The Boreal-Arctic Wetland and Lake Dataset (BAWLD): **Supplementary information**

David Olefeldt¹, Mikael Hovemyr², McKenzie A. Kuhn¹, David Bastviken³, Theodore J. Bohn⁴, John Connolly⁵, Patrick Crill⁶, Eugénie S. Euskirchen^{7,8}, Sarah A. Finkelstein⁹, Hélène Genet⁸, Guido

- Grosse¹⁰, Lorna I. Harris¹, Liam Heffernan¹¹, Manuel Helbig¹², Gustaf Hugelius^{2,13}, Ryan Hutchins¹⁴, Sari Juutinen¹⁵, Mark J. Lara^{16,17}, Avni Malhotra¹⁸, Kristen Manies¹⁹, A. David McGuire⁹, Susan M. Natali²⁰, Jonathan O'Donnell²¹, Frans-Jan W. Parmentier^{22,23}, Aleksi Räsänen²⁴, Christina Schaedel²⁵, Oliver Sonnentag²⁶, Maria Strack²⁷, Suzanne Tank²⁸, Claire Treat¹⁰, Ruth K. Varner^{2,29}, Tarmo Virtanen³⁰, Rebecca K. Warren³¹, Jennifer D. Watts²⁰.
- ¹ Department of Renewable Resources, University of Alberta, Edmonton, AB, T6G 2G7, Canada 10 ² Department of Physical Geography, Stockholm University, 10691 Stockholm, Sweden ³ Department of Thematic Studies – Environmental Change, Linköping University, 58183 Linköping, Sweden ⁴WattIQ, 400 Oyster Point Blvd. Suite 414, South San Francisco, CA, 94080, USA ⁵ Department of Geography, School of Natural Sciences, Trinity College Dublin, Dublin 2, Ireland
- ⁶Department of Geological Sciences, Stockholm University, 10691 Stockholm, Sweden 15 ⁷ Department of Biology and Wildlife, University of Alaska Fairbanks, Fairbanks, AK 99775, USA ⁸ Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775, USA ⁹ Department of Earth Sciences, University of Toronto, Toronto, ON, M5S 3B1, Canada ¹⁰ Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, 14473 Potsdam, Germany
- ¹¹ Department of Ecology and Genetics, Uppsala University, 752 36 Uppsala, Sweden 20
 - ¹² Department of Physics and Atmospheric Science, Dalhousie University, Halifax, NS, B3H 4R2, Canada
 - ¹³ Bolin Centre for Climate Research, Stockholm University, 10691 Stockholm, Sweden
 - ¹⁴ Department of Earth and Environmental Sciences, University of Waterloo, Waterloo, ON, N2L 3G1, Canada
 - ¹⁵ Ecosystems and Environment Research Program, University of Helsinki, FI-00014 Helsinki, Finland
- ¹⁶ Department of Plant Biology, University of Illinois, Urbana, IL 61801, USA 25
 - ¹⁷ Department of Geography, University of Illinois, Urbana, IL 61801, USA
 - ¹⁸ Department of Earth System Science, Stanford University, Stanford, CA 94305, USA
 - ¹⁹ U.S. Geological Survey, Menlo Park, CA, USA
 - ²⁰ Woodwell Climate Research Center, Falmouth, MA 02540, USA
- ²¹ Arctic Network, National Park Service, Anchorage, AK 99501 USA 30
 - ²² Centre for Biogeochemistry in the Anthropocene, Department of Geosciences, University of Oslo, 0315 Oslo, Norway
 - ²³ Department of Physical Geography and Ecosystem Science, Lund University, 223 62 Lund, Sweden

²⁴ Ecosystems and Environment Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, Finland

- ²⁵ Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ 86011, USA 35
 - ²⁶ Département de Géographie, Université de Montréal, Montréal, QC, Canada
 - ²⁷ Department of Geography, University of Waterloo, Waterloo, ON, N2L 3G1, Canada
 - ²⁸ Department of Biological Sciences, University of Alberta, Edmonton, AB, T6G 2E9, Canada

²⁹ Department of Earth Sciences and Institute for the Study of Earth, Oceans and Space, University of New Hampshire,

Durhan, NH 03824, USA 40

³⁰ Ecosystems and Environment Research Programme, Faculty of Biological and Environmental Sciences, 00014 University of Helsinki, Helsinki, Finland

³¹ National Boreal Program, Ducks Unlimited Canada, Edmonton, AB, T5S 0A2, Canada

