Problems with the shoreline development index - a widely used metric of lake shape

Research Letter for consideration by *Geophysical Research Letters*

D. Seekell[1,*] B.B. Cael[2], and P. Byström[1]

[1] Climate Impacts Research Centre, Department of Ecology and Environmental Science, Umeå University, Sweden
[2] National Oceanography Centre, United Kingdom
[*] Corresponding author, email: david.seekell@umu.se

Abstract

The shoreline development index – the ratio of a lake's shore length to the circumference of a circle with the lake's area – is a core metric of lake morphometry used in Earth and planetary sciences. In this paper, we demonstrate that the shoreline development index is scale-dependent and cannot be used to compare lakes with different areas. We show that large lakes will have higher shoreline development index measurements than smaller lakes of the same characteristic shape, even when mapped at the same scale. Specifically, the shoreline development index increases by about 14% for each doubling of lake area. These results call into question a wide variety of previously reported patterns and relationships. We provide several suggestions to improve the application of this index, including a bias-corrected formulation for comparing lakes with different surface areas.

Plain Language Summary

Lakes vary in shape from nearly perfect circles to the almost comically convoluted. These shapes reflect their geologic origins and influence within-lake ecological and chemical processes. As a consequence, the shapes of lakes are often compared, both among lakes on Earth and between Earth's lakes and those on other planetary bodies, to provide context when measuring and interpreting other characteristics. In this paper, we show that a widely-used metric of lake shape – the shoreline development index – is biased and produces false patterns when comparing the shape of lakes with different areas, a common analysis and primary purpose of the metric. In general, we suggest not using the shoreline development index. If it must be used, we suggest: 1) reporting the scale at which lakes are mapped, 2) only comparing index values for lakes mapped at the same scale, and 3) reporting a bias-corrected index in addition to the original index.

Key Points:

- The shoreline development index is scale dependent and cannot be used to compare the shape of lakes with different surface areas
- Patterns of lake shape reported in global hydrographic studies are artefacts of scale dependence
- Bias-corrections are possible, but introduce additional uncertainties

Keyword: Shoreline Development Index, scale-dependence, lake morphometry

1. Introduction

The shoreline development index – the ratio of a lake's shore length to the circumference of a circle with the lake's area – is a core metric of lake morphometry, presented in the early chapters of both introductory (e.g., Wetzel and Likens, 2000; Wetzel, 2001) and specialist text books (e.g., Håkanson, 1981; Timms, 1992), and widely applied to describe the planar shape of lakes in hydrographic surveys (e.g., Verpoorter et al., 2014; Messager et al., 2016), as an explanatory factor in statistical analyses (e.g., Dolson et al., 2009; Casas-Ruiz et al., 2021), and as a basis for comparing lakes on planetary bodies to Earth analogs (e.g., Fasset & Head, 2008; Sharma & Byrne, 2011). In this paper, we show that the shoreline development index is scale-dependent, such that index values increase when calculated based on progressively higher resolution maps. Additionally, we demonstrate that this property translates to comparative analyses of lakes – large lakes have higher index values than small lakes, even when they share the same shape. We discuss implications for previous reports based on this index, and provide several suggestions to improve the application of this index including a bias-corrected formulation for comparing lakes with different surface areas.

2. Theory

The shoreline development index (D_L) is calculated

1) $$D_L = \frac{L}{2\sqrt{\pi A}}$$

where L is the shore length and A is the surface area, in the same units (e.g., m and m^2, or km and km^2) (Wetzel, 2000). The minimum value is $D_L = 1$, indicating a perfectly circular lake. Higher values indicate deviation from a circle, for example due to elongation or shoreline irregularity. One method for measuring lake surface area and shore length is by overlaying gridded transparent paper on a map (Goodchild, 1980). Because the length of the grid boxes sides and map scale are known, each length of the grid box edges (δ) is represented in terms of relevant measurements units (ie. meters, kilometers). The number of grid boxes occupied by the lake (N) is used to calculate area ($A=N\delta^2$) and the number occupied by the shoreline is used to calculate shore length ($L=N\delta$).

