

Supporting Information for "Imperfect slope measurements drive overestimation in geometric models of lake depth"

J. Stachelek¹, P. J. Hanly¹, and P. A. Soranno¹

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

Contents of this file

1. **Figure S1:** Diagram showing expectations regarding depth prediction bias
 2. **Figure S2:** Hypsography classification by state
 3. **Figure S3:** Lake geometry comparison among shape and reservoir classes
 4. **Figure S4:** Lake geometry comparison among shape classes by depth and area
 5. **Figure S5:** Effect of calculation technique on in-lake/nearshore slope comparison.
 6. **Figure S6:** Importance plot for random forest variables
 7. **Figure S7:** Characteristics of lakes with bathymetry versus non-bathymetry lakes
 8. **Figure S8:** Lake characteristics comparison among shape and reservoir classes
 9. **Figure S9:** Comparison between reported depth and bathymetry depth
 10. **Table S1:** Summary of lake characteristics for the present study
 11. **Table S2:** Effect of calculation technique on depth model fit and accuracy.
-

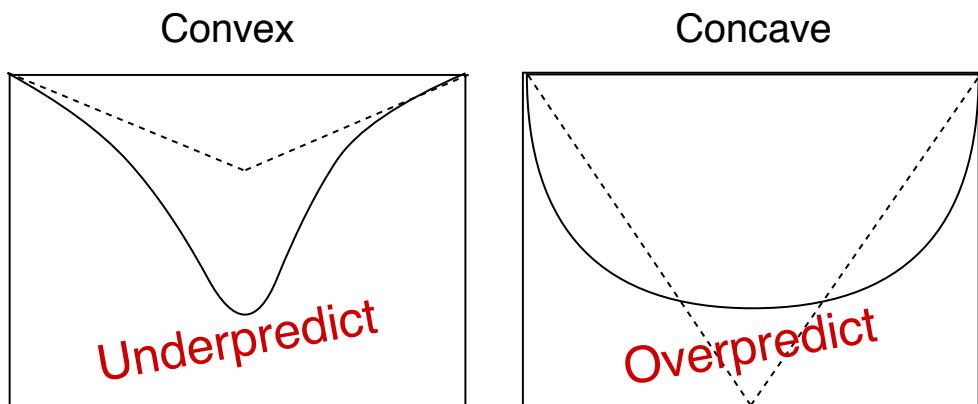


Figure S1. Diagram showing our expectation that slope-based models of lake depth will under predict true depth in convex lakes (left) and over predict true depth in concave lakes (right). Dashed lines represent extrapolated nearshore land slope while solid lines represent the lake bottom.

