permafrost landscape.

Table 1. Characteristics of waterbody maps. A "clean" map state indicates that the map only includes ponds or lakes. A "raw" map state indicates that no manual editing was conducted and the map may contain rivers, streams, partial waterbodies or cloud shadows. References indicate whether the map has already been published or was produced specifically for PeRL.

Map ID	Map extent [km²]	Reso-lution [m]	Accuracy	Classes	Min size [m²]	State	Reference
abi0012010xxxx	38.4	1.0	>0.71	open water	30.0	clean	this paper
arg00120110829	195.9	4.0	N/A	open water	64.0	clean	this paper
arg0022009xxxx	5023.6	5.0	0.85	open water	100.0	raw	Bartsch and Seifert (2012)
arg00320110711	223.7	4.0	N/A	open water	64.0	clean	this paper
bar00119480804	19.0	0.7	N/A	open water	16.0	clean	Andresen and Lougheed (2015)
bar00120020802	19.0	0.7	N/A	open water	16.0	clean	Andresen and Lougheed (2015)
bar00120080730	19.0	0.7	N/A	open water	16.0	clean	Andresen and Lougheed (2015)
bar00120100810	19.0	0.5	0.93	open water	16.0	clean	Andresen and Lougheed (2015)
byk00120060709	170.2	2.5	N/A	open water	30.0	clean	Grosse et al. (2008)
byl00120160728	45.3	0.5	0.77	open water	16.0	clean	this paper
che00120020709	220.8	1.0	N/A	open water	30.0	clean	Grosse et al. (2008)
che00220090724	340.3	3.0	0.97	open water	36.0	clean	Widhalm et al. (2014a);Widhalm et al. (2014b)
elc00120090825	126.0	2.5	N/A	open water, water with emersed vegetation	25.0	clean	this paper
elc00220020801	143.2	0.6	0.89 to 0.95	open water	7.8	clean	Lara et al. (2015)
elc00320090802	51.4	4.0	N/A	open water	64.0	clean	Muster et al. (2013)
elc004200808xx		0.6	>0.75	troughs with open water	1.5	clean	Lara et al. (2015)
esk00120090727	923.4	2.5	N/A	open water	36.0	clean	this paper
fir00120090906	150.0	4.0	N/A	open water	64.0	clean	this paper
fir0022009xxxx	9825.7	5.0	0.85	open water	100.0	raw	Bartsch and Seifert (2012)
fis00120020715	236.8	1.3	N/A	open water	100.0	clean	Jones et al. (2013)
grp00119590707	0.2	0.4	N/A	open water	17.0	clean	Bouchard et al. (2014)
grp00120060707	0.2	0.61	N/A	open water	13.0	clean	Bouchard et al. (2014)
hbl00119540701	4.0	1.0	N/A	open water	51.0	clean with partial waterbodies	Sannel and Brown (2010);Sannel and Kuhry (2011)
hbl00119740617	4.0	1.0	N/A	open water	53.0	clean with partial waterbodies	Sannel and Brown (2010);Sannel and Kuhry (2011)
hbl00120060706	4.0	0.6	N/A	open water	55.0	clean with partial waterbodies	Sannel and Brown (2010);Sannel and Kuhry (2011)
ice0032009xxxx	788.1	5.0	0.85	open water	100.0	raw	Bartsch and Seifert (2012)
imc00120040725	1309.9	0.7	N/A	open water	100.0	clean	this paper
ind00120090907	654.0	0.5	N/A	open water	100.0	clean	this paper
kol00119650721	2382.0	5.0	N/A	open water	200.0	clean	this paper
kol00219650721	2638.1	5.0	N/A	open water	200.0	clean	this paper

Map ID	Map extent [km²]	Reso- lution [m]	Accuracy	Classes	Min size [m²]	State	Reference
kcp001201007xx	20.7	N/A	N/A	open water	1.0	clean	Walker et al. (1986);Raynolds et al. (2014)
kcp002201007xx	20.1	N/A	N/A	open water	1.0	clean	Walker et al. (1986);Raynolds et al. (2014)
kcp003201007xx	18.9	N/A	N/A	open water	1.0	clean	Walker et al. (1986);Raynolds et al. (2014)
ksl00119620628	558.8	6.0	N/A	open water	144.0	clean	this paper
ksl0012012xxxx	558.8	2.5	N/A	open water	25.0	clean	this paper
kur00120100805	55.5	1.0	N/A	open water	4.0	clean	this paper
kur00220080926	251.5	2.5	N/A	open water	187.0	clean	this paper
kyt00120070728	262.3	3.0	0.97	open water	36.0	clean	Widhalm et al. (2014a);Widhalm et al. (2014b)
log00120110811	69.7	2.4	0.90	open water	23.0	clean	Palmtag et al. (2016)
mdn00120100716	1510.3	2.5	N/A	open water	64.0	clean	this paper
mdw00120090921	1614.8	2.5	N/A	open water	36.0	clean	this paper
ole00120060708	75.1	2.5	N/A	open water	30.0	clean	Grosse et al. (2008)
pbp00120090813	68.6	2.5	N/A	open water	12.0	clean	Muster et al. (2013)
ric001201209125	587.4	2.5	N/A	open water	8.0	clean	this paper
rog00120070626	10.0	0.6	N/A	open water	30.0	clean	Sjöberg et al. (2013)
rog00219740726	3.4	1.0	N/A	open water	28.0	clean with partial waterbodies	Sannel and Kuhry (2011)
rog00220070707	3.4	0.6	N/A	open water	28.0	clean with partial waterbodies	Sannel and Kuhry (2011)
rog00320070626	59.6	2.4	0.68	open water	11.0	clean	this paper
rog00420070704	62.4	2.4	0.83	open water	29.0	clean	this paper
rog00520070704	62.6	2.4	0.83	open water	11.5	clean	this paper
sam001200808xx	1.6	0.3	N/A	open water	1.0	clean	Muster et al. (2012)
sei00120070706	6.7	0.6	N/A	open water	30.0	clean	Sjöberg et al. (2013)
sei00220070706	82.9	2.4	0.68	open water	11.5	clean	this paper
sei00320080629	91.1	3.0	N/A	open water	36.0	clean	this paper
sur00120130802	1765.6	2.0	N/A	open water	16.0	clean	this paper
tav00119630831	0.8	0.5	N/A	open water	28.0	clean with partial waterbodies	Sannel and Kuhry (2011)
tav00119750810	0.8	0.9	N/A	open water	28.0	clean with partial waterbodies	Sannel and Kuhry (2011)
tav00120030702	0.8	1.0	N/A	open water	28.0	clean with partial waterbodies	Sannel and Kuhry (2011)
tbr00120100901	694.3	2.5	N/A	open water	36.0	clean	this paper
tea00120100901	462.9	2.5	N/A	open water	36.0	clean	this paper
tuk00120120723	477.6	2.5	N/A	open water	36.0	clean	this paper
wlc00120090825	1400.2	2.5	N/A	open water, water	25.0	clean	this paper

Map ID	Map extent [km²]	Reso-lution [m]	Accuracy	Classes	Min size [m²]	State	Reference
				with emersed vegetation			
wlc00220020801	153.8	0.6	0.89 to 0.95	open water	7.8	clean	Lara et al. (2015)
wlc00320090802	297.3	4.0	N/A	open water	64.0	clean	Muster et al. (2013)
yak0012009xxxx	2035.5	5.0	0.85	open water	75.0	raw	Bartsch and Seifert (2012)
yam00120080824	1294.3	2.5	N/A	open water	36.0	clean	this paper
yam00220100820	1006.6	2.5	N/A	open water, water with emersed vegetation	100.0	clean	this paper
yfl0012011xxxx	100.0	1.0	N/A	open water	4.0	clean	this paper
yuk00120090812	1078.7	2.5	N/A	open water	36.0	clean	this paper
yuk00220090812	575.3	2.5	N/A	open water	36.0	clean	this paper

5 PeRL database features ~~and accuracy~~

The database provides two different map products: (i) site level waterbody maps and (ii) an extrapolated circum-arctic waterbody map. The database also provides different tables which present statistical parameters for each individual waterbody map (Appendix B) and aggregated statistics for PL unit in the circum-arctic map (Table 3).

5.1 Site-level waterbody maps

5.1.1 Data set structure

The database features altogether ~~69~~ 70 individual waterbody maps as ESRI shape-files. Each waterbody shape-file is named according to a map ID. The map ID consists of a three letter abbreviation of the site name, followed by a running three-digit number and the acquisition date of the base ~~imagery~~ imagery (YYYY-MM-DD). Vector files were projected to the NPLAEA projection. Area and perimeter of each waterbody and site ~~was~~ were calculated in ArcGIS 10.4 in square meters. Each vector file is accompanied by an xml-file which lists metadata about classification and references as presented in table 1 and supplementary ~~T~~tables ~~S~~B1 ~~and B2~~. Each map has a polygon associated with it that ~~describes~~ contains the study area, *i.e.*, the total land area of the waterbody map. All study area~~s~~ boundaries are stored in the file *PeRL_study_areas.shp* and can be identified via the map ID. The study area shape-file also includes the site characteristics listed in Table ~~1~~2.

Fifty-eight maps are considered "clean", i.e. they have been manually edited to include only ponds and lakes (Table 1). Eight maps are "clean with partial waterbodies". Those are multi-temporal maps with very small map extents where partial waterbodies along the study area boundary were not deleted in order to retain information for change detection analysis. Four maps were not manually edited due to their very large map extent and may include partial waterbodies, streams, rivers or shadows.

5.1.2 Spatial and environmental characteristics

PeRL study areas are widely distributed throughout Arctic lowlands in Alaska, Canada, Russia, and Europe and cover a latitudinal gradient of about 20° (55.3°N to 75.7°N), including tundra to boreal regions and located in continuous,

discontinuous and sporadic permafrost zones (Fig. 1). Mean annual temperature ranges from 0 °C to -20 °C and average annual precipitation ranges from 97 mm to 650 mm (Table 1).

Twenty-one sites are located in Alaska covering a~~n~~ total area of 7.3E+03 km². Canada has 14 sites covering 6.4E+03 km² and Russia has 30 sites covering 2.9E+03 km² in total. Four sites are located in Sweden with a total mapped area of 41 km². Individual map extents range from 0.2 km² to 9825.7 km² with a mean of 622.8 km² (Table 1). The database includes six multi-temporal classifications in the Kotzebue Sound Lowlands and on the Barrow Peninsula, Alaska (Andresen et al., 2015), on the Grande Rivière de la Baleine Plateau (Bouchard et al., 2014) and Hudson Bay Lowlands in Canada, Lapland in Sweden, and the Usa River basin in Russia (Hugelius et al.; 2011, Sannel and Kuhry, 2011). Ponds contributed about 45 to 99 % of the total number of waterbodies with a mean of 85 ± 14 % and up to 34 % to the total water surface area with a mean of 12 ± 8.3 % (Fig. 2 and ~~Table E1, E2, E3, E4~~Appendix B). Mean surface area ranged from 7.9E+01 m² with a standard deviation ~~(sd)~~ of 1.2E+01 m² on Bylot Island, Canada, to 7.1E+04 m² ± 1.5E+05 m² on Eskimo Islands, Canada (Table ~~E~~B2). Water fraction of the total mapped area ranged from about 1% to 21% for all waterbodies and from >1 % up to 6% for ponds. Waterbody density per km² ranged from 1.0E+00 per km² in the Indigirka Lowlands ~~(RUS)~~, Russia, to 9.4E+01 per km² in Olenek Channel of the~~,~~ Lena Delta (~~RUS~~Table B4).

Table 2: Climate and permafrost characteristics for each study area. Latitude (Lat) and longitude (Long) coordinates are reported in decimal degrees (WGS84). Mean annual air temperature (MAAT), total precipitation (TP), permafrost extent (PE~~,~~: C-continuous, D-discontinuous, S-sporadic~~, I isolated~~), permafrost depth (P~~D~~T). References to all data sources are listed in ~~the~~Supplementary table S7.

Map ID	Country	Site name	Lat	Long	MAAT [°C]	TP [mm]	PE	P~~D~~T [m]
abi0012010xxxx	Sweden	Abisko	68.3	19.1	0	362	S	>16
arg00120110829, arg0022009xxxx, arg00320110711	Russia	Arga Island	73.5	123.6	-13	124	C	400-600
bar00119480804, bar00120020802 bar00120080730, bar00120100810	Alaska	Barrow Peninsula	70.9	-156.2	-11	115	C	>400
byk00120060709	Russia	Bykovsky Peninsula	71.8	129.3	-13	427	C	500-600
byl00120160728	Canada	Bylot Island	73.2	-80.0	-15	190	C	>200
che00120020709, che00220090724	Russia	Cherskii	68.8	161.6	-12	294	C	400-500
elc00120090825, elc00220020801, elc00320090802	Alaska	Elson Lagoon Coast Plain	71.2	-156.4	-11	115	C	>400
esk00120090727	Canada	Eskimo Lakes	69.2	-133.3	-10	161	C	750
fir00120090906, fir0022009xxxx	Russia	First Terrace, Lena Delta	72.9	127.3	-13	124	C	400-600
fis00120020715	Alaska	Fish-Judy Creek Floodplain	70.3	-151	-10	97	C	260
grp00119590707, grp00120060707	Canada	Grande Rivière de la Baleine Plateau	55.3	-77.5	-4	650	S	10-50
hbl00119540701, hbl00119740617, hbl00120060706	Canada	Coastal Hudson Bay Lowlands	57.9	-94.2	-6	430	C	NA
ice0032009xxxx	Russia	Ice complex, Lena Delta	72.8	124.7	-13	124	C	400-600
imc00120040725	Alaska	Ikpikpuk Middle Coastal Plain	70.2	-153.3	-10	97	C	260
ind00120090907, kyt00120070728	Russia	Indigirka Lowlands, Kytalik	69.7	148.8	-14	232	C	>300

Map ID	Country	Site name	Lat	Long	MAAT [°C]	TP [mm]	PE	P~~T~~D [m]
kol00119650721	Russia	Kolyma Lowlands	70.0	159.1	-10	110	C	500-600
kol00219650721	Russia	Kolyma Lowlands	69.5	156.3	-13	144	C	500-600
kpc001201007xx, kpc002201007xx, kpc003201007xx	Alaska	Kuparuk Coastal Plain	70.3	-148.5	-10	97	C	260
ksl00119620628, ksl0012012xxxx	Alaska	Kotzebue Sound Lowlands	66.2	-165.8	-3	427	C	<50
kur00120100805, kur00220080926	Russia	Kurungnakh, Lena Delta	73.2	125.1	-13	124	C	400-600
log00120110811	Russia	Logata	73.4	98.5	-13	270	C	NA
mdn00120100716	Canada	Mackenzie Delta ~~West~~North	69.1	-135.2	-8	241	D	<100-500
mdw00120090921	Canada	Mackenzie Delta ~~North~~West	68.5	-134.7	-8	241	D	<100-500
ole00120060708	Russia	Olenek channel, Lena Delta	72.9	122.9	-15	206	C	200-600
pbp00120090813	Canada	Polar Bear Pass	75.7	-98.5	-16	161	C	>500
ric001201209125	Canada	Richards Island	69.5	-134.3	-8	241	C	>400
rog00120070626, rog00219740726, rog00220070707, rog00320070626, rog00420070704, rog00520070704	Russia	Rogovaya	62.3	62.1	-6	538	C	ca. 50
sam001200808xx	Russia	Samoylov Island	72.4	126.5	-13	124	C	400-600
sei00120070706, sei00220070706	Russia	Seida	67.1	62.9	-6	470	C	ca. 50
sur00120130802	Russia	Surgut	62.3	74.6	-17	400	S	50-300
tav00119630831, tav00119750810, tav00120030702	Sweden	Tavvavuoma	68.5	20.9	-3	451	C	<25
tbr00120100901, tea00120100901	Canada	Tanzin Upland Beaulieu River	62.7	-115.2	-4	289	D	NA
tuk00120120723	Canada	Tuktoyaktuk Peninsula	69.9	-130.4	-10	161	C	750
wlc00120090825, wlc00220020801, wlc00320090802	Alaska	Wainwright Lower Coastal Plain	70.9	-156.2	-11	115	C	>400
yak0012009xxxx	Russia	Yakutsk	62.1	130.3	-10	228	C	200-300
yam00120080824, yam00220100820	Russia	Yamal Peninsula	71.5	70.0	-6	260-400	C	100-500
yfl0012011xxxx	Alaska	Yukon Flats Basin	66.2	-145.9	-5	309	D	90
yuk00120090812, yuk00220090812	Alaska	Yukon-Kuskokwim Delta		-162	61	471	C	100-200

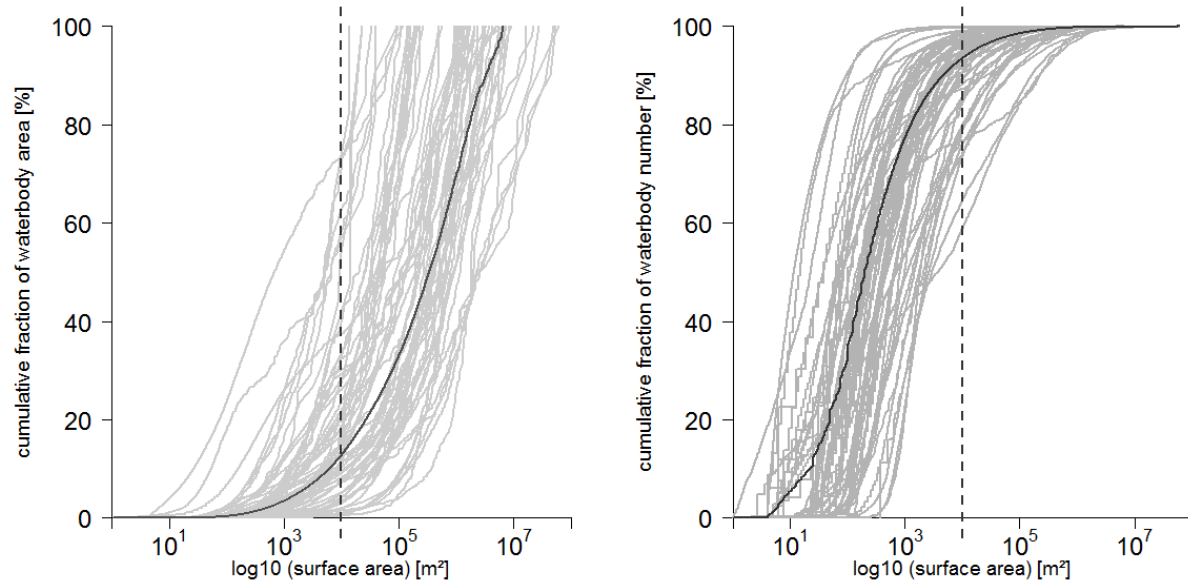



Fig. 2: Empirical cumulative distribution function of waterbody area (left hand side) and waterbody number (right hand side). Grey lines represent individual sites across all regions. Black lines represent the mean function averaged over all sites. Vertical dashed line in each panel represents the pond-lake size threshold used in this paper.

