Water Research 178 (2020) 115801

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Making Waves

Rethinking discretization to advance limnology amid the ongoing information explosion

B.M. Kraemer [1]

IGB Leibniz Institute for Freshwater Ecology and Inland Fisheries, Berlin, Germany

A R T I C L E I N F O

Article history:
Received 24 October 2019
Received in revised form
31 March 2020
Accepted 4 April 2020
Available online 8 April 2020

Keywords:
Classification
Management
Big data
Computing
Statistics
Trophic state
Zonation
Discretization evaluation framework

A B S T R A C T

Limnologists often adhere to a discretized view of waterbodies—they classify them, divide them into zones, promote discrete management targets, and use research tools, experimental designs, and statistical analyses focused on discretization. By offering useful shortcuts, this approach to limnology has profoundly benefited the way we understand, manage, and communicate about waterbodies. But the research questions and the research tools in limnology are changing rapidly in the era of big data, with consequences for the relevance of our current discretization schemes. Here, I examine how and why we discretize and argue that selectively rethinking the extent to which we must discretize gives us an exceptional chance to advance limnology in new ways. To help us decide when to discretize, I offer a framework (discretization evaluation framework) that can be used to compare the usefulness of various discretization approaches to an alternative which relies less on discretization. This framework, together with a keen awareness of discretization's advantages and disadvantages, may help limnologists benefit from the ongoing information explosion.

Contents

1. Introduction 2
 1.1. Why rethink discretization? 2
 1.2. Why do we discretize? 2
 1.3. Discretization amid an information explosion 2
2. Advantages and disadvantages of discretization 3
 2.1. Waterbody discretization (the division of waterbodies into types and zones) 3
 2.2. Discretization in water management 3
 2.3. Discretization in sampling and data analysis 4
3. Discretization evaluation framework 4
4. The bottom line 6
5. The future of limnological discretization 6
6. Conclusions 6
 Author contribution statement 6
 Data availability statement 6
 Declaration of competing interest 6
 Acknowledgements 7
 References 7

E-mail addresses: ben.m.kraemer@gmail.com, ben.m.kraemer@gmail.com.
[1] Present Address: 310 Müggelseedam, Berlin, Germany, 12589.

https://doi.org/10.1016/j.watres.2020.115801

1. Introduction

1.1. Why rethink discretization?

One of the most broadly-shared assumptions about nature is that it can be neatly categorized into objectively-defined, discrete groups. This assumption has played a central role in human history from the philosophies of Plato and Aristotle all the way to modern science, governance, computing, and communication. But, the erosion of this broadly-shared assumption about nature has been the subtext to many of the major scientific advances of the last several hundred years. In biology, Charles Darwin and Alfred Russel Wallace rejected the assumption that all organisms belong to discrete, unrelated species. In his famous “theory of relativity,” Albert Einstein powerfully demonstrated that time and space are not discrete, further arguing that “physical reality must be described in terms of continuous functions” (Einstein, 1979). In social psychology, Judith Butler radically disrupted the binary view of sex, gender, and sexuality (Butler, 1990). Over the course of the 20th century, and especially since the rapid growth of human genome sequencing, we’ve learned that standard racial categories do not reflect actual genetic structure in humans (Yudell et al., 2016). These key advances demonstrate that rethinking this fundamental assumption, that nature is discrete, can promote inspiring scientific progress.

Rethinking the assumption that nature can be neatly categorized into discrete groups has also led to major scientific advances, specifically in limnology—the study of inland waters. Limnology was founded on the premise that waterbodies, lakes in particular, are distinct from the terrestrial ecosystems around them. Lake ecosystems were considered superb habitats for ecological experiments because their biological, chemical, and physical processes were considered isolated. In 1887, Stephen Forbes, a founder of limnology, went so far as to doubt whether annihilating all terrestrial animals would have any important effect on lake ecosystems at all (Forbes, 1925). But, we know today that lake ecosystems are intricately connected to surrounding ecosystems and vice versa. The boundary where lake ecosystems end and the next ecosystem begins has become increasingly blurred in recent decades. During the research boom on “ecosystem subsidies” beginning in the 2000’s, limnologists showed that a large proportion of lake ecosystem carbon can be derived from terrestrial production (Pace et al., 2004). We recognize today that lakes emit greenhouse gases that have widespread effects beyond their boundaries (Raymond et al., 2013). The emergent recognition that lake ecosystems are not discrete or isolated has been a substantial limnological advance (Tranvik et al., 2018).

Limnologists have begun to critically reflect on many other aspects of discretization, leading to a variety of limnological advances. For example, in recognition of the subjectivity of defining discrete mixed layers, some argue that an idealised concept of a well-defined mixed layer does not necessarily reflect the reality of aquatic physics (Gray et al., 2019). In recognition of the substantial amount of subsurface flow across catchments, “discrete” hydrogeological units are increasingly recognized as having no real boundaries to water movement (Fan, 2019). By moving beyond research focused on single, discrete waterbody types, we have also begun to find the remarkable commonalities among seemingly disparate aquatic ecosystems leading to more general theories for how waterbodies function. For instance, waterbodies across size and flow gradients have been shown to have similar controls on their nutrient limitation (Elser et al., 2007), metabolism (Yvon-Durocher et al., 2012), trophic cascades (Shurin et al., 2002), and responses to human activity (King et al., 2019).

1.2. Why do we discretize?

Yet the ubiquity of limnological discretization today is testament to its enduring, broadscale appeal and usefulness. We classify waterbodies based on their size (pond, lake, Great Lake, or Ocean), flow rates (creek, stream, river, or large river), trophic status (oligotrophic, mesotrophic, or eutrophic), mixing (stratified or unstratified), salinity (freshwater or saline), latitude (tropical, temperate, or arctic), and human hydrological influence (lake or reservoir). We also divide waterbodies into different zones according to their mixing (epilimnion, metalimnion, or hypolimnion), light climate (photic or aphotic), distance from shore (littoral or limnetic), and distance from the bottom (benthic or pelagic). The publications which codify these waterbody discretizations are often very well-cited (e.g. Lewis Jr. 1983).

