Aquatic Conservation: Marine and Freshwater Ecosystems

WILEY

SHORT COMMUNICATION OPEN ACCESS

The Methods for Estimating Lake Volume, Mean Depth, and Maximum Depth in European Standard EN 16039:2011 Are Flawed and Should Not Be Used

David Seekell[1] | Cristian Gudasz[2]

[1]Atle Investment Services, Stockholm, Sweden | [2]Climate Impacts Research Centre, Department of Ecology and Environmental Science, Umeå University, Umeå, Sweden

Correspondence: Cristian Gudasz (cristian.gudasz@umu.se)

Received: 12 August 2024 | **Revised:** 19 September 2024 | **Accepted:** 20 September 2024

Funding: The author received no specific funding for this work.

Keywords: EN 16039:2011 | hydromorphology | maximum depth | mean depth | volume | Water Framework Directive

ABSTRACT

The European Standard EN 16039:2011 provides guidelines for assessing lake hydromorphology within the Water Framework Directive and includes methods for estimating volume, mean depth, and maximum depth based on statistical models applied to data from topographic maps. We tested the predictive accuracy of these models using independently collected bathymetric data from 35 Swedish lakes. The models had no predictive power, and the maximum depth predictions were inversely correlated with the observed values. The mean absolute percent error was 46% for volume and mean depth, and 85% for maximum depth. The models are flawed and should not be used.

1 | Introduction

Bathymetric surveys are time consuming and prohibitively expensive for large numbers of lakes—a major challenge for evaluating lake ecological status because volume and maximum depth are key constraints on many ecosystem characteristics (Boon et al. 2019; Cael, Heathcote, and Seekell 2017; Cael and Seekell 2022). This has motivated a series of empirical studies seeking to predict lake-specific volume and depth based on surface area and other easily mapped characteristics, typically some metric of vertical relief within the catchment (Cael and Seekell 2022). Empirical relationships of this nature are often criticized for being conceptually weak and not being useful when applied to specific cases because the use of logarithmic scales results in wide prediction intervals (Pace, Findlay, and Lints 1991; Cael and Seekell 2022). Despite this, these empirical relationships are rarely subjected to independent validation (Pace, Findlay, and Lints 1991; Muñoz et al. 2020; Archer et al. 2021).

The European Committee for Standardization (CEN) Standard EN 16039:2011 contains guidance on assessing the hydromorphological features of lakes. The purpose of this standard is to facilitate comparisons between lakes during habitat assessments, especially those made within and among European countries to comply with the European Water Framework Directive, a major international environmental policy that commits all European countries to achieve good water quality in lakes, rivers, and coastal waters (Boon et al. 2019). The CEN standard is the official technical standard for the Water Framework Directive (Boon et al. 2019).

The CEN standard includes methods for estimating volume, mean, and maximum depth based on the sequential application of three statistical models to data derived from topographic maps. However, the methods are suspect—the researcher that originally published the methods also reported that they had little predictive power when applied to the calibration dataset (Håkanson 1999). Unfortunately, this was not clearly disclosed

---

https://doi.org/10.1002/aqc.4260

in the same researcher's book (Håkanson 2004), which is cited by the CEN standard as the source of the method. Additionally, despite being adopted as a European standard, the models have apparently never been evaluated with independent data, which has long been recognized as the ideal way to validate predictive models (Snee 1977; Pace, Findlay, and Lints 1991).

In this paper, we test the method for estimating volume, mean depth, and maximum depth in CEN Standard EN 16039:2011 using independently collected bathymetric and catchment data for 35 Swedish lakes. We use several metrics of prediction error to compare our direct measurements to predictions from the statistical approach. Finally, we discuss the practical implications of our findings in the context of the Water Framework Directive.