5.2 Circum-arctic waterbody map

5.2.1 Data set structure

The unified vector file *PeRL_perma_land.shp* contains the permafrost landscapes and the extrapolated waterbody statistics (Table 3). Average statistics were calculated for 10 x 10 km boxes within large maps or when four or more maps were present in the permafrost landscapes. Average statistics are reported with their relative standard error (RE). The RE is the standard error expressed as a percentage. The standard error is the standard deviation of the sampling distribution of a statistic. The permafrost landscapes are also provided as separate vector files for each region (*alaska_perma_land.shp, canada_perma_land.shp, and russia_perma_land.shp*) and contains the landscape characteristic of each permafrost landscape as individual attributes (Supplementary tables SX, SX, SX). The unified vector file (*PeRL_perma_land.shp*) and the regional files can be joined using the common PERMID.

Table 3: Attributes of ESRI shape-file *PeRL_perma_land.shp*

Field name	Description
PERMA_LAND	permafrost landscape: [permafrost extent]/[ground ice volume]/[surficial geology]/[texture]
PERMID	Each permafrost landscape in the vector file is assigned a unique ID (PERMID). The first digit stands for the region (1 – Alaska, 2 – Canada, 3 – Russia), digits 2–6 identify the single polygon, and the last three digits identify the ecozone.
AREA	area of polygon in square meters
PERIMETER	perimeter of polygon in square meters
Map_ID	Short ID of waterbody map used for extrapolation of statistics
confidence	1: high confidence, 2: low confidence
frac	areal fraction of waterbodies (1.0E+02 m² to 1E+06m² in surface area) in %
frac_re	relative standard error of areal fraction of waterbodies (1.0E+02 m² to smaller than 1E+06m² in surface area) in %
dens	density: number of waterbodies (1.0E+02 m² to 1E+06m² in surface area) per square kilometer
dens_re	Relative standard error of density of waterbodies (1.0E+02 m² to 1E+06m² in surface area) in %
frac_ponds	areal fraction of waterbodies (1.0E+02 m² to smaller than 1E+04m² in surface area) in %
frac_po_re	relative standard error of areal fraction of waterbodies (1.0E+02 m² to smaller than 1E+04m² in surface area) in %
dens_ponds	ponds density: number of ponds (1.0E+02 m² to 1E+04m² in surface area) per square kilometre
dens_po_re	relative standard error of pond density (1.0E+02 m² to smaller than 1E+04m² in surface area) in %

5.2.2 Spatial and environmental characteristics

Altogether we identified 230 different permafrost landscapes in the Russian lowlands, 160 in the Canadian lowlands, and 51 in the lowlands of Alaska. PeRL waterbody maps were located in 28 different permafrost landscapes (Table 3) covering which cover a total area of 1.4E+06 km² across the Arctic; thereof 1.0E+06 km² in Russia, 2.1E+05 km² in Canada and 1.7E+05 km² in Alaska. About 65% of the extrapolated area was classified as high confidence (Fig. 3). The Lhighest landscape average areal fraction of water surface reached maxima ofwas 21 % (Fig. 4 and Table 3) and density waterbody density per km² of 57 (Fig. 5 and Table 3).

Relative error (RE) for different subsets or maps within a permafrost landscape was about 7 % on average with a maximum of 30 % (Table 2). RE for waterbody density was 8% on average with a maximum of 50 %. Our extrapolated area (1.4E+6 km²) representsed 17.0 % of the current Arctic permafrost lowland area (below 300 m a.s.l. elevation). PeRL providsed pond and lakes estimates for about 29 % (in area) of the Alaskan lowlands, 7 % of the Canadian lowlands, and 21 % of the Russian lowlands. Together all extrapolated landscapes contributed about

7% to the current estimated Arctic permafrost area (Brown et al., 1998). In Alaska, waterbody maps were missing for permafrost landscapes with isolated permafrost (16% of total area) or rocky lithology (36% of total area). Dominant types of surficial geology that were not mapped include colluvial sites, and sites with bedrock or of glacial origin which together contribute 61% to the total area. In Canada, both isolated and sporadic permafrost were not inventoried (22% of total area), as well as areas with ground ice content of 10-20 % or less (23% of total area). Six of the 19 geology classes were inventoried which contribute 75% to the total area. Six of seven lithology types with an areal coverage of 90 % were represented. In Russia, waterbody maps were not available in the discontinuous permafrost zone (13 % of the total area). No maps were present in regions with the geological type "deluvial-coluvial and creep" which accounts for 28 % of the total area.

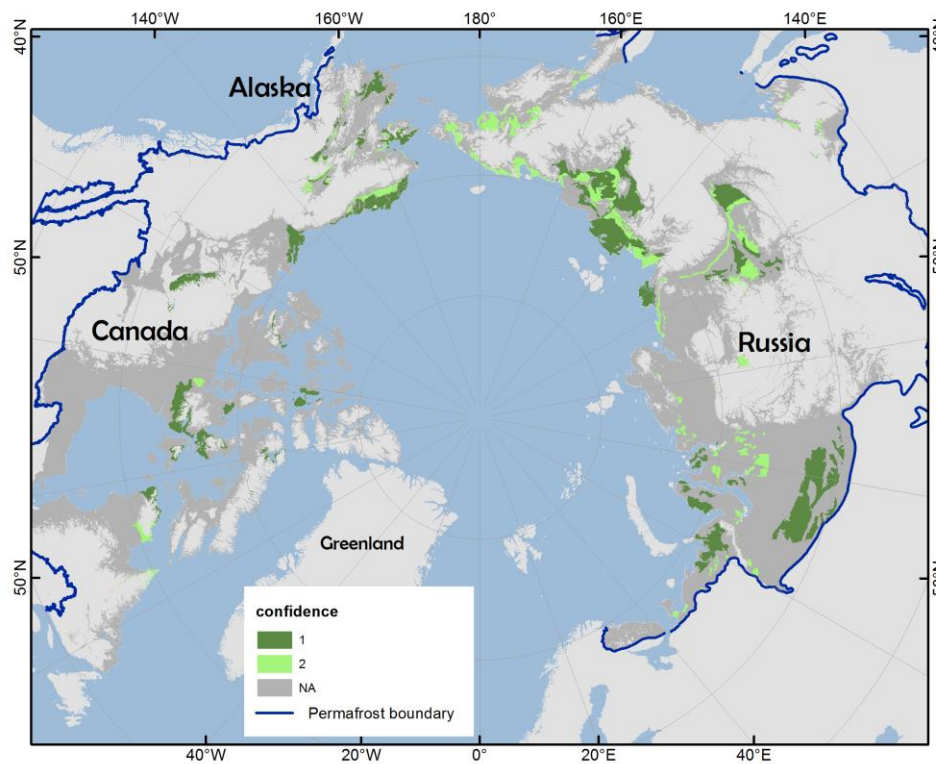



Fig 3: Confidence for permafrost lowland landscapes . Confidence class 1 (high confidence) designates permafrost landscapes where waterbody maps are available in lowland areas. Confidence class 2 (low confidence) represents permafrost landscapes with extrapolated waterbody statistics. No value (dark grey) areas indicate that no maps were available in these permafrost landscapes. Light grey areas indicate terrain with elevations (GTOPO 30, USGS) higher than 300 m a.s.l. which were not considered in the extrapolation. Permafrost boundary was derived from the regional databases.

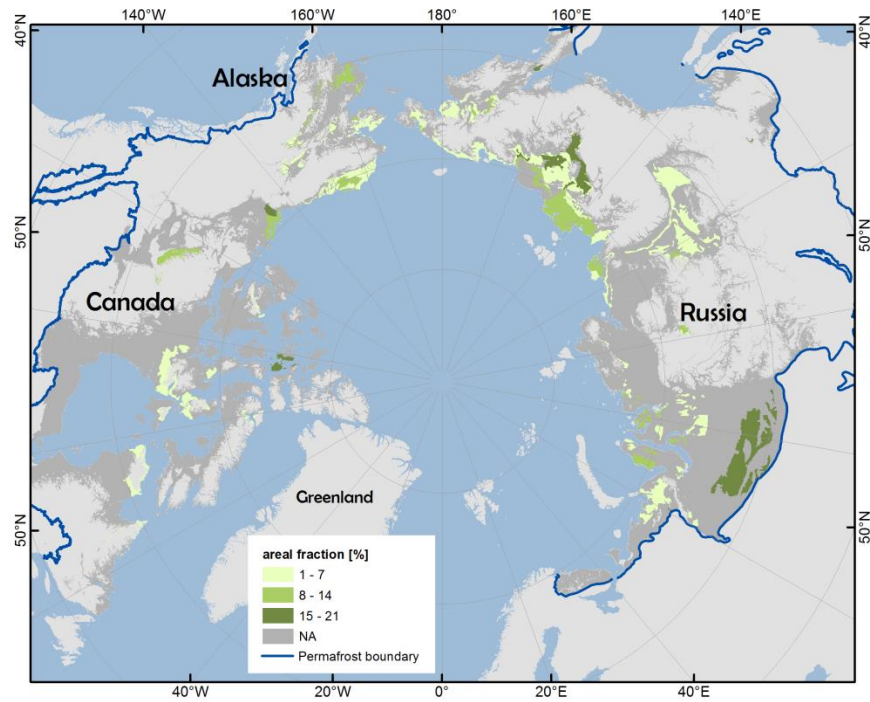



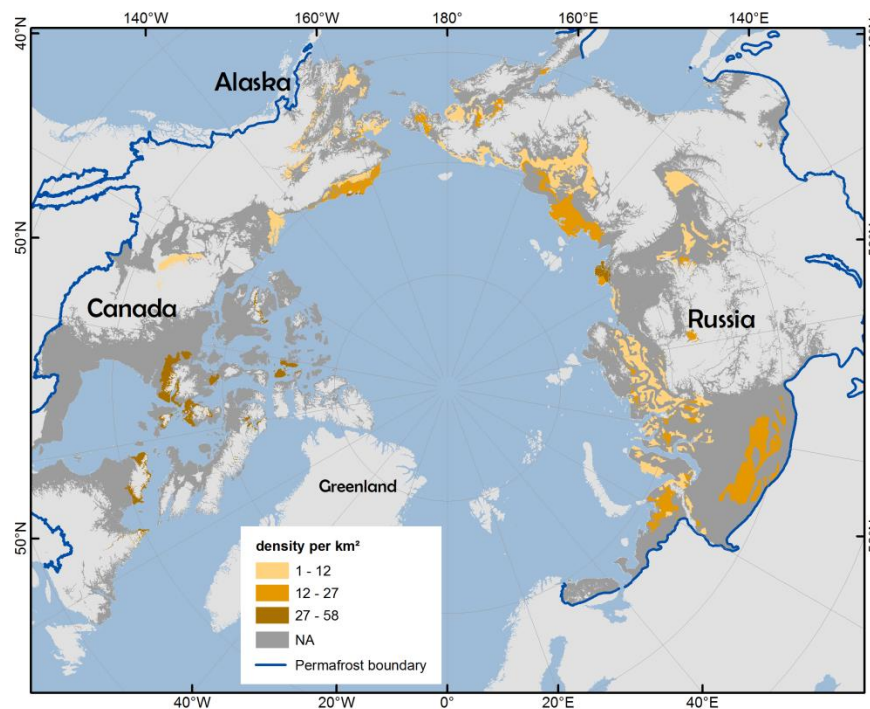
Figure 4: Areal fraction of waterbodies with surface areas between 1.0E+02 m² and 1.0E+06 m². Permafrost boundary was derived from the regional databases.

5   **Figure 5: Waterbody density per km² for waterbodies with surface areas of between 1.0E+02 m² and 1.0E+06 m² within permafrost landscape units.** Permafrost boundary was derived from the regional databases.

Table 3: Extrapolated waterbody statistics for permafrost landscapes. Permafrost extent (PE) is reported as C – continuous, D – discontinuous, and S – sporadic., and I - isolated. Extrapolated statistics include areal fraction and ~~waterbody density~~ number(frequency per km²) for waterbodies (>=1.0E+02 and <=1.0E+06 m²) and ponds (>=1.0E+02 and <1.0E+04 m²). Numbers in brackets denote the relative error in %. The relative standard error (RE) was calculated for 10 x 10 km boxes within large maps or when four or more small maps could be averaged. The RE is the standard error expressed as a percentage. The standard error is the standard deviation of the sampling distribution of a statistic.

Country	Ecozone	PE	Ground ice [vol %]	Surficial geology	Lithology	Fraction	Density	Pond fraction	Pond density
Alaska	Arctic Tundra	C	10-40	alluvial-marine	Sandy	7.3 (4.6)	17.6 (3.4)	1.6 (2.8)	16.9 (3.4)
	Arctic Tundra	C	<10	eolian, sand	Sandy	11.1 (4.9)	21.3 (10.3)	1.8 (11.4)	20.4 (10.4)
	Arctic Tundra	C	>40	glaciomarine	Silty	8 (NA)	28 (NA)	1.8 (NA)	27.3 (NA)
	Bering Taiga	D	>40	eolian, loess	Silty	10.5 (NA)	7.3 (NA)	1.1 (NA)	5.9 (NA)
	Bering Taiga	S	10-40	fluvial, abandoned terrace	Silty	9.9 (5.3)	6.2 (3.8)	0.7 (3.3)	5.1 (3.9)
	Bering Tundra	C	>40	eolian, loess	Silty	5.7 (11.8)	1.6 (18)	0.3 (13.7)	1.1 (28.5)
	Intermontane Boreal	D	10-40	fluvial, abandoned terrace	Silty	7.1 (NA)	2.9 (NA)	0.3 (NA)	2.4 (NA)
Canada	Northern Arctic	C	<20	colluvial fines	Sandy Loam	15.1 (NA)	42.9 (NA)	5.9 (NA)	40.6 (NA)
	Northern Arctic	C	<10	till veneer	NA	3.7 (NA)	57 (NA)	2.1 (NA)	5~~0~~6.7 (NA)
	Southern Arctic	C	>20	glaciofluvial plain	Organic	7.6 (4.5)	5.7 (3.4)	0.5 (3)	5.3 (3.4)
	Southern Arctic	C	>20	till blanket	Clay	13.2 (3.8)	1.9 (3.2)	0.2 (5.3)	1.0 (5.0)
	Southern Arctic	D	<20	alluvial deposits	Loam	7.7 (6.2)	6.9 (5.3)	0.8 (5.2)	6.2 (5.4)
	Taiga Plain	D	<20	alluvial deposits	Loam	21.1 (1.4)	7.6 (1.7)	1.1 (1.8)	5.4 (2.4)
	Taiga Shield	D	<20	till veneer	Sand	8 (13.9)	7.0 (1.9)	1.0 (1.7)	5.8 (1.2)
	Taiga Shield	D	<20	undivided	Sand	10.8 (3.0)	2.7 (2.1)	0.3 (3.4)	1.7 (3.0)
Russia	East Siberian Taiga	C	>40	alluvial-limnetic	Coarse	5.3 (10.1)	1 (8.8)	0 (11.5)	0.7 (9.6)
	East Siberian Taiga	C	>40	alluvial-limnetic	Medium	4.9 (3.8)	2.5 (4.2)	0.4 (4.7)	2.0 (5.0)
	Northeast Siberian Coastal Tundra	C	>40	alluvial-limnetic	Medium	8.3 (NA)	18.2 (NA)	0.8 (NA)	17.6 (NA)
	Northeast Siberian Taiga	C	>40	alluvial-limnetic	Organic	16.7 (10.2)	3.5 (6.3)	0.4 (5.7)	2.3 (6.4)
	Northeast Siberian Taiga	C	>40	alluvial-limnetic	Coarse	5.3 (10.1)	1 (8.8)	0 (11.5)	0.7 (9.6)
	Northeast Siberian Taiga	C	>40	cover, loess like deposits, loess and clays	Medium	1.1 (NA)	3.3 (NA)	0.2 (NA)	3.1 (NA)
	Northwest Russian-Nov. Zem. Tundra	C	<20	glacial	Medium	6.3 (28.1)	12.7 (21.4)	1.4 (24.6)	11.5 (21.5)
	Taimyr-Central Siberian Tundra	C	>40	alluvial-limnetic	Organic	10.3 (10.5)	38.8 (50.1)	2.1 (27.5)	37.8 (51.3)
	Taimyr-Central Siberian Tundra	C	20-40	alluvial-limnetic	Organic	7.2 (NA)	23.8 (NA)	2.5 (NA)	23.1 (NA)
	Taimyr-Central Siberian Tundra	C	20-40	glacial	Medium	3.6 (NA)	4.5 (NA)	0.4 (NA)	4.1 (NA)
	West Siberian Taiga	S	20-40	organic deposits	Organic	16.7 (2.2)	20.2 (2.1)	3 (2.4)	17.6 (2.2)
	Yamal-Gydan Tundra	C	>40	marine	Medium	8.9 (5.0)	3.6 (3.1)	0.4 (2.9)	2.7 (3.2)
	Yamal-Gydan Tundra	C	20-40	marine	Banded	6.1 (2.3)	8.6 (2.7)	0.8 (2.7)	7.8 (2.7)

6 Discussion

6.1 Classification accuracy and ~~uncertainty~~variability

The accuracy of the individual waterbody map~~s~~ depends on the spectral and spatial properties of the remote sensing imagery employed for classification as well as the classification method. In general, open water surfaces show a high contrast to the surrounding land area in all utilized spectral bands, *i.e.*, panchromatic, near-infrared, and X-band, since water absorbs most of the incoming radiation (Grosse et al., 2005; Muster et al., 2013). Ground surveys of waterbody surface area were available for only a few study sites. Accuracy ranged between 89 % for object-oriented mapping of multispectral (Lara et al., 2015), 93 % for object-oriented mapping of panchromatic imagery (Andresen and Lougheed, 2015), and more than 95 % for a supervised maximum likelihood classification of multispectral aerial images (Muster et al., 2012). Errors in the classification may be largely due to commission errors, i.e. the spectral signal is misinterpreted as water where in reality it may be land surface. Many shallow ponds and pond/lake margins are characterized by vegetation growing or floating in the water which cannot be adequately classified from single-band imagery (Sannel and Brown, 2010). PeRL classifications dating from early August are likely most affected since abundance of aquatic plants peaks around that time of year. In some cases, even multispectral imagery cannot distinguish between lake and land because floating vegetation mats fully underlain by lake water may spectrally appear like a land surface (Parsekian et al. 2011).
Seasonal processes, such as snowmelt, progressing thaw depth, evaporation, and precipitation do affect the extent of surface water. Waterbody maps therefore reflect the local water balance at the time of image acquisition. Seasonal reduction of surface water extent, however, is largest in the first 2 weeks following snowmelt (Bowling et al., 2003). All PeRL maps date from the late summer season so that snowmelt and the early summer season are excluded. Changes of water extent in late summer are primarily due to evaporation and precipitation. In a study area on the Barrow Peninsula, Alaska, we find that the open water extent varies between 6 % and 8 % between the beginning and end of August of different years. However, the effect is hard to quantify as other factors such as spectral properties and resolution also impact classifications of different times at the same site. Seasonal variations may be larger in case of heavy rain events right before image acquisition but ultimately depend on local conditions which control surface and subsurface runoff.
~~However, the effect is hard to quantify as other factors such as spectral properties and resolution also impact classifications of different times at the same site.~~

~~Different map products were compared for a common area of the Barrow Peninsula. Waterbody areal fraction was 10 % for a Quickbird derived dataset from on August 1, 2002, 6 % for a Kompsat dataset from August 2, 2009, and 8 % for a TerraSAR-X dataset from August 25, 2009. The multispectral classification used on the Quickbird image also identifies waterbodies with aquatic vegetation with an areal fraction of 3 %.~~

6.2 Uncertainty of circum-arctic map

Uncertainties regarding the ~~upscaling~~ extrapolation of waterbody distributions arise from (i) the combination of different waterbody maps, (ii) the accuracy of the underlying regional permafrost maps, and (iii) the level of generalization inherent in the ~~PL~~ permafrost landscape units.

