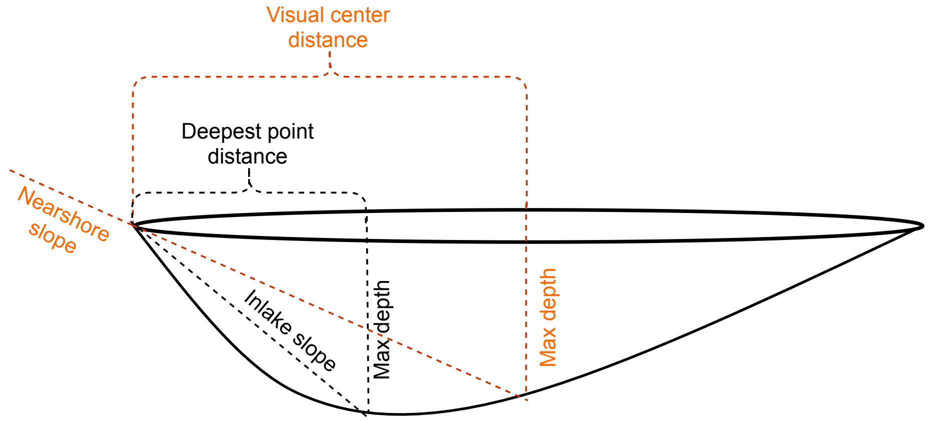

Figure 1: Diagram showing the relations between true (black) and proxy (orange) metrics of lake geometry. Geometric depth calculated via Equation 1 requires a single distance and slope metric.

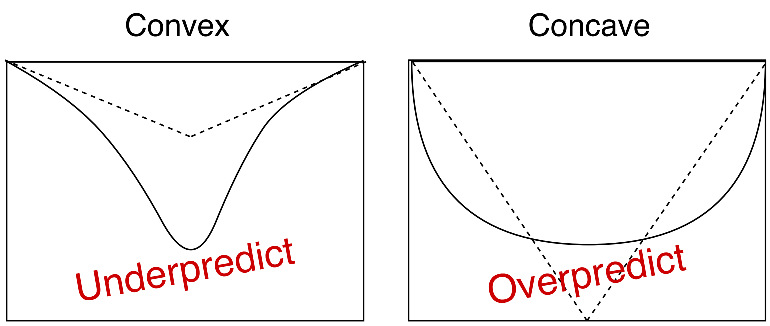

Figure 2: Diagram showing our expectation that slope-based models of lake depth will under predict true depth in convex lakes (left) and over predict true depth in concave lakes (right). Dashed lines represent extrapolated nearshore slope while solid lines represent the lake bottom.

Variable	Median	Q25	Q75	n
Max depth (m)	8.2 (7)	4.6 (3.7)	14 (12)	4850 (17700)
Elevation (m)	300 (340)	180 (210)	400 (460)	4850 (17700)
Area (ha)	55 (33)	21 (11)	140 (100)	4850 (17700)
Island area (ha)	0 (0)	0 (0)	0.18 (0.076)	4850 (17700)
Perimeter (m)	4400 (3500)	2500 (1800)	8100 (7300)	4850 (17700)
Shoreline development	1.7 (1.7)	1.4 (1.4)	2.1 (2.2)	4850 (17700)
Watershed-lake ratio	7.8 (10)	3.9 (4.4)	17 (28)	4850 (17700)
Deepest point distance (m)	180 (-)	110 (-)	290 (-)	4850 (-)
Visual center distance (m)	240 (-)	160 (-)	380 (-)	4850 (-)
Inlake slope (m/m)	0.046 (-)	0.024 (-)	0.079 (-)	4850 (-)
Nearshore slope (m/m)	0.077 (-)	0.051 (-)	0.11 (-)	4850 (-)

Table 1: Summary of lake depth and predictor variables for computing random forest offsets (and for lakes in the contiguous United States from <LAGOSUS-Depth citation>). Predictor variables used in Eq 2 are printed in bold face. Dashes (-) indicate an identical sample size among this study and that of the contiguous United States.