In addition to waterbody discretization, limnologists rely on discretization in many contexts because it offers extraordinarily useful shortcuts that can facilitate limnological progress, management, and communication. In a hypothetical research context, a limnologist could collect total phosphorus samples across depth every millimetre (or even finer) to detect small-scale variation that best reflects the continuous variation in nature. However, such costly, high-resolution sampling might be impractical and unnecessary, especially if a more discrete approximation with one sample each from the “epilimnion,” “metalimnion,” and “hypolimnion” would suffice to answer the research question at hand. In a management context, the ecological integrity of a waterbody could be painstakingly described using the complete nucleotide sequences for all organisms that occupy it. But, simply classifying the waterbody’s ecological integrity as “good” or “bad” based on the presence of a few indicator species may be adequate, depending on the management goal. Discretization can also simplify communication with the public. In the case of public swimming advisories, a dichotomous “safe” or “not safe” advisory may facilitate swimmer decisions about whether or not to get in the water. Alternatively, an advisory stating the exact quantitative probabilities of exposure to all toxins may overburden swimmer decision-making. Thus, limnologists intuitively know that the appropriateness of any specific discretization scheme depends on the objectives and the resources at hand to meet those objectives.

1.3. Discretization amid an information explosion

Our objectives and our resources in limnology are changing rapidly. Compared to previous satellite missions, recently deployed remote sensing platforms have higher resolutions that capture more waterbodies at higher frequencies (Palmer et al., 2015). Continuously profiling cameras can collect underwater images of microscopic organisms, process those images using artificial intelligence, and generate real-time biodiversity profiles (Luo et al., 2018). While further improvements are still needed, automated dissolved nutrient sensors are becoming more accurate and less expensive every day (Beaton et al., 2012; Nightingale et al., 2019). These and other developments in automated sensor technology are inspiring new questions about the drivers of fine-scale variability (Crawford et al., 2015). Numerous computing resources are putting advanced statistical tools at our finger tips for free (Read et al., 2011; Winslow et al., 2016; Woolway et al., 2015). National, international, and global databases containing data from many thousands of waterbodies are inspiring new questions and becoming an increasingly important tool to understand aquatic ecosystems (King et al., 2019; U.S. Environmental Protection Agency, 2009). Global networks (e.g. Global Lake Ecological Observatory Network) are making substantial progress in sharing and interpreting high-resolution sensor data from a broad spectrum of waterbodies to understand their role in and response to environmental change (Hanson et al., 2016). Mobile technologies, social media, citizen science, and open access philosophies are reshaping the ways limnologists communicate with each other and with the public (Hampton et al., 2013; Weyhenmeyer et al., 2017).

These rapid changes will have consequences for how and when limnologists find certain discretizations appropriate. Just as modern genomics has strengthened calls to reconsider the biological concept of race as a scientific categorization (Yudell et al., 2016), limnological modernity may also require rethinking limnological discretizations. The burgeoning availability of data and research tools is likely to lead to new discussions and reinvigorate ongoing arguments about the appropriateness of various discretization schemes. In aggregate, resolving these arguments could substantially influence whether the ongoing information deluge improves our science. Here, I summarise some of the advantages and disadvantages of discretization so that it may inform our arguments about discretization amid the current information explosion. To help us resolve these arguments, I offer a framework that can be used to assess the relative usefulness of various discretization approaches and compare them to the alternatives.

2. Advantages and disadvantages of discretization

2.1. *Waterbody discretization (the division of waterbodies into types and zones)*

Like pixelating an image, waterbody discretization partially masks variability within groups, causing a loss of signal in the gradient (Gray et al., 2019). For instance, limnologists often divide the continuous gradient of human influence on waterbody hydrology into the categories, “lake” and “reservoir.” But these terms mask the variety of waterbodies that fall along the continuous gradient that underlies them. The term, “reservoir” can be used to describe a wide range of waterbodies from those that have been created by humans de novo to those that have slight modifications in their water levels due to a dam. Discretization errors associated with the term “reservoir” are partly reduced by adding more classes (e.g. “run-of-river” reservoir). Adding enough classes to sufficiently reduce discretization errors can take the limnological lexicon down a path toward overbearing complexity (e.g. semi-lacustrine-oligotrophic-tropical-run-of-river reservoir) that impedes communication rather than enhancing it (Biber and Gray, 2010).

Waterbody discretization is widespread, in part, because it can facilitate communication among experts by offering useful linguistic shortcuts—jargon. Single words of limnological jargon can stand for whole paragraphs of text in plain English, so jargon can save time and lead to the development of a unifying scholarly identity. But just as the jargon associated with discretization can facilitate communication among experts, it can also hamper communication with non-experts. Jargon is widely denounced in the public sphere by scientific communication specialists who view it as a key boundary to public understanding of science (Venhuizen et al., 2019). Jargon associated with waterbody discretization can even hamper communication among experts from closely related fields as different experts can have different definitions for the same jargon (Chaloner and Wotton, 2011). For example, the ‘littoral zone’ refers to the ‘shallow illuminated zone’ in freshwater ecology but to the “intertidal zone” in marine ecology. Due to the subjectivity of defining jargon, the number of definitions often proliferate. For example, there are over 20 different definitions of the “mixed depth” in use today in aquatic science (Gray et al., 2019; Kara et al., 2000). Thus, waterbody discretization and its associated jargon can be profoundly useful by expediting communication among experts, but it can hamper communication with non-experts and experts from closely related fields.

Discretizing waterbodies can benefit limnology by guiding expectations for how waterbodies function leading to greater focus in our research and collaborations. For instance, many limnologists inherently expect “discrete” classes and zones to behave in distinct ways. Reservoirs are thought to function in fundamentally different ways compared to lakes (Hayes et al., 2017). Some limnologists encourage developing a unique limnology for very small ponds (Hoverman and Johnson, 2012). And large rivers, it is argued, should be modelled separately from other rivers (Puckridge et al., 1998). These expectations can lead limnologists to form discrete, collaborative teams focused on illuminating exceptionally important research topics even if they are predominantly relevant only for specific waterbody types and zones. As a result, ponds, lakes, wetlands, streams, rivers, and oceans are typically studied in isolation at both fine and broad scales due to their perceived differences (Chaloner and Wotton, 2011; King et al., 2019).

But expectations that waterbody classes and zones represent real structure in nature can also be a disadvantage. The compartmentalization of waterbodies is problematic because it counteracts the formulation of general ecological theory and hypotheses founded on waterbody relatedness. It has been argued that this isolation has slowed the development of a common mechanistic understanding of the drivers of carbon (Hotchkiss et al., 2018), nutrient (Elser et al., 2007) and energy (Chaloner and Wotton, 2011) dynamics in aquatic ecosystems. When limnologists study different waterbody types and zones concurrently along continuous gradients, the basic commonalities among all waterbodies can be more apparent, which promotes synthesis and general theory formulation (Chaloner and Wotton, 2011; Hotchkiss et al., 2018). Trans-disciplinarity is widely touted in the scientific literature (Chaloner and Wotton, 2011), and may be key to merging the understanding generated from studying specific waterbody types and zones.