~~PeRL is a static database that presents late summer inundation conditions only. The effect of image acquisition in relation to rainfall may impact the waterbodies at the site level.~~ ~~However, the effect is hard to quantify as other factors such as spectral properties and resolution also impact classifications of different times at the same site.~~ Waterbody statistics of permafrost landscapes are derived from diverse remote sensing imagery. Imagery dates from different years and months and features different image properties. However, ~~Average size distributions over large regions probably reduce the~~ the importance effect of seasonal variability or image properties on the average statistic is small compared to the natural variability within and between permafrost landscape units. ~~as site level variability within the thaw period is being traded for the large spatial coverage of waterbodies across permafrost landscapes.~~

Permafrost landscapes present a unified circum-arctic categorization to upscale waterbody distributions. Due to the uncertainty and scale of the regional ~~PL maps~~PLM, however, it cannot be expected that non-overlapping waterbody maps within the same ~~PL~~ permafrost landscape have the same size distribution. The regional ~~permafrost landscape maps~~PLM are themselves extrapolated products where finite point sources of information have been used to describe larger spatial domains. No error or uncertainty measure, however, was reported for the regional maps. In addition, the variables used to describe permafrost landscapes present the dominant classes within the landscape unit. Thus, certain waterbody maps may represent landscape subtypes that are not represented by the reported average ~~characteristic~~statistic. For example, two permafrost landscapes have been classified in the Lena Delta in Northern Siberia. The southern and eastern part of the delta is characterized by continuous permafrost with ground ice volumes larger than 40 %, alluvial-limnetic deposits and organic substrate. Local studies differentiate this region further based on geomorphological differences and ground ice content. The Yedoma ice complex in the southern part features much higher ground ice content of up to 80 % and higher elevations than the eastern part which is however not resolved in the Russian ~~PL map~~PLM. These sub-regional landscape variations are also reflected in the waterbody size distributions which are significantly different for the southern and eastern part of the delta. In the averaged statistics this is indicated by a high relative error of 11 % and 28 % for areal fraction of waterbodies and ponds, respectively, and about 50 % for waterbody density estimates. In this case, the ~~PL~~ permafrost landscape unit in that area does not adequately reflect the known distribution of ground ice and geomorphology and demonstrates the need to further improve ~~PL maps~~PLM in the future.

6.3 Potential use of database and future development

Waterbody maps and distribution statistics are most accurate at site-level. At this scale, maps can be used as a baseline to detect changes in surface inundation for seasonal, inter-annual and decadal periods. Site-level size distributions can also be used to validate statistical extrapolation methods which have previously been used to extrapolate from coarser databases to finer scales (Downing et al., 2006; Seekell et al., 2010). Validation of these approaches have since questioned the validity of power laws for smaller lakes and ponds but have also been limited to waterbodies as small as 1.0E+04 m², *i.e.*, 2 orders of magnitude larger than the minimum size in PeRL data sets.

The circum-arctic map provides spatially extrapolated information for larger-scale applications. Coarse-scale global databases such as the Global Lakes and Wetlands database (GLWD) by Lehner and Döll (2004) are used in global Earth System Models to represent the water fraction in model grid cells (Wania et al., 2013). The GLWD renders a reliable inventory of lakes larger 1 km² (Lehner and Döll, 2004). Compared to the GLWD, PeRL inventoried up to 21 % additional waterbody area. Moreover, ponds are the most frequent waterbody type with 45 to 99 %. In light of observed scaling of biogeochemical processes with waterbody surface area (Wik et al., 2016), PeRL results emphasize the need to include

waterbodies of 1.0E+06 m² and smaller in conjunction with their size distributions in physical and biogeochemical models of the high latitude surface. Moreover, the combination of waterbody size distributions with ~~PL~~ landscape properties can motivate further study for process-based predictive simulations both at the site and regional scale. However, users should be aware of the map's uncertainty when using it to upscale landscape properties such as methane or heat fluxes. For this

5   purpose, users should refer to the reported spatial variability, confidence class and extensive metadata.

PeRL's ~~PL~~ permafrost landscape units represent the least common denominator across the Arctic where ~~PL~~ landscape properties have been strongly generalized. More detailed information about landscape properties was available for the Canadian database and Northern Alaska (Jorgensen et al., 2014) but not for central and southern Alaska or Russia. We suggest that more detailed and accurate classes of ground ice as well as further refinement of physiography within the broad

10   lowland zone will likely explain differences in waterbody distributions between different maps in the same ~~PL~~permafrost landscape. Regionally different methodologies currently prohibit to compare permafrost landscapes between regions and to extrapolate across regions. The harmonization of landscape properties, delineation of common terrain units and extrapolation methods for the whole Arctic require a coordinated circum-arctic effort.

Our extrapolated area (1.4E+06 km²) represents only 7.0 % of the current estimated Arctic permafrost area (Brown et al.,

15   1998) but about 17 % of the current Arctic permafrost lowland area (below 300 m a.s.l.) where most of the Arctic lakes are located (Lehner and Döll, 2004; Smith et al., 2007; Grosse et al., 2013). With a few exceptions, the reported sites are predominantly located in coastal areas. Especially the lake-rich permafrost lowlands of Canada and Central Siberia are underrepresented, despite their large spatial coverage. Underrepresented landscape types are areas with discontinuous, isolated or sporadic permafrost, as well as areas in boreal regions. PeRL maps are conservative estimates of surface

20   inundation as most maps capture open water only and do not include ponds smaller than 1.0E+02 m² in size. PeRL maps with resolutions of less than 1 m, however, indicate the presence of many waterbodies smaller than the current threshold of 1.0E+02 m². These very small waterbodies as well as water areas with emersed vegetation are highly productive environments that require attention in future mapping efforts.

7 Data availability

25   Waterbody maps, study area boundaries and maps of regional permafrost landscapes including link to detailed metadata are available at https://doi.pangaea.de/10.1594/PANGAEA.868349.

8 Conclusions

PeRL maps and statistics provide a great resource for a large suite of applications across the Arctic such as resource and habitat management, hydrological and ecological modeling, pond and lake change detection, and upscaling of

30   biogeochemical processes. PeRL maps includes waterbodies with surface areas ~~from 1.0E+06 m² down to~~as small as 1.0E+02 m² which complements available global databases and increases waterbody size resolution by 2−4 orders of magnitude. Ponds, i.e. waterbodies with surface areas smaller than 1.0E+04 m² are the dominant waterbody type found in all study areas across the Arctic. This demonstrates the need to include small waterbodies and parameterize size distributions in global land surface models. Furthermore, PeRL presents a baseline that allows future studies to investigate

35   direction and magnitude of past and future Arctic surface inundation. The current compilation of high-resolution ~~surface inundation~~ waterbody maps underlines the need to produce more: vast areas in all regions are still unmapped regarding

small waterbodies, especially the Canadian lowlands and boreal regions of Russia. Future mapping efforts should therefore focus equally on ~~both~~ filling gaps and ~~also~~ monitoring inventoried sites. The combination of waterbody statistics and landscape properties has great potential to improve our understanding of environmental drivers of surface inundation in permafrost lowlands. However, permafrost landscape maps need to be improved by increasing the level of detail as well by

5   harmonizing   mapping   and   extrapolation   approaches   across   Arctic   regions.

Table A1: Metadata and references for climate data and permafrost depth. MAAT—mean annual air temperature, TP—total precipitation.

map ID	country	MAAT & TP period	climate data source	permafrost thickness source
abi001	Sweden	2006–2011, 1997–2007	MAAT: Johannsson et al. (2013), TP: Abisko Station, www.polar.se/abisko	Akerman and Johansson (2008), Dobinski (2010)
arg001, arg002	Russia	1999–2011	Boike et al. (2013)	Yershov et al. (1991)
bar001, bar002, bar003, bar004	Alaska	1981–2010	National Climatic Data Center (2016): Barrow W Post Rogers Airport AK US	Brown et al. (1980)
byk001	Russia	1984–1994	Rivas-Martínez (2008)	Grosse et al. (2008)
byl001	Canada	N/A	MAAT: Godin et al. (2014), TP: Fortier et al. (2007)	Smith and Burgess (2000)
che001, che002	Russia	1984–1994	Rivas-Martínez (2008)	Grosse et al. (2008)
elc001, elc002, ecl003	Alaska	1981–2010	National Climatic Data Center (2016): Barrow W Post Rogers Airport, AK, US	Sellmann and Brown (1973)
esk001	Canada	1981–2010	Environment Canada (2016), Tuktoyaktuk A	Taylor and Judge (1981)
fis001	Alaska	1981–2010	National Climatic Data Center (2016): Kuparuk AK US	Jorgensen et al. (2008)
fir001, fir002, fir003	Russia	1999–2011	Boike et al. (2013)	Yershov et al. (1991)
gpr001	Canada	1971–2000	Bouchard et al. (2014), Environment Canada (2012)	Smith and Burgess (2002)
hbl001	Canada	1971–2000	Sannel and Kuhry (2011), Environment Canada	N/A
ice003	Russia	1999–2011	Boike et al. (2013)	Yershov et al. (1991)
ind001, kyt001	Russia	1961–1990	Chokurdakh WMO station 21946	N/A
kol001	Russia	1996–2015	Bukhta Ambarchik meteostation (WMO ID 25034)	Yershov et al. (1991)
kol002	Russia	1996–2015	Andrushkino meteostation (WMO ID 25017)	Yershov et al. (1991)
kur001, kur002	Russia	1999–2011	Boike et al. (2013)	Yershov et al. (1991)
log001	Russia	1961–1990	Khatanga WMO station 20891	N/A
mdn001	Canada	1981–2010	Environment Canada (2016), Inuvik A	Burn and Kokelj (2009)
mdw001	Canada	1981–2010	Environment Canada (2016), Inuvik A	Burn and Kokelj (2009)
ole001	Russia	1948–1960	Rivas-Martínez (2008)	Grosse et al. (2008)
pbp001	Canada	1981–2010	Environment Canada (2016), Resolute Cars	Smith and Burgess (2002)

map ID	country	MAAT & TP period	climate data source	permafrost thickness source
imc001, kcp001, kpc002, kpc003	Alaska	1981–2010	National Climatic Data Center (2016), Kuparuk AK US	N/A
ric001	Canada	1981–2010	Environment Canada (2016), Inuvik A	Burn (2002)
rog001, rog002, rog003, rog004, rog005	Russia	1961–1990	Vorkuta, Hugelius et al. (2011)	Rivkin et al. (2008)
sam001	Russia	1999–2011	Boike et al. (2013)	Yershov et al. (1991)
sei001, sei002	Russia	N/A	Sjöberg et al. (2013)	Rivkin et al. (2008)
ksl001	Alaska	1981–2010	National Climatic Data Center (2016): Noma Municipal Airport AK US	Jorgensen et al. (2008)
sur001	Russia	N/A	Krementski et al. (2003)	Krementski et al. (2003)
tbr001, tea001	Canada	1981–2010	Environment Canada (2016), Yellowknife A	N/A
tav001	Sweden	1971–2000	Sannel and Kuhry (2011)	Sannel and Kuhry (2011)
tuk001	Canada	1981–2010	Environment Canada (2016), Tuktoyaktuk A	Taylor and Judge (1981)
yak001	Russia	1930–2010	Fedorov et al. (2014)	Yershov et al. (1991)
yam001, yam002	Russia	2004–2013	Marre-Sale, Leibman et al. (2014)	Yershov et al. (1991)
yfl001	Alaska	1981–2010	National Climatic Data Center (2016): Central Number 2 AK US	Walvoord et al. (2012)
yuk001, yuk002	Alaska	1981–2010	National Climatic Data Center (2016): Bethel Airport AK US	N/A

Appendix B

Table B1: Metadata about base imagery, classification accuracy and classification objects, and associated references. Classifications that include partial lakes, rivers or streams are labelled "no" in the "clean" column.

map ID	filenames	base imagery	acquisition date [YYYY-MM-DD]	reso- lution [m]	accuracy	classes	min size [m²]	Clean	reference
abi0012010xxxx	abi0012010xxxx_ortho_nplaea	Orthophoto (Lantmäteriet, I2014/00691); SPOT5; DEM	2010	1.0	>0.71	open water	21	Yes	this paper
arg00120110829	arg00120110829_k2_nplaea	KOMPSAT-2	2011-08-29	4.0	N/A	open water	64	Yes	this paper
arg0022009xxxx, fir0032009xxxx, ice0032009xxxx	arg0022009xxxx_re_nplaea, fir0032009xxxx_re_nplaea, ice0032009xxxx_re_nplaea	RapidEye	2009–2011	5.0	0.85 in comparison to Kompsat II	open water	100	No	Bartsch and Seifert (2012)
bar00119480804	bar00119480804_ai_nplaea	aerial imagery	1948-08-04	0.7	N/A	open water	16	Yes	Andresen et al. (2015)
bar00120020802	bar00120020802_qb_nplaea	Quickbird	2002-08-02	0.7	N/A	open water	16	Yes	Andresen et al. (2015)
bar00120080730	bar00120080730_qb_nplaea	Quickbird	2008-07-30	0.7	N/A	open water	16	yes	Andresen et al. (2015)
bar00120100810	bar00120100810_wv2_nplaea	WorldView-2	2010-08-10	0.5	0.93	open water	16	yes	Andresen et al. (2015)
byk00120060709	byk00120060709_spot_nplaea	SPOT-5	2006-07-09	2.5	N/A	open water	30	yes	Grosse et al. (2008)
byl00120100902	byl00120100902_geo_nplaea	GeoEye-1	2010-09-02	0.5	N/A	open water	4	yes	this paper
che00120020709	che00120020709_iko2_nplaea	IKONOS-2	2002-07-09	1.0	N/A	open water	30	yes	Grosse et al. (2008)
che00220090724	che00220090724_alos_nplaea	ALOS PRISM	2009-07-24	3.0	0.97	open water	36	yes	Widhalm et al. (2014a, 2014b)
elc00120090825, wlc00120090825,	elc00120090825_tsx_nplaea, bar00220090825_tsx_nplaea	TerraSAR-X	2009-08-25	2.5	N/A	open water, water with emersed vegetation	25	yes	this paper
elc00220020801, wlc00220020801	elc00220020801_qb_nplaea, bar00320020801_qb_nplae	Quickbird	2002-08-01	0.6	0.89 to 0.95	open water	7.8	yes	Lara et al. (2015)
elc00320090802,	elc00320090802_k2_nplaea,	KOMPSAT-2	2009-08-02	4.0	N/A	open water	64	yes	Muster et al. (2013)

map ID	filenames	base imagery	acquisition date [YYYY-MM-DD]	reso-lution [m]	accuracy	classes	min size [m²]	Clean	reference
wlc00320090802	bar00420090802_k2_nplaea								
elc004200808xx	elc004200808xx_tr_qb_nplaea	Quickbird	2008-08	0.6	>0.75	open water	1.5	yes	Lara et al. (2015)
esk00120090727	esk00120090727_tsx_nplaea	TerraSAR-X	2009-07-27	2.5	N/A	open water	36	yes	this paper
fir00120090906	fir00120090906_k2_nplaea	KOMPSAT-2	2009-09-06	4.0	N/A	open water	64	yes	this paper
fir00220110711	fir00220110829_k2_nplaea	KOMPSAT-2	2011-08-29	4.0	N/A	open water	64	yes	this paper
fis00120020715	fis00120020715_orri_nplaea	Airborne Orthorectified Radar Imagery (IFSAR)	2002-07-15	1.3	N/A	open water	100	Yes	Jones et al. (2013)
grp00119590707	grp00119590708_ai_nplaea	aerial imagery	1959-07-07	0.4	N/A	open water	17	yes	Bouchard et al. (2014)
grp00120060707	grp00120060708_qb_nplaea	QuickBird	2006-07-07	0.61	N/A	open water	13	yes	Bouchard et al. (2014)
hbl00119540701	hbl00119540701_ai_nplaea	aerial imagery	1954-07-01	1.0	N/A	open water	51	yes	Sannel and Brown (2010), Sannel and Kuhry (2011)
hbl00119740617	hbl00119740617_ai_nplaea	aerial imagery	1974-06-17	1.0	N/A	open water	53	yes	Sannel and Brown (2010), Sannel and Kuhry (2011)
hbl00120060706	hbl00120060706_qb_nplaea	QuickBird	2006-07-06	0.6	N/A	open water	55	yes	Sannel and Brown (2010), Sannel and Kuhry (2011)
imc00120040725	imc00120040725_qb02_nplaea	QuickBird-2	2004-07-25	0.7	N/A	open water	100	yes	this paper
ind00120090907	ind00120090907_wv1_nplaea	WorldView-1	2009-09-07	0.5	N/A	open water	100	yes	this paper
kol00119650721, kol00219650721	kol00119659721_cor_nplaea, kol00219659721_cor_nplaea	CORONA	1965-07-21	5.0	N/A	open water	200	yes	this paper
kur00120100805	kur00120100805_geo_nplaea	Geoeye	2010-08-05	1.0	N/A	open water	4	yes	this paper
kur00220080926	kur00220080926_ap_nplaea	ALOS PRISM	2008-09-26	2.5	RMSE 5.8 m georeferenc-ing accuracy	open water	187	yes	this paper
kyt00120070728	kyt00120070728_ap_nplaea	ALOS PRISM	2007-07-28	3.0	0.97	open water	36	yes	Widhalm et al. (2014a, 2014b)
log00120110811	log00120110811_qp_nplaea	QuickBird	2011-08-11	2.4	0.90	open water	23	yes	Palmtag et al. (2016)
mdw00120090921	mac00120090921_tsx_nplaea	TerraSAR-X	2009-09-21	2.5	N/A	open water	36	yes	this paper
mdn00120100716	mdn00120100716_tsx_nplaea	TerraSAR-X	2010-07-16	2.5	N/A	open water	64	yes	this paper