45 Table S1. Average percent coverage of each of the 19 BAWLD land cover classes within the 15 identified wetscapes.

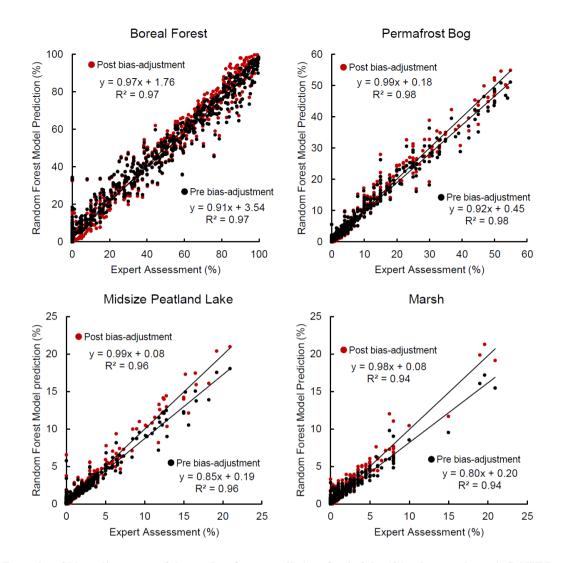
WETSCAPES								BAWL	D LAN	D COV	ER CLA	ASSES							
	GLA	ROC	TUN	BOR	BOG	FEN	MAR	PEB	TUW	LAL	MGL	MPL	MYL	SGL	SPL	SYL	LAR	SPR	SRR
Wetland-rich Tundra	0.0	0.7	59.8	8.3	0.1	0.3	0.3	012.2	10.4	1.2	0.7	0.8	1.4	0.4	1.0	1.7	0.4	0.1	0.1
Wetland and Lake-rich Tundra	0.0	1.4	54.7	4.0	0.4	0.5	0.6	011.3	11.5	2.5	5 .2	4.4	0.1	01.3	1.6	0.0	0.4	0.1	0.1
Wetland and Lake-rich Yedoma Tundra	0.0	0.2	<u> </u>	8.6	0.1	0.2	0.3	011.0	15.4	4.2	2.2	3.0	6.2	0.6	1.2	1.8	0.8	0.1	0.1
Upland Tundra	0.6	07.7	76.3	3.9	0.3	0.3	0.1	3.9	4.0	0.5	0.9	0.3	0.1	0.5	0.4	0.1	0.1	0.2	0.0
Lake-rich Shield	0.0	24.3	38.2	13.3	0.3	0.7	0.2	2.0	2.5	5 .9	8.5	1.0	0.0	2.4	0.3	0.0	0.1	0.1	0.0
Alpine and Tundra Barrens	5 .1	37.7	<u> </u>	9.5	0.3	0.3	0.1	0.4	0.6	0.4	0.7	0.1	0.0	0.2	0.0	0.0	0.0	0.2	0.0
Permafrost Peatlands	0.0	0.6	17.3	030.9	4.3	6.8	1.7	27.1	3.7	2.0	1.1	01.7	0.0	0.3	1.9	0.1	0.3	0.1	0.1
Lake-rich Wetlands	0.0	0.1	2.9	22.1	013.7	016.3	8.1	016.2	2.4	01.5	0.4	4.8	0.2	0.4	8.6	0.3	1.7	0.1	0.2
Dominant Boreal Wetlands	0.0	0.3	0.2	28.6	26.9	28.2	3.9	4.4	0.9	1.1	0.3	01.6	0.0	0.0	2.9	0.0	0.5	0.0	0.2
Common Boreal Wetlands	0.0	1.3	1.7	55.5	017.3	017.0	1.7	1.0	0.1	1.9	0.7	0.8	0.0	0.1	0.6	0.0	0.3	0.1	0.1
Sparse Boreal Wetlands	0.1	5 .0	2.3	69.5	7.8	6.5	01.0	0.8	0.1	3.4	2.0	0.6	0.0	0.2	0.2	0.0	0.2	0.1	0.1
Upland Boreal	0.0	5.5	07.7	82.0	0.8	0.8	0.1	1.4	0.4	0.3	0.2	0.0	0.0	0.0	0.0	0.0	0.3	0.2	0.0
Large Lakes	0.0	4.0	7.0	20.9	2.1	2.3	0.5	2.2	0.7	57.4	1.3	0.5	0.0	0.7	0.2	0.0	0.0	0.0	0.0
Large Rivers	0.3	01.5	21.9	42.8	2.2	2.6	2.0	6.8	6.2	0.6	0.9	1.0	0.5	0.5	1.3	0.5	8.2	0.1	0.1
Glaciers	91.0	4.7	3.7	0.3	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

GLA = Glaciers, ROC = Rocklands, TUN = Dry Tundra, BOR = Boreal Forest, BOG = Bogs, FEN = Fens, MAR = Marshes, PEB = Permafrost Bogs, TUW = Tundra Wetlands, LAL = Large Lakes, MGL = Midsize Glacial Lakes, MPL = Midsize Peatland Lakes, MYL = Midsize Yedoma Lakes, SGL = Small Glacial Lakes, SPL = Small Peatland Lakes, SYL = Small Yedoma Lakes, LAR = Large Rivers, SPR = Small Organic-Poor Rivers, SRR = Small Organic-Rich Rivers.

Green, Yellow, and Red circles indicate relatively high, moderate, and low land cover class coverage within wetscapes, relative to each land cover class.



Figure S1. Spatial distribution of training cells where expert assessment of land cover composition was completed.



60 Figure S2. Examples of bias-adjustments of the random forest predictions for 4 of the 19 land-cover classes in BAWLD.

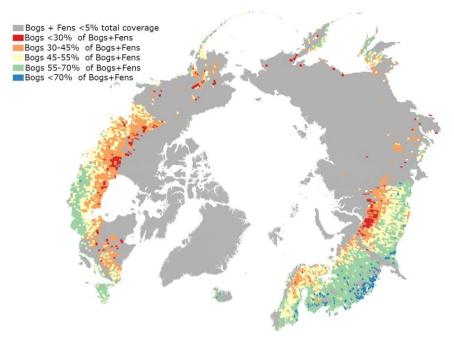


Figure S3. Predicted relative abundance of Bogs and Fens for grid cells within BAWLD where combined Bog and Fen cover was greater than 5% of the total grid cell area.

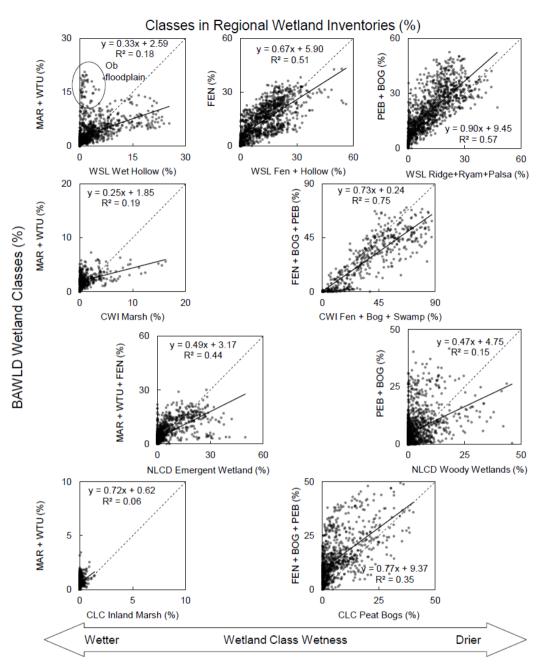


Figure S4. Comparison of the grid cell coverage of BAWLD wetland land cover classes with that of similar wetland land cover classes in four independent regional land cover datasets; the wetland mapping of the West Siberian Lowlands (WSL) (top row), the Canadian Wetland Inventory (CWI) (second row), the National Land Cover Database (NLCD) (third row), and the CORINE

70 Land Cover (CLC) dataset (fourth row). Abbreviations used for the BAWLD wetland classes: Marshes (MAR), Wetland Tundra (WTU), Fens (FEN), Bogs (BOG), and Permafrost Bogs (PEB). Several comparisons required combinations of wetland classes in order to ensure similar definitions. Relatively wetter wetland classes are compared on the left side of the figure, while drier wetland classes are compared on the right hand side. The ranking of relative wetness for BAWLD classes is MAR > WTU > FEN > BOG > PEB.