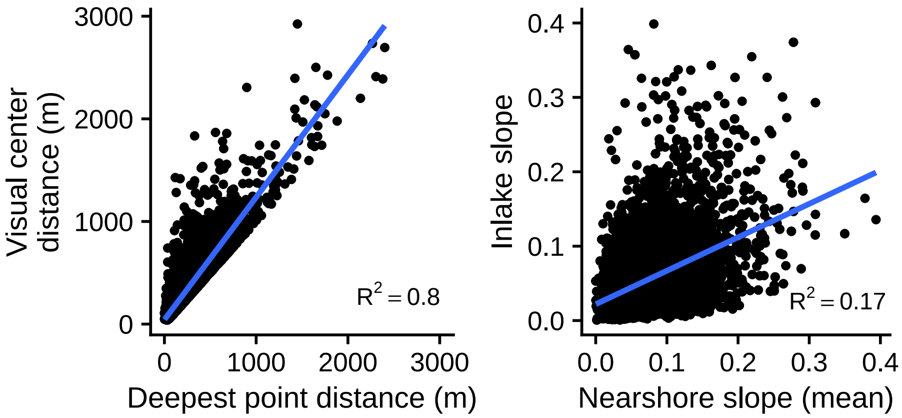

Figure 3: Comparison among proxy and true values of lake geometry for A) distance to deepest point versus distance to lake visual center and B) nearshore slope versus inlake slope. A best-fit line and coefficient of determination is shown to illustrate representativeness.

slope	distance	rmse	rsq
true	true	-	-
true	proxy	4.23 m	0.73
proxy	true	6.87 m	0.26
proxy	proxy	6.61 m	0.31

Table 2: Model fit and predictive accuracy metrics for all combinations of true (inlake slope, deepest point distance) and proxy (nearshore slope, visual center distance) metrics.

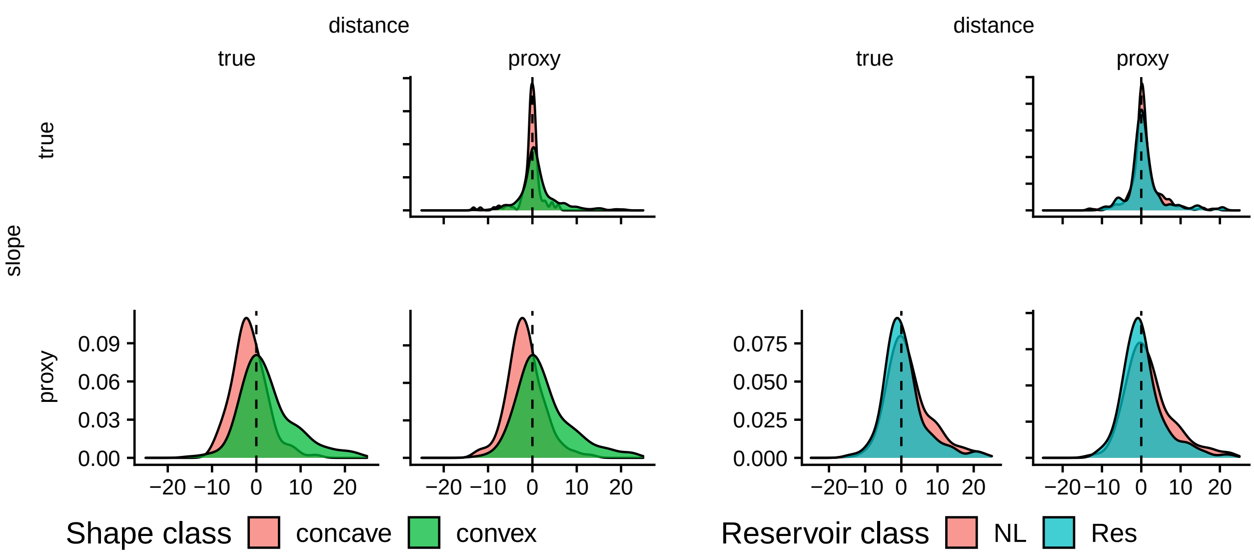

Figure 4: Depth model residuals in meters by cross-section shape and reservoir class.

