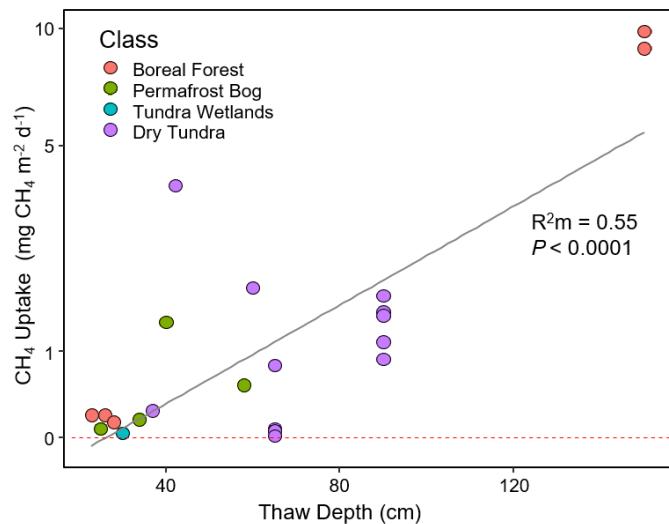
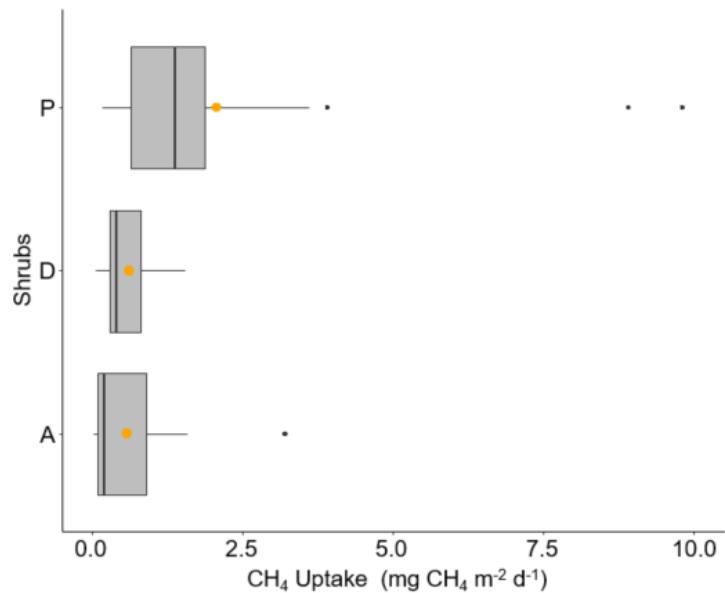