2 | Methods

2.1 | Overview of the CEN Method

The approach in CEN standard comes from Håkanson (1999, 2004). First, a metric of catchment relief C_R (dimensionless) is calculated based on catchment area (C_A, km^2) and the catchment's elevation range (C_z, m):

$$C_R = C_z / \sqrt{C_A}$$

Second, lake-specific volume (L_V, km^3) is estimated based on lake area (L_A, km^2) and the catchment relief metric C_R:

$$\log\left(1000 \times L_V\right) = 0.134 + 1.224 \times \log\left(L_A\right) + 0.332 \times \log\left(C_R\right)$$

Finally, lake-specific maximum depth (z_{max}, m) and mean depth (z_{mean}, m) are estimated based on the estimated lake volume and the observed lake area:

$$\log(z_{max}) = -4.202 + 4.558 \times \left(1000 \times L_V\right)^{0.1} - 1.008 \times \log(L_A)$$

$$z_{mean} = 1000\left(L_V / L_A\right)$$

In practice, mean and maximum depth will always be estimated using the predicted volume. If bathymetric surveys for volume have been completed, mean and maximum depth would be calculated directly rather than with models.

2.2 | Data

Data suitable for validating the CEN approach should meet three criteria:

1. the lakes should cover the typical ranges of morphometric measurements for lakes generally,
2. the bathymetry should be measured using a single widely used method, and
3. the sample size should be sufficient to precisely estimate metrics of prediction error.

Based on these criteria, we selected a bathymetric dataset of 35 lakes in the mountainous region of northern Sweden. The lakes were part of a larger comparative study that used a common set of methods and are distributed between clusters in the southern, middle, and northern portions of the region to cover its major geophysical gradients (see map in Klaus, Karlsson, and Seekell 2021). Sweden's mountains are an example of a region where applying the CEN methods would be especially valuable due to the added difficulty and expense of reaching them for depth sounding.

The lakes cover the typical range of most morphometrics, with the only notable deviation being to smaller surface areas, which are the most abundant and more likely to be unmapped compared with larger surface area lakes (Table 1; Seekell 2018). Additionally, the sample is 52% larger than the minimum ($n = 23$) necessary to validate the CEN method (Text S1). Hence, the dataset meets the criteria needed to validate the CEN method.

The methods and data for locations, surface areas, lake surface elevation, and mean and maximum depths were previously published by Klaus, Karlsson, and Seekell (2021), Norman et al. (2022), Verheijen et al. (2022), and Karlsson et al. (2024). Briefly, the values derived from bathymetric data collected using an echo sounder with an internal GPS antenna (Lowrance HDS-5 Gen2) were interpolated using ordinary kriging (Klaus, Karlsson, and Seekell 2021). Volume and mean depth were calculated based on these maps, and it is assumed that the maximum recorded depth is the maximum depth.

TABLE 1 | The morphological characteristics of the study lakes compared with typical values.

Characteristics	Median in this study	Range in this study	Typical value	Source for typical value
Latitude	66.1°	63.3–68.5°	45–75°	Verpoorter et al. (2014)
Maximum depth	9.2 m	2.3–31.2 m	<20 m	Seekell and Cael (2023)
Mean depth	2.8 m	0.8–10.9 m	<10 m	Cael, Heathcote, and Seekell (2017)
Lake area	0.07 km^2	0.02–0.4 km^2	0.01–10 km^2	Seekell, Cael, and Byström (2022a)
Catchment area	1.17 km^2	0.08–12.4 km^2	0.01–10 km^2	Seekell, Cael, and Byström (2022a)
Drainage area relief ratio	105	40–284	0–200	Håkanson (1999)

We added to these data by calculating catchment areas for the lakes using Sweden's national 1-m digital elevation model and typical hydrological methods based on flow direction and accumulation. The highest point in the catchment, needed to calculate catchment relief C_z, was identified using the digital elevation model. Catchment relief C_z is the maximum elevation in the catchment minus the lake elevation.

2.3 | Empirical Analysis

We predicted volume, mean depth, and maximum depth for each lake using the CEN method and then evaluated the predictions graphically and with deviance measures. Specifically, we first evaluated plots of predicted versus observed values, long understood as one of the simplest but most powerful methods for model evaluation (Mayer and Butler 1993). Next, we calculated mean absolute error and mean absolute percent error as measures of deviance (Loague and Green 1991; Mayer and Butler 1993). Finally, we calculated the Nash–Sutcliffe modeling efficiency index (NSE), which describes the fit of predicted and observed values around a 1:1 line (Loague and Green 1991). For a hypothetical model with perfect predictions, mean absolute error and mean absolute percent error would be 0, whereas modeling efficiency would be $\mathrm{NSE}=1$. In hydrological analyses, $\mathrm{NSE}<0.5$ indicates an unsatisfactory model, and $\mathrm{NSE}<0$ indicates a model that produces predictions that are less accurate than simply using the mean (Moriasi et al. 2007).