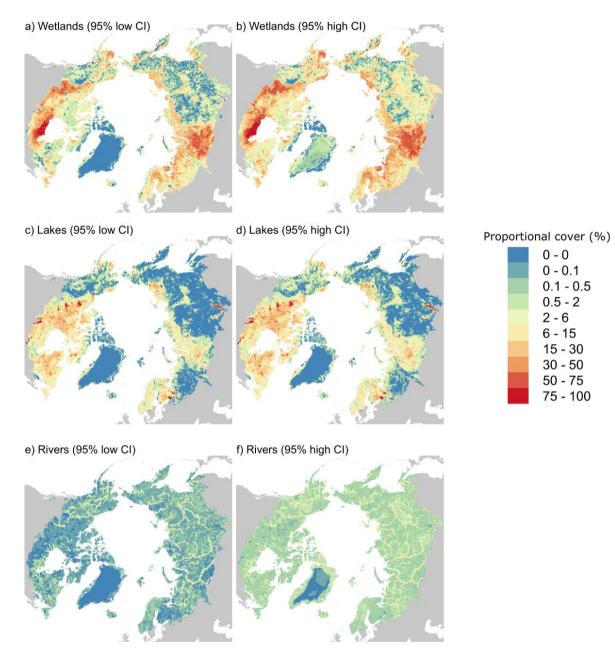
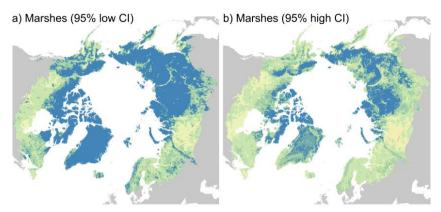


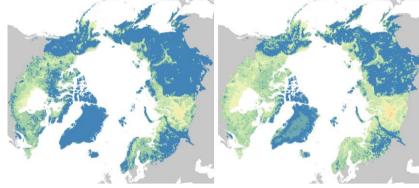


Figure S5. Low and High 95% confidence bounds for the fractional total coverage of wetland classes (sum of 5 classes) (a-b), lake classes (sum of 7 classes) (c-d) and river classes (sum of 3 classes) (e-f).



c) Small Peatland Lakes (95% low CI)

d) Small Peatland Lakes (95% high CI)



Proporti	onal cover (%)
	0 - 0
	0 - 0.1
	0.1 - 0.5
	0.5 - 2
	2 - 6
	6 - 15
	15 - 30
	30 - 50
	50 - 75
	75 - 100

e) Small Organic-Rich Rivers (95% low CI) f) Small Organic-Rich Rivers (95% high CI)

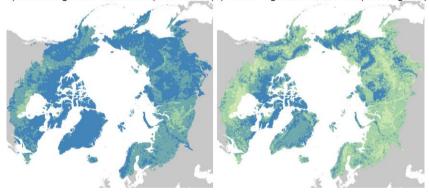
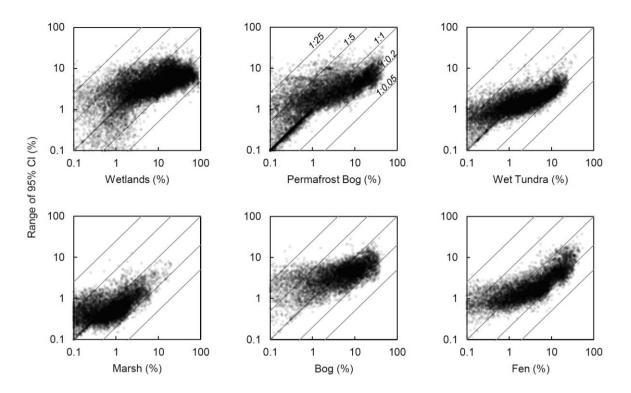


Figure S6. Examples of Low and High 95% confidence bounds for the fractional coverage of 3 of the 19 land cover classes in BAWLD; Marshes (a-b), Small Peatland Lakes (c-d), Small Organic-Rich Rivers (e-f). The chosen land cover classes have a relatively high uncertainty in BAWLD compared to other land cover classes, and are also considered classes with relatively high methane emissions.



85 Figure S7. Scatterplots of central estimates vs the range of the 95% confidence intervals for wetland classes within individual BAWLD grid cells. Diagonal lines indicate the relative size range of the 95% confidence interval in comparison to the central estimate, ranging from the confidence interval being 2500% greater than the central estimate (1:25) to only representing 5% of the central estimate (1:0.05). Top left plate shows the total wetland estimates, while the remaining plates focus on individual wetland classes.

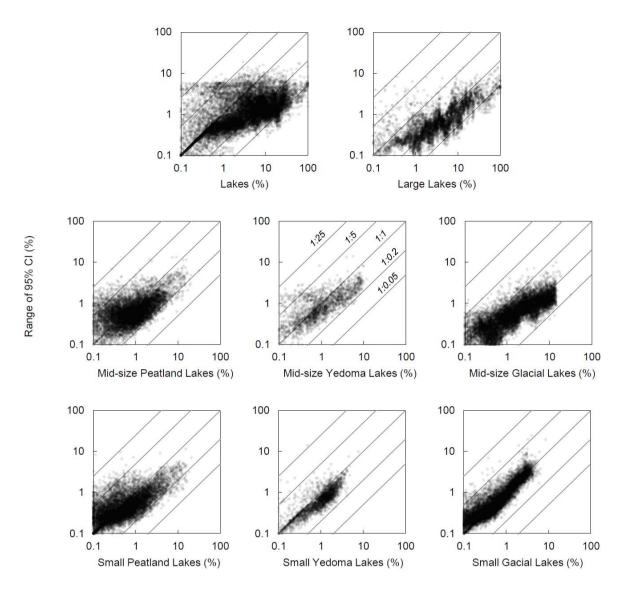




Figure S8. Scatterplots of central estimates vs the range of the 95% confidence intervals for lake classes within individual BAWLD grid cells. Diagonal lines indicate the relative size range of the 95% confidence interval in comparison to the central estimate, ranging from the confidence interval being 2500% greater than the central estimate (1:25) to only representing 5% of the central estimate (1:0.05). Top left plate shows the total lake estimates, while the remaining plates focus on individual lake classes.