2.2. *Discretization in water management*

Discretization in limnology goes far beyond waterbody classification and zonation—discretization is also widely relied on by limnologists when promoting certain resource management targets. For example, discrete pollution limits are promoted as a waterbody management tool which allows polluters to pollute up to a specific threshold without having to pay (Liu et al., 2015). Threshold-based management is widely encouraged in the scientific literature (Liu et al., 2015) and adopted by local, national, and international management authorities. For instance, the Environmental Protection Agency in the United States uses “Total Maximum Daily Loads” (the maximum amount of a pollutant that a waterbody is allowed to receive) to enforce discrete pollution standards. In Europe, the Water Framework Directive requires that waterbody status be classified as “high,” “good,” “moderate,” “poor,” and “bad” with the goal of achieving at least “good” status for all European waterbodies by discrete deadlines. So, discrete management targets have been relied on for decades to control waterbody stressors with some success (Carvalho et al., 2019; Reckhow, 2001).

But in some contexts, management approaches based on discrete management targets can be suboptimal. The effectiveness of discrete management targets partially depends on whether waterbodies have a predictable, well-defined, discrete capacity to withstand stress. There is an abundant limnological literature on thresholds in stressor-response relationships, but this literature shows that strong thresholds in these relationships are rare, uncertain, and difficult to predict (Groffman et al., 2006; Gsell et al., 2016). Furthermore, potential thresholds may be dynamic through time, making them an unrealistic management target even with multitudinous data from a specific system. Discrete stressor targets can also cause defeatism—the thinking that stressor reductions are valuable only if they take stressors below a threshold. Defeatism may prevent incremental stressor reductions which are beneficial but don't meet the threshold. Conversely, discrete stressor targets can cause complacency—the thinking that stressor reductions that occur below a threshold are worthless, when further reductions would still elicit resource benefits.

2.3. Discretization in sampling and data analysis

All data collection is inherently discrete because scientists cannot continuously measure all of nature. As a result, all limnological measurements, even high-frequency measurements, inevitably reflect limnological processes at discrete points in space and in time. Higher frequency sampling is typically more expensive, but can be more representative of the continuous variation in nature (Meinson et al., 2015). This trade-off between cost and realism is widely documented and approaches have been developed for optimizing sampling frequency (Anttila et al., 2012). Recent technological developments such as automated in situ sensors have substantially altered the acceptability of lower frequency monitoring. As a result of these changes, occasional spot samples are often not an acceptable standalone monitoring approach for many aquatic variables (Nõges et al., 2010) except when collected across many spatially distributed waterbodies. Similarly, due to advances in remote sensing of inland waters, measurements taken from the center of waterbodies are no longer considered to be reflective of individual waterbodies as a whole (Mason et al., 2016; Woolway and Merchant, 2018).

Furthermore, limnologists' discretized approach to designing experiments and analysing data has also been the core of limnological progress for decades. Limnologists design experiments using discrete, replicated treatments (Johnson et al., 2009); test for statistical significance using discrete p-values; cluster data using k-means and other clustering approaches (Maberly et al., 2020; Savoy et al., 2019); and use classification trees to explain variability in data by dividing it into classes based on discrete cut-offs in predictor variables (O'Reilly et al., 2015). These discretizations can save time, resources, and can simplify the interpretation of complex statistical findings.

Despite the widespread dependence of limnological progress on analysis using discretization, some approaches rely more on discretization than others. Overdependence on discretization can lead limnologists to design limnological studies with less statistical power that are more difficult to incorporate into ecological theory. For instance, analysis of variance (ANOVA)—a widely used statistical technique in limnology—has less statistical power than linear regressions when both tests' assumptions are met, yet we often design experiments with a discrete, replicated ANOVA design (Cottingham et al., 2005; Kreyling et al., 2018). Furthermore, linear regression can provide quantitative output with fewer parameters that can be more effectively incorporated into ecological models than ANOVA output (Cottingham et al., 2005; Kreyling et al., 2018). Simple classification tree analysis is also commonly used in limnology, but its discretized simplicity belies its many disadvantages. Simple classification trees make highly approximated representations of continuous functions, their output can be extremely unstable when fit with new data, and other methods vastly outperform them according to widely ranging performance metrics (Prasad et al., 2006). Furthermore, statisticians have made widespread calls to end dichotomous significance testing because the practice often leads to misunderstandings and misinterpretation of results (Amrhein et al., 2019; Johnson, 2007). Just as dichotomous significance testing has been widely criticized, so too have the various statistical clustering approaches commonly used in limnology for easily finding discrete structure in un-structured data (Cormack, 1971).

3. Discretization evaluation framework

When limnologists have a choice about the extent to which they rely on discretization, I suggest that they carefully examine the relative value of specific discretization approaches and compare them to the alternatives using the discretization evaluation framework (Fig. 1). The value of any discretization system should be judged based on the extent to which it satisfies three criteria: objectivity (independent researchers make the same conclusion about the number and definitions of discrete boundaries), predictivity (performs well when predicting other variables), and stability (doesn't change when new observations contribute to the system). The objectivity, predictivity, and stability of discretizations should always be compared to alternatives such as analogous continuous gradients when they are available.

To demonstrate the discretization evaluation framework, here I use a data-driven example from the 2007 National Lakes Assessment (NLA) to assess the objectivity, predictivity, and stability of various lake depth classifications (Fig. 2). There is no broadly-accepted and objective lake depth classification in use today, so I created an ad hoc set of discretized gradients by iteratively dividing