map ID	filenames	base imagery	acquisition date [YYYY-MM-DD]	reso-lution [m]	accuracy	classes	min size [m²]	Clean	reference
ole00120060708	ole00120060708_spot_nplaea	SPOT-5	2006-07-08	2.5	N/A	open water	30	yes	Grosse et al. (2008)
pbp00120090813	pbp00120090813_tsx_nplaea	TerraSAR-X	2009-08-13	2.5	N/A	open water	12	yes	Muster et al. (2013)
kpc001201007xx, kpc002201007xx, kpc003201007xx	pru001201007xx_ai_nplaea, pru002201007xx_ai_nplaea, pru003201007xx_ai_nplaea	digital true color aerial imagery	2010-07-09 and 2010-07-25	N/A	N/A	open water	1	yes	Walker et al. (2014), Raynolds et al. (2014)
ric001201209125	ric001201209125_tsx_nplaea	TerraSAR-X	2012-09-25	2.5	N/A	open water	8	yes	this paper
rog00120070626	rog00120070626_qb_nplaea	QuickBird	2007-06-26	0.6	N/A	open water	30	yes	Sjöberg et al. (2013)
rog00219740726	rog00219740726_ai_nplaea	aerial imagery	1974-07-26	1.0	N/A	open water	28	yes	Sannel and Kuhry (2011)
rog00220070707	rog00220070707_qb_nplaea	QuickBird	2007-07-07	0.6	N/A	open water	28	yes	Sannel and Kuhry (2011)
rog00320070626	rog00320070626_qb_nplaea	QuickBird-2	2007-06-26	2.4	Overall Kappa: 0.68	open water	11	yes	this paper
rog00420070704	rog00420070704_qb_nplaea	QuickBird-2	2007-07-04	2.4	Overall Kappa: 0.83	open water	29	yes	this paper
rog00520070704	rog00520070704_qb_nplaea	QuickBird-2	2007-07-04	2.4	Overall Kappa: 0.83	open water	11.5	yes	this paper
sam001200808xx	sam001200808xx_ai_nplaea	aerial imagery	2008-08	0.3	N/A	open water	1	yes	Muster et al. (2012)
sei00120070706	sei00120070706_qb_nplaea	QuickBird	2007-07-06	0.6	N/A	open water	30	yes	Sjöberg et al. (2013)
sei00220070706	sei00220070706_qb_nplaea	QuickBird-2	2007-07-06	2.4	Overall Kappa: 0.68	open water	11.5	yes	this paper
ksl0012012xxxx	ksl0012012xxxx_spot_nplaea	SPOT-5	2009-2010	2.5	N/A	open water	25	yes	this paper
ksl00119620628	ksl00119620628_cor_nplaea	Corona KH4	1962-06-28	6.0	N/A	open water	144	yes	this paper
sur00120130802	sur00120130802_tsx_nplaea	TerraSAR-X	2013-08-02	2.0	N/A	open water	16	yes	this paper
tbr00120100901, tea00120100901	tan00120100901_tsx_nplaea, tan00220100901_tsx_nplaea	TerraSAR-X	2010-09-01	2.5	N/A	open water	36	yes	this paper
tav00119630831	tav00119630831_ai_nplaea	aerial imagery	1963-08-31	0.5	N/A	open water	28	yes	Sannel and Kuhry (2011)
tav00119750810	tav00119750810_ai_nplaea	aerial imagery	1975-08-10	0.9	N/A	open water	28	yes	Sannel and Kuhry (2011)
tav00120030702	tav00120030702_iko2_nplaea	IKONOS	2003-07-02	1.0	N/A	open water	28	yes	Sannel and Kuhry (2011)

map ID	filenames	base imagery	acquisition date [YYYY-MM-DD]	reso-lution [m]	accuracy	classes	min size [m²]	Clean	reference
tuk00120120723	tuk00120120723_tsx_nplaea	TerraSAR-X	2012-07-23	2.5	N/A	open water	36	yes	this paper
yak0012009xxxx	yak0012009xxxx_re_nplaea	RapidEye	2009–2011	5.0	0.85 comparison with Landsat	open water	75	no	Bartsch and Seifert (2012)
yam00120080824	yam00120080824_tsx_nplaea	TerraSAR-X	2008-08-24	2.5	N/A	open water	36	yes	this paper
yam00220100820	yam00220100820_tsx_nplaea	TerraSAR-X	2010-08-20	2.5	N/A	open water, water with emersed vegetation	100	yes	this paper
yfl0012011xxxx	yfl0012011xxxx_ai_nplaea	aerial imagery	2011-06-18 2011-09-05	1.0	N/A	open water	4	yes	this paper
yuk00120090812, yuk00220090812	yuk00120090812_tsx_nplaea	TerraSAR-X	2009-08-12	2.5	N/A	open water	36	yes	this paper

Table B2: Metadata about classification software, algorithm

site ID	software	pre-processing	classification band(s)	spectral range	classification method	post-processing	data repository
abi0012010xxxx	GDAL, SAGA GIS, Orfeo toolbox	1. Image segmentation using Orthophoto RGB (1m) and DEM (2m), minimum object size 130m2; 2. Classification of watermask: SVM classifier using Red band and slope	red (1m) + slope (2m)	610-680 nm	segmentation (SVM)	N/A	
arg00120110829	ENVI 4.8, ArcGIS 12	georeferencing	Near-infrared	760-900 nm	density slice	N/A	
arg0022009xxxx, fir0032009xxxx, ice0032009xxxx	N/A	georeferencing, histogram matching routine for radiometric normalization, mosaiking	B, G, R, RE, NIR	blue: 440-510 nm, green: 520-590 nm, red: 630-685 nm, RE: 690-730 nm, NIR: 760-850 nm	object-oriented classification	N/A	http://geo.tuwien.ac.at/permafrost/
bar00119480804	ENVI 4.4	coregistration and orthorectification using an image-to-image correction (RMSE of 1.0 m)	panchromatic	N/A	object-oriented classification	N/A	

site ID	software	pre-processing	classification band(s)	spectral range	classification method	post-processing	data repository
bar00120020802	ENVI 4.4	coregistration and orthorectification using an image-to-image correction (RMSE of 1.0 m)	panchromatic	445–900 nm	object-oriented classification	N/A	
bar00120080730	ENVI 4.4	coregistration and orthorectification using an image-to-image correction (RMSE of 1.0 m)	panchromatic	445–900 nm	object-oriented classification	N/A	
bar00120100810	ENVI 4.4	coregistration and orthorectification using an image-to-image correction (RMSE of 1.0 m)	panchromatic	450–800 nm	object-oriented classification	N/A	
byk00120060709	ArcGIS	georeferenced to topographic maps of scale 1:100,000	panchromatic	480–710 nm	density slice		
byl00120100902	Geomatica 2015, ArcGIS 10.2	pan-sharpening, orthorectification	B, G, R, NIR	450–510 nm, 510–580 nm, 655–690 nm, 780–920 nm	unsupervised k-means classification	spatial filter on raster to remove classification imperfections,removed: water bodies < 4 m², major rivers, streams, shadows	
che00120020709	ArcGIS	orthorectification	panchromatic	760–850 nm	density slice	removed: shadows, stream water bodies, man made structures; included: lake ice, turbid and shallow water	
che00220090724	N/A	N/A	panchromatic	520–770 nm	density slice	N/A	https://doi.pangaea.de/10.1594/PANGAEA.834200
elc00120090825, bar00220090825	ENVI 4.4	gamma filter (11x11 pixel)	X-Band	wavelength 31 mm, frequency 9.6 GHz	k-means classification	N/A	
elc00220020801, bar00320020801	ArcGIS 10.2, eCognition 9.1	principle component analysis on raster bands, pan-sharpening	R, B, G, IR	450–520 nm, 560–600 nm, 630–690 nm, 760–890 nm	object-oriented classification	N/A	
elc00320090802, bar00420090802	ENVI 4.8, ArcGIS 10	georeferencing	Near-infrared	760–900 nm	density slice	partial lakes, rivers,streams and cloud shadows removed	https://doi.pangaea.de/10.1594/PANGAEA.786073

site ID	software	pre-processing	classification band(s)	spectral range	classification method	post-processing	data repository
elc004200808xx	ArcGIS 10.2, eCognition 9.2	principle component analysis on raster bands, pan-sharpening	R,B,G,IR	450-520 nm, 560-600 nm, 630-690 nm, 760-890 nm	object-oriented classification	N/A	
esk00120090727	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (11x11 pixel)	X-Band	wavelength 31 mm, frequency 9.6 GHz	k-means classification	river, streams, shadows removed	
fir00120090906	ENVI 4.8, ArcGIS 10	georeferencing	Near-infrared	760-900 nm	density slice	river, streams, shadows removed	
fir00220110711	ENVI 4.8, ArcGIS 10	georeferencing	Near-infrared	760-900 nm	density slice	river, streams, shadows removed	
fis00120020715	ENVI 4.5	N/A	panchromatic	135 or 270 MHz	density slice	N/A	
grp00119590707	ArcGIS 10.0	Scanning (1814 dpi); georeferencing (RMSE = 1.8m (4.2pix))	black and white	N/A	manual	N/A	
grp00120060707	ArcGIS 10.0	geometric correction (cubic convolution resampling)	panchromatic	450-900 nm	manual	N/A	
hbl00119540701	ENVI 4.5, ENVI 4.7, ArcGIS 9	resampling of pixel resolution to 0.6 m, georeferenced to QuickBird image from 2006 with RMSE of 0.38-1.42	panchromatic	N/A	manual delineation	N/A	
hbl00119740617	ENVI 4.5, ENVI 4.7, ArcGIS 9	resampling of pixel resolution to 0.6 m, georeferenced to QuickBird image from 2006 RMSE of -0.38-1.42	panchromatic	N/A	manual delineation	N/A	
hbl00120060706	ENVI 4.5, ENVI 4.7, ArcGIS 9	N/A	panchromatic	-405-1053 nm	manual delineation	N/A	
imc00120040725	ENVI 5.0, ArcGIS 10.3	orthorectification	panchromatic	760-850 nm	density slice	river, streams, shadows removed	
ind00120090907	ENVI 5.0, ArcGIS 10.3	orthorectification	panchromatic	400-900 nm	density slice, opening filter 3x3	river, streams, shadows removed	
kol00119650721, kol00219650721	ArcGIS 9	georeferencing	panchromatic	N/A	manual delineation	N/A	
kur00120100805	ENVI 4.8, ArcGIS 10	pan-sharpening, orthorectification (RMSE: 0.36m)	Near-infrared	760-900 nm	unsupervised k-means classification	river, streams, shadows removed	
kur00220080926	PCI Geomatica, ArcGIS 9	orthorectification based on own stereo DEM	panchromatic	520-770 nm	manual delineation	N/A	

site ID	software	pre-processing	classification band(s)	spectral range	classification method	post-processing	data repository
kyt00120070728	ENVI 4.8, ArcGIS 10	N/A	panchromatic	520–770 nm	density slice	N/A	https://doi.pangaea.de/10.1594/PANGAEA.834200
log00120110811	ENVI 5.1	N/A	NIR, R, G, B	485–830 nm	supervised maximum likelihood classification	N/A	
mdw00120090921	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (11x11 pixel)	X-Band	wavelength 31 mm, frequency 9.6 GHz	k-means classification	river, streams, shadows removed	
mdn00120100716	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (11x11 pixel)	X-Band	wavelength 31 mm, frequency 9.6 GHz	density slice, DN: -28.06 to -14.79	river, streams, shadows removed	
ole00120060708	ArcGIS	georeferenced to topographic maps of scale 1:100,000	panchromatic	480–710 nm	density slice	shadows, stream water bodies, man made structures removed, (lake ice, turbid and shallow water included	
pbp00120090813	ENVI 4.8, ArcGIS 10	gamma filter (11x11 pixel)	X-Band	wavelength 31 mm, frequency 9.6 GHz	density slice	majority filter (7x7 pixel) to reduce spurious pixels in classification	
kpc001201007xx, kpc002201007xx, kpc003201007xx	N/A	N/A	R, G, B	N/A	manual delineation	N/A	http://geobotanical.portal.gina.alaska.edu/catalogs/10609-prudhoe-bay-cumulative-impact-map-a-raynolds-2014
ric001201209125	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (11x11 pixel)	X-Band	wavelength 31 mm, frequency 9.6 GHz	k-means classification	river, streams, shadows removed	
rog00120070626	ArcGIS	N/A	panchromatic	450–900 nm	manual delineation	N/A	
rog00219740726	ENVI 4.5, ENVI 4.7, ArcGIS 9	resampling of pixel resolution to 0.6 m, georeferenced to QuickBird image from 2007 with RMSE of 0.20–1.25, aerial photographs, georeferenced to QuickBird image from 2007 (RMSE of 0.17–1.07)	panchromatic	405–1053 nm	manual delineation	N/A	
rog00220070707	ENVI 4.5, ENVI 4.7, ArcGIS 9	N/A	panchromatic	405–1053 nm	manual delineation	N/A	

site ID	software	pre-processing	classification band(s)	spectral range	classification method	post-processing	data repository
rog00320070626	Classification in Definiens Professional 5.0, post processiong with ArcGIS	re-georeferenced using field measured GPS-points	B, G, R, Near-IR	blue: 450-520 nm, green: 520-600 nm, red: 630-690 nm, NIR: 760-890 nm	supervised, see more in INFO_Qbird_classification_version4.pdf	N/A	
rog00420070704	Classification in Definiens Professional 5.0, post processiong with ArcGIS	re-georeferenced using field measured GPS-points	B, G, R, Near-IR	blue: 450-520 nm, green: 520-600 nm, red: 630-690 nm, NIR: 760-890 nm	supervised, see more in INFO_Qbird_classification_version4.pdf	N/A	
rog00520070704	Classification in Definiens Professional 5.0, post processing with ArcGIS	re-georeferenced using field measured GPS-points	B, G, R, Near-IR	blue: 450-520 nm, green: 520-600 nm, red: 630-690 nm, NIR: 760-890 nm	supervised, see more in INFO_Qbird_classification_version4.pdf	N/A	
sam001200808xx	ENVI 4.8, ArcGIS 10	georeferencing	Near-infrared	N/A	N/A	river, streams, shadows removed	
sei00120070706	ArcGIS	N/A	panchromatic	450-900 nm	manual delineation	N/A	
sei00220070706	Classification in Definiens Professional 5.0, post processing with ArcGIS	re-georeferenced using field measured GPS-points	B, G, R, Near-IR,	blue: 450-520 nm, green: 520-600 nm, red: 630-690 nm, NIR: 760-890 nm	supervised, see more in INFO_Qbird_classification_version4.pdf	N/A	
ksal0012012xxxx	eCognition 8	Principal Component Analysis on visible and IR bands	G, R, NIR	N/A	object-oriented classification	river, streams, shadows removed	
ksl00119620628	ENVI 4.8, ArcGIS 10	georeferencing to Landsat ETM & TM (RMSE=10 m)	panchromatic	N/A	density slice	river, streams, shadows removed	
sur00120130802	ENVI 4.8, ArcGIS 10	orthorectification, gamma filter (10x10m)	X-Band	wavelength 31 mm, frequency 9.6 GHz	threshold classification	artefacts from streets and airport were removed manually	PeRL database
tbr00120100901, tea00120100901	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (11x11 pixel)	X-Band	wavelength 31 mm, frequency 9.6 GHz	k-means classification	river, streams, shadows removed	
tav00119630831	ENVI 4.5, ENVI 4.7, ArcGIS 9	resampling of pixel resolution to 0.6 m, georeferenced to IKONOS image from 2003	panchromatic	N/A	manual delineation	N/A	
tav00119750810	ENVI 4.5, ENVI 4.7, ArcGIS 9	resampling of pixel resolution to 0.6 m, georeferenced to IKONOS image from 2003 with RMSE of 0.60-2.38	panchromatic	N/A	manual delineation	N/A	

site ID	software	pre-processing	classification band(s)	spectral range	classification method	post-processing	data repository
tav00120030702	ENVI 4.5, ENVI 4.7, ArcGIS 9	resampling of pixel resolution to 0.6 m	panchromatic	760–850 nm	manual delineation	N/A	
tuk00120120723	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (7x7 pixel)	X-Band	wavelength 31 mm, frequency 9.6 GHz	k-means classification	river, streams, shadows removed	
yak0012009xxxx		georeferencing, histogram matching routine for radiometric normalization, mosaicking	B, G, R, RE, NIR	blue: 450–520 nm, green: 520–600 nm, red: 630–690 nm, NIR: 760–890 nm	object-oriented classification	N/A	http://geo.tuwien.ac.at/permafrost/
yam00120080824	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (7x7 pixel)	X-Band	wavelength 31 mm, frequency 9.6 GHz	density slice, DN: -2.7 to -1.9	river, streams, shadows removed	
yam00220100820	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (11x11 pixel)	X-Band	wavelength 31 mm, frequency 9.6 GHz	density slice, DN: 0 to 40	river, streams, shadows removed	
yfl0012011xxxx	eCognition 8, ArcGIS 10.2	orthorectification	IR, R, G, B	450–520 nm, 560–600 nm, 630–690 nm, 760–890 nm	object-oriented classification	N/A	
yuk00120090812, yuk00220090812	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (11x11 pixel)	X-Band	wavelength 31 mm, frequency 9.6 GHz	k-means classification: 15 classes, 5 iterations	river, streams, shadows removed	

Appendix C

Table C1: Description of attributes contained in the polygon attribute table of Alaskan permafrost landscapes (*alaska_perma_land.shp*)

Field name	Description	Source
ECOZONE	ecozone	AK2008
GEN_GEOL	generalized geology	AK2008
LITHOLOGY	texture	AK2008
GROUND_ICE	ground ice content [vol %]	AK2008
PF_EXTENT	permafrost extent	AK2008
PERMA_LAND	combined label of PF_EXTENT / GROUND_ICE/GEN_GEOL/LITHOLOGY	PeRL
ECOZID	ecozone ID	PeRL
PERMID	ID for each polygon in the vector file. The first digit stands for the region (1 – Alaska, 2 – Canada, 3 – Russia), digits 2 – 6 identify the single polygon, and the last three digits identify the ecozone.	PeRL
AREA	area of polygon in square meters	PeRL
PERIMETER	perimeter of polygon in square meters	PeRL

Table C2: Description of attributes contained in the polygon attribute table of Canadian permafrost landscapes (*canada_perma_land.shp*).

field name	description	source
ECOZONE	ecozone	NEF
ECOREGION	ecoregion	NEF
ECODISTRIC	ecodistrict	NEF
GEN_GEOL	dominant fraction of generalized (surficial) geology	NEF
LITHOLOGY	dominant fraction of texture	NEF
GROUND_ICE	dominant fraction of ground ice content in vol%	NEF
PF_EXTENT	dominant fraction of permafrost extent	NEF
PERMA_LAND	combined label of PF_EXTENT / GROUND_ICE/GEN_GEOL/LITHOLOGY	PeRL
ECOZID	ecozone ID	PeRL
PERMID	ID for each polygon in the vector file. The first digit stands for the region (1 – Alaska, 2 – Canada, 3 – Russia), digits 2 – 6 identify the single polygon, and the last three digits identify the ecozone.	PeRL
AREA	area of polygon in square meters	PeRL
PERIMETER	perimeter of polygon in square meters	PeRL

Table C3: Description of attributes contained in the polygon attribute table of Russian permafrost landscapes (*russia_perma_land.shp*).

field name	description	source
ECOZONE	Metadata: http://maps.tnc.org/files/metadata/TNC_Lands.xml	Olson et al., 2001 Downloaded at http://maps.tnc.org/gis_data.html#TerrEcos
GEN_GEOL	Surficial geology	LRR, Stolbovoi et al. (2002c)
LITHOLOGY	Texture	LRR, Stolbovoi et al. (2002c)
GROUND_ICE	ground ice content in vol%	LRR
PF_EXTENT	permafrost extent	LRR
PERMA_LAND	combined label of PF_EXTENT / GROUND_ICE/GEN_GEOL/LITHOLOGY	LRR
ECOZID	ecozone ID	PeRL
PERMID	ID for each polygon in the vector file. The first digit stands for the region (1 – Alaska, 2 – Canada, 3 – Russia), digits 2 – 6 identify the single polygon, and the last three digits identify the ecozone.	PeRL
AREA	area of polygon in square meters	PeRL
PERIMETER	perimeter of polygon in square meters	PeRL

Table C4: Description of attributes contained in the polygon attribute table of circum-arctic permafrost landscapes (*PeRL_study_area.shp*)

field name	description
country	country
Map_ID	ID of individual waterbody map
site	site name
MAAT	mean annual air temperature [°C]
TP	mean annual total precipitation [mm]
PE_DEPTH	Permafrost depth [m]
lat	latitude coordinate of polygon centroid in decimal degrees (WGS84)
long	longitude coordinate of polygon centroid in decimal degrees (WGS84)
AREA	area of polygon in square metres
AREA_SQKM	area of polygon in square kilometres

Appendix A~~D~~

Figure A̶D̶1: Study areas and associated permafrost landscapes in Alaska. Legend lists type of ~~physiography,~~ permafrost extent (C-continuous, D-discontinuous, S-sporadic), ground ice content [vol%], surficial geology and lithology. Shadowed labels name study areas with waterbody maps. Black lines and labels denote ecozones.