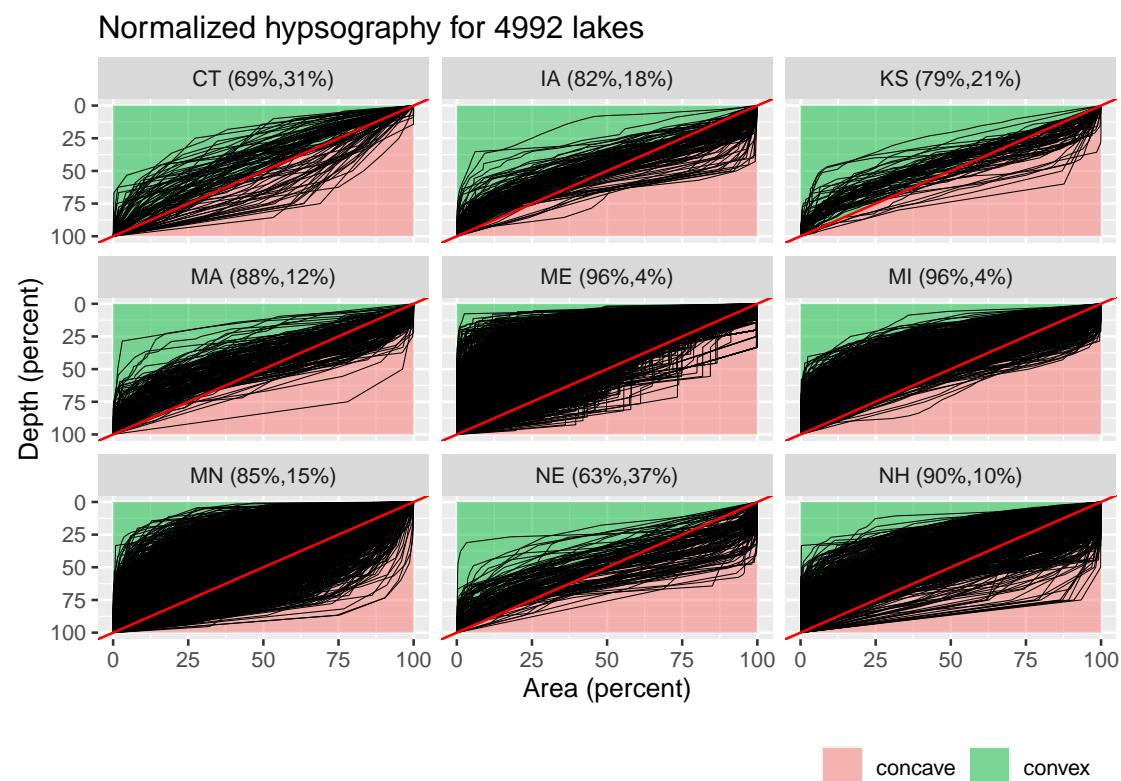


Figure S2. Hypsography classification by state. Numbers on panel labels indicate the percentage of lakes in each state with a convex versus a concave cross-section shape.

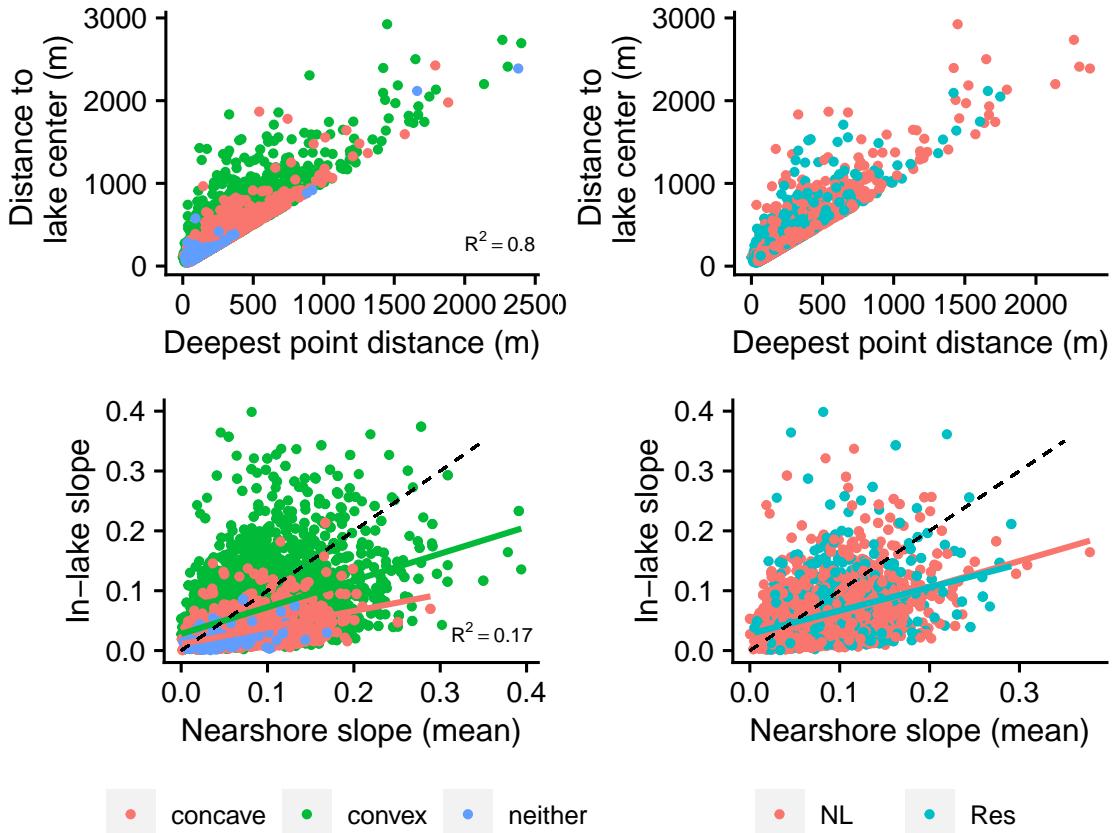


Figure S3. Comparison among lake shape and reservoir classes for A-B) distance to deepest point versus distance to lake visual center and C-D) nearshore slope versus inlake slope. A dashed 1:1 line is shown for comparison. Cross-section shape and reservoir class plots are not identical because not all lakes had a reservoir classification exceeding a 0.75 probability confidence level.