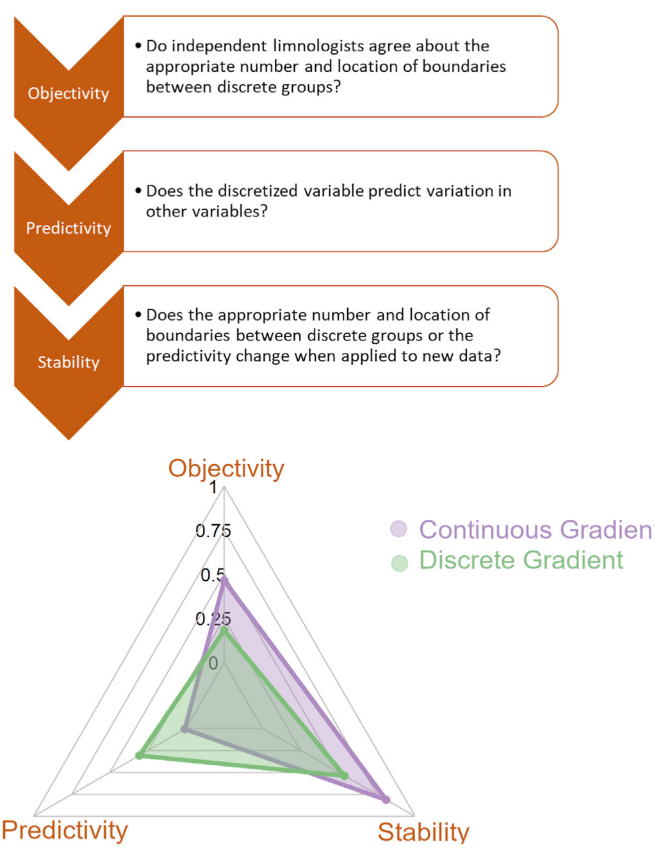

Fig. 1. The discretization evaluation framework. Three criteria by which discrete gradients in limnology can be evaluated for their usefulness as compared to other gradients. The performance of the gradients can be quantified and should be compared in a holistic way to the alternatives. For instance, comparing the relative area of the triangles in the radar plot (where each component is quantified and scaled from 0 to 1) could be one way to facilitate decision-making about whether to use a discretized gradient.

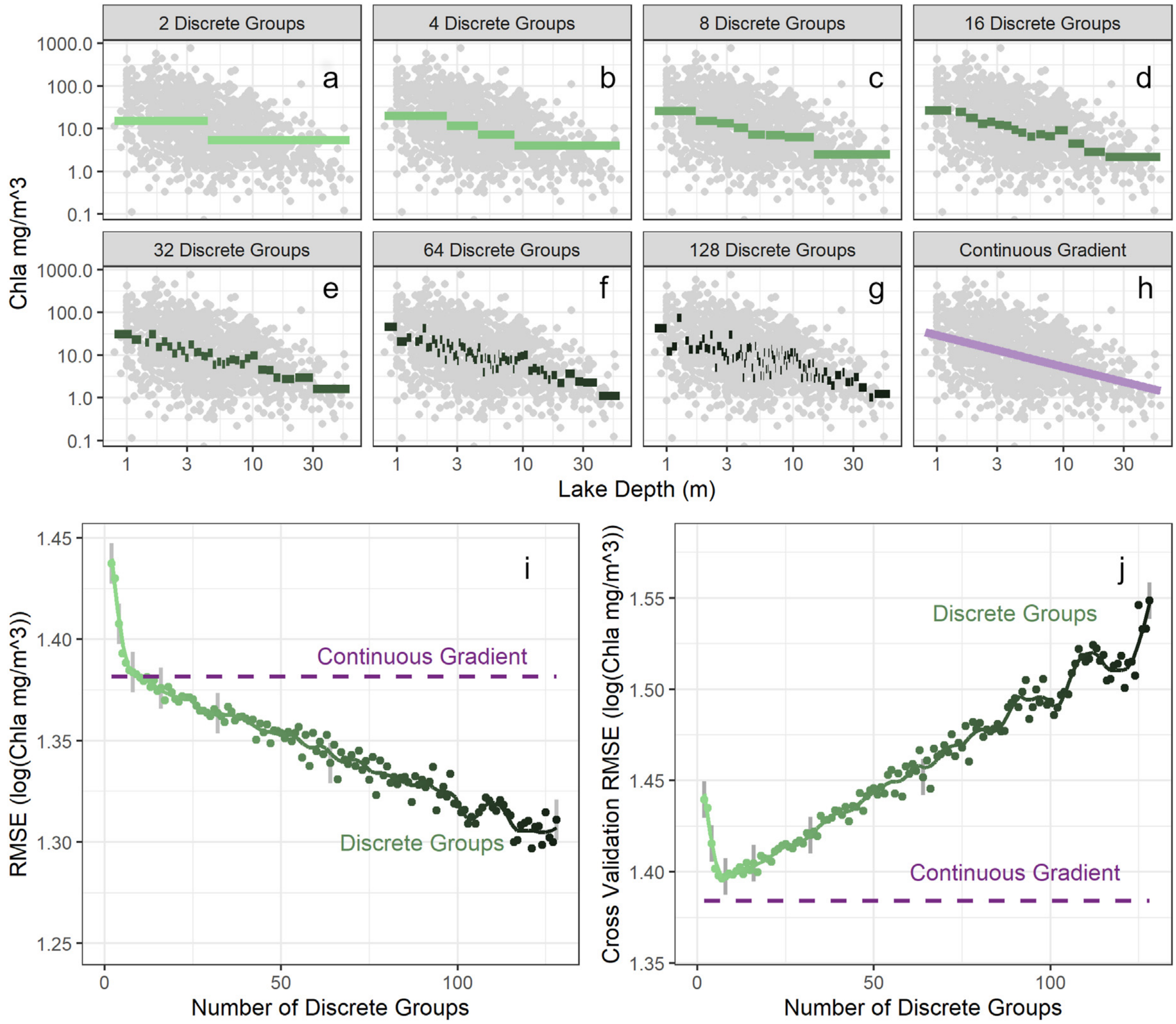

Fig. 2. Data-driven example of the discretization evaluation framework. This example analysis could be used to support decision-making about whether to rely on lake depth classification systems in a specific context. In the discretization evaluation framework presented here, analogous discrete and continuous gradients are evaluated for their objectivity (a–h), predictivity (i), and stability (j). The discretized lake depth gradients used here are fully subjective because no established theory can scientifically demonstrate a priori which discretization scheme (a–g) best reflects reality. For large numbers of groups, discrete gradients are more predictive because they have lower root mean squared errors (RMSE) when used to predict chl-a concentrations (i), although this higher level of predictivity could be considered model over-fitting. Discrete gradients are less stable (higher cross validation RMSE) because all discrete gradients are outperformed by standard linear regression with a continuous depth gradient when evaluated through cross-validation (j).