We used a variety of Python packages for geospatial data processing and hydrological analysis including “geopandas” (Jordahl et al. 2020), “rasterio” (Gillies 2019), “shapely” (Gillies et al. 2024), NumPy (Harris et al. 2020), and gdal, ogr, and osr (GDAL/OGR contributors 2024). We also used WhiteboxTools for hydrological processing (Lindsay 2016). Our statistical analyses were conducted with R using the “metrica” package (R Core Team 2023; Correndo et al. 2022). The data used in our analysis are archived on Zenodo (doi: 10.5281/zenodo.12655243) and an R script for our analysis included in the Supporting Information.

3 | Results

The CEN methods performed poorly. The model efficiency was unsatisfactory ($\mathrm{NSE}<0.5$) for all three metrics, and the models for mean and maximum depth performed worse than simply taking the mean ($\mathrm{NSE}<0$) (Table 2). In all cases, the mean absolute error is a substantial portion of the typical range of these measurements, and the mean absolute percent error was 46% for mean depth and volume and 85% for maximum depth (Table 2).

The predicted and observed volumes were correlated and generally clustered around the 1:1 line, although with substantial errors in both absolute and relative terms that contribute to the weak predictive performance (Figure 1A). Predicted mean depths appeared unrelated to observed mean depths, although, by definition, the relative errors are identical to those for volume (Figure 1B; Table 2). This weak correlation is consistent with the Nash–Sutcliffe efficiency being approximately zero (Table 2). The maximum depth predictions were the worst of the three metrics and were inversely correlated with the observed depths (Figure 1C). The CEN method greatly underestimated the maximum depth of the deepest lakes and greatly overestimated the depth of the shallowest lakes.

Prediction errors for volume and mean depth were closely related ($r=-0.87$) to the dynamic ratio—the ratio of the square root of surface area to mean depth. The volumes of lakes with high dynamics ratios, those that are relatively large but shallow, are overestimated by up to 279% (Table 2). In contrast, the CEN method underestimates the volumes of small but relatively deep lakes by as much as 71%. Dynamic ratio increases, on average, with lake surface area; hence, we generally expect the CEN method to overestimate the volume of larger lakes and underestimate the volumes of smaller lakes (Cael and Seekell 2023). The largest prediction errors for volume were typically responsible for the largest prediction error for maximum depth. However, there was not a strong relationship between prediction error and any landscape or lake metric beyond this.

4 | Discussion and Recommendations

Predictive models are best evaluated based on how large an error can be expected compared with how much uncertainty can be accepted when making decisions. The CEN method clearly fails this evaluation when confronted with independent data—errors so large that the predictions are worse than the mean are unacceptable for decision-making. The CEN method is flawed and should not be used as basis for evaluating or comparing habitat quality in the context of the Water Framework Directive.

Our results have the most direct implications for the Water Framework Directive's “System A” hydromorphologic typology. This divides lakes into three classes based on mean depth. Only 57% of our study lakes are correctly classified when using the

TABLE 2 | Deviance metrics from the validation analysis. NSE is the Nash–Sutcliffe model efficiency index.

		Absolute error				Absolute percent error			
Measurement	**NSE**	**Mean**	**Median**	**Minimum**	**Maximum**	**Mean**	**Median**	**Minimum**	**Maximum**
Volume	0.49	187,166 m^3	69,914 m^3	10,876 m^3	986,518 m^3	46%	41%	6%	279%
Mean depth	−0.002	1.6 m	0.84 m	0.17 m	7.7 m	46%	41%	6%	279%
Maximum depth	−1.12	7.2 m	5 m	0.1 m	23 m	85%	42%	1.8%	391%