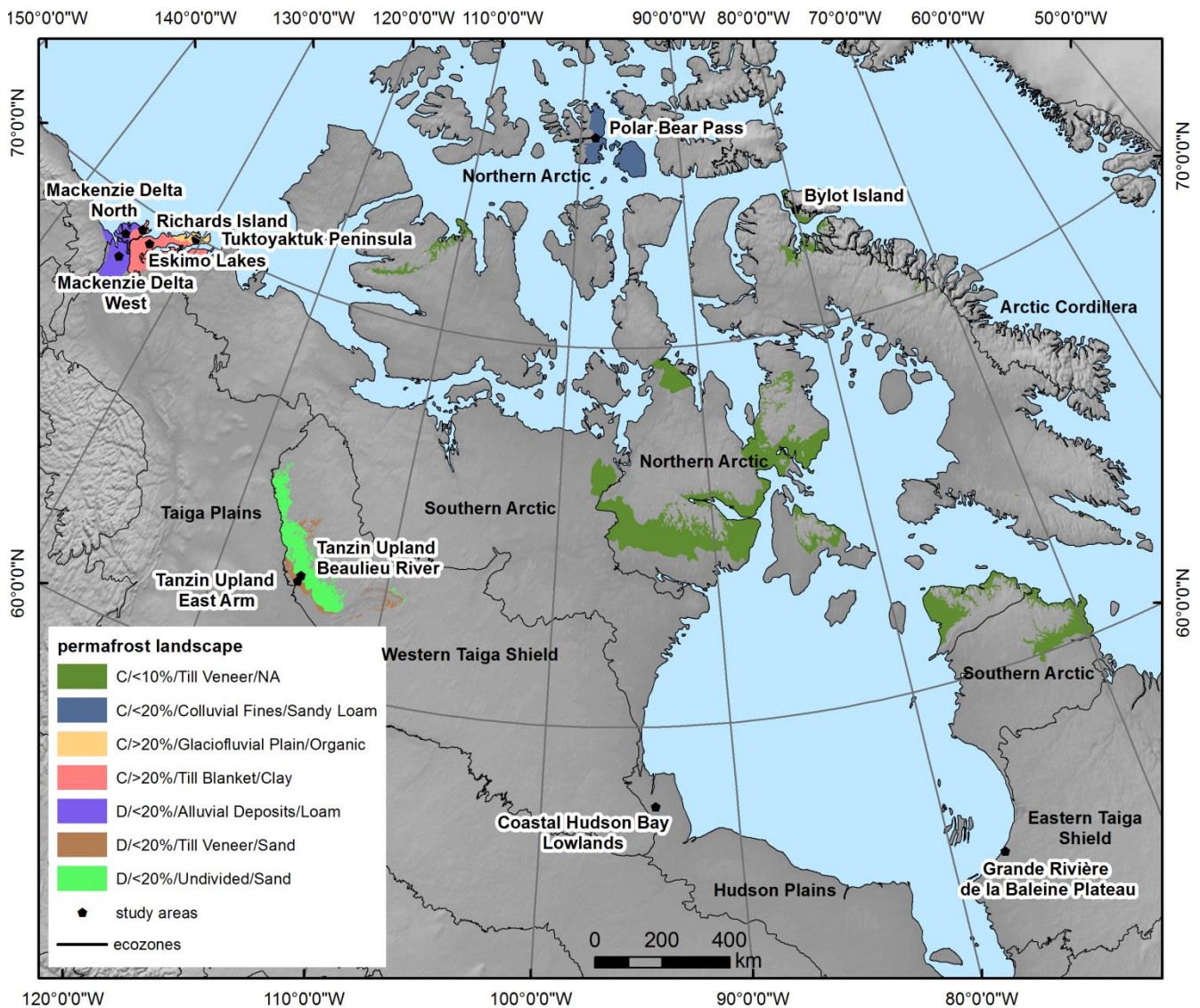



Figure A̶D̶2: Study areas and associated permafrost landscapes in Canada. Legend lists type of ~~physiography,~~ permafrost extent (C-continuous, D-discontinuous, S-sporadic), ground ice content [vol%], surficial geology and lithology. Shadowed labels name study areas with waterbody maps. Black lines and labels denote ecozones.

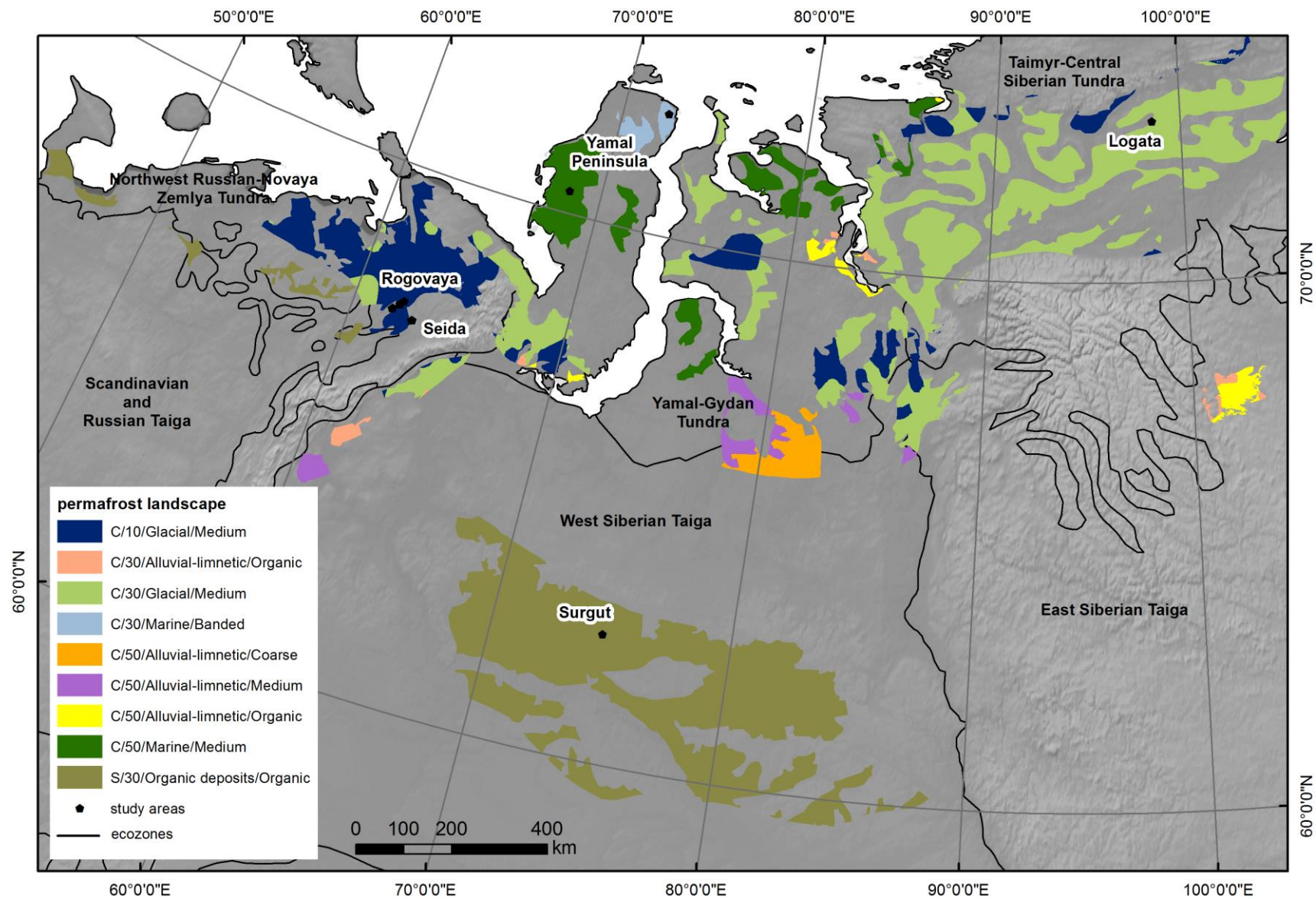



Figure ~~AD~~3: Study areas and associated permafrost landscapes in East Russia. Legend lists type of ~~physiography,~~ permafrost extent (C-continuous, D-discontinuous, S-sporadic), ground ice content [vol%], surficial geology and lithology. Shadowed labels name study areas with waterbody maps. Black lines and labels denote ecozones.

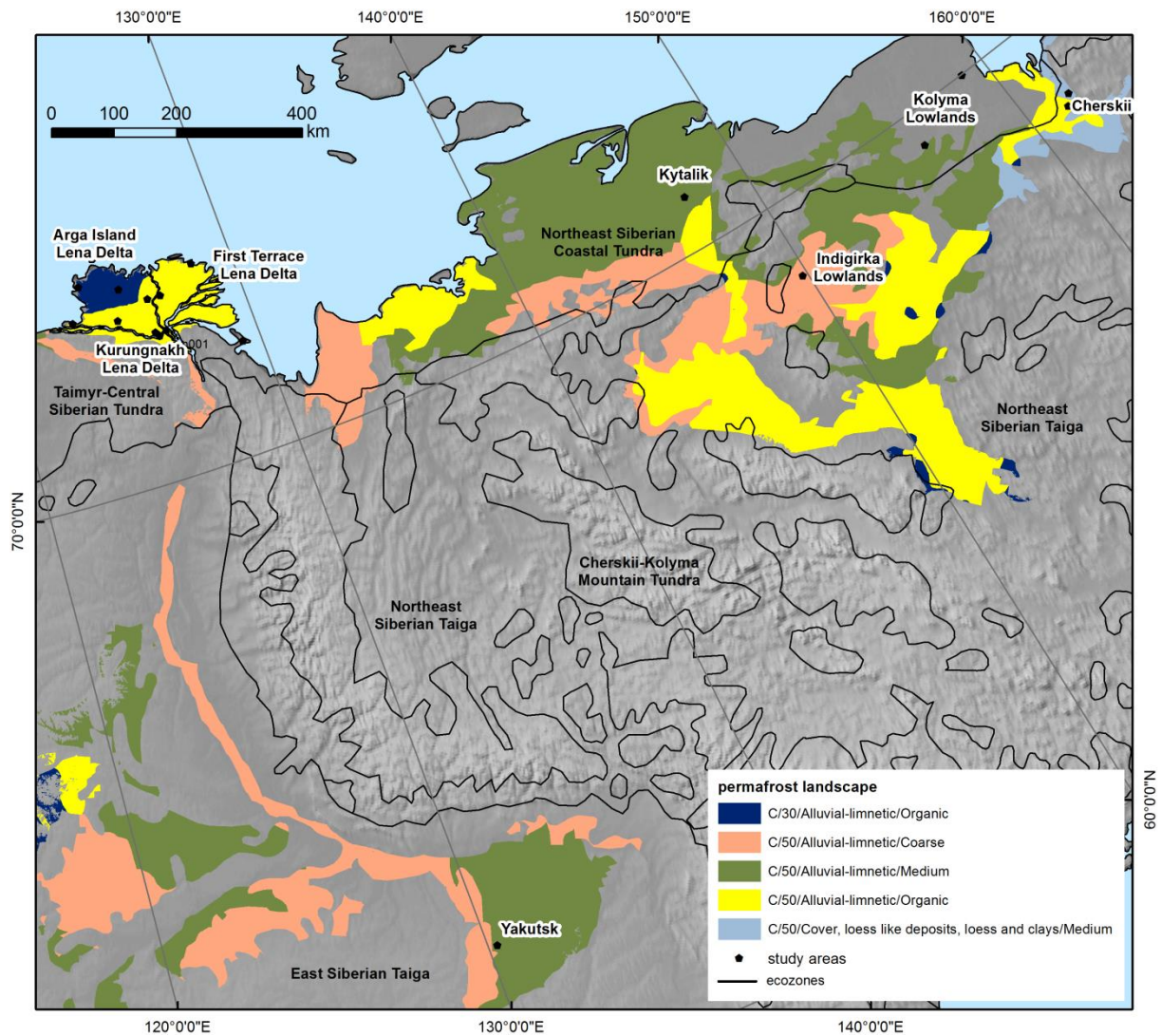



Figure AD4: Study areas and associated permafrost landscapes in West Russia. Legend lists type of ~~physiography,~~ permafrost extent (C-continuous, D-discontinuous, S-sporadic), ground ice content [vol%], surficial geology and lithology. Shadowed labels name study areas with waterbody maps. Black lines and labels denote ecozones.

Appendix BE

Table BE1: Areal fraction and density per waterbody map in Alaska. Map IDs with an asterisk were not used for extrapolation. F – areal fraction of waterbodies from 1.0 E+02 m² to 1.0 E+06 m² in size, REF – relative error of fraction for map subsets of 10x10 km. D – waterbody density per km², RED – relative error of density, PF – pond areal fraction for waterbodies from 1.0 E+02 m² to smaller than 5   1.0 E+04 m², REPF – relative error of pond fraction, PD – pond density, REPD – relative error of pond density.

Mmap ID	Mmap areaexten t [km²]	F [%]	REF [%]	D [# per km²]	RED [%]	PF [%]	REPF [%]	PD [# per km²]	REPD [%]
bar00119480804	1.9E+01	14	N/A	114	N/A	8	N/A	112	N/A
bar00120020802	1.9E+01	10	N/A	78	N/A	5	N/A	76	N/A
bar00120080730	1.9E+01	11	N/A	78	N/A	6	N/A	76	N/A
bar00120100810	1.9E+01	10	N/A	77	N/A	6	N/A	76	N/A
bar002wlc00120090 825	1.4E+03	7	5	17	3	2	3	17	3
bar00320020801wlc 00220020801	1.5E+02	11	N/A	39	N/A	3	N/A	38	N/A
bar00420090802wlc 00320090802	3.0E+02	6	N/A	13	N/A	1	N/A	12	N/A
elc00120090825	1.3E+02	7	N/A	18	N/A	1	N/A	17	N/A
elc00220020801	1.4E+02	9	N/A	48	N/A	3	N/A	47	N/A
elc00320090802	5.1E+01	7	N/A	18	N/A	1	N/A	17	N/A
fis00120020715	2.4E+02	13	N/A	24	N/A	2	N/A	23	N/A
imc00120040725	1.3E+03	11	5	21	10	2	11	20	10
kcp001201007xx	2.1E+01	22	N/A	49	N/A	7	N/A	46	N/A
kcp002201007xx	2.0E+01	121	N/A	15	N/A	0	N/A	1	N/A
kcp003201007xx	1.9E+01	24	N/A	58	N/A	7	N/A	55	N/A
ksl00119620628	5.6E+02	11	7	12	11	1	10	11	11
ksl0012012xxxx	5.6E+02	6	12	2	18	0	14	1	29
ksl0012012xxxx	5.6E+02	6	12	2	18	0	14	1	29
ycb0012011xxxx	1.0E+02	7	N/A	3	N/A	0	N/A	2	N/A
yuk00120090812	1.1E+03	10	5	6	4	1	3	5	4
yuk00220090812	5.8E+02	10	N/A	7	N/A	1	N/A	6	N/A

Table B~~E~~2: Areal fraction and density per waterbody map in Canada. Map IDs with an asterisk were not used for extrapolation. F – areal fraction of waterbodies from 1.0 E+02 m² to 1.0 E+06 m² in size, REF – relative error of fraction for map subsets of 10x10 km. D – waterbody density per km², RED – relative error of density, PF – pond areal fraction for waterbodies from 1.0 E+02 m² to smaller than 1.0 E+04 m², REPF – relative error of pond fraction, PD – pond density, REPD – relative error of pond density.

Map ID	Map extent [km²]	F [%]	REF [%]	D [# per km²]	RED [%]	PF [%]	REPF [%]	PD [# per km²]	REPD[%]
~~map ID~~	~~map area [km²]~~	~~F [%]~~	~~REF [%]~~	~~D [# per km²]~~	~~RED [%]~~	~~PF [%]~~	~~REPF [%]~~	~~PD # per km²]~~	~~REPD [%]~~
byl001 20160728~~20100902~~	3.6~~4.5~~E+01	4	N/A	51~~7~~	N/A	2	N/A	57~~1~~	N/A
esk00120090727	9.2E+02	14	4	2	3	0	4	1	5
gpr00119590707*	1.8E-01	11	N/A	359	N/A	11	N/A	359	N/A
gpr00120060707*	1.8E-01	11	N/A	326	N/A	11	N/A	326	N/A
hbl00119540701*	4.0E+00	36	N/A	60	N/A	6	N/A	57	N/A
hbl00119740617*	4.0E+00	38	N/A	73	N/A	7	N/A	69	N/A
hbl00120060706*	4.0E+00	35	N/A	60	N/A	6	N/A	56	N/A
mdn00120100716	1.5E+03	8	6	7	5	1	5	6	5
mdw00120090921	1.6E+03	21	1	8	2	1	2	5	2
pbp00120090813	6.9E+01	15	N/A	43	N/A	6	N/A	41	N/A
ric00120120925	5.9E+02	11	8	2	10	0	12	1	11
tbr00120100901	6.9E+02	11	3	3	2	0	3	2	3
tea00120100901	4.6E+02	8	14	7	2	1	2	6	1
tuk00120120723	4.8E+02	8	5	6	3	1	3	5	3

Table B~~E~~3: Areal fraction and density per waterbody map in Scandinavia. Map IDs with an asterisk were not used for extrapolation. F – areal fraction of waterbodies from 1.0 E+02 m² to 1.0 E+06 m² in size, REF – relative error of fraction for map subsets of 10x10 km. D – waterbody density per km², RED – relative error of density, PF – pond areal fraction for waterbodies from 1.0 E+02 m² to smaller than 1.0 E+04 m², REPF – relative error of pond fraction, PD – pond density, REPD – relative error of pond density.

Map ID	Map extent [km²]	F [%]	REF [%]	D [# per km²]	RED [%]	PF [%]	REPF [%]	PD [# per km²]	REPD[%]
~~map ID~~	~~map area [km²]~~	~~F [%]~~	~~REF [%]~~	~~D [# per km²]~~	~~RED [%]~~	~~PF [%]~~	~~REPF [%]~~	~~PD # per km²]~~	~~REPD [%]~~
abi0012010xxxx	3.8E+01	6	N/A	5	N/A	1	N/A	4	N/A
tav00119630831	8.5E-01	15	N/A	40	N/A	4	N/A	34	N/A
tav00119750810	8.5E-01	17	N/A	64	N/A	7	N/A	59	N/A
tav00120030702	8.5E-01	12	N/A	53	N/A	6	N/A	50	N/A

Table BE4: Areal fraction and density per waterbody map in Russia. Map IDs with an asterisk were not used for extrapolation. F – areal fraction of waterbodies from 1.0 E+02 m² to 1.0 E+06 m² in size, REF – relative error of fraction for map subsets of 10x10 km. D – waterbody density per km², RED – relative error of density, PF – pond areal fraction for waterbodies from 1.0 E+02 m² to smaller than 1.0 E+04 m², REPF – relative error of pond fraction, PD – pond density, REPD – relative error of pond density.