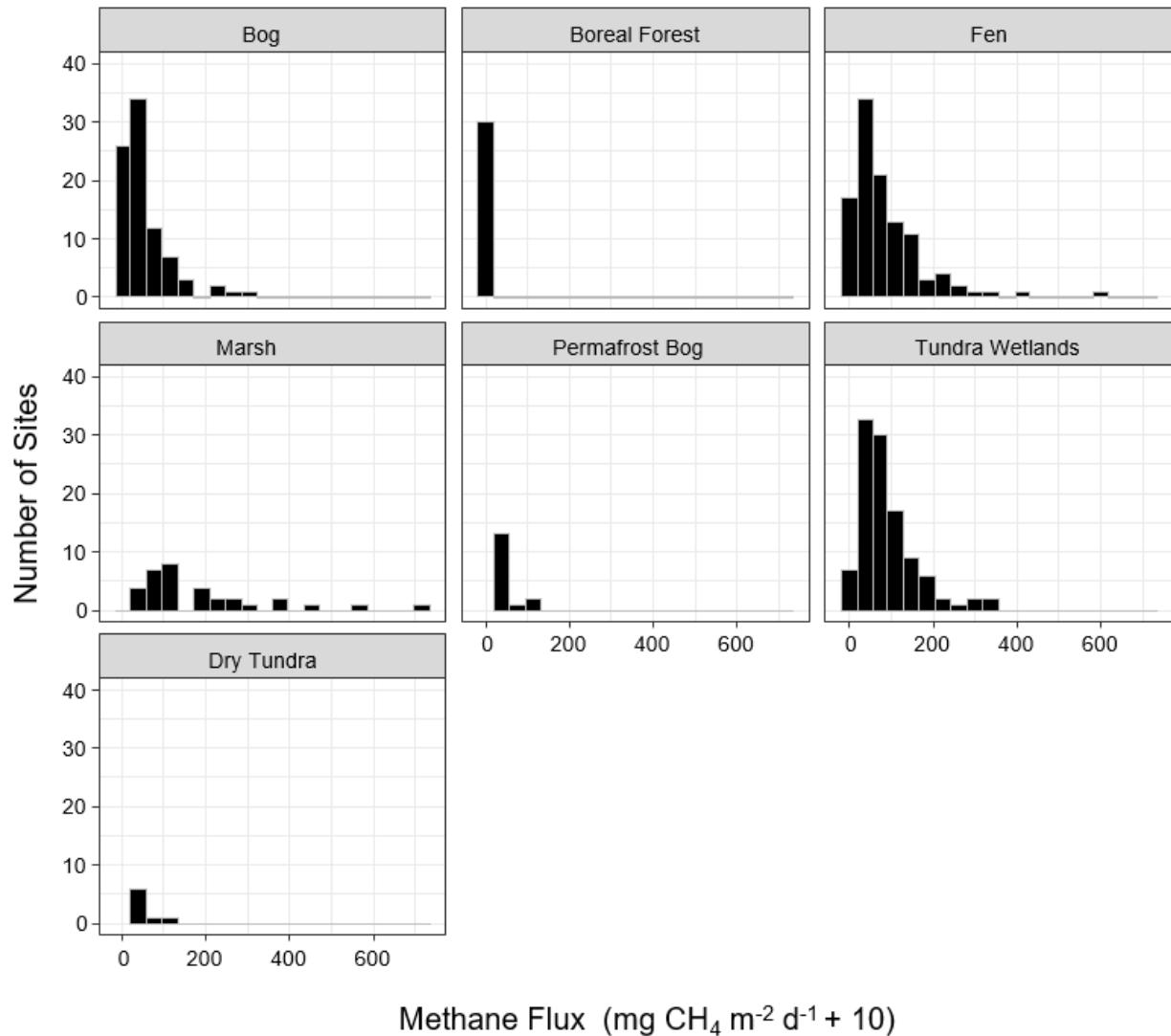
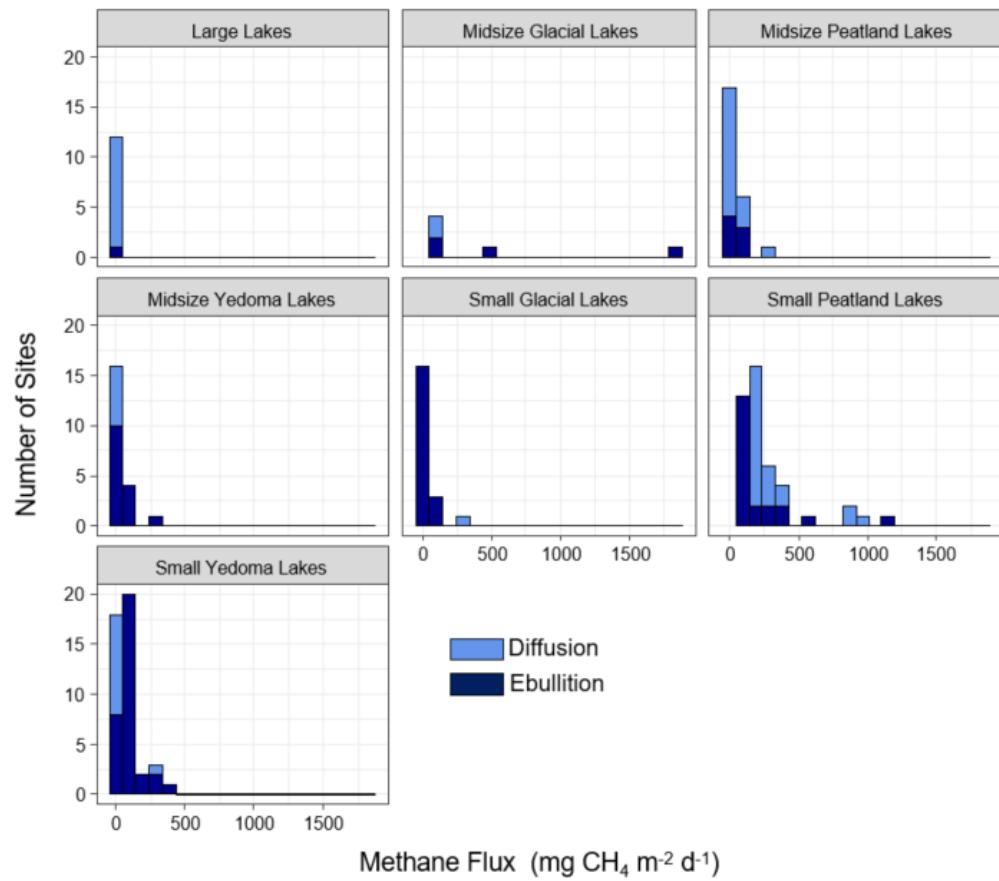