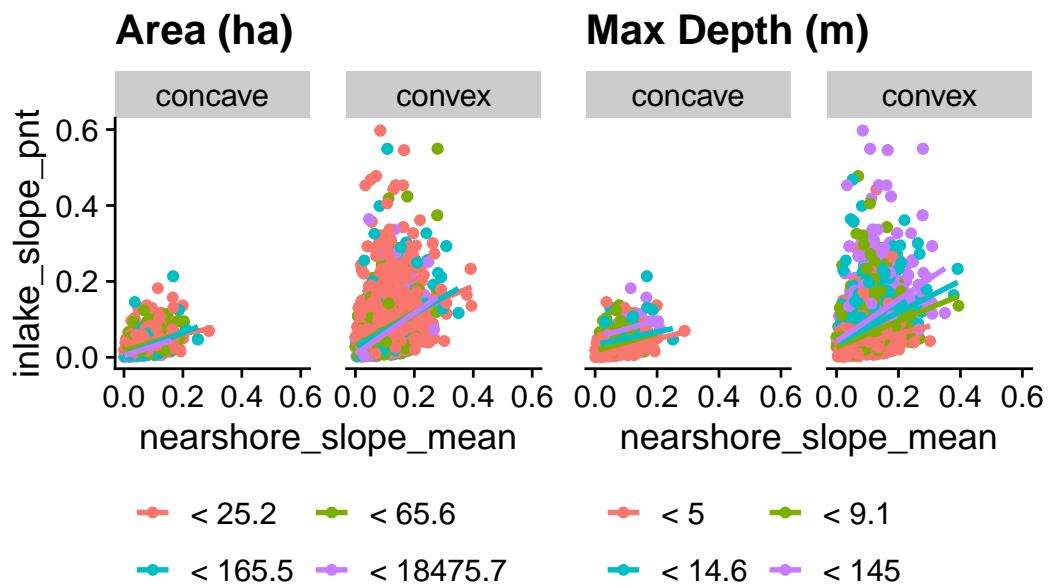


Figure S4. Comparison between in-lake and nearshore slopes in concave and convex lakes of the same size and max depth. Categories are quantile bins (< 25%, 25-50%, 50-75%, and 75-100%).