Water Resources Research

Supporting information for

Bathymetry data from thousands of lakes show that lake depth prediction is confounded by difficulty modeling inlake slope

J. Stachelek[1], P. Hanly[1], and P.A. Soranno[1]

[1]Department of Fisheries and Wildlife, Michigan State University, 480 Wilson Rd., East Lansing, Michigan 48824 USA

Contents of this file

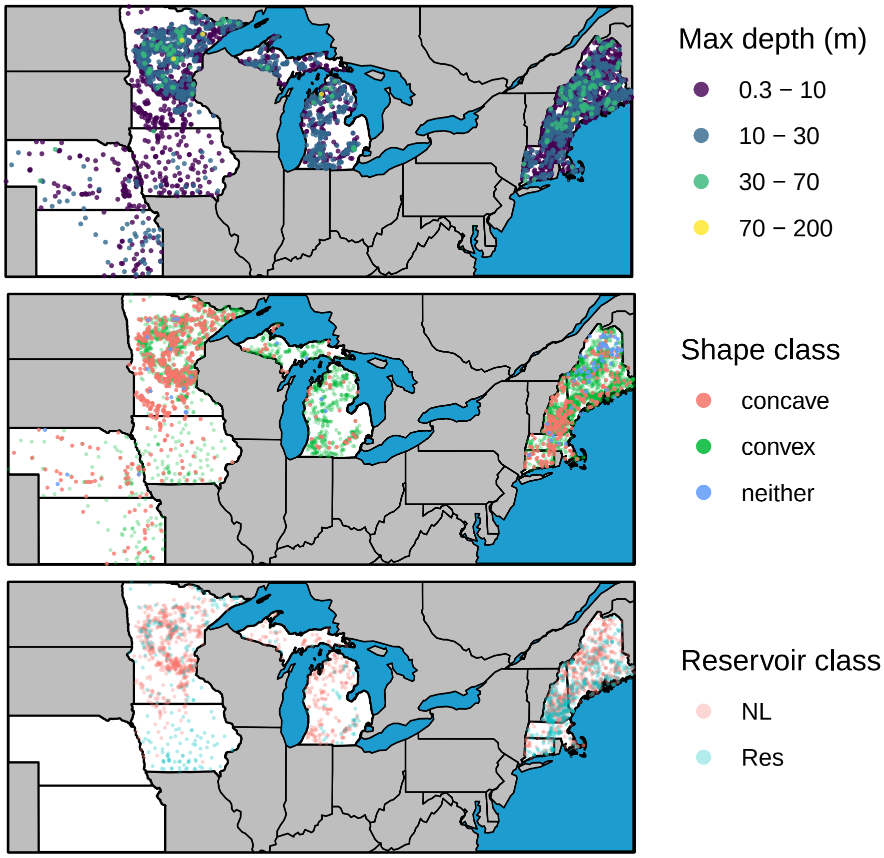

Figure S1: Map of study lakes showing A) lake maximum depth measurements, B) cross-section shape class, and C) reservoir classification.