Map ID	Map extent [km²]	F [%]	REF [%]	D [# per km²]	RED [%]	PF [%]	REPF [%]	PD [# per km²]	REPD[%]
~~map ID~~	~~map area~~ [km²]	~~F~~ [%]	~~REF~~ [%]	~~D~~ [# per km²]	~~RED~~ [%]	~~PF~~ [%]	~~REPF~~ [%]	~~PD~~ # per km²	~~REPD~~ [%]
arg00120110829	2.0E+02	7	N/A	24	N/A	2	N/A	23	N/A
arg0022009xxxx*	5.0E+03	9	2	10	2	1	2	9	2
byk00120060709*	1.7E+02	8	N/A	29	N/A	1	N/A	28	N/A
che00120020709	2.2E+02	1	N/A	3	N/A	0	N/A	3	N/A
che00220090724	3.4E+02	17	10	4	6	0	6	2	6
fir00120090906	1.5E+02	9	N/A	36	N/A	2	N/A	35	N/A
~~fir00220110829~~arg00320110829	2.2E+02	13	N/A	19	N/A	1	N/A	18	N/A
fir002~~3~~2009xxxx*	9.8E+03	12	4	21	4	2	4	20	4
ice0032009xxxx*	7.9E+02	1	29	0	22	0	27	0	22
ind00120090907	6.5E+02	5	10	1	9	0	12	1	10
kol00119659721*	2.4E+03	6	4	2	10	0	11	2	12
kol00219659721*	2.6E+03	7	4	1	4	0	4	1	5
kur00120100805*	5.5E+01	8	N/A	14	N/A	1	N/A	13	N/A
kur00220080926	2.5E+02	8	N/A	6	N/A	1	N/A	5	N/A
kyt00120070728	2.6E+02	8	N/A	18	N/A	1	N/A	18	N/A
log00120110811	7.0E+01	4	N/A	5	N/A	0	N/A	4	N/A
ole00120060708	7.5E+01	11	N/A	94	N/A	4	N/A	93	N/A
rog00120070626*	1.0E+01	19	N/A	15	N/A	3	N/A	12	N/A
rog00219740726*	3.4E+00	32	N/A	33	N/A	3	N/A	28	N/A
rog00220070707*	3.4E+00	26	N/A	28	N/A	2	N/A	24	N/A
rog00320070626	6.0E+01	7	N/A	11	N/A	1	N/A	9	N/A
rog00420070704	6.2E+01	10	N/A	11	N/A	2	N/A	9	N/A
rog00520070704	6.3E+01	8	N/A	21	N/A	2	N/A	19	N/A
sam001200808xx*	1.6E+00	14	N/A	116	N/A	5	N/A	114	N/A
sei00120070706*	6.7E+00	8	N/A	49	N/A	4	N/A	48	N/A
sei00220070706	8.3E+01	1	N/A	9	N/A	1	N/A	8	N/A
sur00120130802	1.8E+03	17	2	20	2	3	2	18	2
yak0012009xxxx	2.0E+03	5	4	3	4	0	5	2	5
yam00120080824	1.3E+03	9	5	4	3	0	3	3	3
yam00220100820	1.0E+03	6	2	9	3	1	3	8	3

Author contributions

S.M. compiled the database, made the analysis and figures, and wrote the manuscript. S.L. assisted in data analysis. All authors contributed to the database compilation and assisted in the writing of the manuscript.

Acknowledgements

5   This work was supported by the Helmholtz Association through a grant (VH-NG 203) awarded to Sina Muster. WJR and CDK were supported by the US Department of Energy, BER, under the NGEE-Arctic project under contract no. DE-AC02-05CH11231. Guido Grosse was supported by ERC #338335. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the US Government.

TerraSAR-X data of Polar Bear Pass, Canada, were acquired through the German Space Agency (DLR) via the project
10 HYD0546. All other TerraSAR -X data in Alaska, Canada, and Russia were made available by DLR via PI agreement LAN1747 within the framework of the ESA funded DUE GlobPermafrost project, FP7 PAGE21 and the Austrian-Russian joint FWF project COLD Yamal (I 1401). RapidEye classifications have been available through the information system of the ESA DUE Permafrost project. Classifications of Rogovaya and Seida sites in Russia were conducted within the CARBONorth project (contract 036993).

15 Quickbird-2 and WorldView-1/-2 images (© DigitalGlobe) of Ikpikpuk Middle Coastal Plain, Alaska, and Indigirka Lowlands, Russia, were provided by the Polar Geospatial Center at the University of Minnesota through NSF AON project 1107481. The Geographic Information Network for Alaska (GINA) provided SPOT imagery of the Kotzebue Sound Lowlands, Alaska.

References

20 Abnizova, A., Siemens, J., Langer, M., and Boike, J.: Small ponds with major impact: The relevance of ponds and lakes in permafrost landscapes to carbon dioxide emissions, Global Biogeochem. Cycles, 26, 10.1029/2011gb004237, 2012.
Andresen, C. G., and Lougheed, V. L.: Disappearing Arctic tundra ponds: Fine-scale analysis of surface hydrology in drained thaw lake basins over a 65 year period (1948–2013), J Geophys Res-Biogeo, 120, 466–479, 2015.
Bartsch, A., and Seifert, F. M.: The ESA DUE Permafrost project-A service for high latitude research, Geoscience and
25 Remote Sensing Symposium (IGARSS), 2012 IEEE International, 2012, 5222-5225,
Boike, J., Grau, T., Heim, B., Günther, F., Langer, M., Muster, S., Gouttevin, I., and Lange, S.: Satellite-derived changes in the permafrost landscape of central Yakutia, 2000–2011: Wetting, drying, and fires, Global Planet. Change, 139, 116-127, 10.1016/j.gloplacha.2016.01.001, 2016.
Bouchard, F., Francus, P., Pienitz, R., Laurion, I., and Feyte, S.: Subarctic thermokarst ponds: Investigating recent
30 landscape evolution and sediment dynamics in thawed permafrost of northern Québec (Canada), Arct. Antarct. Alp. Res., 46, 251-271, 2014.
Bouchard, F., Laurion, I., Preskienis, V., Fortier, D., Xu, X., and Whiticar, M.: Modern to millennium-old greenhouse gases emitted from ponds and lakes of the Eastern Canadian Arctic (Bylot Island, Nunavut), Biogeosciences, 12, 7279-7298, 2015.
35 Bowling, L. C., Kane, D. L., Gieck, R. E., Hinzman, L. D., and Lettenmaier, D. P.: The Role of Surface Storage in a Low-Gradient Arctic Watershed, Water Resour. Res., 39, 1087, 10.1029/2002WR001466, 2003.
CAFF: Arctic Biodiversity Assessment 2013, Arctic Council, Akureyri, Iceland, 2013.
Carroll, M., Townshend, J., DiMiceli, C., Loboda, T., and Sohlberg, R.: Shrinking lakes of the Arctic: Spatial relationships and trajectory of change, Geophys. Res. Lett., 38, L20406, 2011.
40 Feng, M., Sexton, J. O., Channan, S., and Townshend, J. R.: A global, high-resolution (30-m) inland water body dataset for 2000: First results of a topographic–spectral classification algorithm, Int J Digit Earth, 9, 113-133, 2015.

Grosse, G., Romanovsky, V., Walter, K., Morgenstern, A., Lantuit, H., and Zimov, S.: Distribution of Thermokarst Lakes and Ponds at Three Yedoma Sites in Siberia, Ninth International Conference on Permafrost, University of Alaska, Fairbanks, USA, 2008, 551-556,

Grosse, G., Jones, B., and Arp, C. D.: Thermokarst Lakes, Drainage, and Drained Basins, 8, 29, 10.1016/B978-0-12-374739-6.00216-5, 2013.

Hinzman, L. D., Deal, C. J., McGuire, A. D., Mernild, S. H., Polyakov, I. V., and Walsh, J. E.: Trajectory of the Arctic as an integrated system, Ecol. Appl., 23, 1837-1868, 10.1890/11-1498.1, 2013.

Jones, B., Grosse, G., Arp, C., Jones, M., Anthony, K. M. W., and Romanovsky, V.: Modern thermokarst lake dynamics in the continuous permafrost zone, northern Seward Peninsula, Alaska, J. Geophys. Res., 116, G00M03, 10.1029/2011JG001666, 2011.

Jones, B. M., Gusmeroli, A., Arp, C. D., Strozzi, T., Grosse, G., Gaglioti, B. V., and Whitman, M. S.: Classification of freshwater ice conditions on the Alaskan Arctic Coastal Plain using ground penetrating radar and TerraSAR-X satellite data, Int. J. Remote Sens., 34, 8267-8279, 2013.

Jorgenson, M., Yoshikawa, K., Kanevskiy, M., Shur, Y., Romanovsky, V., Marchenko, S., Grosse, G., Brown, J., and Jones, B.: Permafrost characteristics of Alaska, Proceedings of the Ninth International Conference on Permafrost, 2008, 121-122,

Klonus, S., and Ehlers, M.: Pansharpening with TerraSAR-X and optical data, 3rd TerraSAR-X Science Team Meeting, Darmstadt, Germany, 2008, 25-26,

Kravtsova, V. I., and Rodionova, T. V.: Variations in size and number of thermokarst lakesin different permafrost regions: Spaceborne evidence, Earth's Cryosphere, XX, 75-81, 2016.

Labrecque, S., Lacelle, D., Duguay, C. R., Lauriol, B., and Hawkings, J.: Contemporary (1951-2001) evolution of lakes in the Old Crow Basin, Northern Yukon, Canada: Remote sensing, numerical modeling, and stable isotope analysis, Arctic, 225-238, 2009.

Langer, M., Westermann, S., Walter Anthony, K., Wischnewski, K., and Boike, J.: Frozen ponds: production and storage of methane during the Arctic winter in a lowland tundra landscape in northern Siberia, Lena River Delta, Biogeosciences, 12, 977-990, 10.5194/bg-12-977-2015, 2015.

Lara, M. J., McGuire, A. D., Euskirchen, E. S., Tweedie, C. E., Hinkel, K. M., Skurikhin, A. N., Romanovsky, V. E., Grosse, G., Bolton, W. R., and Genet, H.: Polygonal tundra geomorphological change in response to warming alters future CO_2 and CH_4 flux on the Barrow Peninsula, Global change biology, 21, 1634-1651, 2015.

Laurion, I., Vincent, W. F., MacIntyre, S., Retamal, L., Dupont, C., Francus, P., and Pienitz, R.: Variability in greenhouse gas emissions from permafrost thaw ponds, Limnol. Oceanogr., 55, 115–133, 2010.

Lehner, B., and Döll, P.: Development and validation of a global database of lakes, reservoirs and wetlands, J. Hydrol., 296, 1–22, 10.1016/j.jhydrol.2004.03.028, 2004.

Liao, A., Chen, L., Chen, J., He, C., Cao, X., Chen, J., Peng, S., Sun, F., and Gong, P.: High-resolution remote sensing mapping of global land water, Science China Earth Sciences, 57, 2305-2316, 2014.

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

Marshall, I., Schut, P., and Ballard, M.: A national ecological framework for Canada: Attribute data, Agriculture and AgriFood Canada, Research Branch, Centre for Land and Biological Resources Research, and Environment Canada, State of the Environment Directorate, Ecozone Analysis Branch, Ottawa/Hull, Canada, 1999.

Muster, S., Langer, M., Heim, B., Westermann, S., and Boike, J.: Subpixel heterogeneity of ice-wedge polygonal tundra: a multi-scale analysis of land cover and evapotranspiration in the Lena River Delta, Siberia, Tellus B, 64, 10.3402/tellusb.v64i0.17301, 2012.

Muster, S.: Decomposing Arctic Land Cover. Implications of heterogeneity and scale for the estimation of energy fluxes in Arctic tundra landscapes, Universität Heidelberg, 121 pp., 2013.

Muster, S., Heim, B., Abnizova, A., and Boike, J.: Water body distributions across scales: A remote sensing based comparison of three Arctic tundra wetlands, Remote Sensing, 5, 1498–1523, 10.3390/rs5041498, 2013.

Palmtag, J., Ramage, J., Hugelius, G., Gentsch, N., Lashchinskiy, N., Richter, A., and Kuhry, P.: Controls on the storage of organic carbon in permafrost soil in northern Siberia, European Journal of Soil Science, 67, 478-491, 2016.

Paltan, H., Dash, J., and Edwards, M.: A refined mapping of Arctic lakes using Landsat imagery, Int. J. Remote Sens., 36, 5970–5982, 10.1080/01431161.2015.1110263, 2015.

Ramsar Convention Secretariat: Designating ramsar sites: strategic framework and guidelines for the future development of the list of wetland for international importance, 4 ed., Ramsar handbooks for the wise use of wetlands, Ramsar Convention Secretariat, Gland, Switzerland, 2010.

Rautio, M., Dufresne, F., Laurion, I., Bonilla, S., Vincent, W., and Christoffersen, K.: Shallow freshwater ecosystems of the circumpolar Arctic, Ecoscience, 18, 204–222, 2011.

Raynolds, M. K., Walker, D. A., Ambrosius, K. J., Brown, J., Everett, K. R., Kanevskiy, M., Kofinas, G. P., Romanovsky, V. E., Shur, Y., and Webber, P. J.: Cumulative geoecological effects of 62 years of infrastructure and climate change in ice-rich permafrost landscapes, Prudhoe Bay Oilfield, Alaska, Global Change Biology, 20, 1211-1224, 10.1111/gcb.12500, 2014.

5 Sannel, A. B. K., and Brown, I. A.: High-resolution remote sensing identification of thermokarst lake dynamics in a subarctic peat plateau complex, Canadian Journal of Remote Sensing, 36, S26-S40, 10.5589/m10-010, 2010.

Sannel, A. B. K., and Kuhry, P.: Warming-induced destabilization of peat plateau/thermokarst lake complexes, Journal of Geophysical Research, 116, 10.1029/2010JG001635, 2011.

Schuur, E. A. G., Bockheim, J. G., Canadell, J. G., Euskirchen, E. S., Field, C. B., Goryachkin, S. V., Hagemann, S., Kuhry,
10 P., Lafleur, P. M., Lee, H., Mazhitova, G., Nelson, F. E., Rinke, A., Romanovsky, V. E., Shiklomanov, N., Tarnocai, C., Venevsky, S., Vogel, J. G., and Zimov, S. A.: Vulnerability of Permafrost Carbon to Climate Change: Implications for the Global Carbon Cycle, Bioscience, 58, 701-714, 10.1641/B580807, 2008.

Sjöberg, Y., Hugelius, G., and Kuhry, P.: Thermokarst lake morphometry and erosion features in two peat plateau areas of northeast European Russia, Permafr. Periglac. Proc., 24, 75-81, 2013.

15 Smith, L. C., Sheng, Y., and MacDonald, G. M.: A First Pan-Arctic Assessment of the Influence of Glaciation, Permafrost, Topography and Peatlands on Northern Hemisphere Lake Distribution, Permafr. Periglac. Proc., 18, 201–208, 10.1002/ppp.581, 2007.

Veremeeva, A. A., and Glushkova, N. V.: Relief formation in the regions of the ice complex deposit occurrence: Remote sensing and gis□studies in the kolyma lowland tundra, Earth's Cryosphere, XX, 15-25, 2016.

20 Verpoorter, C., Kutser, T., Seekell, D. A., and Tranvik, L. J.: A global inventory of lakes based on high-resolution satellite imagery, Geophys. Res. Lett., 41, 6396–6402, 10.1002/2014GL060641, 2014.

Walker, D. A., Webber, P. J., Walker, M. D., Lederer, N. D., Meehan, R. H., and Nordstrand, E. A.: Use of geobotanical maps and automated mapping techniques to examine cumulative impacts in the Prudhoe Bay Oilfield, Alaska, Environ. Conserv., 13, 149-160, 1986.

25 Watts, J., Kimball, J., Jones, L., Schroeder, R., and McDonald, K.: Satellite Microwave remote sensing of contrasting surface water inundation changes within the Arctic–Boreal Region, Remote Sens. Environ., 127, 223–236, 2012.

Watts, J. D., Kimball, J. S., Bartsch, A., and McDonald, K. C.: Surface water inundation in the boreal-Arctic: potential impacts on regional methane emissions, Environ. Res. Lett., 9, 075001, 2014.

Widhalm, B., Hogström, E., Ressl, C., Trofaier, A., Heim, B., Biasi, C., and Bartsch, A.: PAGE21 WP5-Land surface
30 hydrology from remotely sensed data at PAGE21 sites, Department of Geodesy and Geoinformation (GEO), Research Groups Photogrammetry and Remote Sensing, Vienna University of Technology, Vienna, Austria, 15, 2014a.

Wik, M., Varner, R. K., Anthony, K. W., MacIntyre, S., and Bastviken, D.: Climate-sensitive northern lakes and ponds are critical components of methane release, Nat. Geosci., 9, 99–105, 10.1038/ngeo2578, 2016.

Supporting Information for

PeRL: A Circum-Arctic Permafrost Region Pond and Lake Database

Sina Muster[1], Kurt Roth[2], Moritz Langer[3], Stephan Lange[1], Fabio Cresto Aleina4, Annett Bartsch5, Anne Morgenstern[1], Guido Grosse[1], Benjamin Jones6, A. Britta K. Sannel7, Ylva Sjöberg7, Frank Günther[1], Christian Andresen8, Alexandra Veremeeva9, Prajna R. Lindgren10, Frédéric Bouchard[11],[13], Mark J. Lara[12], Daniel Fortier13, Simon Charbonneau13, Tarmo A. Virtanen14, Gustaf Hugelius7, Juri Palmtag7, Matthias B. Siewert7, William J. Riley15, Charles D. Koven15, and Julia Boike[1]

[1] Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Telegrafenberg A43, 14473 Potsdam, Germany
[2] Institute for Environmental Physics, Heidelberg University , Germany
[3] Humboldt University, Berlin, Germany
[4] Max Planck Institute for Meteorology, Hamburg, Germany
[5] Zentralanstalt für Meteorologie and Geodynamik, Vienna , Austria
[6] U.S. Geological Survey - Alaska Science Center, Anchorage, AK 99508, USA
[7] Stockholm University, Department of Physical Geography and the Bolin Centre for Climate Research, 10691 Stockholm, Sweden
[8] Los Alamos National Laboratory, Los Alamos, NM, USA
[9] Institute of Physicochemical and Biological Problems in Soil Science, Russian Academy of Sciences, Pushchino, Russia
[10] Geophysical Institute, University of Alaska Fairbanks, Fairbanks, AK, USA
[11] Institut national de la recherche scientifique (INRS), Centre Eau Terre Environnement (ETE), Québec QC, G1K 9A9, Canada
[12] Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[13] Geography Department, University of Montréal, Montréal QC, H3C 3J7, Canada
[14] Department of Environmental Sciences, University of Helsinki, Finland
[15] Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, USA

Correspondence to: Sina Muster (sina.muster@awi.de)

Contents of this file

Introduction

This supporting information document includes additional information on image processing and classification as well as the processing of regional permafrost maps.

Processing of TerraSAR-X imagerz is described in Text S1. Detailed metadata of base imagery, image preprocessing, classification method and post-processing are listed in Table S1.

