

Figure 1: Diagram showing the relations between true (black) and proxy (orange) metrics of lake geometry. Geometric depth calculated via Equation 1 requires a single distance and slope metric.

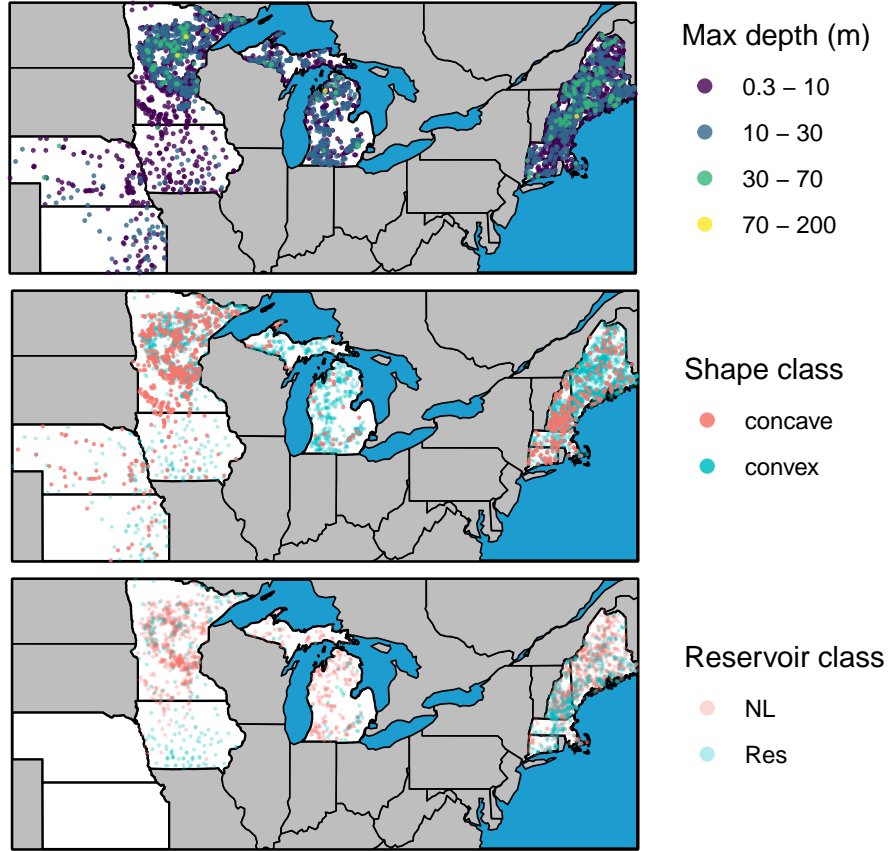


Figure 2: Map of study lakes showing A) lake e maximum depth measurements, B) cross-section shape class, and C) reservoir classification. The distribution of lake depths from panel A is reported in Figure S4.

Variable	Median	Q25	Q75	n
Max depth (m)	8.2 (7)	4.6 (3.7)	14 (12)	4820 (17700)
Area (ha)	55 (33)	21 (11)	140 (100)	4820 (17700)
Island area (ha)	0 (0)	0 (0)	0.18 (0.076)	4820 (17700)
Perimeter (m)	4400 (3500)	2500 (1800)	8100 (7300)	4820 (17700)
Shoreline development	1.7 (1.7)	1.4 (1.4)	2.1 (2.2)	4820 (17700)
Elevation (m)	300 (340)	180 (210)	400 (460)	4820 (17700)
Watershed-lake ratio	7.8 (10)	3.8 (4.4)	17 (29)	4820 (17700)
Deepest point distance (m)	180 (-)	110 (-)	290 (-)	4820 (-)
Mean deepest point distance (m)	140 (-)	87 (-)	230 (-)	4820 (-)
Visual center distance (m)	240 (-)	160 (-)	390 (-)	4820 (-)
Inlake slope (m/m)	0.05 (-)	0.02 (-)	0.08 (-)	4820 (-)
Inlake slope online (m/m)	0.06 (-)	0.03 (-)	0.14 (-)	4800 (-)
Inlake slopes (m/m)	0.06 (-)	0.03 (-)	0.1 (-)	4820 (-)
Inlake slopes online (m/m)	0.07 (-)	0.03 (-)	0.15 (-)	4800 (-)
Mean inlake slope (m/m)	0.04 (-)	0.02 (-)	0.09 (-)	4820 (-)
Nearshore mean slope (m/m)	0.08 (-)	0.05 (-)	0.11 (-)	4820 (-)
Nearshore slope online (m/m)	0.08 (-)	0.04 (-)	0.13 (-)	4590 (-)
Nearshore slopes online (m/m)	0.08 (-)	0.04 (-)	0.13 (-)	4540 (-)