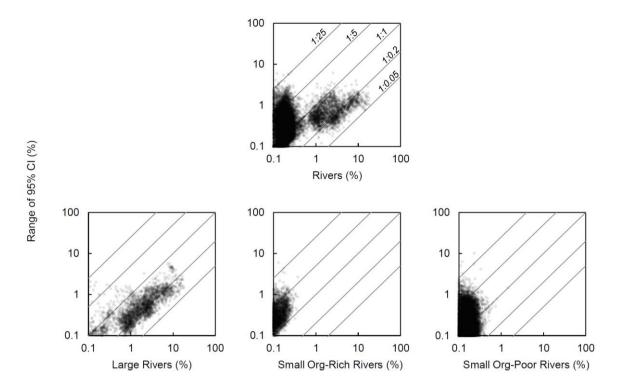


Figure S9. Scatterplots of central estimates vs the range of the 95% confidence intervals for river classes within individual BAWLD grid cells. Diagonal lines indicate the relative size range of the 95% confidence interval in comparison to the central estimate, ranging from the confidence interval being 2500% greater than the central estimate (1:25) to only representing 5% of the central estimate (1:0.05). Top plate shows the total river estimates, while the remaining plates focus on individual river classes.

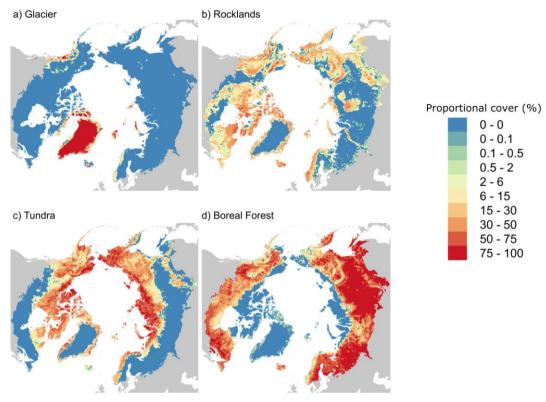


Figure S10. Predicted distributions of a) Glaciers, b) Rocklands, c) Tundra, d) Boreal Forests in BALWD.

Expert Assessment Instructions

Below we have compiled information from three files that were sent out to experts with instructions to complete the land cover fractional assessment. The first file has the step-by-step instructions for how to carry out the assessment. The second file has descriptions and examples of the 19 land cover classes that the experts were asked to assess fractional land coverage

110 of. Note that some of the class names were changed in the final version of BAWLD. The third file shows the available spatial datasets which could be used in combination with Google Earth satellite imagery and expert knowledge of the region to make the assessments. Experts also had access to this data in a .kml file that could be opened in Google Earth.

Suggested workflow for the expert assessment

115 1. Check out the 20 grid cells we have assigned you for your assessment.

With this mail, we have attached 2 files; an excel file "*Expert Assessment – LastName.xlsx*" and a kml file "*Expert Assessment – LastName.kml*". In order to open the kml file, you will need to download *Google Earth Pro on desktop*, scroll down on this webpage for the download link if you don't already have it installed:

https://www.google.com/earth/versions/

130

120 In both the .xlsx and .kml files, you have the same 20 grid cells indicated for which we ask you to give your assessment of the relative coverage of 19 land cover classes. You should be able to cross-reference the cells based on their lat/long and their cell ID's.

Both files further include information for each cell from ~50 available data sources. This data is shown in column to the right in the .xlsx file, and should pop up when you click on a file in Google Earth.

125 Of the 20 cells, 10 have been randomly chosen in regions you specified familiarity with, and 10 have been chosen in the wider boreal-arctic domain. Each expert has been given 20 unique cells to assess. Before you start to populate the .xlsx with your assessment, go through the steps below.

2. Get to know the 19 land cover classes which will be in the final land cover database.

These are presented and defined in the pdf file "*1_Methane Land Cover Classes.pdf*" which was sent to you in the last email. Let me know if you have any questions on the definitions.

3. Get to know the ~50 available spatial data layers which you can use to guide your assessment.

You assessment of the relative coverage of the methane classes in the assessment cells will be based on you knowledge of the region, visual assessment of satellite imagery provided by Google Earth, and already available spatial data from multiple sources. These ~50 available spatial data layers have been compiled into a common 0.5 x 0.5 $^{\circ}$ grid cell framework. I have

135 provided images of the data in each layer in a pdf file "2_Available data layers.pdf", you can explore all the data in detail in Google Earth if you open the kml file "3_Available data layers.kml", and each layer is described in "4_Available data layers.xlsx". You should have these files from my previous mail.

Let me know if you have any questions regarding the available data layers, .e.g. if you are not clear on what they indicate. At the bottom of this instruction are a few tips on how these layers may help your assessment of the 19 land cover classes.

140 4. Fill out your assessment in the "Expert assessment – LastName.xlsx" file.

For each of the 20 cells you have been assigned, give your assessment of the coverage of the 19 land cover classes. The unit is %, and each row of 19 classes should add up to 100%. Use up to two decimals for your assessment (e.g. 21.54%). We use two decimals since it may be required for classes that have very small coverage – e.g. small rivers and marshes where I expect many cells to have <0.10% coverage. Two decimals is also the precision of the available datasets, so you could

145 directly transfer some data if you think it represents one of our land cover classes well (e.g. Large Lakes from the Hydrolakes Large Lake data layer).

View of excel sheet:

$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	Normal Bad Good Check Cell Exploratory T Input Styles	Neutral Calculation Source Date	
ダ Formit Parter			e Format Sont & Find &
board 13 Font 13 Alignment 15 Number 13	2gles		
123 • (* <i>k</i>			Eding
V 0	5 Y U V V X MARKAN AND AND AND AND AND AND AND AND AND A	Decisión en forma service NOT PET DOD PET DOD DOD <thdod< th=""> <thdod< th=""></thdod<></thdod<>	

150 As mentioned above, in order to arrive at estimates, you can use your knowledge of the region, visual assessment of satellite imagery provided by Google Earth, and the available spatial data compiled into ~50 data layers.