The fundamental problem with the shoreline development index is that shore length measurements are scale dependent – shore length is longer when measured on high resolution maps than when measured on low resolution maps (Håkanson, 1979; Goodchild, 1980; Kent & Wong, 1982). This scale-dependence is demonstrated by measuring shore length repeatedly with differently sized grids (or the same sized grid on differently scaled maps):

2) $$L_\delta \propto \delta^{1-d}$$

where L is the shore length in the same units as δ and d is the fractal dimension of the shoreline. Shore length measurements are scale-independent if $d = 1$, but empirical measurements always reveal $d > 1$, with a typical value of $d = 1.28$ (Goodchild, 1980; Sharma and Byrne, 2011; Seekell et al., 2021). As a consequence, the shoreline development index for an individual lake is also scale dependent such that it increases when calculated based on measurements from progressively higher-resolution maps:

3) $$D_L \propto \frac{\delta^{1-d}}{2\sqrt{\pi A}}$$

For example, the shore length of Lake Vänern, the largest lake in Sweden (A = 5,893 km^2), is L = 1,012 km with the shoreline development index D_L = 3.72 when measured on a 1:1,000,00 scale map, but L = 2,007 km and D_L = 7.38 when measured on a 1:10,000 scale map (Håkanson, 1978; Håkanson, 1981). It is clear that shoreline development index cannot be applied to compare, and should not be presented in ways that facilitate comparison, among lakes mapped at different scales.

Scale-dependence also impacts the shoreline development index when used to compare lakes with different surface areas, even if mapped at the same scale (cf. Cheng, 1995). Consider two hypothetical lakes, Lake 1 and Lake 2, with similar shape, but different surface areas. Lakes 1 and 2 are measured with grid cells sized a and b, which are different but can be subdivided into smaller boxes with the same size (δ). The estimated shore lengths and areas for the two lakes are:

4)
$$L_1 \propto \left(\frac{\delta}{a}\right)^{-d} \delta; \quad A_1 \propto \left(\frac{\delta}{a}\right)^{-2} \delta^2$$
$$L_2 \propto \left(\frac{\delta}{b}\right)^{-d} \delta; \quad A_2 \propto \left(\frac{\delta}{b}\right)^{-2} \delta^2$$

It follows that:

5)
$$\frac{L_1}{L_2} \propto \left(\frac{b}{a}\right)^{-d}; \quad \frac{A_1}{A_2} \propto \left(\frac{b}{a}\right)^{-2}$$

Therefore:

6)
$$\frac{L_1}{L_2} \propto \left(\frac{A_1}{A_2}\right)^{d/2}$$

This is equivalent to a power-law regression of shore length by surface area when examining the average pattern for many lakes at once, with $d/2$ being the power exponent and the regression constant describing the average lake shape (Seekell et al., 2021). Because $d > 1$, shore length increases with surface area more rapidly than the circumference of a circle increases with the circle's area (ie. $L_1/L_2 \propto (A_1/A_2)^{0.5}$). As a consequence, large lakes have higher shoreline development index than smaller lakes, even if they have the same characteristic shape and are measured at the same scale:

7)
$$\frac{D_{L1}}{D_{L2}} \propto \left(\frac{A_1}{A_2}\right)^{(d/2)-0.5}$$

Equation 7 is equivalent to a power-law relationship with the exponent $(d/2)$-0.5, when comparing the averages of many lakes at once. Based on the typical fractal dimension of lake shorelines (d = 1.28), this functional form indicates that the shoreline development index increases by 14% for each doubling of lake area.