Table 1: Summary of lake characteristics for the present study (and for lakes in the contiguous United States). Predictor variables for computing random forest offsets (Equation 2) are printed in bold face. Dashes (-) indicate an identical sample size among this study and that of the contiguous United States from the National Hydrography Dataset. The total number lakes is reported as n.

filter	slope	distance	rmse	rsq	mape
all	true	true	-	-	-
	true	proxy	4.8 m	0.73	27 %
	proxy	true	7.3 m	0.31	64 %
	proxy	proxy	7.1 m	0.36	61 %
reservoir	true	true	-	-	-
	true	proxy	5.3 m	0.66	36 %
	proxy	true	7 m	0.40	61 %
	proxy	proxy	7 m	0.41	60 %
natural lake	true	true	-	-	-
	true	proxy	4.1 m	0.74	22 %
	proxy	true	6.7 m	0.26	68 %
	proxy	proxy	6.6 m	0.29	64 %
convex	true	true	-	-	-
	true	proxy	4.7 m	0.74	30 %
	proxy	true	7.2 m	0.34	59 %
	proxy	proxy	6.9 m	0.39	58 %
concave	true	true	-	-	-
	true	proxy	1.6 m	0.78	20 %
	proxy	true	3.1 m	0.14	46 %
	proxy	proxy	3 m	0.17	45 %

Table 2: Model fit and predictive accuracy metrics (RMSE = root mean square error, R2 = coefficient of determination) for all combinations of true (in-lake slope, distance to the deepest point of the lake) and proxy (nearshore land slope, distance to lake center) metrics.

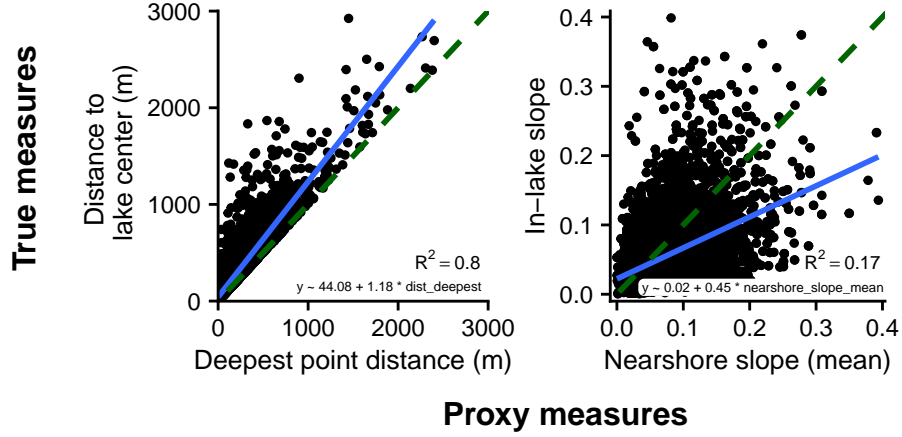


Figure 3: Comparison among proxy and true values of lake geometry for A) distance to deepest point versus distance distance to lake center and B) nearshore land slope versus in-lake slope. A best-fit line and equation is shown to shown to facilitate computation of correction factors for proxy values of lake geometry. Coefficients of determination are shown to illustrate representativeness.