the lakes from the NLA into evenly sized groups from a total of 2 groups (binary classification) up to a maximum of 562 groups (given 1124 lakes in the dataset, 562 was the largest number of groups such that each group had at least 2 lakes). I evaluated these discretized lake depth classifications for their capacity to predict variation in chl-a concentrations using ordinary least squares regression (Fig. 2). I compared the performance of the model based on discretized gradients (computationally equivalent to ANOVA) to that using the full continuous gradient of depth in the original dataset according to the models' root mean squared error (RMSE). I evaluated the stability of the results through model cross validation with a random selection of 50% of the data used for model training and the remaining 50% of the data used for testing the model with 10,000 repetitions. The stability was characterized by the mean RMSE from the test dataset in cross validation across all 10,000 repetitions. I found that the discretized gradients with the most groups achieved the highest model predictivity (lowest RMSE), although this could be interpreted as a result of model overfitting (Fig. 2). However, the continuous gradient outperformed the discretized gradients with 9 or fewer groups, and the continuous gradient was consistently the most stable in cross validation (lowest RMSE in cross validation). Holistically, the higher objectivity and stability (but partially lower predictivity) of the continuous depth gradient could justify using the continuous depth gradient when predicting chl-a concentrations instead of a discretized depth gradient.

Many other analyses similar to those presented here could be used to assess continuous versus discrete gradients in other contexts. For example, the HydroLakes database, containing information on 1.4 million lakes worldwide, now includes data on the proportion of the waterbody's volume which has been impounded (continuous gradient) in addition to the discretized categories, "lake" and "reservoir" (Messager et al., 2016). Tests of the relative objectivity, predictivity, and stability of these two variables could be illuminating. In a management context, the discretization evaluation framework could also be used to rigorously compare the management outcomes when using discrete pollution targets versus per unit pollution taxes or pollution trading that affect all levels of pollution regardless of whether it is above or below a discrete threshold. Pollution taxes and trading may lead to more optimal outcomes in resource management in some contexts (Muller and Mendelsohn, 2009), but the relative objectivity, predictivity, and stability of these two management approaches needs to be tested for waterbodies at broad scales. Rigorous testing of discretization schemes according to the discretization evaluation framework may even lead us toward objective, scientifically testable discretizations that avoid potentially unproductive arguments about the appropriate number and definitions of boundaries between groups.

4. The bottom line

Regardless of a discretized gradient's performance in the discretization evaluation framework, any calls for rejecting waterbody discretization schemes are likely to be met with fierce resistance because limnological progress to date has relied heavily on them. For instance, early in limnology's history, lakes were divided into low, medium, and high trophic classes based on their productivity (Carlson, 1977). So-called "trophic state classifications" expanded rapidly following this early work, with new limnologists frequently reinventing the scale. But in the 1970's, limnologists began to find the contradictions among trophic state classification systems problematic because the class boundaries were scientifically untestable, preventing the development of a single general classification system (Carlson, 1977). Robert E. Carlson argued for replacing a classification-based system with a less arbitrary continuous gradient from 0 to 100 called the "trophic state index" (Carlson, 1977). While Carlson's paper receives many citations, the discretized, classification-based systems are still widely-used in limnology today. Similar calls for replacing other waterbody classifications or zonations are likely to be met with similar resistance. Just as Darwin and Wallace did not eliminate the species concept with their theory of evolution, limnologists will continue to rely on waterbody discretizations for the foreseeable future because of the useful shortcuts that they offer.

More fundamentally, discretization is partially unassailable because it is essential to science, language, and modern computing. After all, science itself advances through discrete datasets, analyses, and publications. Publications are written using language which is constructed with discrete letters and words. Moreover, computers are essentially discretization machines—turning precise information into 1's and 0's so that it can be represented by on-off transistors in the computer's hardware. Thus, discretization is perpetually engrained in limnology as it is in virtually all human pursuits. Aside from our general human reliance on discretization for science, language, and computing, discretization may simply be unavoidable in specific contexts. Our data may have been prediscretized, we may lack sufficient funds, sampling power, or computational power to fully capture continuous gradients and implement regression-based experimental designs. Or, we may be temporarily required by law to discretize (e.g. section 314 of the United States Clean Water Act requires that lakes be classified according to their "eutrophic" character). Nevertheless, in cases where limnologists do have a choice, rigorously questioning the extent to which we must rely on discretization is appropriate and timely, especially given the ongoing changes to the research questions we ask and the resources we have at hand to address them.

5. The future of limnological discretization

More than 100 years since Stephen Forbes' publication of "The Lake as a Microcosm," we have moved beyond the idea that lake ecosystems are discrete. Limnologists may be poised to further transform their reliance on discretization in the next 100 years. Radical innovations in sensing technology could largely supplant the need for discrete water samples. Unforeseen technical, statistical, and data visualization approaches could allow us to better capture and communicate the full signal in continuous gradients. And quantum computing could even move us one big step past binary computing's limitations (Arute et al., 2019). The extent to which we benefit from these future changes will depend on our mindfulness of discretization and the role it plays in our research, management, and communication. Selectively rethinking discretization and applying the discretization evaluation framework could open up new interactions, new questions, and could even lead to the next major advance. So for now, let's carefully weigh the advantages against the disadvantages and compare to the alternatives when they are available. Doing so will improve our science.

6. Conclusions

- Discretization has profoundly benefitted the way we study, understand, manage, and communicate about aquatic ecosystems. But, limnologists sometimes have a choice over the extent to which they rely on discretization. In these cases, discretization's advantages should be carefully weighed against their disadvantages.
- The core advantage of discretization is that it can provide extraordinarily useful shortcuts, especially when faced with limited resources. These shortcuts can facilitate data collection, data analysis, data interpretation, communication, decision-making, and management while guiding expectations for how ecosystems function.
- But, discretization also has several key disadvantages in specific contexts which are sometimes overlooked. Discretization can inhibit communication, distract from general theory formation, introduce unnecessary subjectivity, mask the relatedness between discrete groups, mask the variability within discrete groups, and lead to suboptimal management approaches and research designs.
- Discretization is partially unassailable because it is the foundation of modern science, language, and computing. But recent changes to the field of limnology including big data and high-resolution sensors have challenged specific aspects of the way we discretize, leading to substantial limnological advances. In light of the rapid and ongoing changes to the field of limnology, I encourage the careful and selective examination of limnological discretization in terms of its objectivity, predictivity, and stability following the discretization evaluation framework. This examination may help limnology stay relevant amid the ongoing information explosion.

Author contribution statement

BMK wrote the manuscript, completed the analyses, and made the figures.