3. Empirical Analysis

We tested the relationship between the shoreline development index and area for 106 Scandinavian lakes, primarily from the mountainous border region between Sweden and Norway which is populated by many glacial lakes (Table 1). Specifically, we extracted lake surface areas and perimeters from digitized 1:50,000 scale maps from the Swedish Mapping Agency Lantmäteriet and the Norwegian Water Resource and Energy Directorate (Lindmark, 2021). We calculated the fractal dimension of the shorelines based on the regression of the logarithm of shore length by the logarithm of area. We then evaluated the relationship between the logarithm of shoreline development index and logarithm of area. Specifically, we tested if the power-exponent was equal to the theoretical expectation $d/2$-0.5. Our analysis was conducted using R version 4.0.2 with the ‘boot’, ‘foreign’, and ‘CAR’ packages (Fox & Weisberg, 2019; Canty & Ripley, 2020; R Core Team,2020). We report confidence intervals based on bootstrapping (n = 9,999 replications).

Shore length scaled to the $d/2$ = 0.63 power of area (95% CI = 0.59-0.66), which is within the theoretical range and similar to reports from other regions (Figure 1A). The regression intercept (2.07, 95% CI = 1.98-2.15) is typical of glacial lakes (Seekell et al. 2021). There was a significant positive correlation between shoreline development and area (Kendall’s τ = 0.37, 95% CI = 0.25-0.48). More specifically, the shoreline development index scaled to the 0.13 power of area (95% CI = 0.09-0.16). This value matches our theoretical prediction ($d/2$-0.5 = 0.13) exactly (Figure 1B). Hence, the statistically significant relationship between the shoreline development index and area is explained completely by biases originating from the scale-dependence of shore lengths, rather than patterns of shape across the lake size spectrum.

Table 1. Morphometry of the study lakes.

Parameter	Median	Range
Area (m^2)	135,003	9,086 - 3,781,505
Shore length (m)	2,265	487 - 18,003
Shoreline Development Index	1.67	1.11 - 4.54

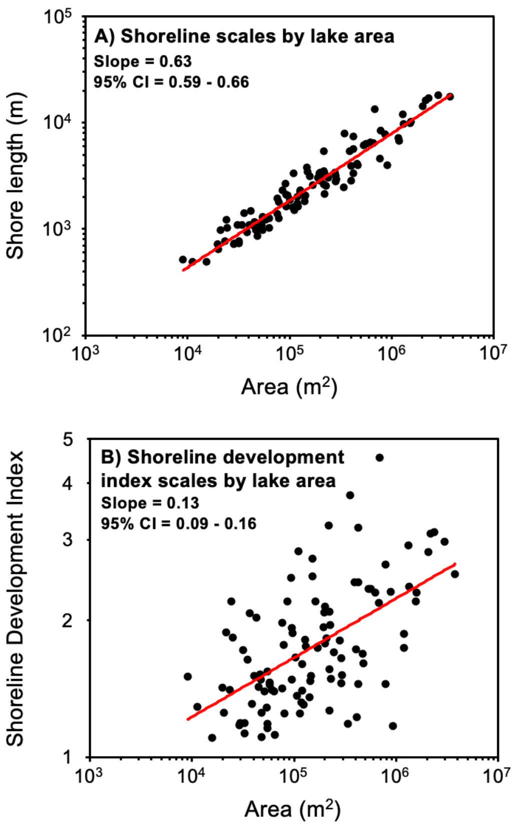

Figure 1. Scaling relationships for 106 Scandinavian lakes. A) The relationship between shore length and area B) The relationship between the shoreline development index and area.

4. Discussion

Our analysis demonstrates that the shoreline development index is a flawed metric, and casts doubts on a variety of results based on comparisons using this metric. Cautionary messages about the shoreline development index have been published several times (e.g., Håkanson, 1981; Kent & Wong, 1982; Timms, 1992), however these have been incompletely developed and were focused on variations in index values for individual lakes due to map scale. Our study provides a complete explanation of the implications of scale-dependence for the shoreline development index, including biases related to comparing lakes with different sizes, which is the most common use of the index.

