



EPA Public Access

Author manuscript

Ecol Indic. Author manuscript; available in PMC 2023 August 01.

About author manuscripts

Submit a manuscript

Published in final edited form as:

Ecol Indic. 2022 August ; 141: 1–13. doi:10.1016/j.ecolind.2022.109131.

Genus-level, trait-based multimetric diatom indices for assessing the ecological condition of rivers and streams across the conterminous United States

Luisa Riato^{a,*}, Ryan A. Hill^b, Alan T. Herlihy^c, David V. Peck^b, Philip R. Kaufmann^{b,c}, John L. Stoddard^b, Steven G. Paulsen^b

^aOak Ridge Institute for Science and Education (ORISE) Post-Doctoral Fellow c/o U.S. Environmental Protection Agency, Center for Public Health and Environmental Assessment, Pacific Ecological Systems Division, 200 SW 35th St., Corvallis, OR 97333, USA

^bU.S. Environmental Protection Agency, Center for Public Health and Environmental Assessment, Pacific Ecological Systems Division, 200 SW 35th St., Corvallis, OR 97333, USA

^cDepartment of Fisheries, Wildlife and Conservation Sciences, Oregon State University, Corvallis, OR 97333, USA

Abstract

Taxonomic inconsistency in species-level identifications has constrained use of diatoms as biological indicators in aquatic assessments. We addressed this problem by developing diatom multimetric indices (MMIs) of ecological condition using genus-level taxonomy and trait-based autecological information. The MMIs were designed to assess river and stream chemical, physical and biological condition across the conterminous United States. Trait-based approaches have the advantage of using both species-level and genus-level data, which require less effort and expense to acquire than traditional species-based approaches and eliminate the persistent taxonomic biases introduced over vast geographic extents. For large-extent assessment programs that require multiple taxonomic laboratories to process samples, such as the United States Environmental Protection Agency's (U.S. EPA's) National Rivers and Streams Assessment (NRSA), the trait approach can eliminate discrepancies in species-level identification or nomenclature that hinder diatom data interpretation. We developed trait-based MMIs using NRSA data for each of the three

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

*Corresponding author. riato.luisa@epa.gov (L. Riato).

CRedit authorship contribution statement

Luisa Riato: Conceptualization, Data curation, Formal analysis, Project administration, Investigation, Visualization, Methodology, Writing – original draft. **Ryan A. Hill:** Conceptualization, Validation, Methodology, Visualization, Writing – review & editing.

Alan T. Herlihy: Resources, Conceptualization, Validation, Supervision, Methodology, Visualization, Writing – review & editing.

David V. Peck: Resources, Conceptualization, Validation, Methodology, Visualization, Writing – review & editing. **Philip R.**

Kaufmann: Resources, Conceptualization, Validation, Methodology, Visualization, Writing – review & editing. **John L. Stoddard:**

Conceptualization, Validation, Supervision, Funding acquisition, Methodology, Visualization, Writing – review & editing. **Steven G.**

Paulsen: Conceptualization, Validation, Supervision, Methodology, Visualization, Writing – review & editing.

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ecolind.2022.109131>.

large ecoregions across the U.S. - the East, Plains, and West. All three MMIs performed well in discriminating least-disturbed from most-disturbed sites. The MMI for the East had the greatest discrimination ability, followed by MMIs for the Plains and West, respectively. The performance of the MMIs was comparable to that observed in existing NRSA fish and macroinvertebrate MMIs. Our research shows that trait-based diatom indices constructed on genus-level taxonomy can be effective for large-scale assessments, and may also allow programs such as NRSA to assess trends in freshwater condition retrospectively, by revisiting older diatom datasets. Moreover, our genus-based approach facilitates including of diatoms into other assessment programs that have limited monitoring resources.

Keywords

Diatoms; Taxonomic inconsistency; Ecological assessment; Traits; Genus taxonomy; National Rivers and Streams Assessment

1. Introduction

Diatoms are a highly diverse group of algae that are found in almost all aquatic habitats, and have well-established taxon-specific tolerances and preferences for a broad range of environmental factors such as pH, nutrients and salinity (Dixit et al., 1992; Smol and Stoermer, 2010). Because of these characteristics, diatom indices are commonly used to assess environmental changes in freshwater aquatic resources, in which diatoms identified to the lowest taxonomic level serve as a basis for the analyses (e.g., Coste et al., 2009; Lavoie et al., 2006; Whitton and Kelly, 1995). Diatom indices based on species are routinely used in the European Water Framework Directive (WFD; European Union, 2000) for river ecological assessments. These include the Pollution Sensitivity Index (IPS, Coste, 1982), the Trophic Diatom Index (TDI, Kelly and Whitton, 1995), and the Biological Diatom Index (BDI, Prygiel and Coste, 1998). In the United States, studies have shown the potential use of species-level diatom metrics in aquatic assessments using large-scale datasets from federal surveys such as the United States Environmental Protection Agency's (U.S. EPA's) Environmental Monitoring and Assessment Program (EMAP) pilot survey (e.g., Hill et al., 2003, 2000; Pan et al., 1996), and the U.S. Geological Survey National Water-Quality Assessment (e.g., Potapova and Charles, 2007). While species-level diatom-based assessments have met with some success, they also has many challenges. The utility of diatoms to evaluate ecological conditions can be constrained by inconsistencies in taxonomic identification (Tapolczai et al., 2016; Tyree et al., 2020), particularly when the use of a single taxonomist is not practical. For large-scale and long-term assessment programs that require multiple taxonomic laboratories to process samples, such as the U.S. EPA's National Rivers and Streams Assessment (NRSA), discrepancies in species-level identification or nomenclature can preclude the use of diatom data in large-scale regional and national assessments (Cao et al., 2007; Kahlert et al., 2012, 2009; Lee et al., 2019). While the taxonomic intercalibration (i.e., harmonization) among different analysts and laboratories was recognized in past assessments, it is difficult to achieve. This has contributed to sufficiently compromising the quality and confidence in diatom datasets to

delay or prohibit the use of diatoms in the National Aquatic Resource Surveys (Lee et al., 2019).