Data availability statement

All data and code used here are freely available at https://doi.org/10.5281/zenodo.3731973.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The author would like to thank Catalina Munteanu, Karan Kakouei, and two anonymous reviewers for their feedback on this manuscript. I also thank the National Lakes Assessment 2007 whose data reflect the collective efforts of dedicated field crews, laboratory staff, data management and quality control staff, analysts, and many others from the Environmental Protection Agency, states, tribes, federal agencies, universities, and other organizations. Please contact nars-hq@epa.gov with any questions about the NLA data. The author was supported by funding from the Leibniz Institute for Freshwater Ecology and Inland Fisheries' International Postdoctoral Research Fellowship (Germany), and from the German Research Foundation through the LimnoScenES project (91/22-1) (Germany).

References

Amrhein, V., Greenland, S., McShane, B., 2019. Scientists rise up against statistical significance. Nature 567, 305–307. https://doi.org/10.1038/d41586-019-00857-9.

Anttila, S., Ketola, M., Vakkilainen, K., Kairesalo, T., 2012. Assessing temporal representativeness of water quality monitoring data. J. Environ. Monit. https://doi.org/10.1039/c2em10768f.

Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J.C., Barends, R., Biswas, R., Boixo, S., Brandao, F.G.S.L., Buell, D.A., Burkett, B., Chen, Y., Chen, Z., Chiaro, B., Collins, R., Courtney, W., Dunsworth, A., Farhi, E., Foxen, B., Fowler, A., Gidney, C., Giustina, M., Graff, R., Guerin, K., Habegger, S., Harrigan, M.P., Hartmann, M.J., Ho, A., Hoffmann, M., Huang, T., Humble, T.S., Isakov, S.V., Jeffrey, E., Jiang, Z., Kafri, D., Kechedzhi, K., Kelly, J., Klimov, P.V., Knysh, S., Korotkov, A., Kostritsa, F., Landhuis, D., Lindmark, M., Lucero, E., Lyakh, D., Mandrà, S., McClean, J.R., McEwen, M., Megrant, A., Mi, X., Michielsen, K., Mohseni, M., Mutus, J., Naaman, O., Neeley, M., Neill, C., Niu, M.Y., Ostby, E., Petukhov, A., Platt, J.C., Quintana, C., Rieffel, E.G., Roushan, P., Rubin, N.C., Sank, D., Satzinger, K.J., Smelyanskiy, V., Sung, K.J., Trevithick, M.D., Vainsencher, A., Villalonga, B., White, T., Yao, Z.J., Yeh, P., Zalcman, A., Neven, H., Martinis, J.M., 2019. Quantum supremacy using a programmable superconducting processor. Nature 2019, 1–7. https://doi.org/10.1038/s41586-019-1666-5.

Beaton, A.D., Cardwell, C.L., Thomas, R.S., Sieben, V.J., Legiret, F.E., Waugh, E.M., Statham, P.J., Mowlem, M.C., Morgan, H., 2012. Lab-on-chip measurement of nitrate and nitrite for in situ analysis of natural waters. Environ. Sci. Technol. https://doi.org/10.1021/es300419u.

Biber, D., Gray, B., 2010. Challenging stereotypes about academic writing: complexity, elaboration, explicitness. J. Engl. Acad. Purp. 9, 2–20. https://doi.org/10.1016/j.jeap.2010.01.001.

Butler, J., 1990. Gender Trouble: Feminism and the Subversion of Identity. Routledge.

Carlson, R.E., 1977. A trophic state index for lakes. Limnol. Oceanogr. 22, 361–369. https://doi.org/10.4319/lo.1977.22.2.0361.

Carvalho, L., Mackay, E.B., Cardoso, A.C., Baattrup-Pedersen, A., Birk, S., Blackstock, K.L., Borics, G., Borja, A., Feld, C.K., Ferreira, M.T., Globevnik, L., Grizzetti, B., Hendry, S., Hering, D., Kelly, M., Langaas, S., Meissner, K., Panagopoulos, Y., Penning, E., Rouillard, J., Sabater, S., Schmedtje, U., Spears, B.M., Venohr, M., van de Bund, W., Solheim, A.L., 2019. Protecting and restoring Europe's waters: an analysis of the future development needs of the Water Framework Directive. Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2018.12.255.

Chaloner, D.T., Wotton, R.S., 2011. Overview: the links that bind aquatic ecosystems. J. North Am. Benthol. Soc. 30, 751–761. https://doi.org/10.1899/10-125.1.

Cormack, R.M., 1971. A review of classification. J. R. Stat. Soc. Ser. A 134, 321. https://doi.org/10.2307/2344237.

Cottingham, K.L., Lennon, J.T., Brown, B.L., 2005. Knowing when to draw the Line : designing more informative ecological experiments. Front. Ecol. Environ. 3, 145–152.

Crawford, J.T., Loken, L.C., Casson, N.J., Smith, C., Stone, A.G., Winslow, L.A., 2015. High-speed limnology: using advanced sensors to investigate spatial variability in biogeochemistry and hydrology. Environ. Sci. Technol. https://doi.org/10.1021/es504773x.

Einstein, A., 1979. Autobiographical notes. Open Court Publishing, La Salle, Illinois.

Elser, J.J., Bracken, M.E.S., Cleland, E.E., Gruner, D.S., Harpole, W.S., Hillebrand, H., Ngai, J.T., Seabloom, E.W., Shurin, J.B., Smith, J.E., 2007. Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. Ecol. Lett. 10, 1135–1142. https://doi.org/10.1111/j.1461-0248.2007.01113.x.

Fan, Y., 2019. Are catchments leaky? WIREs Water 6. https://doi.org/10.1002/wat2.1386.

Forbes, S.A., 1925. The lake as a microcosm. Ill Nat. Hist. Surv. Bull. 15.

Gray, E., Mackay, E.B., Elliott, J.A., Folkard, A.M., Jones, I.D., 2019. Wide-spread inconsistency in estimation of lake mixed depth impacts interpretation of limnological processes. Water Res. 115136 https://doi.org/10.1016/j.watres.2019.115136.

Groffman, P.M., Baron, J.S., Blett, T., Gold, A.J., Goodman, I., Gunderson, L.H., Levinson, B.M., Palmer, M.A., Paerl, H.W., Peterson, G.D., Poff, N.L., Rejeski, D.W., Reynolds, J.F., Turner, M.G., Weathers, K.C., Wiens, J., 2006. Ecological thresholds: the key to successful environmental management or an important concept with No practical application? Ecosystems 9, 1–13. https://doi.org/10.1007/s10021-003-0142-z.