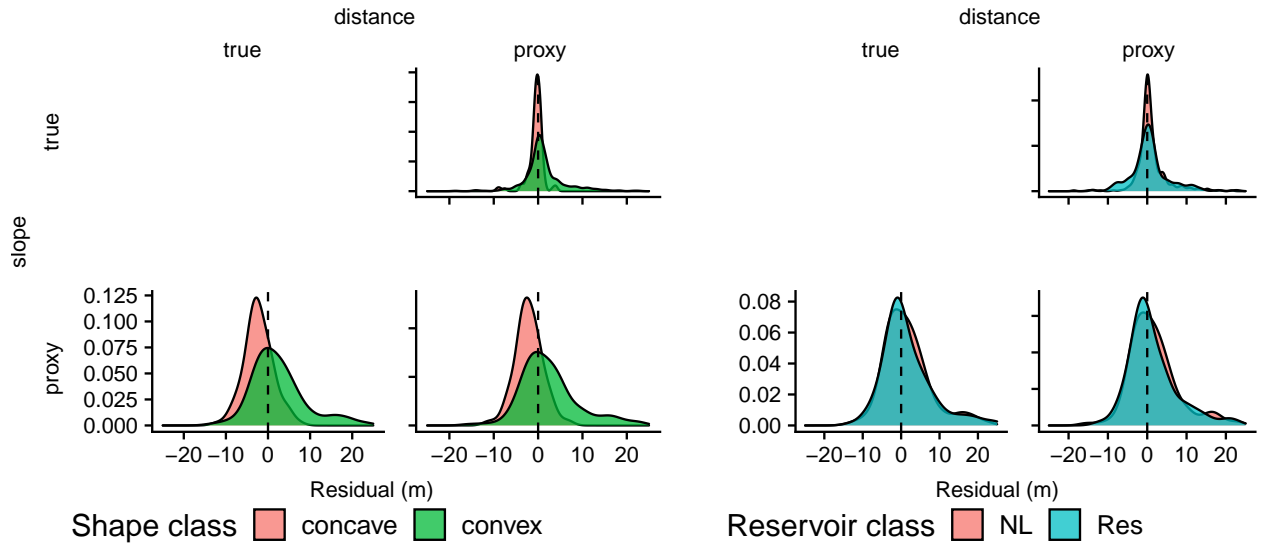


Figure 4: Depth model residuals (residual = observed - predicted) in meters by cross-section shape and reservoir class indicating overprediction of concave and reservoir lakes.

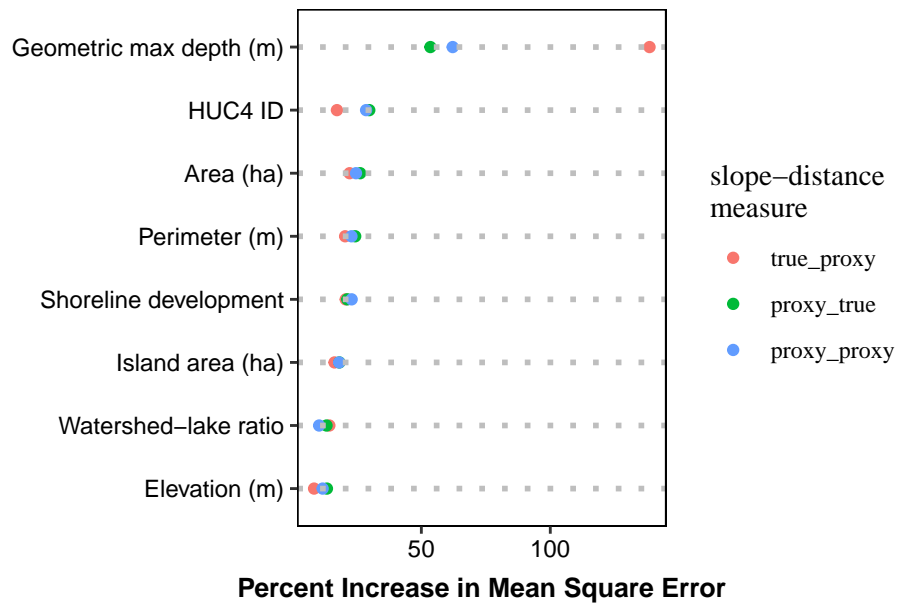


Figure 5: Importance plot for random forest variables showing increase in mean square error. Higher values indicate greater importance to model predictions. See Equation 1 for a definition of geometric max depth. HUC4 ID is a 'dummy' variable of geographic (hydrologic subbasin) location.