In the U.S., the lack of taxonomic consistency is mostly attributed to four analytical practices (Alers-García et al., 2021): (1) Many North American references are incomplete, prompting analysts to rely on a more comprehensive set of European references to assign names to North American species (Kociolek and Spaulding, 2000). While many diatom species are shared between North America and Europe, many others are not shared, and in such cases, the use of European floras can result in the misapplication of names to North American taxa. (2) Access to taxonomic literature can vary among laboratories, where one laboratory may have a comprehensive set of references, another laboratory may have limited resources to support accurate species identification. Thus, the use of different taxonomic references among laboratories can increase the risk of assigning different names to the same species (Tyree et al., 2020). (3) In North America, there are numerous species that are not yet formally described or have incomplete descriptions (e.g., Potapova and Charles 2003), and this increases the potential for taxonomic inconsistencies. (4) Accurate species identification is time-consuming and requires great effort and expertise, yet there is typically insufficient time for analysts to complete the species identifications and enumerations, increasing the likelihood of mistaken identifications and reduced harmonization among laboratories (Bishop et al., 2017). These issues with species-based analysis practices are not exclusive to U.S. programs, but are encountered by environmental programs worldwide (e.g., Bere 2016). In the U.K. and Europe, taxonomic inconsistencies from misidentified diatom species among different analysts and laboratories is a major concern, and, as such, intercalibration exercises have been used as one of the solutions for resolving taxonomic issues (Kahlert et al., 2016; Kelly et al., 2009a; Straile et al., 2013; Werner et al., 2016).

Trait-based approaches are increasingly used as an alternative in aquatic studies because traits can be assigned to both species-level and to genus-level data, which is simpler and less costly to obtain (Rimet and Bouchez, 2012a). These approaches can also provide insight into assemblage responses to environmental changes since morphologically similar species may respond to the environment in the same way, and therefore relationships with environmental factors can be stronger than those with single species (Lange et al., 2016; Rimet and Bouchez, 2012a). In addition, trait-based methods can help us understand the structure and functioning of assemblages (Verberk et al., 2013) that may provide deeper insights into the ecological condition of a stream than traditional species-based methods (Kelly et al., 2009b). Traits are a basic unit of classifying organisms, in which morphological, physiological, or functional features can be assigned to taxa based on species- or genus-level attributes (Riato et al., 2017; Stevenson et al., 2013; Tapolczai et al., 2017). The abundance of taxa associated with a particular trait can be responsive to environmental stressors. For example, several studies have shown that diatom traits (e.g., ecological guilds, life-forms and size) are responsive to nutrient and pesticide contamination (e.g., Lavoie et al., 2010; Passy 2007a; Rimet and Bouchez, 2011a), organic pollution (Berthon et al., 2011), and changing hydrology (B-Béres et al., 2014). More recently, Tapolczai et al. (2017) showed that trait-based indices had strong correlations with a nutrient and organic matter/turbidity gradient, and that traits could be useful in water quality assessments. In this regard, trait-based approaches using genus-level taxonomy could offer a suitable alternative

when reliable species-level data are not available because of taxonomic inconsistencies. While the level of taxonomic resolution required for diatom-based assessments depends on the research questions and goals of specific assessment programs, earlier studies show genus-level indicators can provide a robust biotic assessment because of the large similarity in ecological preferences and tolerances of species within a genus (Chessman et al., 1999; Hill et al., 2001; Rimet and Bouchez, 2012b). Moreover, Lee et al. (2019) found that using genus-level taxonomy improved taxonomic consistency in the 2008–2009 NRSA diatom data.

In this paper, we develop multimetric indices (MMIs) using genus-level diatom taxonomy and trait-based autecological information, effectively circumventing the problem of taxonomic inconsistencies in the NRSA diatom datasets from 2008–2009 and 2013–2014. We designed an MMI to assess river and stream ecological condition across the conterminous United States, maximizing the discrimination of least-disturbed sites from most-disturbed sites. An MMI is composed of several metrics that are intended to represent different perspectives on how an assemblage uses the resources of a habitat, and that these dimensions also respond to anthropogenic disturbances (Barbour et al., 1996; Stoddard et al., 2008). Over the years, the NRSA surveys have used MMIs based on assemblages of macroinvertebrates and fish to assess the ecological condition of rivers and streams across the conterminous U.S. (CONUS) (e.g., Stoddard et al., 2008; USEPA, 2020a). The NRSA has routinely used genus-level data for macroinvertebrates, assigning traits at the genus-level, and the macroinvertebrate MMIs have been extremely effective for large-scale assessments (USEPA, 2016a, 2020a). Even though diatom assemblage samples were collected, the NRSA does not yet include an MMI for diatoms in the suite of biological indicators used routinely for national and regional assessments. Our goals were to: (1) develop trait-based diatom MMIs from the 2008–2009 NRSA genus-level diatom data to assess river and stream condition at the national-scale and within three large ecoregions used by the NRSA; (2) evaluate the robustness of the final MMIs by comparing the performance of the MMIs on a validation dataset (2013–2014 NRSA data) to the calibration dataset (2008–2009 NRSA data); and (3) further evaluate the utility of the final MMIs by comparing their performance with that of existing NRSA MMIs for other biological assemblages. The process used in this study provides an approach to developing diatom indices that could be applied to diatom assessment datasets at any spatial scale, in any geographical area.

2. Material and methods

2.1. Sample collection and treatment

Each NRSA survey was conducted over a two-year period with surveys in 2008–2009 and 2013–2014. The 2008–2009 sample data was used as a calibration data set to develop the MMI; and the 2013–2014 data was used as a validation data set (USEPA, 2020b, 2016b). Repeat site visits were included in the MMI development to calculate the signal- to-noise ratio (S/N), as part of the step-wise screening procedure (refer to 2.2.3 Best-candidate metric selection). Datasets are available at: <https://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys>.

The protocols to sample, process, and standardize diatom, water chemical, and physical habitat data for the 2008–2009 and 2013–2014 NRSA surveys are described in USEPA (2008, 2016a, 2020a). Briefly, for the 2008–2009 NRSA survey that we used as the calibration dataset, diatoms, physical habitat parameters, and chemical parameters were sampled in the summer from 2123 river and stream sites across the CONUS (USEPA, 2016a). These included 1924 sites that were randomly selected from a pool of sites using a probabilistic sampling design, and 199 hand-picked sites selected to increase the number of potentially least-disturbed sites. For the 2013–2014 NRSA survey that we used as the validation dataset, similar samples were collected in the summer from 2080 sites (1853 randomly selected and 227 hand-picked sites) across the CONUS (USEPA, 2020a).

As part of the MMI development process, all sites were classified as either least-disturbed, intermediate, or most-disturbed by screening a set of disturbance metrics using the approach outlined in (Herlihy et al., 2008). Both least-disturbed and most-disturbed thresholds were set for each metric by NARS ecoregion (refer to 2.2 Development of MMIs). The list of disturbance metrics included measures of water chemistry (total nitrogen, total phosphorus, sulfate, chloride, turbidity, and acid neutralizing capacity), physical habitat (riparian zone disturbance index, % fine substrate), and local (1 km circle around each site) land cover data (% agriculture and % developed). A site had to pass every one of the ecoregional least-disturbed metric thresholds for the site to be considered least-disturbed (USEPA, 2016a). If a site exceeded any one of the ecoregional most-disturbed metric thresholds, then it was considered most-disturbed.