An empirical regularity of large-scale hydrographic studies is that, on average, the shoreline development index is higher for larger lakes than smaller lakes (e.g., Schiefer & Klinkenberg, 2004; Verpoorter et al., 2014; Messager et al., 2016), indicating that large lakes are either more elongated or otherwise have more irregular shorelines than smaller lakes. Our empirical analysis demonstrates that this pattern reflects bias in the shoreline development index rather than a true change in shape across the lake size spectrum. This result holds when examining larger datasets. For example, $d/2 = 0.63$ for the 1.4 million lakes in the widely used HydroLakes database developed by Messager et al. (2016). The power-exponent related shoreline development index to area is 0.13, exactly the theoretically specified value. Hence, the global pattern of lake shape by area is completely attributable to bias in the shoreline development index and our result casts doubt on mechanistic interpretations of this pattern (e.g., Schiefer & Klinkenberg, 2004).

The shoreline development index is sometimes used as an explanatory factor in statistical analyses on the basis that it provides a metric of shape independent of area (e.g., Dolson et al., 2009). Our analyses have demonstrated that this reasoning is incorrect. Other studies have recognized that the shoreline development index is scale dependent, but apply it anyway based on the argument that errors are minor (e.g., Sharma and Byrne, 2011). This is also not true. The typical range of shoreline development index values is D_L = 1.5-10 (Timms, 1992). The average shoreline development index for different size classes in the HydroLakes database across is D_L = 1.6-7.8, which matches the typical range of variation for shoreline development index, and can be completely attributed to bias. Hence, the magnitude of bias is significant.

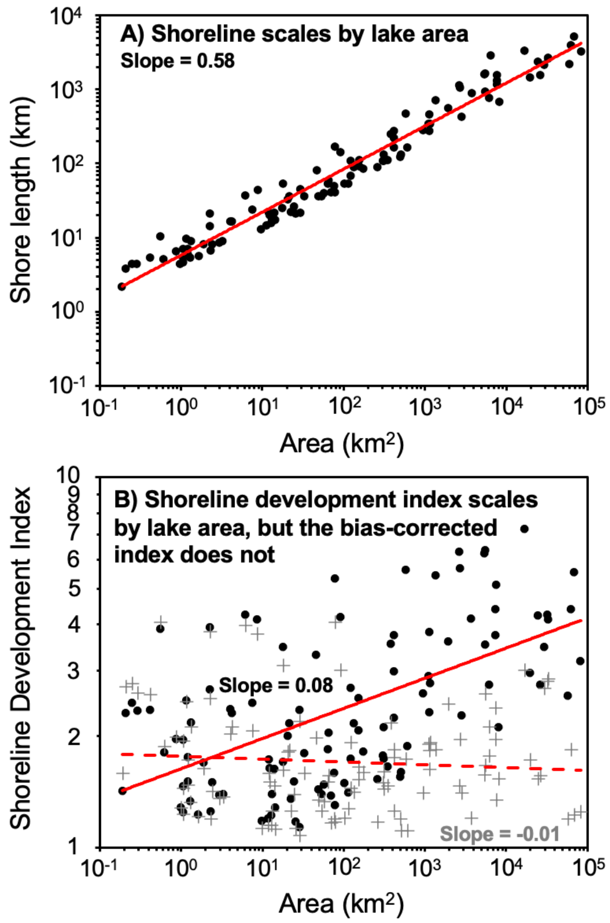

Figure 2. Scaling relationships for 111globally distributed lakes. A) The relationship between shore length and area B) The relationship between the shoreline development index and area (black circles, solid red line). This slope is matches theoretical expectations (the slope from panel A minus 0.5) exactly. The bias-corrected index is not correlated with area (grey crosses, dashed red line).