Optional resources:

160

If you are ambitious, and used to working in GIS, you can download a few of the key available datasets to see their original

155 data, rather than the data compiled into the 0.5 x 0.5 cells. For example, I found it useful sometimes to overlay the lake polygons in HydroLakes over the satellite imagery, in order to see how much of the visible smaller open water areas that were not included as lakes in the HydroLakes data:

https://www.hydrosheds.org/page/hydrolakes

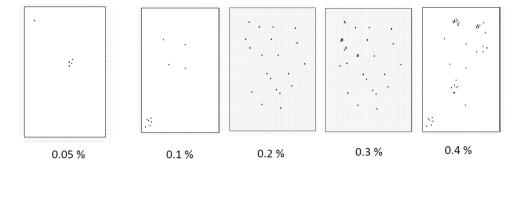
Another dataset of potential interest is the Northers Circumpolar Soils Database V2, which is probably the most reliable included datasource for distribution of peatland soils (both permafrost – histels, and non-permafrost – histosols):

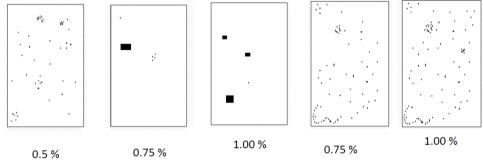
https://bolin.su.se/data/ncscd/shape.php

I also sometimes found it useful to look at the Global Surface Water data, which has an online explorer, as it often well indicates areas that are continuously vs temporarily inundated, and how much of open water areas (peatland pools) within wetland complexes that it captures:

165 https://global-surface-water.appspot.com/map

That said, using these additional resources is not required.

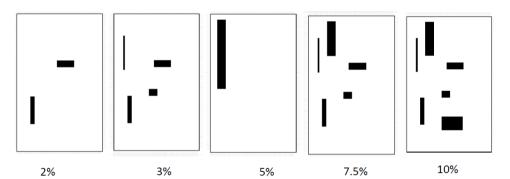


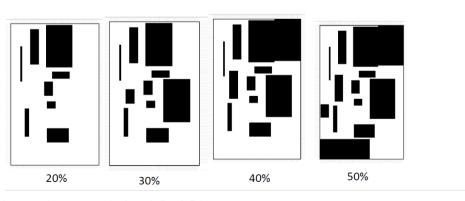


0.5 %



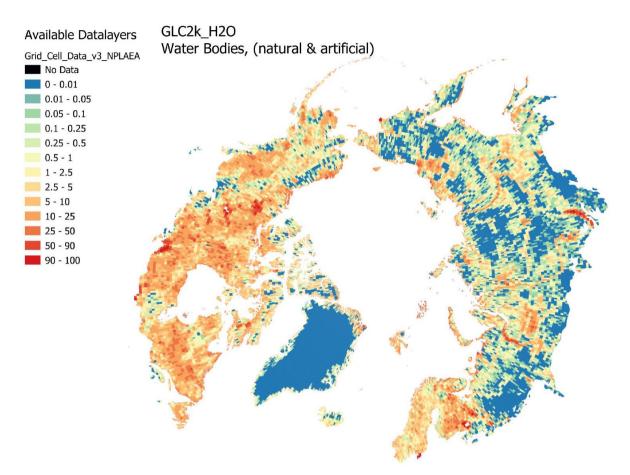


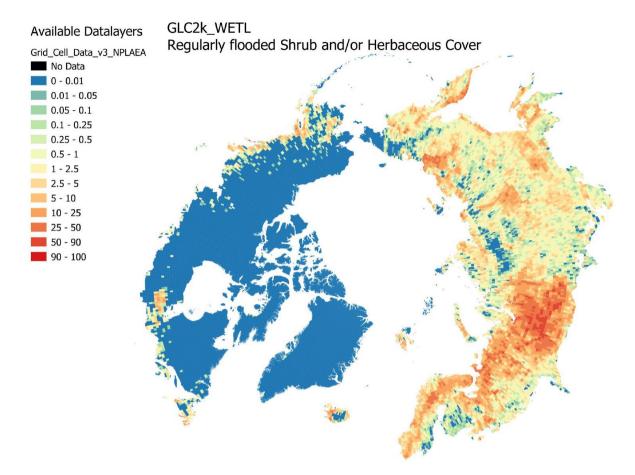


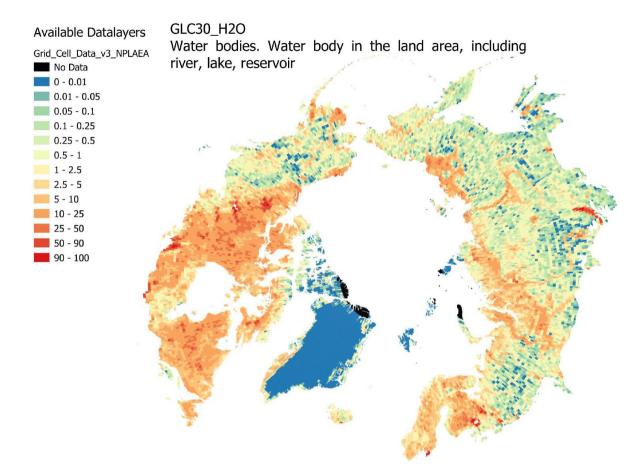


Guide to assessing spatial cover within a 0.5 x 0.5 $^\circ$ cell.

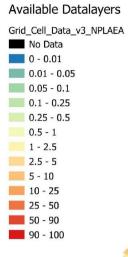
170 Available spatial datasets that were be used to inform the expert assessment.