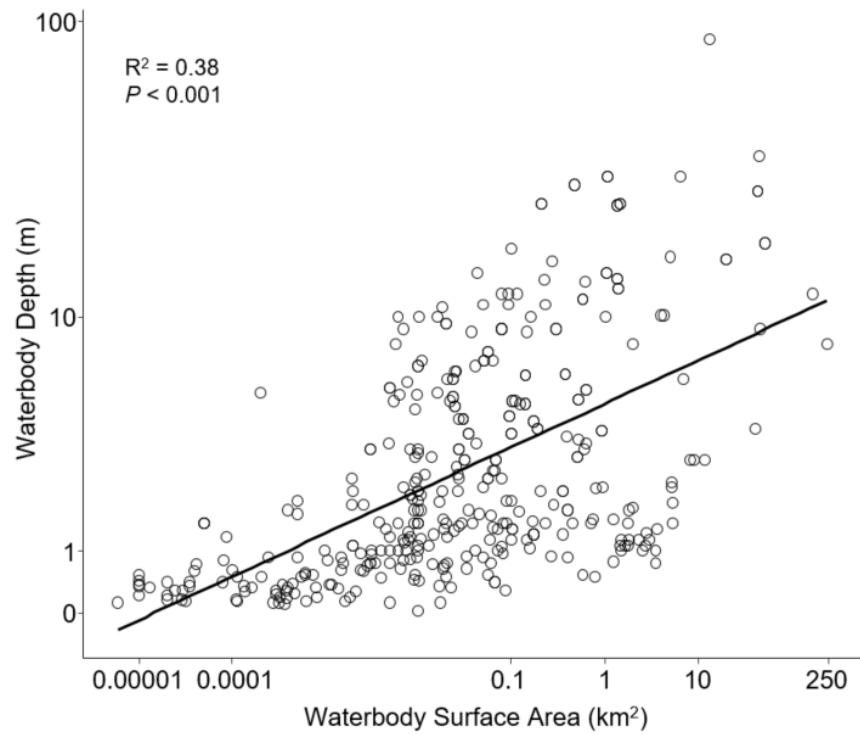
Supplementary Information