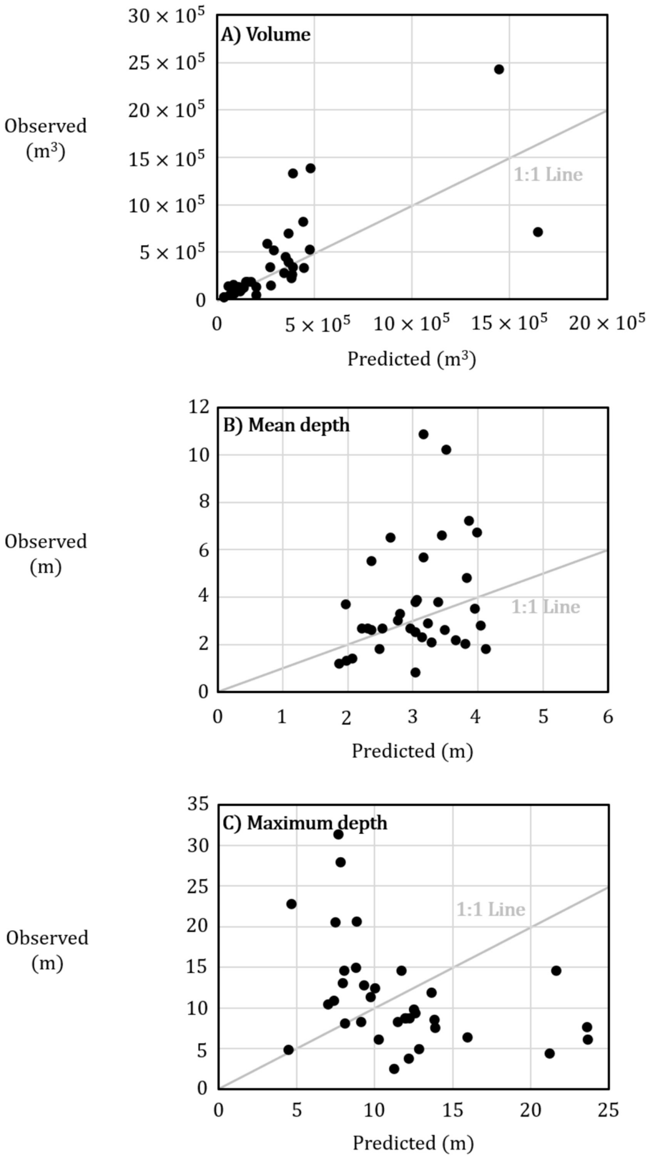

FIGURE 1 | Predicted versus observed values for (A) volume, (B) mean depth, and (C) maximum depth.

CEN method. The status of biological communities is evaluated relative to the expected community composition in an unimpacted natural state, which depends on depth class. Therefore, incorrect baselines for assessing ecological status may be adopted when using the CEN method. Of course, errors also impact other aspects of ecological status for which volume or depth is a key constraint such as trophic status and carbon cycling. Because these processes relate to morphology nonlinearly, we expect prediction errors for ecological processes to accumulate rather than cancel out when conducting assessments and comparisons at large scales.

Predictive models should be calibrated based on samples representative of the target populations for model application. The model used in the CEN standard were calibrated to a sample 95 Swedish lakes even though the intended use is for lakes across Europe. Hence, this was never an appropriate selection for a European Standard. However, our observation that these models do not even work for other Swedish lakes shows that the flaws are more fundamental than lack representative calibration data alone. The regression parameters in the CEN method represent population level characteristics. Inferring individual lake characteristics from these relationships is a fundamentally improper use of these relationships, akin to the ecological fallacy, and has never provided useful levels of predictive power (Muñoz et al. 2020; Cael and Seekell 2022).

The CEN method's failure to develop predictive power is consistent with theoretical and empirical findings that volume and depth are expected vary by an order of magnitude for lakes with the same surface area and landscape (Cael, Heathcote, and Seekell 2017; Cael and Seekell 2022). These studies show that the statistical distributions of lake morphometrics are predictable based on map-derived features but provide no reason to believe that these same features should also be able to accurately predict the volumes and depths of individual lakes, which are essentially random values from these distributions (Cael, Heathcote, and Seekell 2017; Cael and Seekell 2022).