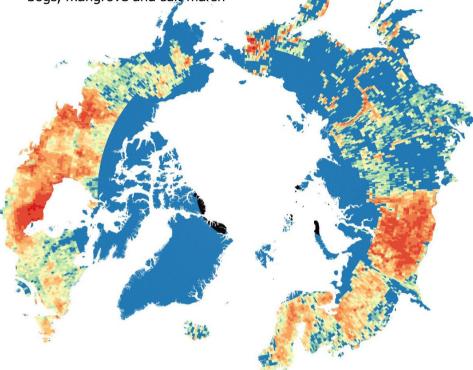




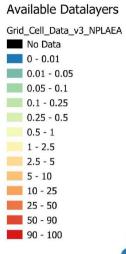
GLC30_WETL



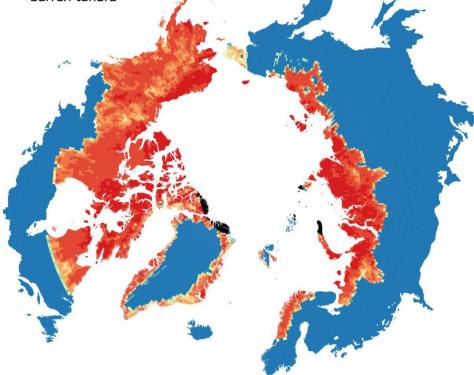
Wetland. Lands covered with wetland plants and water bodies, including inland marsh, lake marsh, river floodplain wetland, forest/shrub wetland, peat bogs, mangrove and salt marsh



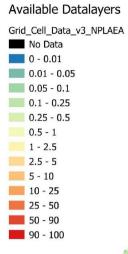
GLC30_TUND



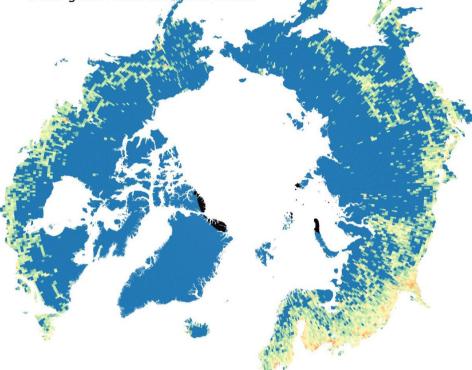
Tundra. Lands covered by lichen, moss, hardy perennial herb and shrubs in the polar regions, including shrub tundra, herbaceous tundra, wet tundra and barren tundra

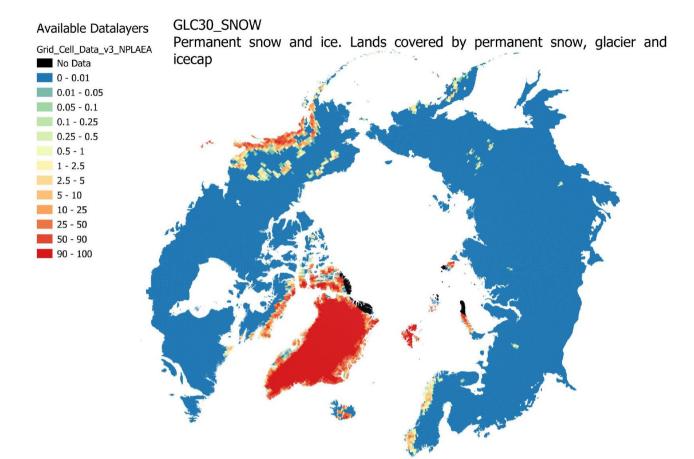


GLC30_ARTI

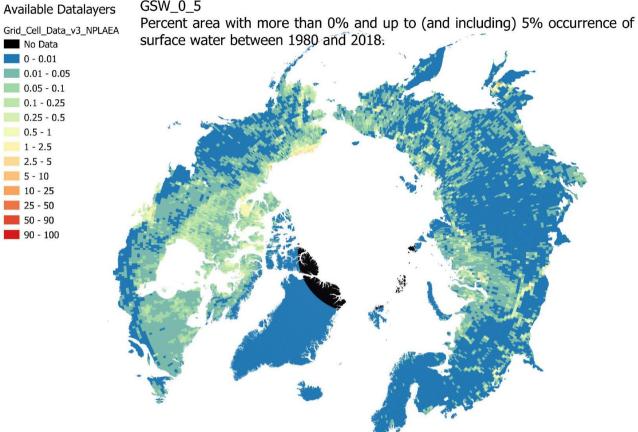


Artificial Surfaces. Lands modified by human activities, including all kinds of habitation, industrial and mining area, transportation facilities, and interior urban green zones and water bodies

