



Supplement of

The Boreal–Arctic Wetland and Lake Dataset (BAWLD)

David Olefeldt et al.

Correspondence to: David Olefeldt (olefeldt@ualberta.ca)

The copyright of individual parts of the supplement might differ from the article licence.

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

WETSCAPES		BAWLD LAND COVER CLASSES																	
	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	0 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	6 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	01.1	01.7	0.0	0.3	01.9	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	01.9	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	1.0	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	01.3	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)



Propor	tional 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:

File	Home	Insert	Page Lay	put Formul	las Dal	a Review	View	Add-Br	s Acro	tude																																		0	0 - 5
B *	Cut	1	Calibri	- 11 -	A' A'		*-	≣w	trap Text		Number	r			1			Iormal	I.	B	lad		Go	od		Neutral	Calc	alation		3	-	×	1	AutoSur	m - 1	7 6	8								
Paste 4	Copy - Format Pak	oter	в / ц	1 = - 4	· <u>A</u> ·		1 98 9	1 III M	lerge & Ce	nter -	s - 1	s .	54.4	1 Co For	nditional	Forma Table	t as	heck (Cell	E	xplana	tory T.	Inp	ut		Linked Cell	Note			Inser	t Del	ete Form	hat d	Clear -	Si H	ter * Sel	J&. ect *								
Clipt	oard	14		Font	6		Align	ment		6	2	wimber.		6		-	_	_	_	_	_	_	Styles						_		0	Ri .			Editio	9		_							
	23			in .																																									
- A	8 (5	D Dete in	E cells to right	Y	6	Ĥ I	J	ĸ	L	м	N	0	p	0	1	1	5	Ť	U	V	0	×	SUM of	z	AA Dats in Column	ns to Sh	AB AC	A GL	0 AE	ak GL	C30 GLC	30 GL	H A	30 GL	C30 65V	AL 6SW	AM 65¥	AN 6SW	AO GLDV		AQ HL_LA	AA LM B	AS /	N, BRN.
1 D	Area Lat Y 5	6.25 -	are sou assess 10.75	end	GLA	ROC TU	N BOR	PEB	PEW	MAR	806	FEN	LAL	MPL	MYL	MGL	. SP	L 51	YL 1	SGL	LAR	SRR	SRP	VERS	a	evailable datas	sets in	0_PC1	т_н 00	0.41	, "н	12 D	D 7.21	. AR	TI W	0_5	5_50	50_95	95_10 21 0.3	B_RIV	1	BGE 0.2	122	DNT DIS	C SPOR
3 6308 4 3828	Y 6 Y 5	825	04.25 54.75				-			_										_				00	0			0	100	7.24	14	36 2 0.43 50	0.11	1	0	0	61 DE 0 DI	2 26	65 1 25 0/	6 36 N	0 514	0	0.99	0	0 0
0 0.000 6 9014 7 2418	Y 6	872 325 - 525 -	80.00 101.25 101.25				-	-		-		-	-	-	+	+	-	-	-	-	-	-		00	000			00	100 1	5.47 3.04	0	147 1 147 1 162 29	182 40	23.4	0	0 0	10 01 10 01	5 411 29 1	21 34	2	0 - 599	7362	0.74	10	0 0
8 12771 9 8744	Y 6 Y 6	8.25 2.75	72.25 12.25								-													00	0			0 50	勝 10 ;	6.06 60 0.57 S	281	13.91 71	0	1	0	0 0	10 162	86 13.5 28 4	33 14 16 8.6	13 17	0 999	11.57	107	482	0 0
11 1200	N S	2.25	-/4 /5 -9175 -7275				_	-		-		-	-	+	+	+	+	-	-	-	_	-	-	00				00	100	7.62	0 0	7.62 11	100	1	0	8 1	81 07	59 41 11 21	01 30 07 0	6	0 19 9	264	472	-3899	9990 -9990
13 515 14 13143	N A N G	875 -	47.25 40.25							-					-	-		-		-	_			0.0	00			1 74	21	5.72 2.71	0	51 1 51	0	A S	0.15	0	0 03	9 28 59 57	83 1 48 D.	и 11	0 -191	131 0.42	2.29	3035	3030 -3030 0 0
15 304 16 13940	N 4 N 6	12	77.25 18.75					-								-					-			0.0	0			0	10	0.73 2.44 M	35	3.21 2 10.69 4	158 4	564	0	0.02	0 01	11 21 59 85	16 0.0 99 1	17 21 16	17 -991 0 -995	122	0.62	D	111 80.85
15 18199 15 279	N 1 N 4	175 -	63.75 43.75				-	-		-		-	-	-	+	-	+		-	_	_	-		00				00	10.0	0	0	6.32 2.35 3	0	· .	0	8 0	82 D.	23 84	43 0. 56 0.0	2	0 499	0	6.27	10	0 0
20 4533 21 333	N 5	875 - 825 -	10.75 72.75							_								1				_		0.0	0			0	100 8	4.11	29 8 0.1	336 5 505 0	1.76		0	0	01 21	61 3 51 3	7 50.8	16 17	0 -999	53.53 0	2.29 3.78	-9899	0 100
	Fill in all these cells													ŝ	1	Ensure	eea	ach	ro	w	0/																								
																										aaas u	ip t	0 1	00	.00	70														

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