We recommend that the EN 16039:2011 standard be reconsidered or withdrawn. Other methods in the standard are also faulty. For example, the shoreline development index, included in the CEN standard as a metric of planar shape, is biased in a way that creates false patterns including when relating lake morphology to ecological process—the very purpose of collecting morphometrics in the Water Framework Directive (Seekell, Cael, and Byström 2022b; Seekell and Cael 2024). Therefore, we also recommend that CEN adopt procedures to prevent the inclusion of unvalidated methods in its standards.

Morphometry is so fundamental to the evaluation of lake ecosystems that it seems trivial, but our analysis demonstrates there are no shortcuts to collecting bathymetric data. Bathymetric surveys are time-consuming, expensive, and necessary to make credible assessments and comparisons in the context of the Water Framework Directive or other large-scale assessments.

Ethics Statement

No approvals were required for this research.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data are archived on Zenodo (doi: 10.5281/zenodo.12655243).

References

Archer, L., K. I. E. Snell, J. Ensor, M. T. Hudda, G. S. Collins, and R. D. Riley. 2021. "Minimum Sample Size for External Validation of a Clinical Prediction Model With a Continuous Outcome." *Statistics in Medicine* 40: 133–146. https://doi.org/10.1002/sim.8766.

Boon, P., C. Argillier, A. Boggero, et al. 2019. "Developing a Standard Approach for Assessing the Hydromorphology of Lakes in Europe." *Aquatic Conservation: Marine and Freshwater Ecosystems* 29: 655–669. https://doi.org/10.1002/aqc.3015.

Cael, B. B., A. J. Heathcote, and D. A. Seekell. 2017. “The Volume and Mean Depth of Earth's Lakes.” *Geophysical Research Letters* 44: 209–218. https://doi.org/10.1002/2017GL071378.

Cael, B. B., and D. Seekell. 2022. “A Theory for the Relationship Between Lake Surface Area and Maximum Depth.” *Limnology and Oceanography Letters* 7: 527–533. https://doi.org/10.1002/lol2.10269.

Cael, B. B., and D. A. Seekell. 2023. “Simple Model of Morphometric Constrain on Carbon Burial in Boreal Lakes.” *Frontiers in Environmental Science* 11: 1101332. https://doi.org/10.3389/fenvs.2023.1001332.

Correndo, A. A., L. H. M. Rosso, C. H. Hernandez, et al. 2022. “Metrica: An R Package to Evaluate Prediction Performance of Regression and Classification Point-Forecast Models.” *Journal of Open Source Software* 7: 4655. https://doi.org/10.21105/joss.04655.

GDAL/OGR Contributors. 2024. GDAL/OGR Geospatial Data Abstraction Software Library. Open Source Geospatial Foundation. https://doi.org/10.5281/zenodo.5884351.

Gillies, S., van der Wel, C., Van den Bossche, J. et al. 2024. “Shapely (Version 2.0.4) [Software]”. https://doi.org/10.5281/zenodo.5597138.

Gillies, S. 2019. “Rasterio: Geospatial Raster I/O for Python Programmers [Software].” Mapbox. https://github.com/rasterio/rasterio.

Håkanson, L. 1999. “Error Propagations in Step-By-Step Predictions: Examples for Environmental Management Using Regression Models for Lake Ecosystems.” *Environmental Modelling & Software* 14: 49–58.

Håkanson, L. 2004. *Lakes—Form and Function*. Caldwell, New Jersey, USA: Blackburn Press.

Harris, C. R., K. J. Millman, S. J. van der Walt, et al. 2020. “Array Programming With NumPy.” *Nature* 585: 357362. https://doi.org/10.1038/s41586-020-2649-2.

Jordahl, K., J. Van den Bossche, M. Fleischmann, et al. 2020. geopandas/geopandas: v0.8.1 [Software]. Zenodo. https://doi.org/10.5281/zenodo.3946761.

Karlsson, J., H. A. Verheijen, D. A. Seekell, D. Vachon, and M. Klaus. 2024. “Ice-Melt Period Dominates Annual Carbon Dioxide Evasion From Clear-Water Arctic Lakes.” *Limnology and Oceanography Letters* 9: 112–118. https://doi.org/10.1002/lol2.10369.