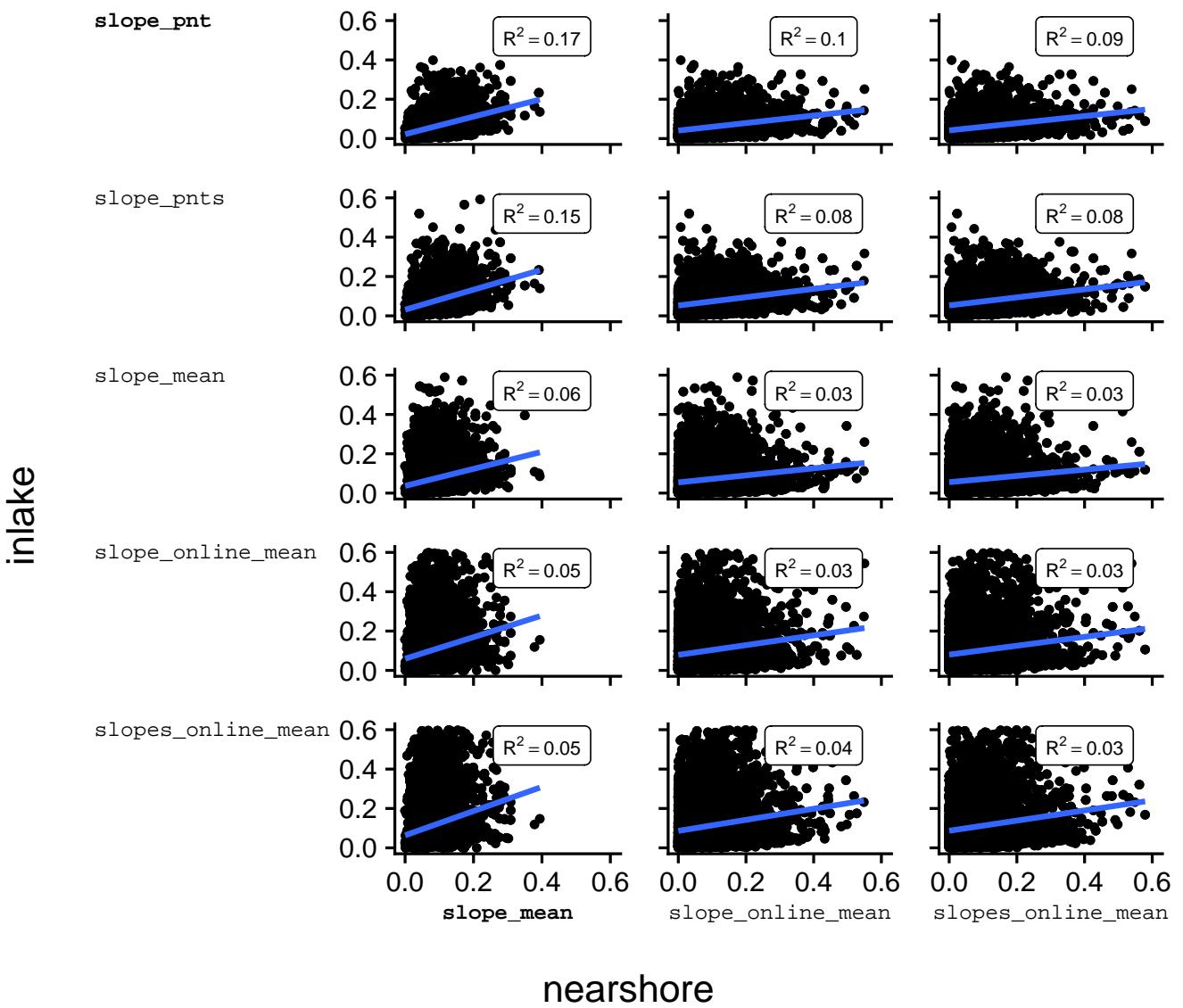


Figure S5. Comparison between in-lake and nearshore slope using different calculation techniques. The techniques used in the main text analyses are bolded and the combination of these techniques (top-left corner) produces the strongest relationship between the two metrics. **slope_mean** is the mean slope of all inlake or nearshore buffer points. **slope_pnts** is the average slope (i.e. **slope_pnt**) of all points at maximum depth. **slope_online_mean** is the mean pixel-to-pixel slope of each pixel lying on a straight line either from the single deepest point to the lake shoreline (in the case of inlake slope) or from the lake shoreline point extending to the buffer exterior (in the case of nearshore slope). **slopes_online_mean** is the same as **slope_online_mean** except it uses all inlake points at maximum depth.