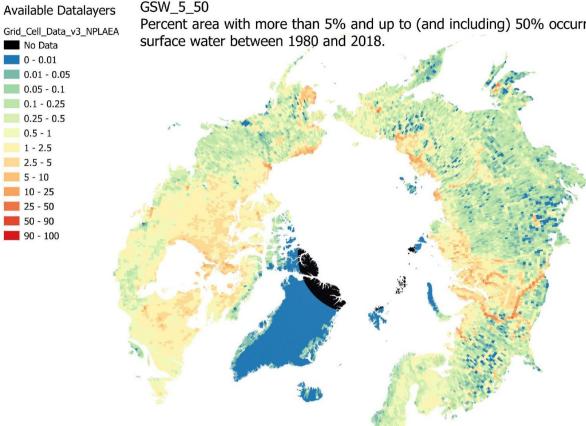




GSW_0_5

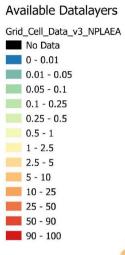


GSW_5_50

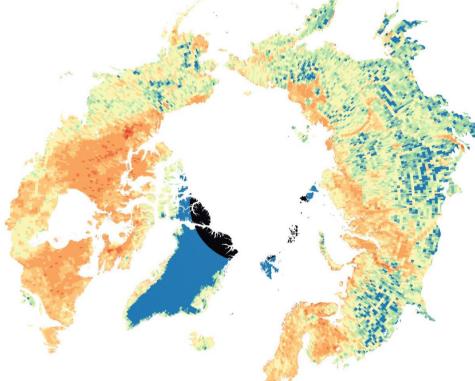


Percent area with more than 5% and up to (and including) 50% occurrence of

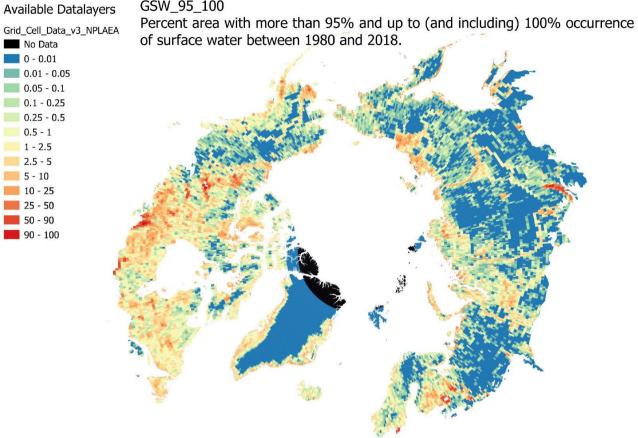
GSW_50_95

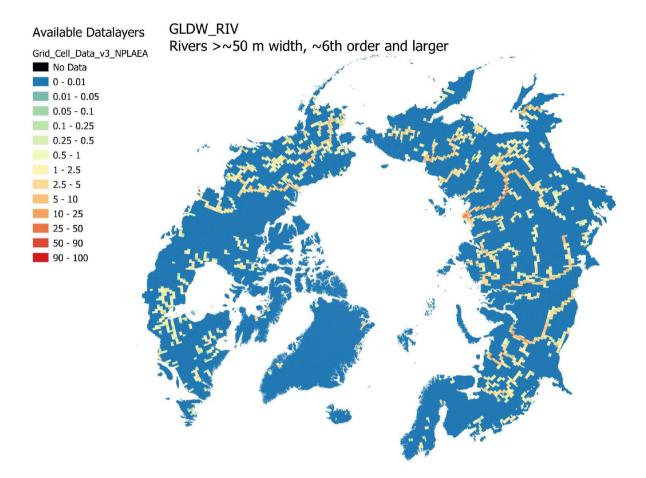


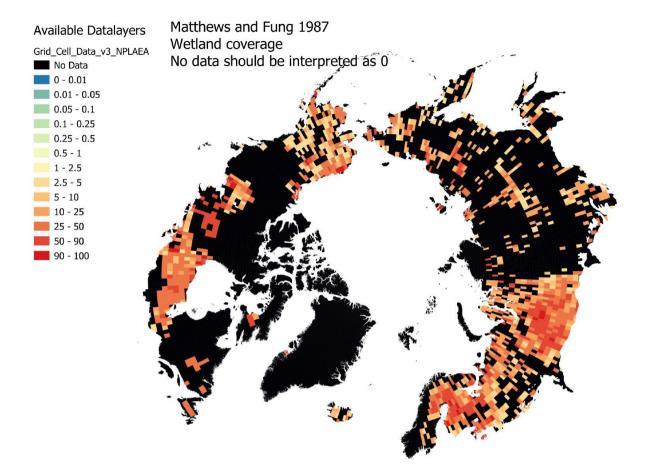
Percent area with more than 50% and up to (and including) 95% occurrence of surface water between 1980 and 2018.