Klaus, M., J. Karlsson, and D. Seekell. 2021. “Tree Line Advance Reduces Mixing and Oxygen Concentrations in Arctic-Alpine Lakes Through Wind Sheltering and Organic Carbon Supply.” *Global Change Biology* 27: 4238–4253. https://doi.org/10.1111/gcb.15660.

Lindsay, J. B. 2016. “Whitebox GAT: A Case Study in Geomorphometric Analysis.” *Computers & Geosciences* 95: 75–84. https://doi.org/10.1016/j.cageo.2016.07.003.

Loague, K., and R. E. Green. 1991. “Statistical and Graphical Methods for Evaluating Solute Transport Models: Overview and Application.” *Journal of Contaminant Hydrology* 7: 51–73.

Mayer, D. G., and D. G. Butler. 1993. “Statistical Validation.” *Ecological Modelling* 68: 21–32.

Moriasi, D. N., J. G. Arnold, M. W. Van Liew, et al. 2007. “Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations.” *Transactions of the ASABE* 50: 885–900.

Muñoz, R., C. Huggel, H. Frey, A. Cochachin, and W. Haeberli. 2020. “Glacial Lake Depth and Volume Estimation Based on a Large Bathymetric Dataset From the Cordillera Blanca, Peru.” *Earth Surface Processes and Landforms* 45: 1510–1527. https://doi.org/10.1002/esp.4826.

Norman, S., K. A. Nilsson, M. Klaus, D. Seekell, J. Karlsson, and P. Byström. 2022. “Effects of Habitat-Specific Primary Production on Fish Size, Biomass and Production in Northern Oligotrophic Lakes.” *Ecosystems* 25: 1555–1570. https://doi.org/10.1007/s10021-01-00733-6.

Pace, M. L., S. E. G. Findlay, and D. Lints. 1991. “Variance in Zooplankton Samples: Evaluation of a Predictive Model.” *Canadian Journal of Fisheries and Aquatic Sciences* 42: 146–151.

R Core Team. 2023. *R: A Language and Environment for Statistical Computer*. Vienna, Austria: R Foundation for Statistical Computing. https://www.r-project.org.

Seekell, D. 2018. “Fractal Characteristics of Lakes.” In *Thule: Kungl. Skytteanska Samfundets Årsbok 2018*, edited by R. Jacobsson, 109–120. Umeå, Sweden: Kungliga Skytteanska Samfundet.

Seekell, D., B. Cael, and P. Byström. 2022a. “The Scaling Relationship for the Length of Tributaries to Lakes.” *Geophysical Research Letters* 49: e2022GL098183. https://doi.org/10.1029/2022GL098183.

Seekell, D., and B. B. Cael. 2023. “Why Does the Relationship Between Benthic Primary Production and Lake Morphometry Vary Regionally?.” *Aquatic Sciences* 85: art87. https://doi.org/10.1007/s00027-023-00986-4.

Seekell, D., and B. B. Cael. 2024. “Bias in the Shoreline Development Index: Ecological Implications Illustrated With an Analysis of Littoral-Pelagic Habitat Coupling.” *Limnologica* 106: art126173. https://doi.org/10.1016/j.limno.2024.126173.

Seekell, D., B. B. Cael, and P. Byström. 2022b. “Problems With the Shoreline Development Index—A Widely Used Metric of Lake Shape.” *Geophysical Research Letters* 49: e2022GL098499. https://doi.org/10.1029/2022GL098499.

Snee, R. D. 1977. “Validation of Regression Models: Methods and Examples.” *Technometrics* 19: 415–428.

Verheijen, H. A., M. Klaus, D. A. Seekell, and J. Karlsson. 2022. “Magnitude and Origin of CO_2 Evasion From High-Latitude Lakes.” *Journal of Geophysical Research – Biogeosciences* 127: e2021JG006768. https://doi.org/10.1029/2021JG006768.

Verpoorter, C., T. Kutser, D. A. Seekell, and L. J. Tranvik. 2014. “A Global Inventory of Lakes Based on High-Resolution Satellite Imagery.” *Geophysical Research Letters* 41: 6396–6402. https://doi.org/10.1002/2014GL060641.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.