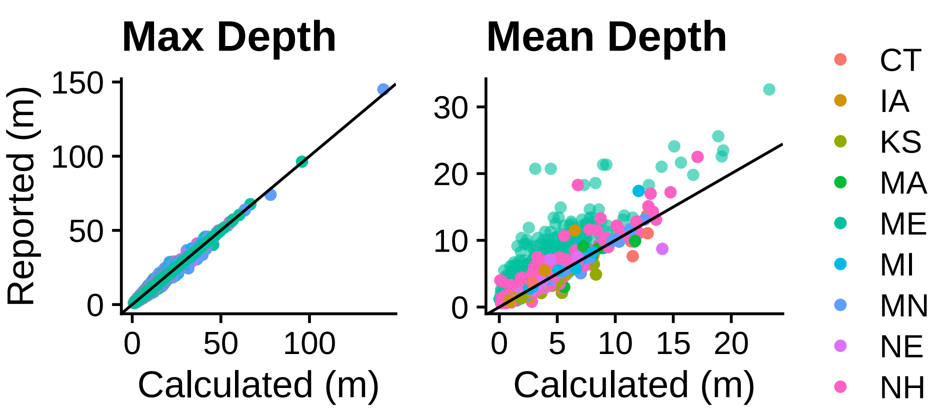

Figure S2: Comparison between reported depth and depth estimated from bathymetry surfaces by US State where reported depths come from the LAGOSUS-Depth product (citation). For this figure, no reported depth values originated from the same source as its corresponding bathymetry-derived value.

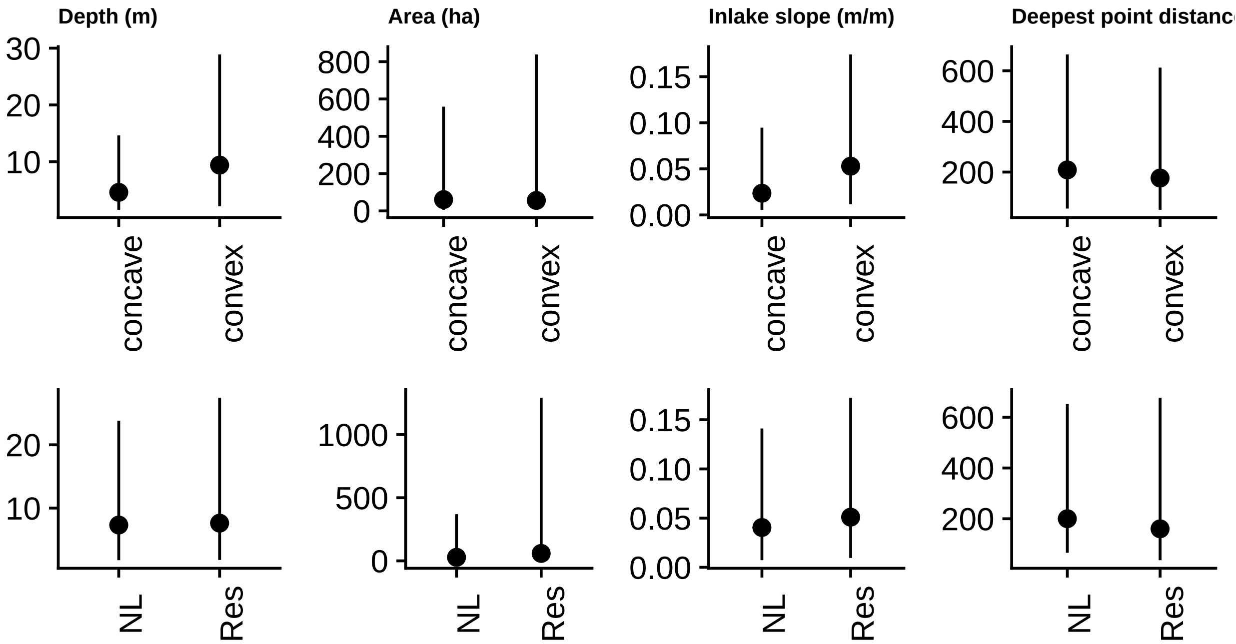

Figure S3: Lake characteristics by categorical variables.