The target population for the NRSA surveys includes all flowing waters of the U.S.—both wadeable and boatable sites were sampled, from headwater streams less than a meter wide to the Mississippi River. At each site, the sampled reach was equivalent in length to 40 times the mean wetted channel-width, with a minimum/maximum length of 150–4000 m, respectively, for very small or very wide systems. For both wadeable and boatable sites, periphyton samples were collected from 11 evenly-spaced transects along a sample reach around the randomly selected site (USEPA, 2013a, 2013b). For wadeable sites, periphyton samples were collected while wading from alternating, left, center, right areas of each transect, whereas for boatable sites, samples were collected at the near shore from boats and onshore (USEPA, 2013a, 2013b). At each of the 11 transects, a sample of hard substrate (rock or wood < 15 cm diameter) (or soft sediment if large substrates were unavailable) was collected at a depth of 0.5 m. Periphyton was removed from the upper surface of the hard substrate using a toothbrush, or, for soft sediment samples, the top 1 cm of sediment using a syringe. Periphyton material from all 11 transects were composited in one bottle to produce a single sample for each site and preserved in 2 mL of 10% formalin per 50 mL of sample. Samples were returned to the laboratory where they were acid-cleaned and mounted on microscope slides for identification and enumeration of diatoms (USEPA, 2008). Diatom samples were analyzed by three different laboratories in the U.S. The target count for taxonomists was 600 diatom valves per sample which was not always attainable due to a high quantity of debris or small quantity of diatoms in some samples (Lee et al., 2019). We retained samples with at least 400 valves for the construction and validation of the MMIs. Table 1 shows the number of sample sites included in both the calibration and validation datasets of the MMIs.

2.2. Development of MMIs

We constructed trait-based diatom MMIs at two spatial scales (i.e., national and regional) to determine the most effective scale for diatom-based assessment of river and stream condition across the CONUS. We built a national MMI based on all sampled sites across the U.S. For the regional-scale, we developed a separate MMI for each of three major ecoregions used by the National Aquatic Resources Survey reports - the East, Plains, and West (Fig. 1). Each major region is comprised of finer-scale ecoregions – 9 ecoregions across all three regions aggregated from Omernik (1987) level III ecoregions (Herlihy et al., 2008; USEPA, 2006). The East includes 4 of those 9 ecoregions: Northern Appalachians (NAP), Southern Appalachians (SAP), Coastal Plains (CPL), and Upper Midwest (UMW); the Plains MMI is comprised of 3 ecoregions: Temperate Plains (TPL), Northern Plains (NPL), and Southern Plains (SPL); and the West includes 2 ecoregions: Western Mountains (WMT) and Xeric (XER) (Fig. 1). In the sections below, we describe the methods to develop trait-based diatom MMIs at both the national- and regional-scale.

2.2.1. Candidate diatom metrics—To develop candidate trait metrics, we assigned traits to each taxon with genus-level taxonomy. A total of 19 traits from two trait categories were assigned to taxa (Table 2):

1. Functional traits, including fourteen life-form classes (e.g., Tube-living, Rosette or Ribbon colonies) and four ecological guilds (High-profile, Low-profile, Motile or Planktic), that were assigned to taxa from an existing database developed by Rimet and Bouchez (2012a; 2019). These authors assigned species and genera to the different life-forms and guilds based on existing classifications found in Passy (2007a, 2007b).
2. One physiological trait comprised of N-fixing diatoms. This trait is assigned to *Epithemia* and *Rhopalodia* taxa that have nitrogen-fixing cyanobacterial endosymbionts (Lange et al., 2016; Prechtel, 2004).

Several genera in the NRSA data were not in the traits database and subsequently were dropped from the analysis. However, taxa belonging to these genera were rare (occurred in < 1 % of sites) and unlikely to have an important contribution to the metrics. Candidate metrics were also developed from all diatom genera. Each trait and genus metric was calculated as % taxa and % individuals. For example, the metric % *High-profile guild taxa* is the percentage of the total number of taxa identified in a sample that are members of the High-profile guild, and % *High-profile individuals* is the percentage of all individuals in a sample that are members of the High-profile guild. A total of 304 candidate metrics were developed and evaluated to identify the best set of metrics for the MMIs.

2.2.2. Assigning traits to taxa at genus-level vs. species-level—While some genera are comprised of species that possess the same trait (e.g., all *Navicula* are motile), there can be a diversity of traits among species within other genera. For example, many species of the *Fragilaria* genus belong to the High-profile guild but some *Fragilaria* belong to the Planktic guild. Thus, assigning traits to diatoms using genus-level taxonomy and not species-level could potentially influence the effectiveness of the trait metrics to indicate condition. For this reason, we determined whether species that had been assigned traits at the

genus-level in the final MMI metrics had been assigned the correct traits at the species-level. For each trait metric in the final MMI, we used trait information at the species-level based on Rimet and Bouchez (2012a; 2019), to examine the traits of every single species that occurred in 1% of sites in the calibration dataset (2008–2009 NRSA) that were included in the scoring of the trait metric, and the validation dataset (2013–2014 NRSA). This evaluation would help determine whether genus-level taxonomy was sufficient to apply reliable trait metrics that were included in the final MMIs in NRSA assessments. For those species that were assigned traits incorrectly at the genus level, species that had relatively low occurrences (<10% of sites) were included in the MMI development since they would have little influence on the overall metric performance, and species that had higher occurrences that could affect metric performance were dropped from the analysis.

2.2.3. Best-candidate metric selection—Best-candidate metrics were selected from the pool of 304 candidate metrics using the following step-wise screening procedure detailed in (Stoddard et al., 2008).

Step 1 Range test—We discarded metrics that resulted in a ‘FAIL’ in any part of either range Test 1 or Test 2 as described below and in Magee et al. (2019); this eliminates metrics that have poor ability to distinguish among sites in different ecological conditions.

Test 1. Identify metrics with many 0 values or strongly skewed distributions:

- a. If the 75th percentile = 0, i.e., 75% or more of values are zero, then FAIL.
- b. If the 75th percentile = the minimum OR the 25th percentile = max (indicating 75% of values identical), then FAIL.