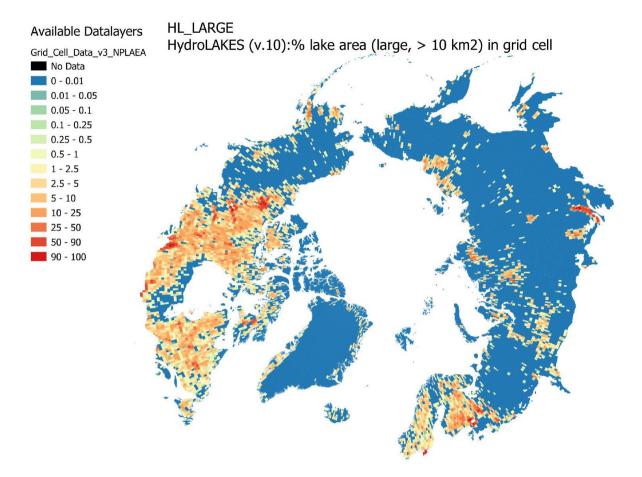


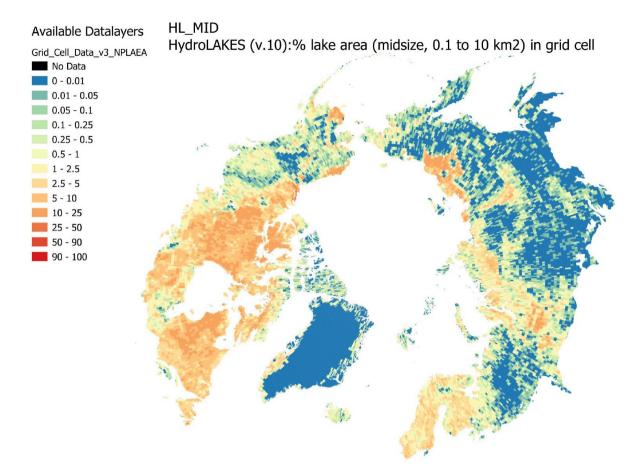
GSW_95_100

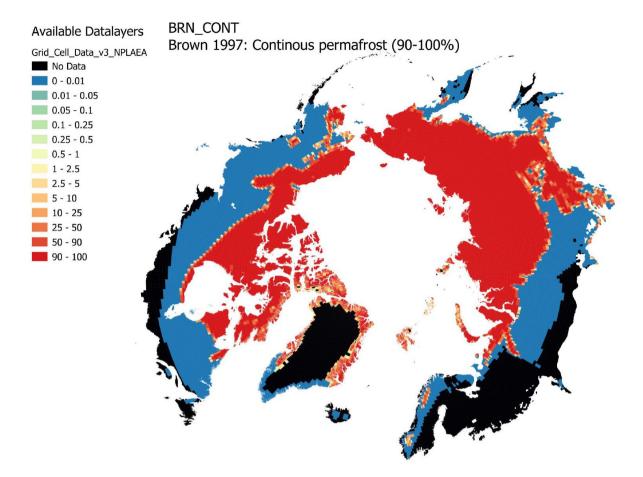


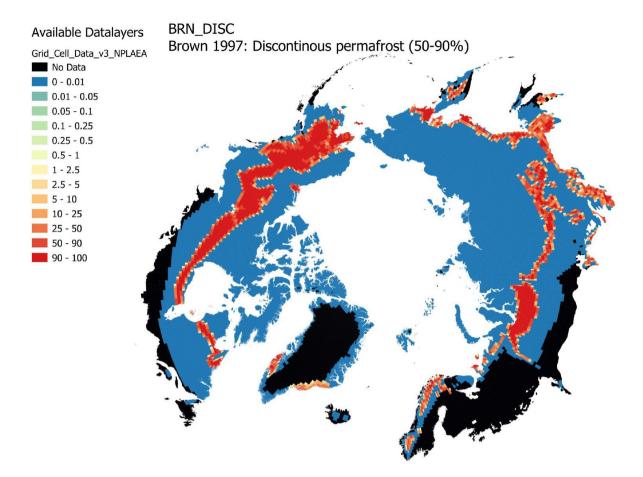


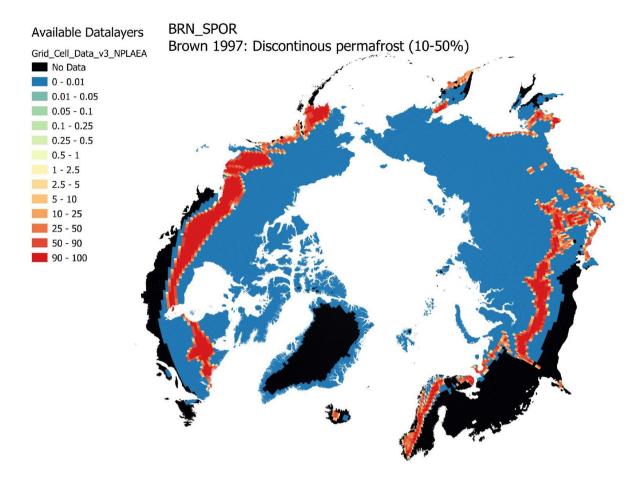


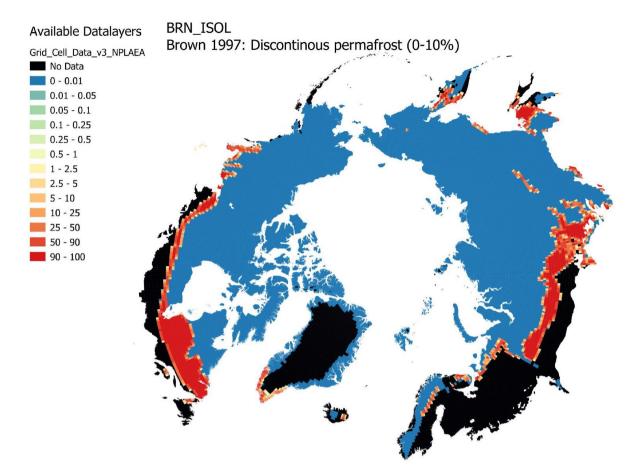


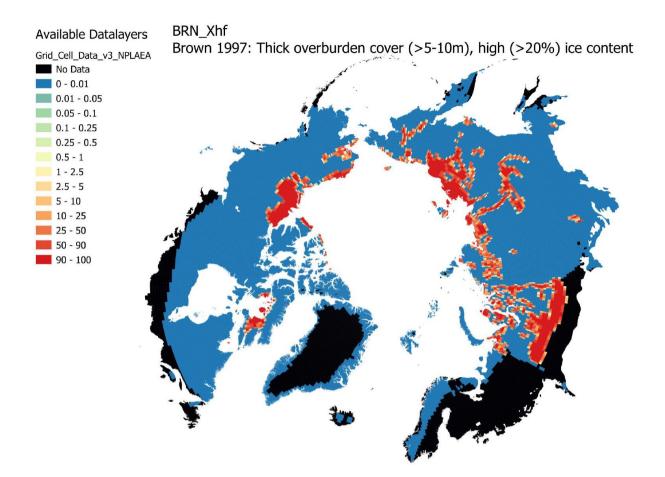


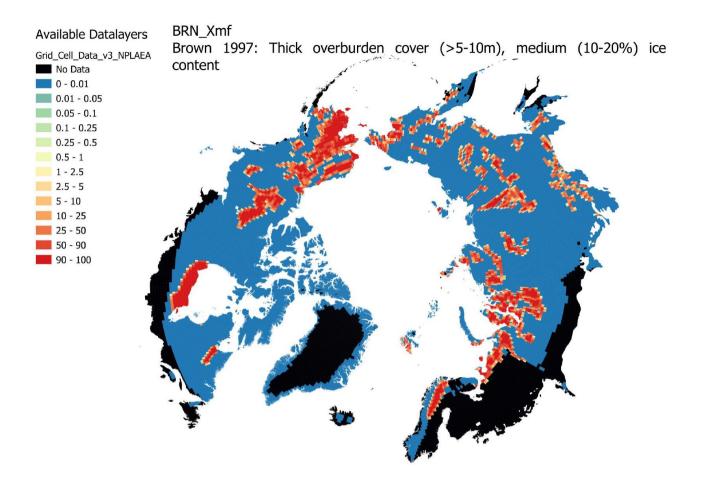


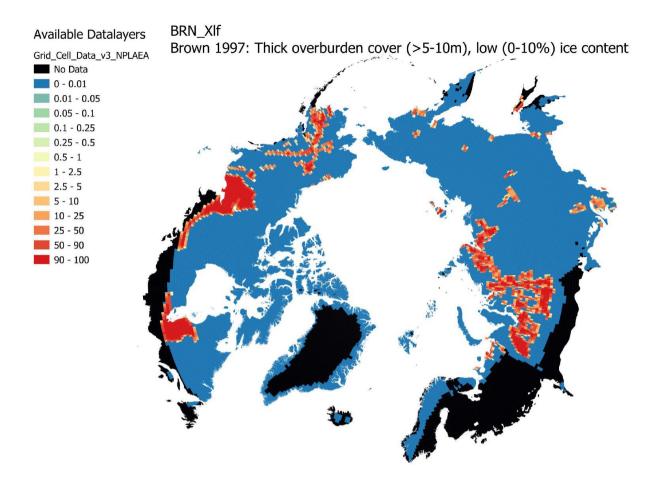












BRN_Xhr