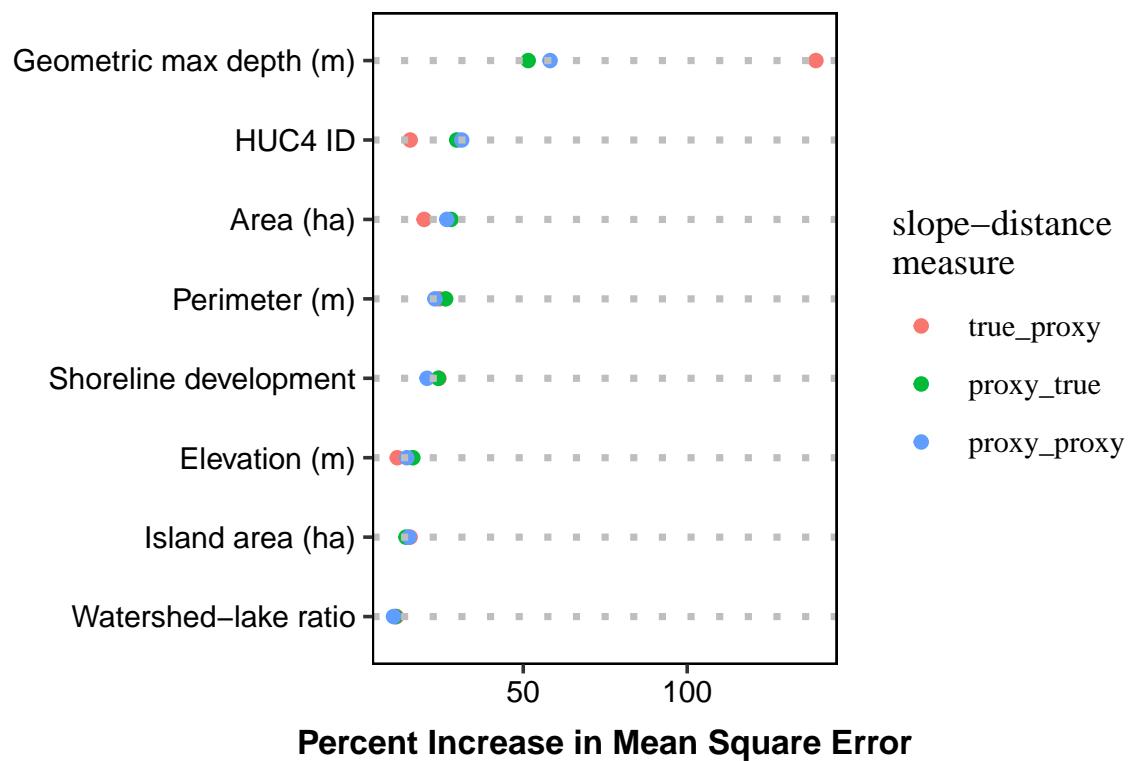


Figure S6. Importance plot for random forest variables showing increase in mean square error. Higher values indicate greater importance to model predictions. See Equation 1 for a definition of geometric max depth. HUC4 ID is a 'dummy' variable of geographic (hydrologic subbasin) location.

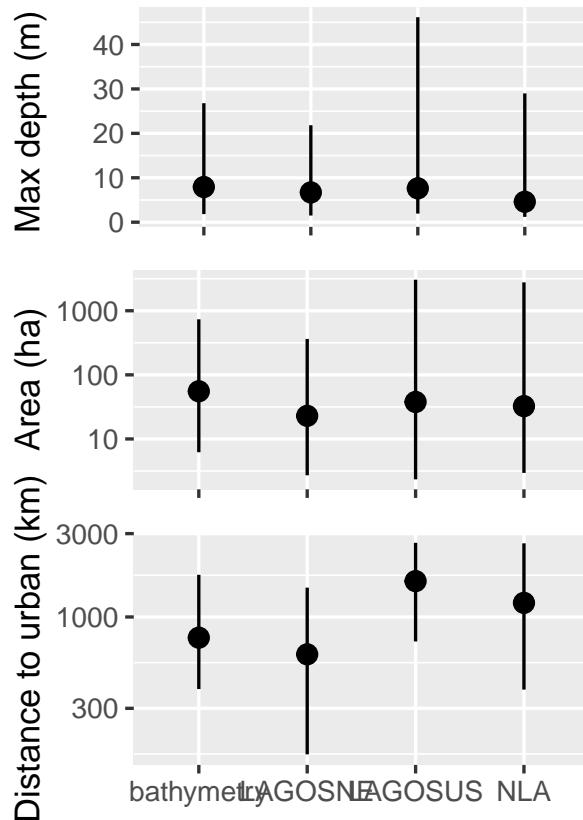


Figure S7. Comparison between characteristics of lakes with bathymetry data against lakes with depth from other sources in the LAGOSUS-Depth product. The distance to urban area metric is calculated using data from the 2018 US Census Urban and Rural Classification.