Test 2. Identify metrics with very narrow ranges:

- a. If the metric is a percent variable and $(\text{max} - 25\text{th percentile}) < 15\%$, then FAIL;
- b. If the metric is not a percent variable and $(\text{max} - 25\text{th percentile}) < (\text{max}/3)$, then FAIL.

Step 2 Repeatability—We calculated the signal-to-noise ratio (S/N), which is the ratio of the variance of a metric among sites (signal) to the variance within a site (noise) where the noise is based on within-season revisits to resampled sites (i.e., sites that were sampled in both the 2008–2009 and 2013–2014 survey, which were visited twice in each survey) (Kaufmann et al., 1999). For each sample year of the NRSA, 10% of the sites that were selected from the previous survey to be resampled were selected at random to be sampled twice during the summer index period. Variance components were calculated using R statistical software (R Development Core Team, 2017) package “lme4” version 1.1–7 (Kaufmann et al., 1999). The higher the S/N ratio of a metric, the smaller the proportion of measurement error and within-season variability, and the greater the amount of variance among sites that is potentially explainable by differences in geoclimatic and anthropogenic driving variables. We eliminated metrics with a $S/N < 3$, which was a higher threshold for rejection than was used for periphyton metrics (<1.5) and macroinvertebrate metrics (<2) (Stoddard et al., 2008).

Step 3 Responsiveness—This test evaluates the metric’s ability to discriminate least-disturbed sites from most-disturbed sites. We tested responsiveness by using *t*-scores to compare mean values between least- and most-disturbed sites; the higher the *t*-score, the greater the metric responsiveness. For the national MMI and East MMI, we eliminated metrics with a *t*-score < 3 (the standard in NRSA MMI development), and for the Plains MMI and West MMI, *t*-score < 2 were deleted. Using all the site data, only 29 metrics out of a pool of 304 candidate metrics passed all three screening steps and were considered for inclusion in the national MMI. Similarly, 35 out of 274 metrics passed for further evaluation for just the East MMI development; 21 out of 266 metrics passed for the Plains MMI development; and 28 out of 240 metrics for the West MMI development.

2.2.4. Metric re-scaling—To standardize the range of final candidate metrics that passed the initial screening process, each metric was rescaled to a score between 0 and 10, based on metric values between the 5th (P_5) and 95th (P_{95}) percentile across all sites (Blocksom, 2003). Each metric value was rescaled using the formula:

$$\text{Rescaled metric value} = \frac{\text{original metric value} - P_5}{P_{95} - P_5}$$

For metrics that decreased at most-disturbed sites and increased at least-disturbed sites, scores were represented on a 0–10 scale, whereas metrics that increased at most-disturbed sites and decreased at least-disturbed sites, the scale was reversed so that metric scores were on a 10–0 scale. If rescaled metric values were less than zero, these values were adjusted to 0, and if rescaled values were greater than 10, values were adjusted to 10.

2.2.5. Final metric selection—The final candidate metrics were further screened by evaluating the performance of all possible MMIs using all metric combinations following a modified approach of Van Sickle (2010). We examined thousands of candidate MMIs based on random combinations of 4, 5, 6, 8, and 10 metrics. Evaluating the performance of numerous candidate MMIs based on all possible metric subsets avoids selection of metrics with spuriously good performance (Mazor et al., 2016). The performances of each candidate MMI was evaluated according to four measures:

1. *Redundancy*—We required that candidate MMIs had a maximum Pearson correlation among component metrics (in a single MMI) of < 0.75, while the mean Pearson correlation among all of the metrics (in a single MMI) had to be < 0.5. These cutoffs were based on methodology used in developing other national biotic MMIs in NARS (e.g., Stoddard et al., 2008) and we found that they worked equally as well for our diatom MMI.
2. *Sensitivity*—This is measured as the percentage of most-disturbed sites that were evaluated as being in poor condition based on the MMI; the higher the percentage, the more sensitive the MMI. To determine this, we used an interval test (Kilgour et al., 1998) that establishes whether the MMI score for a given site is significantly lower ($P < 0.05$) than the 5th percentile of MMI scores for least-disturbed sites. The test assumed a normal distribution MMI scores within

the set of least-disturbed sites, and used a non-central F distribution to model the uncertainty in the 5th percentile of that distribution. In this way, this test allowed us to compare an individual sample to a critical F value to determine whether impact on diatoms was detected. This is a more conservative approach than simply comparing the MMI at a site to a 5th percentile threshold because it considers that uncertainty around the estimate of the 5th percentile. (Kilgour et al., 1998).

3. *Repeatability*—We evaluated each candidate MMI using its S/N ratio and selected candidate MMIs with higher S/N ratios.
4. *Precision*—We calculated the Standard Deviation (SD) of scaled MMI values among the least-disturbed sites, with lower SDs indicating higher precision.

A list of the highest performing MMIs was then produced, ranked in descending order of their sensitivity. From this list of most sensitive MMIs, we selected MMIs that had the highest S/N, lowest SD, and low mean and maximum correlations among component metrics. Those top-performing MMIs ranged in metric set size from 4 to 8 metrics. Finally, we compared the responsiveness of the top performing MMIs, as evidenced by their ability to distinguish least- from most-disturbed sites – a critical test of the effectiveness of an MMI (Stoddard et al., 2008). To test responsiveness, we evaluated boxplot separation distances of MMI scores between least- and most-disturbed sites, and t -tests to compare mean scores between the two disturbance classes of sites. The boxplot and t -test analyses were based on re-scaled MMI scores for each site on a scale of 0 to 100 by summing the scored values for all the metrics in the MMI, then dividing that value by the number of metrics in the MMI and multiplying by 100 following (Stoddard et al., 2008). We then selected the final national MMI and final MMIs for the East, Plains and West based on an MMI having the best overall responsiveness (i.e., highest t -scores and greatest boxplot separation distance between least- and most-disturbed sites). For the national MMI, we examined the overall responsiveness of the MMI within each of the 9 NRSA ecoregions; and for the MMIs for the East, West and Plains, we assessed overall responsiveness within each of its component ecoregions. We compared the overall responsiveness of the final national MMI with the responsiveness of the final MMIs for the East, Plains and West, to select the MMIs that were most effective (i.e., MMIs with overall highest t -scores and greatest boxplot separation between least- and most-disturbed sites).

2.3. Validation of MMIs

We tested the robustness of the selected MMIs by comparing the MMI's ability to distinguish least- from most-disturbed sites in the calibration data (2008–2009 NRSA), with their ability to separate these two classes of sites in the validation data (2013–2014 NRSA). For a robust MMI, the ability to discriminate between the two classes of sites within the calibration data should be comparable with the ability to discriminate within the validation data. To test this, we developed the MMI component metrics based on the validation data, and then re-scaled MMI scores for each site on a scale of 0 to 100 in the same way as the calibration data, and then used boxplot separation differences to compare MMI scores between least- and most-disturbed sites. We also compared their

responsiveness (i.e., *t*-scores) with the responsiveness of existing NRSA MMIs for stream fish and macroinvertebrate condition.