BAWLD-CH4: A Comprehensive Dataset of Methane Fluxes from Boreal and Arctic ecosystems

Kuhn, M.A.[1]., Varner, R.K.[2,3]., Bastviken, D.,[4] Crill, P.[5,6], MacIntyre, S.[7], Turetsky, M.[8], Walter Anthony, K.[9], McGuire, A.D.[10] and Olefeldt, D[1].

[1] Department of Renewable Resources, University of Alberta. T6E 1V6, Edmonton, Alberta, Canada
[2] Department of Earth Sciences and Earth System Research Center, Institute for the Study of Earth, Oceans and Space, University of New Hampshire, Durham, NH 03824 USA
3 Department of Physical Geography, Stockholm University, 10691 Stockholm, Sweden
[4] Department of Thematic Studies – Environmental Change, Linköping University, SE-581 83 Linköping, Sweden
[5] Department of Geological Sciences, Stockholm University, Stockholm, Sweden
[6] Bolin Centre for Climate Research, Stockholm, Sweden
[7] Marine Science Institute, University of California at Santa Barbara, Santa Barbara, USA
[8] Institute of Arctic and Alpine Research (INSTAAR), University of Colorado Boulder, Boulder, CO, USA.
[9] Water and Environmental Research Center, University of Alaska Fairbanks, PO Box 755860, Fairbanks, Alaska 99775-5860, USA.
[10] Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, Alaska 99775, USA

Correspondence to: McKenzie A. Kuhn (kuhn.mckenzie@gmail.com)

SI Figure 1. Correlation between thaw depth and methane uptake. Positive numbers represent net uptake from the atmosphere. Colors represented different ecosystem class. Neutral (i.e. zero) fluxes were not included in the regression analyses.

SI Figure 2. Boxplots of CH₄ uptake as a factor of shrub cover. Positive numbers represent net uptake from the atmosphere. Orange dots represent mean uptake within a given category. P = Present; D = Dominant; A = Absent. Neutral fluxes were not included in the boxplots.

SI Figure 3. Non-transformed flux frequencies across the terrestrial ecosystem classes. A constant of 10 was added to include CH_4 uptake.

SI Figure 4. Non-transformed flux frequencies across the aquatic land cover classes.

SI Figure 5. Relationship between waterbody surface area and water column depth.

SI Table 1. Temperature sensitivities of methane fluxes (Q_{10} values) across terrestrial and aquatic ecosystem classes. Individual classes without Q_{10} values did not have significant relationships with temperature when analyzed on their own or did not have a large enough sample size (n = 15).

Class or group of classes	Q10
All terrestrial class	2.88
"Wet" terrestrial classes (Marshes, Tundra Wetlands, Fens, Bogs)	2.82
"Dry" terrestrial classes (Dry Tundra and Boreal Forest, Permafrost Bogs)	3.71
Marshes	-
Tundra Wetlands	2.57
Fens	1.99
Bogs	3.39
Permafrost Bogs	-
Dry Tundra	2.63
Boreal Forest	-
All aquatic classes diffusion	4.27
All peatland lakes diffusion	2.63
All yedoma lakes diffusion	3.89
All glacial lakes diffusion	-
All aquatic class ebullition	2.40

SI Table 2: Model selection for terrestrial CH$_4$ emissions. "Site" represents the best model using site level predictor variables (biophysical variables measured directly by the authors). "Region" represents the best model using predictor variables that can be attributed across larger spatial scales and extracted from gridded or mapped products. Tests with "site and region" represent the model models that include both site level and regional level predictors. The null model includes only the random effect of SiteID. The best models for each test represented here were picked through forward model selection. K = number of fixed terms the model, AICc = size-corrected Akaike information criterion, DeltaAICc = change in AICc between a given model and the best model, AICcwt = AICc weights indicating the probability a given model is the most parsimonious model in the group of models tested, R^{2}m = marginal R^2 for the fixed terms for mixed models. R^{2}c = conditional R^2 for fixed and random terms for mixed effects models. See main text for explanation of fixed effects short names. Non-significant fixed terms that were tested include: MAAT, MAP, Permafrost Zone, Permafrost Presence or Absence, and Biome. TsoilB = soil temperature at 2-25 cm. WTAv – average water table position. Sedge = graminoid cover.