Gsell, A.S., Scharfenberger, U., Özkundakci, D., Walters, A., Hansson, L.-A., Janssen, A.B.G., Nõges, P., Reid, P.C., Schindler, D.E., Donk, E. Van, Dakos, V., Adrian, R., 2016. Evaluating early-warning indicators of critical transitions in natural aquatic ecosystems. Proc. Natl. Acad. Sci. Unit. States Am. https://doi.org/10.1073/PNAS.1608242113, 201608242.

Hampton, S.E., Strasser, C.A., Tewksbury, J.J., Gram, W.K., Budden, A.E., Batcheller, A.L., Duke, C.S., Porter, J.H., 2013. Big data and the future of ecology. Front. Ecol. Environ. https://doi.org/10.1890/120103.

Hanson, P.C., Weathers, K.C., Kratz, T.K., 2016. Networked lake science: how the Global Lake Ecological Observatory Network (GLEON) works to understand, predict, and communicate lake ecosystem response to global change. Inl. Waters. https://doi.org/10.5268/IW-6.4.904.

Hayes, N.M., Deemer, B.R., Corman, J.R., Razavi, N.R., Strock, K.E., 2017. Key differences between lakes and reservoirs modify climate signals: a case for a new conceptual model. Limnol. Oceanogr. Lett. 2, 47–62. https://doi.org/10.1002/lol2.10036.

Hotchkiss, E.R., Sadro, S., Hanson, P.C., 2018. Toward a more integrative perspective on carbon metabolism across lentic and lotic inland waters. Limnol. Oceanogr.: Lett. 3, 57–63. https://doi.org/10.1002/lol2.10081.

Hoverman, J.T., Johnson, P.T.J., 2012. Ponds and lakes: a journey through the life aquatic | learn science at scitable. Nat. Educ. Knowl. 3, 17.

Johnson, D.H., 2007. The insignificance of statistical significance testing. J. Wildl. Manag. 63, 763. https://doi.org/10.2307/3802789.

Johnson, P.T.J., Olden, J.D., Solomon, C.T., Vander Zanden, M.J., 2009. Interactions among invaders: community and ecosystem effects of multiple invasive species in an experimental aquatic system. Oecologia. https://doi.org/10.1007/s00442-008-1176-x.

Kara, A.B., Rochford, P.A., Hurlburt, H.E., 2000. An optimal definition for ocean mixed layer depth. J. Geophys. Res. Ocean. 105, 16803–16821. https://doi.org/10.1029/2000JC900072.

King, K., Cheruvelil, K.S., Pollard, A., 2019. Drivers and spatial structure of abiotic and biotic properties of lakes, wetlands, and streams at the national scale. Ecol. Appl. https://doi.org/10.1002/eap.1957.

Kreyling, J., Schweiger, A.H., Bahn, M., Ineson, P., Migliavacca, M., Morel-Journel, T., Christiansen, J.R., Schtickzelle, N., Larsen, K.S., 2018. To replicate, or not to replicate – that is the question: how to tackle nonlinear responses in ecological experiments. Ecol. Lett. https://doi.org/10.1111/ele.13134.

Liu, J., Kattel, G., Arp, H.P.H., Yang, H., 2015. Towards threshold-based management of freshwater ecosystems in the context of climate change. Ecol. Model. 318, 265–274. https://doi.org/10.1016/j.ecolmodel.2014.09.010.

Luo, J.Y., Irisson, J.O., Graham, B., Guigand, C., Sarafraz, A., Mader, C., Cowen, R.K., 2018. Automated plankton image analysis using convolutional neural networks. Limnol Oceanogr. Methods 16, 814–827. https://doi.org/10.1002/lom3.10285.

Maberly, S.C., O'Donnell, R.A., Woolway, R.I., Cutler, M.E.J., Gong, M., Jones, I.D., Merchant, C.J., Miller, C.A., Politi, E., Scott, E.M., Thackeray, S.J., Tyler, A.N., 2020. Global lake thermal regions shift under climate change. Nat. Commun. 11, 1–9. https://doi.org/10.1038/s41467-020-15108-z.

Mason, L.A., Riseng, C.M., Gronewold, A.D., Rutherford, E.S., Wang, J., Clites, A., Smith, S.D.P., McIntyre, P.B., 2016. Fine-scale spatial variation in ice cover and surface temperature trends across the surface of the Laurentian Great Lakes. Climatic Change 138. https://doi.org/10.1007/s10584-016-1721-2.

Meinson, P., Idrizaj, A., Nõges, P., Nõges, T., Laas, A., 2015. Continuous and high-frequency measurements in limnology: history, applications, and future challenges. Environ. Rev. https://doi.org/10.1139/er-2015-0030.

Messager, M.L., Lehner, B., Grill, G., Nedeva, I., Schmitt, O., 2016. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. Nat. Commun. 7 https://doi.org/10.1038/ncomms13603.

Muller, N.Z., Mendelsohn, R., 2009. Efficient pollution regulation: getting the prices right. Am. Econ. Rev. 99, 1714–1739. https://doi.org/10.1257/aer.99.5.1714.

Nightingale, A.M., Hassan, S.U., Warren, B.M., Makris, K., Evans, G.W.H., Papadopoulou, E., Coleman, S., Niu, X., 2019. A droplet microfluidic-based sensor for simultaneous in situ monitoring of nitrate and nitrite in natural waters. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.9b01032.

Nõges, P., Poikane, S., Kõiv, T., Nõges, T., 2010. Effect of chlorophyll sampling design on water quality assessment in thermally stratified lakes. Hydrobiologia. https://doi.org/10.1007/s10750-010-0237-4.