3. Results and discussion

3.1. National MMI vs regional MMIs

Using the calibration data, the top-performing MMIs included a 4-metric national MMI, and 3 separate 5-metric MMI's for the East, Plains, and West ecoregions. Table 3 shows the component metrics for each MMI and their expected positive or negative response to anthropogenic disturbances. The national MMI was comprised of trait metrics only. Each regional MMI was comprised of a combination of trait metrics and genus metrics. As one approach to reduce taxonomic inconsistencies in the 2008–2009 NRSA diatom data, (Lee et al., 2019) elevated the species-level taxonomy to the genus level but did not find improvements in taxonomic consistency for the genera *Sellaphora*, *Mayamaea*, and *Psammodictyon*. However, this was not an issue in the current study for the genus metrics included in the regional MMIs (e.g., *Cyclotella*, *Melosira* and *Surirella % individuals*; Table 3) because they did not include the genera that were problematic in (Lee et al., 2019).

All three regional MMIs included a metric that combined *Achnantheidium* and *Achnanthes* individuals because the combined metric improved the MMI responsiveness relative to including either metric individually. Some taxa in *Achnantheidium* and other monoraphid genera were not distinguished from *Achnanthes* in many of the references available to taxonomists during the past NRSA surveys, resulting in taxonomic inconsistency in diatom datasets created during this time period. In that regard, applying a combined metric could offer a practical advantage for using existing data from past broad-scale diatom assessments since it requires less effort to make corrections to outdated synonyms, and is therefore less prone to taxonomic errors.

Further analysis is required to determine what stream conditions are reflected by the metrics included in the final MMIs, which was beyond the scope of this study. However, we can hypothesize the type of stream conditions that the metrics could reflect based on existing empirical relationships of traits to nutrient enrichment and their resistance to physical disturbance. For example, both the *high-profile % individuals/taxa* and the *planktic % taxa* metric, that is comprised of planktic diatoms that settle into the biofilm, may indicate reasonably good stream condition (Rimet and Bouchez, 2011a), that has lower conductivity (Stenger-Kovács et al., 2018), and a low level of physical disturbances (e.g., shear stress caused by water velocities) (B-Béres et al., 2016; Passy, 2007a; Stenger-Kovács et al., 2013). The *planktic % taxa* may reflect some level of sedimentation (Rimet and Bouchez, 2012a). The *tube-living colony % individuals* may also reflect good stream condition since tube-living diatoms are generally considered to be pollution sensitive (e.g., Berthon et al., 2011; Leira et al., 2009; Rumeau and Coste, 1988). Although Rimet and Bouchez (2011a) found the opposite, where tube-living diatoms increased as dissolved nutrients increased in pesticide-contaminated mesocosms, hypothesizing that the tubule, composed of exopolysaccharide matrices, may protect the cells living within from dissolved chemicals in the water. We hypothesized that genus metrics included in the MMIs such *Navicula %*

individuals, *Nitzschia % individuals*, and *Surirella % individuals* would reflect poor stream condition (e.g., from nutrient and organic pollution; Passy, 2007a).

Overall, the national MMI and all three regional MMIs performed well based on having low redundancy (i.e., maximum and mean correlation among metrics; range = 0.32–0.48 and range = 0.13–0.24, respectively), high repeatability (S/N; range = 4.3–7.8), and high responsiveness (*t*-scores; range 8.5–18.9) (Table 4). Although the sensitivity values, particularly in the Plains, seemed low (sensitivity range, 10.8–43.4%) these sensitivity values were comparable to those reported for MMIs for other biological assemblages. For example, stream MMIs for macroinvertebrates, vertebrates, and fish where sensitivity for MMIs ranged from 10 to 40% (Van Sickle, 2010); and a wetland MMI for vegetation where sensitivity was 48% (Magee et al., 2019).

Greater separation of the distributions of MMI scores between least- and most-disturbed sites within each of the 9 NRSA ecoregions in *t*-scores and boxplots (Fig. 2) clearly showed the superiority of the regionally-derived MMIs compared with the national MMI (Fig. 2; Table 5). For example, compared to the Plains MMI, the national MMI showed very poor separation between least- and most-disturbed sites within each of the Plains component ecoregions, particularly in TPL, which had the most overlap of interquartile ranges for the least- and most-disturbed sites (Fig. 2b). However, both the national MMI and Plains MMI poorly distinguished the two disturbance classes in the NPL. One explanation for the less distinct separation is that, as a group, least- and most-disturbed sites are not as distinctly different in terms of the physical/chemical environment as in the other ecoregions because the reference sites are relatively disturbed (Herlihy et al., 2008). In the West ecoregions, the West MMI had better discriminatory ability than the national MMI in both WMT and XER (Fig. 2c). However, the West MMI did not perform as well as those in the other regions. This could be because there are strong natural gradients in this region (e.g., slope and elevation) that cause variance in the MMI metrics, and can mask to some extent the changes associated with human activities, which are largely increases in fine sediment, water temperature, and nutrients (Kaufmann et al., in press). Sediment size is mostly driven by shear-stress, along with lithology and climate (Kaufmann et al., 2008); while water temperature is largely driven by climate, which is related to latitude and elevation, and nutrient concentrations are largely anthropogenic (Herlihy et al., 2020).

In the East, the national MMI was comparable to the East MMI – both MMIs showed the greatest responsiveness in this region by having the strongest ability to separate least- from most-disturbed sites among all three regions. However, the East MMI had relatively better power to discriminate, particularly in CPL and UMW, where there was a much larger separation of interquartile ranges for the least- and most-disturbed sites (Fig. 2a). Moreover, the regional MMIs generally had higher *t*-scores comparing least- and most-disturbed sites within each of the component ecoregions than did the national MMI (Table 5). Based on these comparisons of responsiveness between the national MMI and regional MMIs, we concluded that a single set of metrics in a national MMI did not perform well in all 9 ecoregions, and that a separate MMIs for the East, Plains and West would be a more effective approach for assessing the biological condition in ecoregional assessments. One explanation why the MMIs developed at the ecoregion level performed better than those

at the national level for certain ecoregions is that that natural characteristics, landscape alteration, and the response to alteration vary greatly among different ecoregions of the US (Herlihy et al., 2008; USEPA, 2006). Thus, the most discriminating set of metrics at the national scale are likely not the most discriminating at a regional scale, and sets of metrics will likely differ among regions (Stoddard et al., 2008). Indeed, we found that most of the MMI components differed across the East, Plains and West, which is possibly due to large-scale differences in natural geoclimatic gradients and anthropogenic disturbance gradients across regions. For example, the Cymbella % taxa metric that discriminates well for the Plains region only, could be an indicator of natural hardwaters with calcium-based salinity, that is fairly common in this region, particularly in the NPL. Whether these differences in MMI components among the East, Plains and West are the result of large-scale variation in environmental conditions among regions awaits further study.