Based on our analysis, we suggest not using the shoreline development index. However, if application is strictly necessary, we suggest 1) disclosing the scale of measurement for each lake, 2) only making comparisons among lakes measured at the same scale, and applying a 3) bias-corrected shoreline development index (D_{BC}). For example,

8) $$D_{BC} = \frac{L}{2\pi^{0.5} A^{(d/2)}}$$

where $2\pi^{0.5}$ is the normalization constant for a circle (Cheng, 1995; Seekell et al., 2021), area is A, and d is the shoreline fractal dimension. For example, we calculated the bias-corrected index for 111 globally distributed lakes (A = 0.2 – 83,512 km^2), which represent a wide variety of originating processes and for which shoreline fractal dimensions have been individually measured by Sharma and Byrne (2011). For these lakes, the shoreline development index scales with area by the theoretically predicted exponent (Figure 2). However, the bias-corrected index is not correlated with area (Figure 2; slope = -0.008, 95% CI = -0.026 to 0.010). With this formulation, the shore length (L) and normalization (ie. the denominator) change at the same rate with surface area, eliminating the bias. An average value can be substituted for d (ie. $\bar{d}$ = 1.28) and applied when d is not known for individual lakes. This can be expected to accurately produce average patterns for many lakes, though $D_{BC} < 1$ is possible for sub-circular lakes with relatively smooth shorelines (ie. if $d < \bar{d}$). In particular $D_{BC} < 1$ for a given lake requires its fractal dimension $d < \bar{d}$ and that it is nearly circular (specifically $D_L < A^{\{(\bar{d}-1)/2\}}$). For instance, for the 111 lakes in Figure 2, if we instead use the lakes' median D value of 1.10 as an average $\bar{d}$, only two lakes have $D_{BC} < 1$, both of which are near-circular karst lakes with $D_L < 1.18$. There is also uncertainty introduced from the estimate of *d*, both for individual and groups of lakes. Another approach without *d* for individual lakes is to regress the logarithm of shore length by the logarithm of area and then to use the residuals as a metric of lake shape (e.g., Eloranta et al., 2016). A primary limitation of this approach is that it may be difficult to make comparisons among studies because the relationships may be variable among regions (Cael et al., 2017; Sjöberg, et al., 2022). Additionally, lakes may be selected in ways such that they are not representative and the residuals unreliable for making comparisons (e.g., Dolson et al. 2009).

When necessary, it is also possible to correct for differences in map scales, though this also introduces further uncertainty. Equation 3 specifies that $D_L \propto \delta^{1-d}$, so the effect due to the different map scales δ_1 and δ_2 can be accounted for by rescaling $D_{L2} = \left(\frac{\delta_2}{\delta_1}\right)^{1-d} D_{L1}$. One may also use an average $\bar{d}$ for this correction as well; note however that uncertainty in *d,* whether lake-specific or an average value, leads to into uncertainty in the map-scale-corrected D_L. For instance, using the average $\bar{d} = 1.10$ for the 111 lakes in Figure 2 instead of the measured $d = 1.20$ for Lake Winnipeg (see Sharma and Bryne, 2011) results in an error of 21% when upscaling or downscaling the map scale by a factor of ten.

Despite its substantial limitations, the shoreline development index retains some usefulness as an internal control on data quality. Specifically, values $D_L < 1$ are not possible and searching for these values is a simple way to screen for unreliable morphometric data that should be excluded from analyses. In our experience, these values typically arise for small lakes due to rounding errors, which are small in absolute terms, but significant for these systems. These errors can also occur if shore length and area are measured using different methods, for example if the shore length were measured with a map measurer but the area was measured with the transparent grid technique, although disparate techniques are rarely applied today due to the accessibility of digital analyses through geographic information system software. While the shoreline development index can be used to screen out erroneous data, we note that passing this screening does not confirm the quality of data.

5. Conclusion

We demonstrated that the shoreline development index is scale dependent and cannot be used to make comparisons among lakes of different size. We demonstrated that bias from this scale dependence underlies previously reported hydrographic patterns, casting doubt on a variety of results based on this metric. To enhance comparisons, merging of data sets, and evaluation of data quality, we recommend that all studies disclose the scale at which perimeter and area measurements are made. Comparisons of shoreline development should only be made for lakes measured at the same scale, or a map scale correction should be applied if this is not possible. Finally, we provide a bias-corrected index that should be used when comparing lakes of different size.

Acknowledgements

This paper is based on research supported by Umeå University, the Knut and Alice Wallenberg Foundation, and the Swedish Research Council Formas (FR-2019/0007). Elin Lindmark provided data for the Scandinavian lakes.