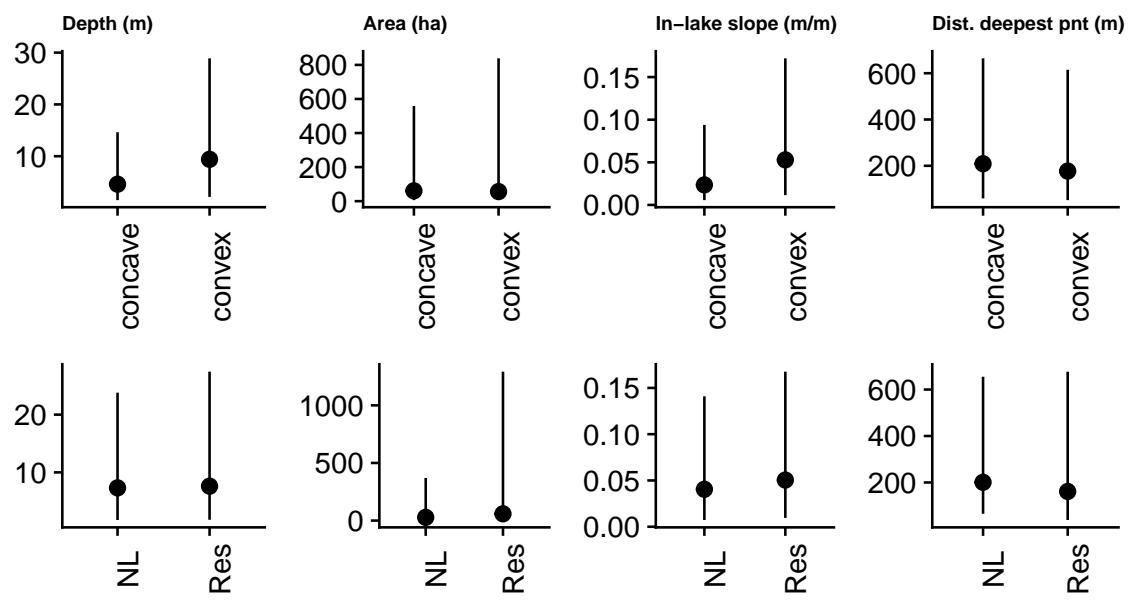


Figure S8. Comparison of lake characteristics according to differences in lake cross-section shape or reservoir status.

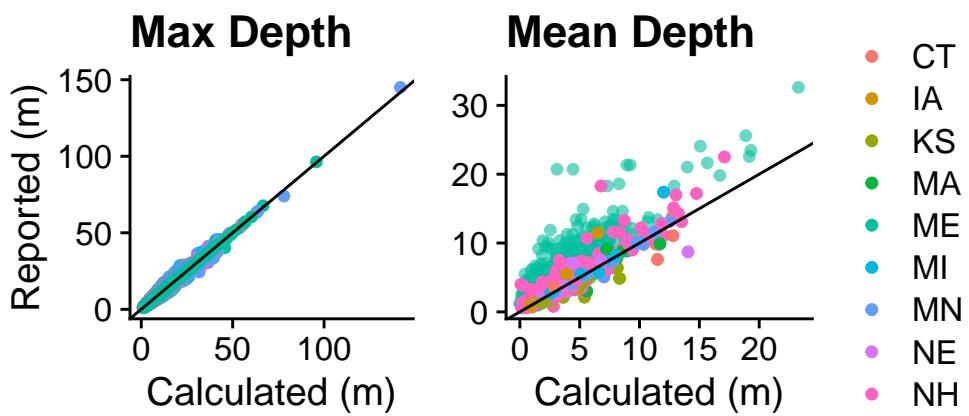


Figure S9. Comparison between reported depth and depth extracted from bathymetry surfaces by US State where reported depths come from the LAGOSUS-Depth product. For this figure, no reported depth values originated from the same source as its corresponding bathymetry-derived value.

Table S1. 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.

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 S2. Model fit and predictive accuracy metrics (RMSE = root mean square error, R^2 = coefficient of determination, MAPE = mean absolute percent error) for the proxy - proxy combination of geometry metrics (see main text Table 1). Each row shows model metrics when proxy and "true" measures are calculated with slight differences from the default (bolded) used in the main text. `slope_mean` is the mean slope of all inlake or nearshore buffer points. `slope_pnts` is the average slope (i.e. `slope_pnt`) of all points at maximum depth. `slope_online_mean` is the mean pixel-to-pixel slope of each pixel lying on a straight line either from the single deepest point to the lake shoreline (in the case of inlake slope) or from the lake shoreline point extending to the buffer exterior (in the case of nearshore slope). `slopes_online_mean` is the same as `slope_online_mean` except it uses all inlake points at maximum depth. `dists_deepest` is the same as `dist_deepest` except distance is calculated for all points at maximum depth.