An important point in the MMI construction is that we combined both wadeable and boatable sites in the development of the MMIs. The terms “wadeable” and “boatable” may give the impression that NRSA are sampling two very different habitats. However, the distinction between the NRSA wadeable and boatable protocol is not a habitat difference but a difference in location. They both are reachwide samples, sampling habitat proportional to what is present along the entire study reach. The difference is that the wadeable protocol considers the entire wetted area of the stream, while the boatable samples were collected from only the near shore littoral zone of the river. Thus, combining wadeable and boatable protocol samples may not make much difference in the periphyton collection. The macroinvertebrate sampling follows the same philosophy, and NRSA have successfully used a single macroinvertebrate MMI for rivers and streams for more than a decade (Stoddard et al., 2008; USEPA, 2020a). We evaluated the individual metrics in the East MMI and West MMI using *t*-tests for least- and most- disturbed sites in wadeable and boatable sites, separately, and found that each metric performed well in both wadeable and boatable sites, and performance of these MMI’s was not degraded by combining both site types. We were unable to evaluate whether the component metrics of the Plains MMIs performed well in boatable sites alone due to a small sample size (e.g., only 6 least-disturbed boatable sites). Given the similarity of protocols between wadeable and boatable sites, we may expect similar performances in the Plains at wadeable and boatable sites, as we found in the East and West.

3.2. Validation of regional MMIs

We evaluated the robustness of each regional MMI by assessing its performance on a separate validation data set of sites that were sampled four years later. For all three ecoregions, the distribution of MMI values for the least- and most-disturbed sites in the validation data were comparable to those for calibration data (Fig. 3). There was good separation of the two classes of sites in both the calibration and validation datasets, with even a slight improvement in the West MMI’s ability to differentiate least- from most-disturbed sites in the validation data compared to the calibration data (Fig. 3). The three regional MMIs also showed good ability to separate intermediate disturbance sites from least- and most-disturbed sites within each ecoregion for both the calibration and validation datasets (as shown by high *F*-scores from *F*-test of one-way ANOVA of the

three classes; Table A.1). Certain authors (e.g., Smol and Stoermer, 2010) have suggested that genus-level indicators do not work as reliably across larger geographic scales as do species-level indicators. In contrast, our regional-scale MMIs showed good responsiveness on both the calibration and validation data, and this consistently good performance suggests that genus-based indices may in fact be reliable indicators of river and stream ecological condition across larger spatial scales.

3.3. Regional diatom MMIs vs MMIs of other biological groups

To evaluate the responsiveness of the diatom MMIs for the East, Plains and West further, we compared their responsiveness with that of existing NRSA MMIs for macroinvertebrates and fish condition, and U.S. EPA's National Lake Assessment (NLA) MMIs for macroinvertebrates and zooplankton condition. Overall, our diatom MMI responsiveness (evaluated using *t*-scores to compare mean values of each MMI within each region between least- and most-disturbed sites) were comparable, if not better than those of NRSA and NLA MMIs based on other assemblages (Table 6.). The East diatom MMI was more responsive (*t*-score = 18) than the MMIs for stream fish and macroinvertebrates (USEPA, 2020a, 2016a) or lake zooplankton (USEPA, 2016c). Moreover, the ability of our Plains diatom MMI to discriminate least- from most-disturbed sites (*t*-score = 9.6) was comparable to stream MMIs for macroinvertebrates and fish in the Plains (*t*-scores; range = 5.7–9.7 and 5.9–9.8, respectively), and was better than the lake MMIs for macro-invertebrates and zooplankton in this region (*t*-score = 6 and 4.5, respectively). Although the West diatom MMI had a much weaker ability to discriminate (*t*-score = 8.5) than the stream MMI for macro-invertebrates (*t*-scores; range = 14.7–15.7), the responsiveness of the West diatom MMI was comparable to MMIs for stream fish (*t*-scores; range = 7.6–8.7) and lake macroinvertebrates (*t*-score = 9.4) and performed marginally better than the lake MMI for zooplankton (*t*-score = 6.3). We concluded that the MMIs developed in this study could complement existing macroinvertebrate and fish MMIs applied in NRSA surveys. Fish, macroinvertebrates, and diatoms can provide complementary information on stream ecological conditions because they differ in their response to any given anthropogenic stressor (Hering et al., 2006). For example, fish and macroinvertebrate communities are generally more sensitive to changes in channel morphology and physical habitat conditions (e.g., presence of dams or damage to the riparian corridor) while diatom communities are more sensitive to changes in water chemistry (Feio et al., 2007). The metrics included in our diatom-based MMIs may serve as good indicators of water quality degradation from multiple stressors. For example, an increase in *Nitzschia % individuals* and *Navicula % individuals* in the East MMI may indicate eutrophication and organic pollution, although these stressor-response relationships require further testing. How well our regional MMIs discriminate at relatively smaller scales, such as states, will require an evaluation on different state-level diatom datasets across the U.S. Moreover, our results demonstrate the potential application of our genus-level, trait-based MMI to lake assessment programs such as the NLA, which has not included diatoms in its bioassessments because of taxonomic inconsistencies.

3.4. Assigning traits to taxa at genus-level vs species-level

We assessed whether assigning traits to species at the genus-level rather than the species-level could introduce errors that could significantly affect the performance of each trait

metric included in the final MMIs. Based on trait information from Rimet and Bouchez (2012a; 2019), we checked that the species associated with each metric had been assigned the correct trait at the species-level. We examined the following trait metrics in the final MMIs: *Tube-living colony % individuals*, *High-profile % individuals/taxa* and *Planktic % taxa* (Table 3). We did not investigate the *N-fixers % individuals* metric because this metric includes species that belong to just two genera, *Epithemia* and *Rhopalodia*, that all possess the N-fixer trait. Therefore, all species that had been assigned the N-fixer trait at the genus-level had been assigned the correct trait at species-level.