Variable modeled	Test	Fixed effect	K	AICc	DeltaAICc	AICcwt	R^{2}m/R^{2}c
Terrestrial Log.CH4.Flux (n=206)	Site*region	TsoilB*Class+ WTAv*Class + Sedge	25	98.87	0		0.73/0.83
	Site + region	TsoilB + WTAv+ Sedge + Class	13	99.70	0.84		0.69/0.81
	Site	TsoilB + WTAv+ Sedge	7	131.20	32.34		0.54/0.78
	Region	Class	9	172.77	73.91		0.55/0.71
	Null	-	3	275.8	176.92		0/0.72

SI Table 3: Model selection for aquatic diffusive CH$_4$ emissions. "Site" represents the best model using site level predictor variables (biophysical variables measured directly by the authors). "Region" represents the best model using predictor variables that can be attributed across larger spatial scales and extracted from gridded or mapped products. Tests with "site and region" represent the model models that include both site level and regional level predictors. The null model was ran as follows lm(log.CH4.flux) ~1). The best models for each test represented here were picked through forward model selection. K = number of fixed terms the model, AICc = size-corrected Akaike information criterion, DeltaAICc = change in AICc between a given model and the best model, AICcwt = AICc weights indicating the probability a given model is the most parsimonious model in the group of models tested, R^2m = marginal R^2 for the fixed terms for mixed models. See main text for explanation of predictor variable short names. Non-significant predictor terms that were tested include MAP, Permafrost Zone, DOC, Biome, waterbody depth, and Class). SA = waterbody surface area. TYPE = overarching lake type by lake genesis. TEMP = measured water temperature. GRID_T = gridded mean annual temperature. LAT = latitude.

Variable modeled	Test	Predictor variable	K	AICc	DeltaAICc	AICcwt	Adj R^2
Aquatic diffusion Log.CH4.Flux (n=149)	Site*Region	log10(SA)*TYPE + TEMP	8	162.5	0	0.79	0.41
	Site	log10(SA) + TEMP	4	165.8	3.3	0.94	0.36
	Site + Region	log10(SA) + TEMP + GRID_T	7	167.5	5.0	1.0	0.38
	Region	GRID_T + TYPE + LAT	6	201.7	39.2	1.0	0.35
	Null	-	3	215.1	52.6	1.0	-

SI Table 4. Joint analysis of terrestrial and aquatic growing season/ice-free emissions.
The best models for each test represented here were picked through forward model selection. K = number of fixed terms the model, AICc = size-corrected Akaike information criterion, DeltaAICc = change in AICc between a given model and the best model, AICcwt = AICc weights indicating the probability a given model is the most parsimonious model in the group of models tested, R^2m = marginal R^2 for the fixed terms for mixed models. R^2c = conditional R^2 for fixed and random terms for mixed effects models. GRID_P = gridded mean annual precipitation. GRID_T = gridded mean annual temperature.

Variable modeled	Fixed effect	K	AICc	DeltaAICc	AICcwt	R^2m
Log.CH4.Flux (n=793)	Class + GRID_T	17	959.8	0	0.64	0.47
	Class + GRID_T + Biome	18	961.6	1.83	0.25	0.47
	Class + Grid_T + Permafrost Zone	21	963.3	3.53	0.11	0.47
	Class + GRID_P + Permafrost Zone	21	977.3	17.5	0	0.46
	Class + GRID_P	17	979.5	19.7	0	0.44
	Class + Permafrost Zone	20	994.4	34.7	0	0.46
	Class	16	996.4	36.6	0	44
	Null	3	1276.7	316.9	0	-