Text S2 explains the subgrid sampling of larger study areas for the representative calculation of waterbody statistics. Texts S3, S4, S5, and S6 describe original production and properties of permafrost landscapes for Alaska, Canada, and Russia and their harmonization for use in this study. Attributes of the vector files of each permafrost landscape region are listed and explained in tables S2, S3, and S4. Vector files of the regional permafrost landscapes were joined to a harmonized circum-arctic vector file. Table S5 lists the attributes of the circum-arctic vector file.

Table S7 lists metadata and references used for the characterization of climate and permfrost depth for each study area.

Text S1. Processing of TerraSAR-X imagery

Geocoded EEC products obtained from DLR are delivered in radar brightness. They are projected to the best available DEM, i.e., SRTM X band DEMs (30 m resolution) and SRTM C band DEMs (90 m resolution). For the remaining areas, the 1 km resolution GLOBE DEM is used. The EEC is a detected multi-look product with reduced speckle and approximately square cells on ground. The slant-range resolution of the image is 1.2 m, which corresponds to 3.3−3.5 m projected on the ground for incidence angles between 45° and 20° and an azimuth resolution of 3.3 m (Eineder et al., 2008). SSC were geocoded to the DUE Permafrost DEM and no multi-looking was applied.

Text S2. Subgrid sampling

In large study areas we performed a subgrid analysis, i.e. we selected waterbodies within equally sized boxes and averaged statistics from all boxes of the study area. In order to determine a representative box size we compared the variability of waterbody distribution statistics within three study areas in Russia, Canada, and Alaska. In each study area, we selected waterbodies from a minimum of 5 and up to 50 randomly distributed boxes with varying sizes of 5x5 km, 10x10 km, and 20x20 km. We calculated the standard error (SE) of the mean of all statistics across all boxes of the same size. SE of density (waterbody number per km²) and waterbody mean surface area was lowest for 10x10 km boxes. SE increased for 20x20 km boxes which is probably due to the significantly lower number of boxes that could be sampled. Only 12 PeRL sites have a study area larger than 1000 km² that would allow to sample a minimum of 5 boxes of 20x20 km in size. A box size of 10x10 km allows the subsampling of 26 sites with a minimum of 5 boxes. Taking into account the overall variability of distributions and the possible number of subgrid samples, a box size of 10x10 km was chosen for subgrid

analysis. Subgrid analysis was conducted for study areas larger than 300 km².

Text S3. Alaskan permafrost landscape maps

The ~~PL~~ permafrost landscape map of Alaska reports surficial geology, MAAT, primary soil texture, permafrost extent, ground ice volume, and primary thermokarst landforms (Jorgenson et al., 2008b). A rule-based model was used to incorporate MAAT and surficial geology. Permafrost characteristics were assigned to each surficial deposit under varying temperatures using terrain-permafrost relationships and expert knowledge (Jorgenson et al., 2008b).

Text S4. Canadian permafrost landscape maps

Permafrost landscapes of Canada are described in the National Ecological Framework (NEF). The NEF distinguishes four levels of generalization nested within each other. Ecozones represent the largest and most generalized units followed by ecoprovinces, ecoregions, and ecodistricts. Ecodistricts were delineated based mainly on differences in parent material, topography, landform and soil development derived from the (Soil Landscapes of Canada Working Group, 2010) at a map scale of 1:3,000,000 to 1:1,000,000 (Ecological Stratification Working Group, 1995;Marshall et al., 1999) whereas ecoregions and ecoprovinces are generalized based mainly on climate, physiography, and vegetation. Ecodistricts were therefore chosen as most appropriate to delineate permafrost landscapes. NEF reports the areal fraction of the underlying soil landscape units and attributes nested within each ecodistrict. The dominant fraction of surficial geology, lithology, permafrost extent and ground ice volume was chosen to describe each ecodistrict. Ecodistricts with the same permafrost landscape type within the same ecozone were then merged to PL units.

Text S5. Russian permafrost landscape characterization

In Russia, information about permafrost extent, ground ice content, generalized geology and lithology was available only as ~~individual~~ separate vector maps (Stolbovoi and McCallum, 2002). The individual maps were combined in ArcGIS 10.4 to delineate Russian ~~PL~~ permafrost landscape units similar to the Canadian and Alaskan databases. Russian ecozones were mapped using the global-scale map by Olson et al. (2001) which conforms to the Alaskan and Canadian ecozones. The geometric union of ecozone, ground ice content, and permafrost extent was calculated in ArcGIS 10.1 via the tool "intersect". Each unique combination of these three variables was assigned the dominant fraction of geology and lithology type.

Table S1: Metadata about image processing and classification. PREP: pre-processing steps; bands used for classification with their corresponding spectral range; POP: post-processing steps (ME-manual editing includes removal of rivers, streams, partial lakes, shadows, streets and gap-filling).

Map ID	Imagery	Software	PREP	Band(s)	Spectral range	Method	POP	Reference
abi0012010xxxx	Orthophoto (Lantmäteriet, I2014/00691); SPOT5; DEM	GDAL, SAGA GIS, Orfeo toolbox	1. Image segmentation using Orthophoto RGB (1m) and DEM (2m), minimum object size 130m2; 2. Classification of watermask: SVM classifier using Red band and slope	red (1m) + slope (2m)	610-680 nm	segmentation -> SVM	ME	this paper
arg00120110829	KOMPSAT-2	ENVI 4.8, ArcGIS 12	georeferencing	NIR	760-900 nm	density slice	ME	this paper
arg0022009xxxx, fir0022009xxxx, ice0032009xxxx	RapidEye	N/A	georeferencing, histogram matching routine for radiometric normalization, mosaiking	B, G, R, RE, NIR	blue: 440-510 nm, green: 520-590 nm, red: 630-685 nm, RE: 690-730 nm, NIR: 760-850 nm	object-oriented classification	N/A	Bartsch and Seifert (2012)
arg00320110711	KOMPSAT-2	ENVI 4.8, ArcGIS 10	georeferencing	NIR	760-900 nm	density slice	ME	Andresen and Lougheed (2015)
bar00119480804	aerial imagery	ENVI 4.4	coregistration and orthorectification using an image-to-image correction (RMSE of 1.0 m)	PAN	N/A	object-oriented classification	ME	Andresen and Lougheed (2015)
bar00120020802	Quickbird	ENVI 4.4	coregistration and orthorectification using an image-to-image correction (RMSE of 1.0 m)	PAN	445-900 nm	object-oriented classification	ME	Andresen and Lougheed (2015)
bar00120080730	Quickbird	ENVI 4.4	coregistration and orthorectification using an image-to-image correction (RMSE of 1.0 m)	PAN	445-900 nm	object-oriented classification	ME	Andresen and Lougheed (2015)
bar00120100810	WorldView-2	ENVI 4.4	coregistration and orthorectification using an image-to-image correction (RMSE of 1.0 m)	PAN	450-800 nm	object-oriented classification	ME	this paper
byk00120060709	SPOT-5	ArcGIS	georeferenced to topographic maps of scale 1:100,000	PAN	480-710 nm	density slice	ME	Lara et al. (2015)
byl00120100902	GeoEye-1	Geomatica 2015, ArcGIS 10.2	pan-sharpening, orthorectification	B, G, R, NIR	450-510 nm, 510-580 nm, 655-690 nm, 780-920 nm	unsupervised k-means classification	ME	Muster et al. (2013)
che00120020709	IKONOS-2	ArcGIS	orthorectification	PAN	760-850 nm	density slice	ME	Grosse et al. (2008)
che00220090724	ALOS PRISM	N/A	N/A	PAN	520-770 nm	density slice	ME	this paper
elc00120090825, wlc00120090825	TerraSAR-X	ENVI 4.4	gamma filter (11x11 pixel)	X	wavelength 31 mm, frequency 9.6 GHz	unsupervised k-means classification	ME	Grosse et al. (2008)
elc00220020801, wlc00220020801	Quickbird	ArcGIS 10.2, eCognition 9.1	principle component analysis on raster bands, pan-sharpening	R, B, G, IR	450-520 nm, 560-600 nm, 630-690 nm, 760-890 nm	object-oriented classification	ME	Widhalm et al. (2014a);Widhalm et

Map ID	Imagery	Software	PREP	Band(s)	Spectral range	Method	POP	Reference
elc00320090802, wlc00320090802	KOMPSAT-2	ENVI 4.8, ArcGIS 10	georeferencing	NIR	760-900 nm	density slice	ME	Lara et al. 2015
elc004200808xx	Quickbird	ArcGIS 10.2, eCognition 9.2	principle component analysis on raster bands, pan-sharpening	R,B,G,IR	450-520 nm, 560-600 nm, 630-690 nm, 760-890 nm	object-oriented classification	N/A	this paper
esk00120090727	TerraSAR-X	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (11x11 pixel)	X	wavelength 31 mm, frequency 9.6 GHz	unsupervised k-means classification	ME	this paper
fir00120090906	KOMPSAT-2	ENVI 4.8, ArcGIS 10	georeferencing	NIR	760-900 nm	density slice	ME	this paper
fis00120020715	Airborne Orthorectified Radar Imagery (IFSAR)	ENVI 4.5	N/A	PAN	bandwidth of 135 or 270 MHz	density slice	ME	Jones et al. (2013)
grp00119590707	aerial imagery	ArcGIS 10.0	Scanning (1814 dpi); georeferencing (RMSE = 1.8m (4.2pix))	black and white	N/A	manual	N/A	Bouchard et al. (2014)
grp00120060707	QuickBird	ArcGIS 10.0	geometric correction (cubic convolution resampling)	PAN	450-900 nm	manual	N/A	Bouchard et al. (2014)
hbl00119540701	aerial imagery	ENVI 4.5, ENVI 4.7, ArcGIS 9	resampling of pixel resolution to 0.6 m, georeferenced to QuickBird image from 2006 with RMSE of 0.38–1.42	PAN	N/A	manual delineation	N/A	Sannel and Brown (2010);Sannel and Kuhry (2011)
hbl00119740617	aerial imagery	ENVI 4.5, ENVI 4.7, ArcGIS 9	resampling of pixel resolution to 0.6 m, georeferenced to QuickBird image from 2006 RMSE of 0.38–1.42	PAN	N/A	manual delineation	N/A	Sannel and Brown (2010);Sannel and Kuhry (2011)
hbl00120060706	QuickBird	ENVI 4.5, ENVI 4.7, ArcGIS 9	N/A	PAN	405-1053 nm	manual delineation	N/A	Sannel and Brown (2010);Sannel and Kuhry (2011)
imc00120040725	QuickBird-2	ENVI 5.0, ArcGIS 10.3	orthorectification	PAN	760-850 nm	density slice	ME	this paper
ind00120090907	WorldView-1	ENVI 5.0, ArcGIS 10.3	orthorectification	PAN	400-900 nm	density slice, opening filter 3x3	ME	this paper
kol00119650721, kol00219650721	CORONA	ArcGIS 9	georeferencing	PAN	N/A	manual delineation	N/A	this paper
kpc001201007xx, kpc002201007xx, kpc003201007xx	digital true color aerial imagery	N/A	N/A	R, G, B	N/A	manual delineation	N/A	Walker et al. (1986);Raynolds et al. (2014)
ksl00119620628	Corona KH4	ENVI 4.8, ArcGIS 10	georeferencing to Landsat ETM & TM (RMSE=10 m)	PAN	N/A	density slice	ME	this paper
ksl0012012xxxx	SPOT-5	eCognition 8	Principal Component Analysis on visible and IR bands	G, R, NIR	N/A	object-oriented classification	ME	this paper
kur00120100805	Geoeye	ENVI 4.8, ArcGIS 10	pan-sharpening, orthorectification (RMSE: 0.36m)	NIR	760-900 nm	unsupervised k-means classification	ME	this paper
kur00220080926	ALOS PRISM	PCI Geomatica,	orthorectification based on own	PAN	520-770 nm	manual delineation	ME	this paper

Map ID	Imagery	Software	PREP	Band(s)	Spectral range	Method	POP	Reference
		ArcGIS 9	stereo DEM					
kyt00120070728	ALOS PRISM	ENVI 4.8, ArcGIS 10	N/A	PAN	520-770 nm	density slice	ME	Widhalm et al. (2014a);Widhalm et al. (2014b)
log00120110811	QuickBird	ENVI 5.1	N/A	NIR, R, G, B	485-830 nm	supervised maximum-likelihood classification	ME	Palmtag et al. (2016)
mdn00120090921	TerraSAR-X	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (11x11 pixel)	X	wavelength 31 mm, frequency 9.6 GHz	density slice, DN: -28.06 to -14.79	ME	this paper
mdn00120100716	TerraSAR-X	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (11x11 pixel)	X	wavelength 31 mm, frequency 9.6 GHz	k-means classification: 15 classes, 5 iterations, class 1,2,3 as water identified	ME	this paper
ole00120060708	SPOT-5	ArcGIS	georeferenced to topographic maps of scale 1:100,000	PAN	480-710 nm	density slice	ME	Grosse et al. (2008)
pbp00120090813	TerraSAR-X	ENVI 4.8, ArcGIS 10	gamma filter (11x11 pixel)	X	wavelength 31 mm, frequency 9.6 GHz	density slice (brightness values smaller −21.55 σ db were classified as open water)	majority filter (7x7 pixel) to reduce spurious pixels in classification, ME	(Muster et al., 2013)
ric001201209125	TerraSAR-X	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (11x11 pixel)	X	wavelength 31 mm, frequency 9.6 GHz	k-means classification: 15 classes, 5 iterations; classes 2,3,4 were identified as water	ME	this paper
rog00120070626	QuickBird	ArcGIS	N/A	PAN	450-900 nm	manual delineation	N/A	Sjöberg et al. (2013)
rog00219740726	aerial imagery	ENVI 4.5, ENVI 4.7, ArcGIS 9	resampling of pixel resolution to 0.6 m, georeferenced to QuickBird image from 2007 with RMSE of 0.20-1.25, aerial photographs, georeferenced to QuickBird image from 2007 (RMSE of 0.17-1.07)	PAN	405-1053 nm	manual delineation	N/A	Sannel and Kuhry (2011)
rog00220070707	QuickBird	ENVI 4.5, ENVI 4.7, ArcGIS 9	N/A	PAN	405-1053 nm	manual delineation	N/A	Sannel and Kuhry (2011)
rog00320070626	QuickBird-2	Classification in Definiens Professional 5.0, post processiong with ArcGIS	re-georeferenced using field measured GPS-points	B, G, R, NIR	blue: 450-520 nm, green: 520-600 nm, red: 630-690 nm, NIR: 760-890 nm	supervised, see more in INFO_Qbird_classification_ version4.pdf	ME	this paper
rog00420070704	QuickBird-2	Classification in Definiens Professional 5.0, post	re-georeferenced using field measured GPS-points	B, G, R, NIR	blue: 450-520 nm, green: 520-600 nm, red: 630-690 nm,	supervised, see more in INFO_Qbird_classification_ cation_	ME	this paper

Map ID	Imagery	Software	PREP	Band(s)	Spectral range	Method	POP	Reference
		processiong with ArcGIS			NIR: 760-890 nm	version4.pdf		
rog00520070704	QuickBird-2	Classification in Definiens Professional 5.0, post processing with ArcGIS	re-georeferenced using field measured GPS-points	B, G, R, NIR	450-520 nm, 520-600 nm, 630-690 nm, 760-890 nm	supervised, see more in INFO_Qbird_classification_version4.pdf	ME	this paper
sam001200808xx	aerial imagery	ENVI 4.8, ArcGIS 10	georeferencing	NIR	N/A	N/A	ME	Muster et al. (2012)
sei00120070706	QuickBird	ArcGIS	N/A	PAN	450-900 nm	manual delineation	N/A	Sjöberg et al. (2013)
sei00220070706	QuickBird-2	Classification in Definiens Professional 5.0, post processing with ArcGIS	re-georeferenced using field measured GPS-points	B, G, R, NIR	blue: 450-520 nm, green: 520-600 nm, red: 630-690 nm, NIR: 760-890 nm	supervised, see more in INFO_Qbird_classification_version4.pdf	ME	this paper
sur00120130802	TerraSAR-X	ENVI 4.8, ArcGIS 10	orthorectification, gamma filter 10x10m	X	wavelength 31 mm, frequency 9.6 GHz	threshold classification	ME	this paper
tav00119630831	aerial imagery	ENVI 4.5, ENVI 4.7, ArcGIS 9	resampling of pixel resolution to 0.6 m, georeferenced to IKONOS image from 2003	PAN	N/A	manual delineation	N/A	Sannel and Kuhry (2011)
tav00119750810	aerial imagery	ENVI 4.5, ENVI 4.7, ArcGIS 9	resampling of pixel resolution to 0.6 m, georeferenced to IKONOS image from 2003 with RMSE of 0.60-2.38	PAN	N/A	manual delineation	N/A	Sannel and Kuhry (2011)
tav00120030702	IKONOS	ENVI 4.5, ENVI 4.7, ArcGIS 9	resampling of pixel resolution to 0.6 m	PAN	760-850 nm	manual delineation	N/A	Sannel and Kuhry (2011)
tbr00120100901, tea00220100901	TerraSAR-X	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (11x11 pixel)	X	wavelength 31 mm, frequency 9.6 GHz	k-means classification: 15 classes, 5 iterations	ME	this paper
tuk00120120723	TerraSAR-X	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (7x7 pixel)	X	wavelength 31 mm, frequency 9.6 GHz	k-means classification: 15 classes, 5 iterations	ME	this paper
yak0012009xxxx	RapidEye		georeferencing, histogram matching routine for radiometric normalization, mosaicking	B, G, R, RE, NIR	blue: 440-510 nm, green: 520-590 nm, red: 630-685 nm, RE: 690-730 nm, NIR: 760-850 nm	object-oriented classification	ME	Bartsch and Seifert (2012)
yam00120080824	TerraSAR-X	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (7x7 pixel)	X	wavelength 31 mm, frequency 9.6 GHz	density slice, DN: -2.7 to -1.9	ME	this paper
yam00220100820	TerraSAR-X	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (11x11 pixel)	X	wavelength 31 mm,	density slice, DN: 0 to 40	ME	this paper

Map ID	Imagery	Software	PREP	Band(s)	Spectral range	Method	POP	Reference
yfl0012011xxxx	aerial imagery	eCognition 8, ArcGIS 10.2	orthorectification	NIR, R, G, B	frequency 9.6 GHz 450-520 nm, 560-600 nm, 630-690 nm, 760-890 nm	object-oriented classification	ME	this paper
yuk00120090812, yuk00220090812	TerraSAR-X	ENVI 4.8, ArcGIS 10	lee filter (3x3 pixel), gamma filter (11x11 pixel)	X	wavelength 31 mm, frequency 9.6 GHz	k-means classification: 15 classes, 5 iterations	ME	this paper

Table S2. Attributes contained in the polygon attribute table of Alaskan permafrost landscapes (*alaska_perma_land.shp*)

Field name	Description	Source
ECOZONE	ecozone	AK2008
GEN_GEOL	generalized geology	AK2008
LITHOLOGY	texture	AK2008
GROUND_ICE	ground ice content [vol %]	AK2008
PF_EXTENT	permafrost extent	AK2008
PERMA_LAND	combined label of PF_EXTENT / GROUND_ICE/GEN_GEOL/LITHOLOGY	PeRL
ECOZID	ecozone ID	PeRL
PERMID	ID for each polygon in the vector file. The first digit stands for the region (1 – Alaska, 2 – Canada, 3 – Russia), digits 2 – 6 identify the single polygon, and the last three digits identify the ecozone.	PeRL
AREA	area of polygon in square meters	PeRL
PERIMETER	perimeter of polygon in square meters	PeRL

Table S3. Description of attributes contained in the polygon attribute table of Canadian permafrost landscapes (*canada_perma_land.shp*).