Open Research

We use only previously published data, which are available from the original sources. Specifically, the Scandinavian lakes data are in Seekell et al. (2021), globally distributed lakes with individually measured fractal dimensions are in Sharma and Byrne (2011), and the HydroLakes database is described by Messager et al. (2016) and available online at: https://www.hydrosheds.org/pages/hydrolakes

References

Canty, A., & Ripley, B. (2020). *Boot: Bootstrap R (S-Plus) functions*. R package version 1.3-28.

Casas-Ruiz, J.P., et al. (2021). The role of lake morphometry in modulating surface water carbon concentrations in boreal lakes. *Environmental Research Letters, 16*, 074037. https://doi.org/10.1088/1748-9326/ac0be3

Cheng, Q. (1995). The perimeter-area fractal model and its application to geology. *Mathematical Geology, 27*, 69-82. https://doi.org/10.1007/BF02083568

Dolson, R., et al. (2009). Lake morphometry predicts the degree of habitat coupling by a mobile predator. *Oikos, 118*, 1230-1238. https://doi.org/10.1111/j.1600-0706.2009.17351.x

Eloranta, A.P., et al. (2016). Community structure influences species' abundance along environmental gradients. *Journal of Animal Ecology, 85,* 273-282. https://doi.org/10.1111/1365-2656.12461

Fassett, C.I., & Head, J.W. (2008). Valley network-fed, open-basin lakes on Mars: Distribution and implications for Noachian surface and subsurface hydrology. *Icarus, 198,* 37-56. https://doi.org/10.1016/j.icarus.2008.06.016

Fox, J., & Weisberg, S. (2019). *An r companion to applied regression*. Sage.

Goodchild, M.F. (1980). Fractals and the accuracy of geographical measures. *Mathematical Geology, 12*, 85-98. https://doi.org/10.1007/BF01035241

Håkanson, L. (1978). The length of closed geomorphic lines. *Journal of the International Association of Mathematical Geology, 10*, 141-167. https://doi.org/10.1007/BF01032862

Håkanson, L. (1981). *A manual of lake morphometry*. Springer.

Kent, C., & Wong, J. (1982). An index of littoral zone complexity and its measurement. *Canadian Journal of Fisheries and Aquatic Sciences, 39*, 847-853. https://doi.org/10.1139/f82-115

Lindmark, E. (2021). *Habitat availability and ontogenetic niche shifts: The effects on adult size of lake-living brown trout (Salmo trutta)* (Master's thesis). Umeå University.

Messager, M.L., et al. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications, 7,* 13603. https://doi.org/10.1038/ncomms13603

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for statistical Computing. Retrieved from: https://www.R-project.org/

Schiefer, E., & Klinkenberg, B. (2004). The distribution and morphometry of lakes and reservoirs in British Columbia: a provincial inventory. *The Canadian Geographer, 48,* 345-355. https://doi.org/10.1111/j.0008-3658.2004.00064.x

Seekell, D., et al. (2021). The fractal scaling relationship for river inlets to lakes. *Geophysical Research Letters, 48*, e2021GL093366. https://doi.org/10.1029/2021GL093366

Sharma, P., & S. Byrne (2011). Comparison of Titan's north polar lakes with terrestrial analogs. *Geophysical Research Letters, 38,* L24203. https://doi.org/10.1029/2011GL049577

Sjöberg, Y., et al. (2022). Scaling relations reveal global and regional differences in morphometry of reservoirs and natural lakes. *Science of the Total Environment, 822,* 153510. https://doi.org/10.1016/j.scitotenv.2022.153510
Timms, B.V. (1992), *Lake geomorphology*. Gleneagles Publishing.
Verpoorter, C., et al. (2014). A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters, 41,* 6396-6402. https://doi.org/10.1002/2014GL060641
Wetzel, R.G., and G.E. Likens (2000), *Limnological Analyses, 3rd ed.* Springer
Wetzel, R.G. (2001), *Limnology, 3rd ed.* Academic Press.