Introduction

This file contains detailed results of the statistical analyses performed to compare the biogeochemistry of peatland ponds, thermokarst waterbodies and lakes. Table S1 shows GLM results, Table S2 and S3 shows k-means clustering results, Table S4 shows discriminant analysis results, and Table S5 shows GLM results of interacting effects.

Table S1. Results of the generalized linear models (GLMs) performed to compared peatland ponds, thermokarst waterbodies and lakes morphometric and chemical data. *Gamma* (for depth, and DOC, TN and TP concentrations) and *Gaussian* (for area, after log-transformation, and pH) distribution families were used. Reference system for each model is in parentheses. Systems with different level letters are statistically different at a significance level of 5% ($P < 0.05$). GLMs were performed on the whole dataset, from file Arsenault-et-al_All.csv (doi:10.5281/zenodo.5619484).

Response	n	System	t	*P*	System	t	*P*	Level
Area (log)	7 015	(Lakes)	-69.55	-	Lakes	46.43	<0.001	A
	211	Peatland ponds	-46.43	<0.001	(Peatland ponds)	-59.18	-	B
	452	Thermokarst waterbodies	-43.49	<0.001	Thermokarst waterbodies	13.65	<0.001	C
Depth	2 029	(Lakes)	24.86	-	Lakes	-8.90	<0.001	A
	286	Peatland ponds	8.90	<0.001	(Peatland ponds)	9.34	-	B
	248	Thermokarst waterbodies	8.08	<0.001	Thermokarst waterbodies	-2.57	0.010	B
DOC	8 286	(Lakes)	83.82	-	Lakes	22.48	<0.001	A
	278	Peatland ponds	-22.48	<0.001	(Peatland ponds)	15.35	-	B
	656	Thermokarst waterbodies	-34.21	<0.001	Thermokarst waterbodies	-1.10	0.273	B
pH	9 576	(Lakes)	594.82	-	Lakes	34.63	<0.001	A
	233	Peatland ponds	-34.63	<0.001	(Peatland ponds)	57.75	-	B
	600	Thermokarst waterbodies	-26.14	<0.001	Thermokarst waterbodies	15.63	<0.001	C
TN	2 185	(Lakes)	23.32	-	Lakes	-0.88	0.381	A
	237	Peatland ponds	0.88	0.381	(Peatland ponds)	7.71	-	A
	163	Thermokarst waterbodies	-1.26	0.206	Thermokarst waterbodies	-1.57	0.118	A
TP	3 022	(Lakes)	5.49	-	Lakes	-1.49	0.135	A
	237	Peatland ponds	1.49	0.135	(Peatland ponds)	1.54	-	A
	243	Thermokarst waterbodies	1.46	0.145	Thermokarst waterbodies	-0.81	0.418	A

Table S2. Composition of the 5 k-means clusters based on the analysis of the pH and the dissolved organic carbon (DOC), total nitrogen (TN) and total phosphorus (TP) concentrations for 837 freshwater ecosystems. This analysis was performed using a subset of waterbodies, from file Arsenault-et-al_Nutrients.csv (doi:10.5281/zenodo.5619484).

	Peatland ponds	Thermokarst waterbodies	Lakes
K1 (n = 234)	82.0 % (n = 192)	1.3 % (n = 3)	16.7 % (n = 39)
K2 (n = 50)	34.0 % (n = 17)	60.0 % (n = 30)	6.0 % (n = 3)
K3 (n = 85)	0.0 % (n = 0)	1.2 % (n = 1)	98.8 % (n = 84)
K4 (n = 236)	0.9 % (n = 2)	22.0 % (n = 52)	77.1 % (n = 182)
K5 (n = 232)	0.9 % (n = 2)	9.5 % (n =22)	89.6 % (n = 208)

Table S3. Composition of the 3 k-means clusters based on the analysis of the area, depth, dissolved organic carbon (DOC) concentrations and pH for 1235 freshwater ecosystems < 0.1 km^2. This analysis was performed using a subset of waterbodies, from file Arsenault-et-al_Small-waterbodies.csv (doi:10.5281/zenodo.5619484).

	Peatland ponds	Thermokarst waterbodies	Lakes
K1 (n = 290)	56.5 % (n = 164)	41.4 % (n = 120)	2.1 % (n = 6)
K2 (n = 654)	0.0 % (n = 0)	0.5 % (n = 3)	99.5 % (n = 651)
K3 (n = 291)	2.4 % (n = 7)	21.6 % (n = 63)	76.0 % (n = 221)

Table S4. Absolute number of predicted systems and proportion of rightfully predicted systems as determined by the discriminant analysis of 837 freshwater ecosystems based on pH and dissolved organic carbon (DOC), total nitrogen (TN) and total phosphorus (TP) concentrations. The model is based on a random subsample of 30% of the initial sample. This analysis was performed using a subset of waterbodies, from file Arsenault-et-al_Nutrients.csv (doi:10.5281/zenodo.5619484).

		Predicted system		
		Peatland ponds	Thermokarst waterbodies	Lakes
A priori classification	Peatland ponds (n = 149)	**142 (95.3%)**	3	4
	Thermokarst waterbodies (n = 75)	1	**25 (33.3%)**	49
	Lakes (n = 361)	12	9	**340 (94.2%)**

Table S5. Results of the generalized linear models (GLMs) of interacting effects of freshwater system type and their morphometry (area and depth) on DOC concentrations and pH. Reference systems in the models are peatland ponds (in parentheses). GLMs were performed on the whole dataset, from file Arsenault-et-al_All.csv (doi:10.5281/zenodo.5619484).

Response	Interaction	t	*P*
DOC vs Area	(Peatland ponds)	12.57	<0.001
	Area	1.47	0.141
	Thermokarst waterbodies	-1.70	0.090
	Lakes	18.03	<0.001
	Area:Thermokarst waterbodies	-1.47	0.141
	Area:Lakes	-1.47	0.141
DOC vs Depth	(Peatland ponds)	4.93	<0.001
	Depth	0.45	0.655
	Thermokarst waterbodies	-2.87	0.004
	Lakes	5.12	<0.001
	Depth: Thermokarst waterbodies	1.59	0.111
	Depth:Lakes	0.17	0.868
pH vs Area	(Peatland ponds)	55.82	<0.001
	Area	2.18	0.029
	Thermokarst waterbodies	8.99	<0.001
	Lakes	34.10	<0.001
	Area: Thermokarst waterbodies	-2.18	0.029
	Area:Lakes	-2.18	0.029
pH vs Depth	(Peatland ponds)	44.35	<0.001
	Area	-0.52	0.602
	Thermokarst waterbodies	5.84	<0.001
	Lakes	30.21	<0.001
	Depth: Thermokarst waterbodies	3.44	<0.001
	Depth:Lakes	0.51	0.609