Field name	Description	Source
ECOZONE	ecozone	NEF
ECOREGION	ecoregion	NEF
ECODISTRIC	ecodistrict	NEF
GEN_GEOL	dominant fraction of generalized (surficial) geology	NEF
LITHOLOGY	dominant fraction of texture	NEF
GROUND_ICE	dominant fraction of ground ice content in vol%	NEF
PF_EXTENT	dominant fraction of permafrost extent	NEF
PERMA_LAND	combined label of PF_EXTENT / GROUND_ICE/GEN_GEOL/LITHOLOGY	PeRL
ECOZID	ecozone ID	PeRL
PERMID	ID for each polygon in the vector file. The first digit stands for the region (1 – Alaska, 2 – Canada, 3 – Russia), digits 2 – 6 identify the single polygon, and the last three digits identify the ecozone.	PeRL
AREA	area of polygon in square meters	PeRL
PERIMETER	perimeter of polygon in square meters	PeRL

Table S4. Description of attributes contained in the polygon attribute table of Russian permafrost landscapes (*russia_perma_land.shp*).

Field name	Description	Source
ECOZONE	Metadata: http://maps.tnc.org/files/metadata/TNC_Lands.xml	Olson et al., 2001 Downloaded at http://maps.tnc.org/gis_data.html#TerrEcos
GEN_GEOL	Surficial geology	LRR, Stolbovoi and McCallum (2002)
LITHOLOGY	Texture	LRR, Stolbovoi and McCallum (2002)
GROUND_ICE	ground ice content in vol%	LRR
PF_EXTENT	permafrost extent	LRR
PERMA_LAND	combined label of PF_EXTENT / GROUND_ICE/GEN_GEOL/LITHOLOGY	LRR
ECOZID	ecozone ID	PeRL
PERMID	ID for each polygon in the vector file. The first digit stands for the region (1 – Alaska, 2 – Canada, 3 – Russia), digits 2 – 6 identify the single polygon, and the last three digits identify the ecozone.	PeRL
AREA	area of polygon in square meters	PeRL
PERIMETER	perimeter of polygon in square meters	PeRL

Table S5. Terminology for permafrost properties in the regional permafrost databases of Alaska (AK2008), Canada (NEF), and Russia (LRR).

Description	PeRL	AK2008	NEF	LRR
ecozone	ECOZONE	NA	ecozone	NA
surficial geology	GEN_GEOL	AGGRDEPOS	UNIT	PARROCK
lithology	LITHOLOGY	TEXTURE	TEXTURE	TEXTUR
permafrost extent	PF_EXTENT	PF_EXTENT	PERMAFROST	EXTENT_OF_
ground ice	GROUND_ICE	ICECLOWASS	PERMAFROST	MIN_MAX

Table S6. Attributes contained in the polygon attribute table of circum-arctic permafrost landscapes (*PeRL_study_area.shp*)

Field name	Description
country	country
Map_ID	ID of individual waterbody map
site	site name
MAAT	mean annual air temperature [°C]
TP	mean annual total precipitation [mm]
PE_DEPTH	Permafrost depth [m]
lat	latitude coordinate of polygon centroid in decimal degrees (WGS84)
long	longitude coordinate of polygon centroid in decimal degrees (WGS84)
AREA	area of polygon in square metres
AREA_SQKM	area of polygon in square kilometres

Table S7. Metadata and references for climate data and permafrost depth. MAAT – mean annual air temperature, TP – total precipitation.

Map ID	Country	MAAT & TP period	Climate data source, station	Permafrost depth source
abi0012010xxxx	Sweden	2006–2011, 1997–2007	MAAT: Johansson et al. (2013), TP: Abisko, www.polar.se/abisko	Åkerman and Johansson (2008), Dobiński (2010)
arg00120110829, arg0022009xxxx	Russia	1999–2011	Boike et al. (2013)	Yershov et al. (1991)
bar00119480804, bar00120020802 bar00120080730, bar00120100810, wlc00120090825, wlc00220020801, wlc00320090802	Alaska	1981–2010	National Climatic Data Center (2016): Barrow W Post Rogers Airport AK US	Brown et al. (1980)
byk00120060709	Russia	1984–1994	Rivas-Martínez et al. (2011)	Grosse et al. (2008)
byl00120160728	Canada	N/A	MAAT: Godin et al. (2016), TP: Fortier et al. (2007)	Smith and Burgess (2002)
che00120020709, che00220090724	Russia	1984–1994	Rivas-Martínez et al. (2011)	Grosse et al. (2008)
elc00120090825, elc00220020801, elc00320090802	Alaska	1981–2010	National Climatic Data Center (2016): Barrow W Post Rogers Airport, AK, US	Sellmann and Brown (1973)
esk00120090727	Canada	1981–2010	Environment Canada (2016): Tuktoyaktuk A	Taylor and Judge (1981)
fis00120020715	Alaska	1981–2010	National Climatic Data Center (2016): Kuparuk AK US	Jorgenson et al. (2008b)
fir00120090906, fir0022009xxxx, arg00320110711	Russia	1999–2011	Boike et al. (2013)	Yershov et al. (1991)
grp00119590707, grp00120060707	Canada	1971–2000	Bouchard et al. (2014), (Environment Canada, 2016)	Smith and Burgess (2002)
hbl00119540701, hbl00119740617, hbl00120060706	Canada	1971–2000	Sannel and Kuhry (2011), Environment Canada (2016)	N/A
ice0032009xxxx	Russia	1999–2011	Boike et al. (2013)	Yershov et al. (1991)
ind00120090907, kyt00120070728	Russia	1961–1990	Chokurdakh WMO station 21946	N/A
kol00119650721	Russia	1996–2015	Bukhta Ambarchik meteostation (WMO ID 25034)	Yershov et al. (1991)
kol00219650721	Russia	1996–2015	Andrushkino meteostation (WMO ID 25017)	Yershov et al. (1991)
kur00120100805, kur00220080926	Russia	1999–2011	Boike et al. (2013)	Yershov et al. (1991)
log00120110811	Russia	1961–1990	Khatanga WMO station 20891	N/A
mdn00120100716	Canada	1981–2010	Environment Canada (2016): Inuvik A	Burn and Kokelj (2009)
mdw00120090921	Canada	1981–2010	Environment Canada (2016): Inuvik A	Burn and Kokelj (2009)
ole00120060708	Russia	1948–1960	Rivas-Martínez et al. (2011)	Grosse et al. (2008)
pbp00120090813	Canada	1981–2010	Environment Canada (2016): Resolute Cars	Smith and Burgess (2002)

Map ID	Country	MAAT & TP period	Climate data source, station	Permafrost depth source
imc00120040725, kpc001201007xx, kpc002201007xx, kpc003201007xx	Alaska	1981–2010	National Climatic Data Center (2016): Kuparuk AK US	N/A
ric001201209125	Canada	1981–2010	Environment Canada (2016): Inuvik A	Burn (2002)
rog00120070626, rog00219740726, rog00220070707, rog00320070626, rog00420070704, rog00520070704	Russia	1961–1990	Hugelius et al. (2011), Vorkuta	Rivkin et al. (2008)
sam001200808xx	Russia	1999–2011	Boike et al. (2013)	Yershov et al. (1991)
sei00120070706, sei00220070706	Russia	N/A	Sjöberg et al. (2013)	Rivkin et al. (2008)
ksl00119620628, ksl0012012xxxx	Alaska	1981–2010	National Climatic Data Center (2016): Noma Municipal Airport AK US	Jorgenson et al. (2008a)
sur00120130802	Russia	N/A	Kremenetski et al. (2003)	Kremenetski et al. (2003)
tbr00120100901, tea00120100901	Canada	1981–2010	Environment Canada (2016), Yellowknife A	N/A
tav00119630831, tav00119750810, tav00120030702	Sweden	1971–2000	Sannel and Kuhry (2011)	Sannel and Kuhry (2011)
tuk00120120723	Canada	1981–2010	Environment Canada (2016): Tuktoyaktuk A	Taylor and Judge (1981)
yak0012009xxxx	Russia	1930–2010	Fedorov et al. (2014)	Yershov et al. (1991)
yam00120080824, yam00220100820	Russia	2004–2013	Leibman et al. (2015), Marre-Sale	Yershov et al. (1991)
yfl0012011xxxx	Alaska	1981–2010	National Climatic Data Center (2016): Central Number 2 AK US	Walvoord et al. (2012)
yuk00120090812, yuk00220090812	Alaska	1981–2010	National Climatic Data Center (2016): Bethel Airport AK US	N/A

References

Åkerman, H. J., and Johansson, M.: Thawing permafrost and thicker active layers in sub-arctic Sweden, Permafr. Periglac. Proc., 19, 279-292, 2008.

5 Andresen, C. G., and Lougheed, V. L.: Disappearing Arctic tundra ponds: Fine-scale analysis of surface hydrology in drained thaw lake basins over a 65 year period (1948–2013), J Geophys Res-Biogeo, 120, 466–479, 2015.

Bartsch, A., and Seifert, F. M.: The ESA DUE Permafrost project-A service for high latitude research, Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International, 2012, 5222-5225,

Boike, J., Kattenstroth, B., Abramova, K., Bornemann, N., Chetverova, A., Fedorova, I., Fröb, K., Grigoriev, M., Grüber,
10 M., Kutzbach, L., Langer, M., Minke, M., Muster, S., Piel, K., Pfeiffer, E. M., Stoof, G., Westermann, S., Wischnewski, K., Wille, C., and Hubberten, H. W.: Baseline characteristics of climate, permafrost and land cover from a new permafrost observatory in the Lena River Delta, Siberia (1998-2011), Biogeosciences, 10, 2105-2128, 10.5194/bg-10-2105-2013, 2013.

Bouchard, F., Francus, P., Pienitz, R., Laurion, I., and Feyte, S.: Subarctic thermokarst ponds: Investigating recent landscape evolution and sediment dynamics in thawed permafrost of northern Québec (Canada), Arct. Antarct. Alp. Res.,
15 46, 251-271, 2014.

Brown, J., Miller, P. C., Tieszen, L. L., and Bunnell, F.: An arctic ecosystem: the coastal tundra at Barrow, Alaska, Dowden, Hutchinson & Ross, Inc., 1980.

Burn, C.: Tundra lakes and permafrost, Richards Island, western Arctic coast, Canada, Can. J. Earth Sci., 39, 1281-1298, 2002.

20 Burn, C., and Kokelj, S.: The environment and permafrost of the Mackenzie Delta area, Permafr. Periglac. Proc., 20, 83-105, 2009.

Dobiński, W.: Geophysical characteristics of permafrost in the Abisko area, northern Sweden, Pol. Polar Res, 31, 141-158, 2010.

Eineder, M., Fritz, T., Mittermayer, J., Roth, A., Boerner, E., and Breit, H.: TerraSAR-X ground segment, basic product
25 specification document, German Aerospace Agency, Oberpfaffenhofen, Germany, 103, 2008.

Fedorov, A. N., Ivanova, R. N., Park, H., Hiyama, T., and Iijima, Y.: Recent air temperature changes in the permafrost landscapes of northeastern Eurasia, Polar Science, 8, 114-128, 10.1016/j.polar.2014.02.001, 2014.

Fortier, D., Allard, M., and Shur, Y.: Observation of rapid drainage system development by thermal erosion of ice wedges on Bylot Island, Canadian Arctic Archipelago, Permafr. Periglac. Proc., 18, 229-243, 10.1002/ppp.595, 2007.

30 Godin, E., Fortier, D., and Lévesque, E.: Nonlinear thermal and moisture response of ice-wedge polygons to permafrost disturbance increases heterogeneity of high Arctic wetland, Biogeosciences, 13, 1439-1452, 2016.

Grosse, G., Romanovsky, V., Walter, K., Morgenstern, A., Lantuit, H., and Zimov, S.: Distribution of Thermokarst Lakes and Ponds at Three Yedoma Sites in Siberia, Ninth International Conference on Permafrost, University of Alaska, Fairbanks, USA, 2008, 551-556,

35 Hugelius, G., Virtanen, T., Kaverin, D., Pastukhov, A., Rivkin, F., Marchenko, S., Romanovsky, V., and Kuhry, P.: High-resolution mapping of ecosystem carbon storage and potential effects of permafrost thaw in periglacial terrain, European Russian Arctic, J Geophys Res-Biogeo, 116, 2011.

Johansson, M., Callaghan, T. V., Bosiö, J., Åkerman, H. J., Jackowicz-Korczynski, M., and Christensen, T. R.: Rapid responses of permafrost and vegetation to experimentally increased snow cover in sub-arctic Sweden, Environ. Res. Lett., 8,
40 035025, 2013.

Jones, B. M., Gusmeroli, A., Arp, C. D., Strozzi, T., Grosse, G., Gaglioti, B. V., and Whitman, M. S.: Classification of freshwater ice conditions on the Alaskan Arctic Coastal Plain using ground penetrating radar and TerraSAR-X satellite data, Int. J. Remote Sens., 34, 8267-8279, 2013.

Jorgenson, M., Yoshikawa, K., Kanevskiy, M., Shur, Y., Romanovsky, V., Marchenko, S., Grosse, G., Brown, J., and
45 Jones, B.: Permafrost characteristics of Alaska, Proceedings of the Ninth International Conference on Permafrost, 2008a, 121-122,

Kremenetski, K. V., Velichko, A. A., Borisova, O. K., MacDonald, G., Smith, L. C., Frey, K. E., and Orlova, L. A.: Peatlands of the Western Siberian lowlands: current knowledge on zonation, carbon content and Late Quaternary history, Quaternary Science Reviews, 22, 703-723, 2003.

50 Lara, M. J., McGuire, A. D., Euskirchen, E. S., Tweedie, C. E., Hinkel, K. M., Skurikhin, A. N., Romanovsky, V. E., Grosse, G., Bolton, W. R., and Genet, H.: Polygonal tundra geomorphological change in response to warming alters future CO_2 and CH_4 flux on the Barrow Peninsula, Global change biology, 21, 1634-1651, 2015.

Leibman, M., Khomutov, A., Gubarkov, A., Mullanurov, D., and Dvornikov, Y.: The research station" Vaskiny Dachi", Central Yamal, West Siberia, Russia–a review of 25 years of permafrost studies, Fennia, 193, 3-30, 2015.

55 Marshall, I., Schut, P., and Ballard, M.: A national ecological framework for Canada: Attribute data, Agriculture and AgriFood Canada, Research Branch, Centre for Land and Biological Resources Research, and Environment Canada, State of the Environment Directorate, Ecozone Analysis Branch, Ottawa/Hull, Canada, 1999.

Muster, S., Langer, M., Heim, B., Westermann, S., and Boike, J.: Subpixel heterogeneity of ice-wedge polygonal tundra: a multi-scale analysis of land cover and evapotranspiration in the Lena River Delta, Siberia, Tellus B, 64, 10.3402/tellusb.v64i0.17301, 2012.

Muster, S., Heim, B., Abnizova, A., and Boike, J.: Water body distributions across scales: A remote sensing based comparison of three Arctic tundra wetlands, Remote Sensing, 5, 1498–1523, 10.3390/rs5041498, 2013.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V., Underwood, E. C., D'amico, J. A., Itoua, I., Strand, H. E., and Morrison, J. C.: Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity, Bioscience, 51, 933−938, 2001.

Palmtag, J., Ramage, J., Hugelius, G., Gentsch, N., Lashchinskiy, N., Richter, A., and Kuhry, P.: Controls on the storage of organic carbon in permafrost soil in northern Siberia, European Journal of Soil Science, 67, 478-491, 2016.

Raynolds, M. K., Walker, D. A., Ambrosius, K. J., Brown, J., Everett, K. R., Kanevskiy, M., Kofinas, G. P., Romanovsky, V. E., Shur, Y., and Webber, P. J.: Cumulative geoecological effects of 62 years of infrastructure and climate change in ice-rich permafrost landscapes, Prudhoe Bay Oilfield, Alaska, Global Change Biology, 20, 1211-1224, 10.1111/gcb.12500, 2014.

Rivas-Martínez, S., Rivas-Sáenz, S., and Penas, A., S: Worldwide bioclimatic classification system, Global Geobotany, 1, 1-634, 10.5616/gg110001, 2011.

Rivkin, F., Vlasova, J., Popova, A., Mazhitova, G., Kuhry, P., Parmuzin, I., and Chehina, I.: Mesoscale and detailed geocryological mapping as a basis for carbon budget assessment (East European Russian Arctic, CARBO-North project), Proc. 9th Int. Conf. Permafrost, 2008, 1493-1498,

Sannel, A. B. K., and Brown, I. A.: High-resolution remote sensing identification of thermokarst lake dynamics in a subarctic peat plateau complex, Canadian Journal of Remote Sensing, 36, S26-S40, 10.5589/m10-010, 2010.

Sannel, A. B. K., and Kuhry, P.: Warming-induced destabilization of peat plateau/thermokarst lake complexes, Journal of Geophysical Research, 116, 10.1029/2010JG001635, 2011.

Sellmann, P. V., and Brown, J.: Stratigraphy and diagenesis of perennially frozen sediments in the Barrow, Alaska, region, Permafrost: North American Contribution to the Second International Conference. Washington, DC: National Academy of Sciences, 1973, 171-181,

Sjöberg, Y., Hugelius, G., and Kuhry, P.: Thermokarst lake morphometry and erosion features in two peat plateau areas of northeast European Russia, Permafr. Periglac. Proc., 24, 75-81, 2013.

Stolbovoi, V., and McCallum, I.: Land Resources of Russia, International Institute for Applied Systems Analysis and the Russian Academy of Science, Laxenburg, Austria, 2002.

Taylor, A., and Judge, A. S.: Measurement and prediction of permafrost thickness, Arctic Canada, Energy, Mines and Resources Canada, Earth Physics Branch, 1981.

Walker, D. A., Webber, P. J., Walker, M. D., Lederer, N. D., Meehan, R. H., and Nordstrand, E. A.: Use of geobotanical maps and automated mapping techniques to examine cumulative impacts in the Prudhoe Bay Oilfield, Alaska, Environ. Conserv., 13, 149-160, 1986.

Walvoord, M. A., Voss, C. I., and Wellman, T. P.: Influence of permafrost distribution on groundwater flow in the context of climate-driven permafrost thaw: Example from Yukon Flats Basin, Alaska, United States, Water Resour. Res., 48, 2012.

Widhalm, B., Hogström, E., Ressl, C., Trofaier, A., Heim, B., Biasi, C., and Bartsch, A.: PAGE21 WP5-Land surface hydrology from remotely sensed data at PAGE21 sites, Department of Geodesy and Geoinformation (GEO), Research Groups Photogrammetry and Remote Sensing, Vienna University of Technology, Vienna, Austria, 15, 2014a.

Yershov, E. D., Kondrat'yeva, K. A., KLoginov, V. F., and Sychev, I. K.: Geocryological Map of Russia and Neighbouring Republics, Faculty of Geology, Chair of Geocryology, Lomonosov Moscow State University, 1991.