Fig. 4 shows that the species included in the scoring of the *Tube-living colony % individuals* metric in the Plains (based on 2008–2009 NRSA data), had been assigned the correct trait at the genus-level. We found a similar pattern for the *Tube-living colony % individuals*, *High-profile % individuals*, *High-profile % taxa* and *Planktic % taxa* metrics that make up the East and West regional MMIs. Most species assigned traits at the genus-level had been assigned the correct trait at the species-level based on 2008–2009 and 2013–2014 NRSA data (Fig. A.1 and Fig. B.1, respectively). These results imply that genus-level taxonomy was sufficient to develop reliable trait metrics using the NRSA dataset. The results also show that the trait datasets are similar at both resolutions, and therefore genus-level and species-level traits could provide similar information about stream condition. Future analyses should investigate whether other trait metrics used in the screening process in the MMI development (e.g., Motile or Low-profile guilds) were also assigned the correct trait at the genus-level. This type of information could be useful, especially when developing trait-based metrics on other diatom datasets.

3.5. Advantages and limitations of a genus-level, trait-based approach

Our approach and development of a genus-level, trait-based diatom MMI yields several advantages that could benefit large-scale and long-term environmental assessment programs, where taxonomic inconsistencies in the diatom data is a major problem. The most important advantage is that the use of genus-level, trait-based metrics could substantially reduce the effects of taxonomic inconsistencies related to the analysis practices used in assessment programs. The potential for misidentifications and inconsistent taxonomy resulting from the use of references developed for European diatoms to identify North American taxa, and/or because of the use of different taxonomic references among laboratories, will be reduced at the genus-level compared with the species-level for the main reason that species-level taxonomy is more challenging. Species identifications are difficult because of the considerable diversity of diatom species relative to genera (Zampella et al., 2007). Approximately 12,000 diatom species have been described to date compared to only 1,000–1,300 diatom genera, and there are more undescribed species than there are undescribed genera (Guiry, 2012). Although there are continual changes to both species- and genus-level nomenclature that have made taxonomy more challenging and unstable (Tapolczai et al., 2016), this is likely more of an issue with species-level taxonomy because of their large diversity that has resulted in the description of numerous new species and constant taxonomic rearrangements (Rimet and Bouchez, 2012a). Moreover, genus-level identifications may be less prone to taxonomic errors because of reduced time constraints

to complete the identifications and enumerations, since determining diatoms to genus-level requires less time than to species-level (Rimet and Bouchez, 2012a).

The advantage of using traits is that the deleterious effects of taxonomic inconsistencies will likely be avoided regardless of the taxonomic resolution of the diatom data (i.e., whether traits are assigned to taxa at the genus-level as we did here, or at the species-level). Even if there are misidentified genera or species, those misidentifications will be of morphologically similar genera or species that will, for the most part, belong to the same trait category (*Frédéric Rimet*, personal communication). For example, taxa of the *Sellaphora* genus have been misidentified as taxa of the *Eolimna* genus because they are similar in shape (Wetzel et al., 2015), and since they are morphologically similar, they share some of the same traits (e.g., both *Sellaphora* and *Eolimna* taxa belong to the motile guild). Although it is important to mention that for small taxa, misidentified genera or species may not always be of morphologically similar genera or species of the same trait category. For example, if a small taxon was incorrectly assigned to *Navicula* (Motile guild), as opposed to *Achnanthes* (Low-profile guild), its guild would be incorrectly assigned. Nevertheless, studies have shown that diatom functional trait groups (e.g., ecological guilds and life-forms) are generally not affected by misidentifications that can result in disharmony among analysts and laboratories (e.g., Hajnal and Padisák, 2008). Nor are diatom traits affected by the ongoing revisions of species names since the revised species will be assigned to already defined traits (Tapolczai et al., 2017).

Another advantage of using traits and genus-level diatom data, is that by assigning all taxa to traits based on genus-level taxonomy, we include the information of all rare and undescribed taxa in our genus-based MMIs. Such taxa are usually omitted from traditional species-based indices that require precise species determinations, and stable ecological profiles of taxa based on many observations of a given species, rather than unstable profiles of rare taxa that have few data points (Berthon et al., 2011; Tapolczai et al., 2017). Indeed, Tapolczai et al. (2017) demonstrated that a species-based index, where as much as 281 rare taxa from a total of 382 taxa were omitted from the index development, was not as sensitive to a nutrient and organic matter/turbidity gradient compared to a trait-based index that included all the taxa from the same dataset. This implies that information was lost with the removal of those rare taxa and highlights the importance of including all the taxa in bioassessments. This idea is supported in studies of other assemblages; Leitão et al. (2016), for example, reported that rare taxa in stream fish assemblages were generally more sensitive to anthropogenic influences than those common enough to be included in impact assessments. The use of a trait-based approach, where all the taxa are assigned to traits, could be an advantage in countries like the U.S. where around a quarter of all species in North America are still without scientific names (Potapova and Charles, 2003), many of which are restricted to a particular geographical region (Alers-García et al., 2021). Given that the response of diatom traits is thought to be independent of ecoregions, and, hence, trait-based indices may be transferable among ecoregions with very different diatom flora (Soininen et al., 2016), our traits metrics could potentially be applied to diatoms within any ecoregion, including those undescribed taxa with restricted distributions. In this way, trait-based indices could offer an additional advantage over species-based indices, whose applicability is often spatially

restricted because they are based on regional species distributions that can vary considerably among different regions (Stenger-Kovács et al., 2007).

Our approach could also minimize the cost of the diatom analysis, given that identification to genus-level is generally quicker than to species-level, and cost depends on time spent on the identifications. Furthermore, genus-level identifications can be done by personnel with lower level of expertise (lower rate of pay). In that regard, the cost-effectiveness of our approach could benefit state and regional programs that lack the resources to conduct species-level identification of diatoms or expertise to do so but want to include diatoms as part of their bioassessment programs.