O'Reilly, C.M., Sharma, S., Gray, D.K., Hampton, S.E., Read, J.S., Rowley, R.J., Schneider, P., Lenters, J.D., McIntyre, P.B., Kraemer, B.M., Weyhenmeyer, G.A., Straile, D., Dong, B., Adrian, R., Allan, M.G., Anneville, O., Arvola, L., Austin, J., Bailey, J.L., Baron, J.S., Brookes, J.D., de Eyto, E., Dokulil, M.T., Hamilton, D.P., Havens, K., Hetherington, A.L., Higgins, S.N., Hook, S., Izmest'Eva, L.R., Joehnk, K.D., Kangur, K., Kasprzak, P., Kumagai, M., Kuusisto, E., Leshkevich, G., Livingstone, D.M., MacIntyre, S., May, L., Melack, J.M., Mueller-Navarra, D.C., Naumenko, M., Noges, P., Noges, T., North, R.P., Plisnier, P.-D.D., Rigosi, A., Rimmer, A., Rogora, M., Rudstam, L.G., Rusak, J.A., Salmaso, N., Samal, N.R., Schindler, D.E., Schladow, S.G., Schmid, M., Schmidt, S.R., Silow, E., Soylu, M.E., Teubner, K., Verburg, P., Voutilainen, A., Watkinson, A., Williamson, C.E.,

Zhang, G., 2015. Rapid and highly variable warming of lake surface waters around the globe. Geophys. Res. Lett. 42, 10773–10781. https://doi.org/10.1002/2015GL066235.
Pace, M.L., Cole, J.J., Carpenter, S.R., Kitchell, J.F., Hodgson, J.R., Van De Bogert, M.C., Bade, D.L., Kritzberg, E.S., Bastviken, D., 2004. Whole-lake carbon-13 additions reveal terrestrial support of aquatic food webs. Nature 427, 240–243. https://doi.org/10.1038/nature02227.
Palmer, S.C.J., Kutser, T., Hunter, P.D., 2015. Remote sensing of inland waters: challenges, progress and future directions. Remote Sens. Environ. 157, 1–8. https://doi.org/10.1016/j.rse.2014.09.021.
Prasad, A.M., Iverson, L.R., Liaw, A., 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 9, 181–199. https://doi.org/10.1007/s10021-005-0054-1.
Puckridge, J.T., Sheldon, F., Walker, K.F., Boulton, A.J., 1998. Flow variability and the ecology of large rivers. Mar. Freshw. Res. 49, 55–72. https://doi.org/10.1071/MF94161.
Raymond, P.A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C., Kortelainen, P., Dürr, H., Meybeck, M., Ciais, P., Guth, P., 2013. Global carbon dioxide emissions from inland waters. Nature 503, 355–359. https://doi.org/10.1038/nature12760.
Read, J.S., Hamilton, D.P., Jones, I.D., Muraoka, K., Winslow, L.A., Kroiss, R., Wu, C.H., Gaiser, E., 2011. Derivation of lake mixing and stratification indices from high-resolution lake buoy data. Environ. Model. Software 26, 1325–1336. https://doi.org/10.1016/j.envsoft.2011.05.006.
Reckhow, K.H., 2001. Assessing the TMDL Approach to Water Quality Management. Water Resour. Res. Inst. News Univ. North Carolina. https://doi.org/10.1029/01eo00380.
Savoy, P., Appling, A.P., Heffernan, J.B., Stets, E.G., Read, J.S., Harvey, J.W., Bernhardt, E.S., 2019. Metabolic rhythms in flowing waters: an approach for classifying river productivity regimes. Limnol. Oceanogr. https://doi.org/10.1002/lno.11154.
Shurin, J.B., Borer, E.T., Seabloom, E.W., Anderson, K., Blanchette, C.A., Broitman, B., Cooper, S.D., Halpern, B.S., 2002. A cross-ecosystem comparison of the strength of trophic cascades. Ecol. Lett. 5, 785–791. https://doi.org/10.1046/j.1461-0248.2002.00381.x.
Tranvik, L.J., Cole, J.J., Prairie, Y.T., 2018. The study of carbon in inland waters-from isolated ecosystems to players in the global carbon cycle. Limnol. Oceanogr. Lett. 3, 41–48. https://doi.org/10.1002/lol2.10068.
U.S. Environmental Protection Agency, 2009. National Lakes Assessment: a Collaborative Survey of the Nation's Lakes. EPA 841-R-09-001, EPA 841-R-09-001.
Venhuizen, G.J., Hut, R., Albers, C., Stoof, C.R., Smeets, I., 2019. Flooded by jargon: how the interpretation of water-related terms differs between hydrology experts and the general audience. Hydrol. Earth Syst. Sci. https://doi.org/10.5194/hess-23-393-2019.
Weyhenmeyer, G.A., MacKay, M., Stockwell, J.D., Thiery, W., Grossart, H.P., Augusto-Silva, P.B., Baulch, H.M., De Eyto, E., Hejzlar, J., Kangur, K., Kirillin, G., Pierson, D.C., Rusak, J.A., Sadro, S., Woolway, R.I., 2017. Citizen science shows systematic changes in the temperature difference between air and inland waters with global warming. Sci. Rep. 7, 1–9. https://doi.org/10.1038/srep43890.
Winslow, L.A., Zwart, J.A., Batt, R.D., Dugan, H.A., Iestyn Woolway, R., Corman, J.R., Hanson, P.C., Read, J.S., 2016. LakeMetabolizer: an R package for estimating lake metabolism from free-water oxygen using diverse statistical models. Inl. Waters 6, 622–636. https://doi.org/10.5268/IW-6.4.883.
Woolway, R.I., Jones, I.D., Hamilton, D.P., Maberly, S.C., Muraoka, K., Read, J.S., Smyth, R.L., Winslow, L.A., 2015. Automated calculation of surface energy fluxes with high-frequency lake buoy data. Environ. Model. Software. https://doi.org/10.1016/j.envsoft.2015.04.013.
Woolway, R.I., Merchant, C.J., 2018. Intralake heterogeneity of thermal responses to climate change: a study of large northern hemisphere lakes. J. Geophys. Res. Atmos. https://doi.org/10.1002/2017JD027661.
Yudell, M., Roberts, D., DeSalle, R., Tishkoff, S., 2016. Taking race out of human genetics. Science 351, 564–565. https://doi.org/10.1126/science.aac4951 (80-.).
Yvon-Durocher, G., Caffrey, J.M., Cescatti, A., Dossena, M., Giorgio, P. Del, Gasol, J.M., Montoya, J.M., Pumpanen, J., Staehr, P.a., Trimmer, M., Woodward, G., Allen, A.P., del Giorgio, P., Gasol, J.M., Montoya, J.M., Pumpanen, J., Staehr, P.a., Trimmer, M., Woodward, G., Allen, A.P., Giorgio, P. Del, Gasol, J.M., Montoya, J.M., Pumpanen, J., Staehr, P.a., Trimmer, M., Woodward, G., Allen, A.P., 2012. Reconciling the temperature dependence of respiration across timescales and ecosystem types. Nature 487, 472–476. https://doi.org/10.1038/nature11205.