Inlake slope	Nearshore slope	Inlake distance	RMSE	R^2	MAPE
<code>slope_pnts</code>	<code>slope_mean</code>	<code>dists_deepest</code>	6.2 m	0.38	58 %
<code>slope_pnt</code>	<code>slope_mean</code>	<code>dist_deepest</code>	6.4 m	0.35	59 %
<code>slope_pnts</code>	<code>slopes_online_mean</code>	<code>dist_deepest</code>	6.4 m	0.32	61 %
<code>slope_online_mean</code>	<code>slope_mean</code>	<code>dists_deepest</code>	6.5 m	0.41	63 %
<code>slope_pnts</code>	<code>slope_mean</code>	<code>dist_deepest</code>	6.7 m	0.44	58 %
<code>slope_online_mean</code>	<code>slope_mean</code>	<code>dist_deepest</code>	6.7 m	0.36	59 %
<code>slope_online_mean</code>	<code>slopes_online_mean</code>	<code>dist_deepest</code>	6.7 m	0.32	66 %
<code>slope_mean</code>	<code>slope_mean</code>	<code>dists_deepest</code>	6.8 m	0.36	59 %
<code>slope_pnt</code>	<code>slopes_online_mean</code>	<code>dists_deepest</code>	6.8 m	0.25	73 %
<code>slope_pnts</code>	<code>slope_online_mean</code>	<code>dist_deepest</code>	6.9 m	0.3	71 %
<code>slope_online_mean</code>	<code>slope_online_mean</code>	<code>dist_deepest</code>	6.9 m	0.32	68 %
<code>slope_online_mean</code>	<code>slope_online_mean</code>	<code>dists_deepest</code>	6.9 m	0.33	65 %
<code>slope_mean</code>	<code>slopes_online_mean</code>	<code>dists_deepest</code>	7 m	0.24	65 %
<code>slope_mean</code>	<code>slope_mean</code>	<code>dist_deepest</code>	7.1 m	0.4	64 %
<code>slopes_online_mean</code>	<code>slope_mean</code>	<code>dist_deepest</code>	7.1 m	0.37	56 %
<code>slope_mean</code>	<code>slope_online_mean</code>	<code>dist_deepest</code>	7.1 m	0.3	69 %
<code>slopes_online_mean</code>	<code>slopes_online_mean</code>	<code>dists_deepest</code>	7.2 m	0.32	63 %
<code>slopes_online_mean</code>	<code>slopes_online_mean</code>	<code>dist_deepest</code>	7.3 m	0.25	64 %
<code>slope_pnt</code>	<code>slope_mean</code>	<code>dists_deepest</code>	7.3 m	0.35	61 %
<code>slopes_online_mean</code>	<code>slope_mean</code>	<code>dists_deepest</code>	7.3 m	0.36	60 %
<code>slope_online_mean</code>	<code>slopes_online_mean</code>	<code>dists_deepest</code>	7.3 m	0.29	58 %
<code>slope_pnts</code>	<code>slope_online_mean</code>	<code>dists_deepest</code>	7.4 m	0.27	64 %
<code>slopes_online_mean</code>	<code>slope_online_mean</code>	<code>dists_deepest</code>	7.4 m	0.33	67 %
<code>slopes_online_mean</code>	<code>slope_online_mean</code>	<code>dist_deepest</code>	7.5 m	0.26	61 %
<code>slope_pnt</code>	<code>slopes_online_mean</code>	<code>dist_deepest</code>	7.5 m	0.33	69 %
<code>slope_mean</code>	<code>slopes_online_mean</code>	<code>dist_deepest</code>	7.6 m	0.26	64 %
<code>slope_pnt</code>	<code>slope_online_mean</code>	<code>dist_deepest</code>	7.7 m	0.27	68 %
<code>slope_pnts</code>	<code>slopes_online_mean</code>	<code>dists_deepest</code>	7.8 m	0.3	65 %
<code>slope_pnt</code>	<code>slope_online_mean</code>	<code>dists_deepest</code>	7.9 m	0.27	67 %
<code>slope_mean</code>	<code>slope_online_mean</code>	<code>dists_deepest</code>	7.9 m	0.31	60 %