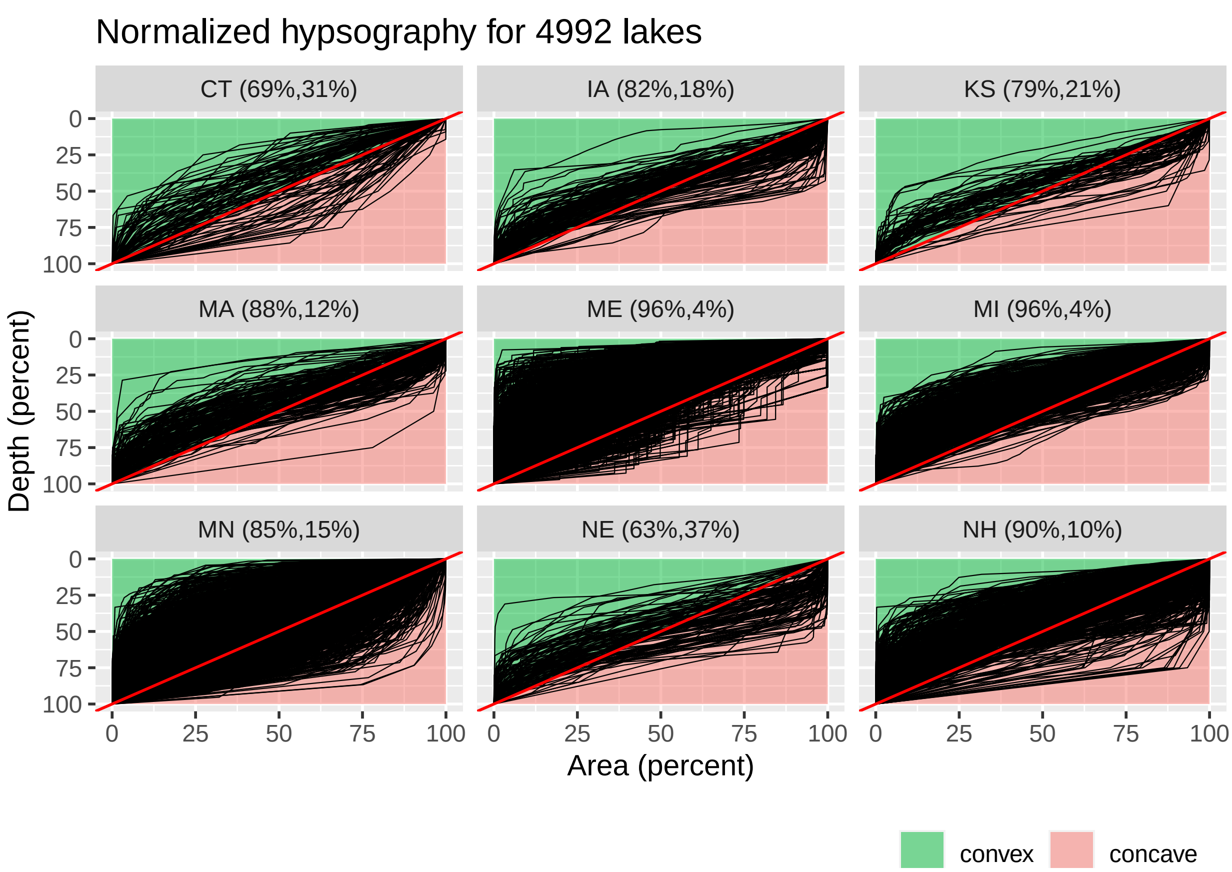

Figure S4: Hypsography classification by state. Numbers on panel labels indicate the percentage of lakes in each state with a convex versus a concave cross-section shape.

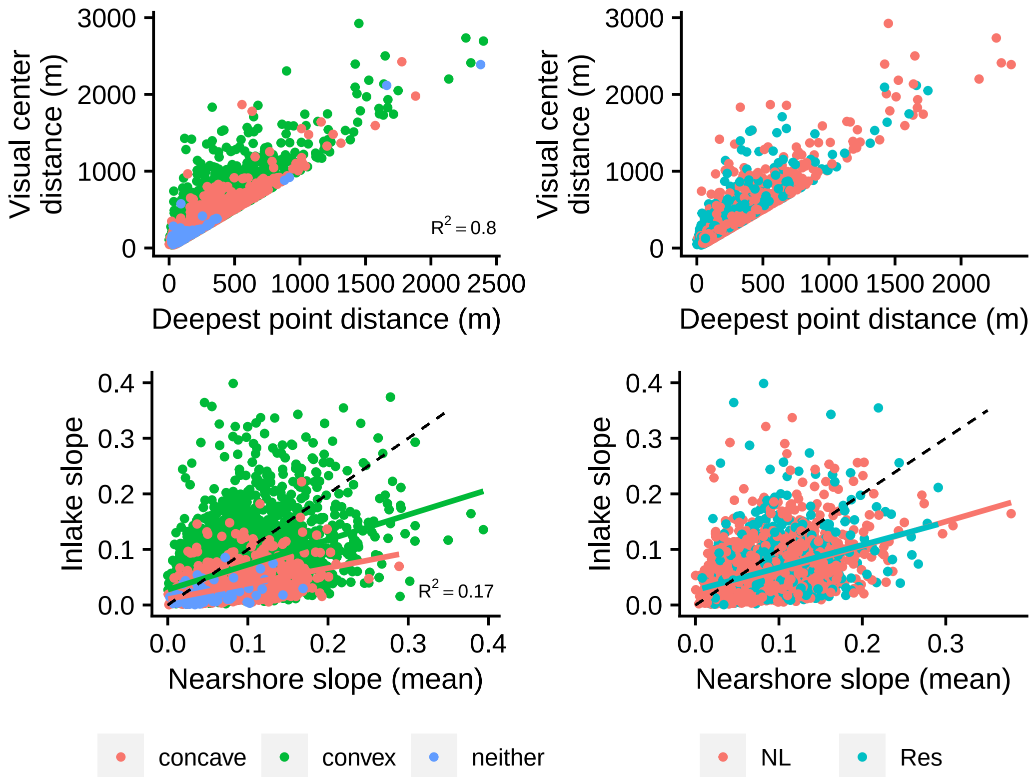

Figure S5: Comparison among lake shape and reservoir classes for A-B) distance to deepest point versus distance to lake visual center and C-D) nearshore slope versus inlake slope. A dashed 1:1 line is shown for comparison. Cross-section shape and reservoir class plots are not identical because not all lakes had a reservoir classification exceeding a 0.75 probability confidence level.

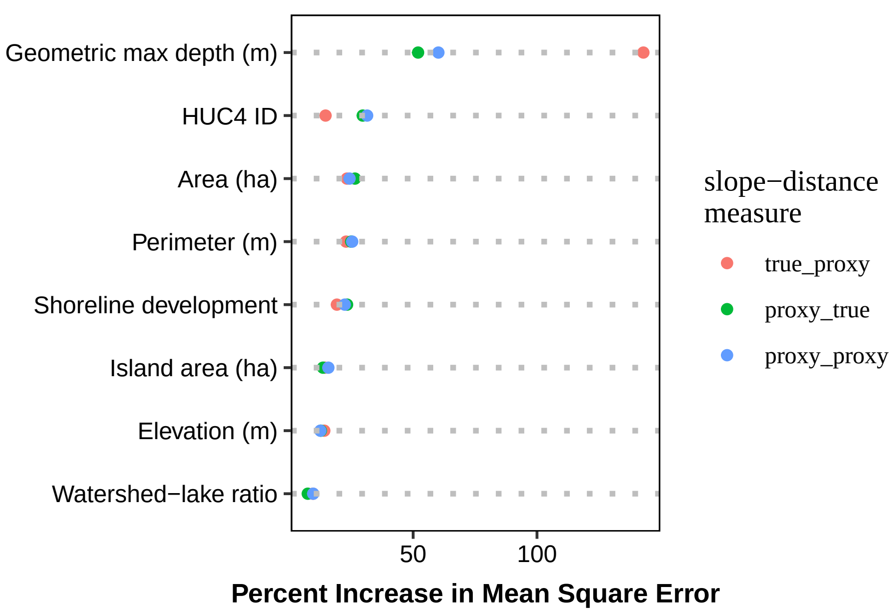

Figure S6: Importance plot for random forest variables showing increase in mean square error. Higher values indicate greater importance to model predictions. See Equation 1 for a definition of geometric max depth. HUC4 ID is a 'dummy' variable of geographic (hydrologic subbasin) location.