While some studies advocate the use of species-level identifications for a more precise bioassessment (e.g., Ponader and Potapova, 2007), other studies have shown that genus-level is a sufficient taxonomical level for a quick and robust assessment (e.g., Keck et al., 2016; Rimet and Bouchez, 2012a), and that improvement in bioassessment precision using determinations at species-level rather than genus-level was negligible (Rimet and Bouchez, 2012b). Kelly et al. (2009b) found that 91% of all diatom samples were assigned the same ecological status classification, regardless of whether species- or genus-level identification was used. Nevertheless, a genus-based index can have limitations. For example, environmental tolerances can vary within some genera (e. g., Potapova and Hamilton, 2007; Poulíková et al., 2008). Moreover, a genus-based index can have limitations in bioassessments at ecoregional scales. Rimet and Bouchez (2012b) found that by using species resolutions as opposed to genus resolutions, there was a stronger relationship between benthic diatom assemblages and the ecoregion classes where the diatoms were sampled. This could be a result of diatom endemism and cosmopolitanism which is largely observed at species-level, and seldom at genus-level (Rimet and Bouchez, 2012b). For that reason, Rimet and Bouchez (2012b) advocate species-level identifications for increased precision in ecoregional bioassessments, which is one of the requirements of the WFD (European Union, 2000). These are just a couple of examples of the value of species-based assessments that have long been used to define the ecological status of European fresh-waters. Given the advantages and limitations with genus- and species-based assessments highlighted here, it is important that environmental programs first consider the level of taxonomic resolution appropriate for their diatom assessment that is seldom questioned in these assessments (Rimet and Bouchez, 2011b). More specifically, whether it is necessary to use data with species-level resolution to achieve the regulatory goals for managing waters or could genus-level resolution suffice. Also, whether a species resolution is even feasible, especially when many species remain undescribed, and additional effort to document (i.e., images in voucher flora) taxa that do not fit current morphological species concepts is needed to maintain consistent taxonomy among different projects or within multi-year projects (Alers-García et al., 2021).

4. Conclusions

We showed that diatom MMIs developed using genus-level taxonomy and trait-based autecological information can be effective for large-scale and long-term assessment programs such as the NRSA, where taxonomic inconsistencies have constrained the use

of diatoms. In contrast to traditional species-level approaches, our approach can use genus-level data, which is simpler and less-expensive to obtain, and can eliminate discrepancies in species-level identification or nomenclature that can reduce confidence and defensibility of the datasets. Thus, the approach we developed requires less labor, and could enable the inclusion of diatoms in state or regional programs with limited time and financial resources. Moreover, the use of genus-level diatom indicators may allow programs such as NRSA to assess historical trends in freshwater condition by revisiting older diatom datasets, where inconsistencies in species-level identification, and recent advances in standardization of identification may have rendered historical datasets unreliable and inconsistent with current standards.

While the level of taxonomic resolution required for diatom-based assessments depends on the specific goals of individual assessment programs, our results are supported by other studies that show genus-level identification can provide a robust biotic assessment. Importantly, we do not suggest that species-based approaches should be replaced by genus-level approaches; instead, our research showed that trait-based MMI's offer a suitable alternative when reliable species-level data are not available. Some scientists have advocated only using species-level data, and that we should expand our efforts to improve taxonomic consistency across taxonomists and establish voucher specimen processes. While we agree that improving the taxonomic processing of diatoms should continue, we are not there yet. To wait for this process to be completed wastes the opportunity (with associated cost-savings) to include diatoms now in national assessments, and the possibility of using older datasets in retrospective analyses.

Based on the successful performance of our MMIs, we also demonstrated that traits assigned to European diatom taxa were transferable to taxa in North America. For future work we will assess the performance of the regional-scale MMIs applied to state-level datasets across the U.S., and will examine how the MMIs and separate metrics respond to individual environmental stressors. We could also test the utility of our trait-based approach on other aquatic systems such as lakes, where programs such as the NLA have met similar challenges because of taxonomic inconsistencies that have prevented the use of diatoms in lake assessments.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Sylvia Lee, Richard Mitchell and Janice Alers-García for their consistent interest and support that greatly assisted the research. Also, thanks to Karen Blocksom for providing data and R code, and Darin Kopp who provided feedback. We also thank the reviewers, including one anonymous reviewer for their valuable comments. The information in this document has been funded entirely by the U.S. Environmental Protection Agency, in part through appointments to the Internship/ Research Participation Program at the Office of Research and Development, U.S. Environmental Protection Agency, administered by the Oak Ridge Institute for Science and Education through an interagency agreement. The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency. Any mention of trade names, products, or services does not imply an endorsement by the U.S. Government or the U.S. Environmental Protection Agency. The EPA does not endorse any commercial products, services, or enterprises. This research

was performed while ATH held a National Research Council Senior Research Associateship award at the USEPA, Pacific Ecological Systems Division, Corvallis, Oregon.

Data availability

I have shared the link to the data and provided citations for it

References

- Alers-García J, Lee SS, Spaulding SA, 2021. Resources and practices to improve diatom data quality. *Limnol. Oceanogr. Bull* 30, 48–53. 10.1002/lob.10433. [PubMed: 34335117]
- Barbour MT, Gerritsen J, Griffith GE, Frydenborg R, McCarron E, White JS, Bastian ML, 1996. A framework for biological criteria for florida streams using benthic macroinvertebrates. *J. North Am. Benthol. Soc* 15, 185–211. 10.2307/1467948.
- B-Béres V, Török P, Kókai Z, Krasznai ET, Tóthmérész B, Bácsi I, 2014. Ecological diatom guilds are useful but not sensitive enough as indicators of extremely changing water regimes. *Hydrobiologia* 738, 191–204. 10.1007/s10750-014-1929-y.
- B-Béres V, Lukács Á, Török P, Kókai Z, Novák Z, T-Krasznai E, Tóthmérész B, Bácsi I, 2016. Combined eco-morphological functional groups are reliable indicators of colonisation processes of benthic diatom assemblages in a lowland stream. *Ecol. Indic* 64, 31–38. 10.1016/j.ecolind.2015.12.031.
- Bere T, 2016. Challenges of diatom-based biological monitoring and assessment of streams in developing countries. *Environ. Sci. Pollut. Res* 23, 5477–5486. 10.1007/s11356-015-5790-y.
- Berthon V, Bouchez A, Rimet F, 2011. Using diatom life-forms and ecological guilds to assess organic pollution and trophic level in rivers: a case study of rivers in south-eastern France. *Hydrobiologia* 673, 259–271.
- Bishop IW, Esposito RM, Tyree M, Spaulding SA, 2017. A diatom voucher flora from selected southeast rivers (USA). *Phytotaxa* 332, 101. 10.11646/phytotaxa.332.2.1.
- Blocksom KA, 2003. A performance comparison of metric scoring methods for a multimetric index for mid-Atlantic highlands streams. *Environ. Manage* 31, 670–682. 10.1007/s00267-002-2949-3. [PubMed: 12719896]
- Cao Y, Hawkins CP, Olson J, Kosterman MA, 2007. Modeling natural environmental gradients improves the accuracy and precision of diatom-based indicators. *J. North Am. Benthol. Soc* 26, 566–585. 10.1899/06-078.1.
- Chessman B, Grown IvoR, Currey J, Plunkett-Cole N, 1999. Predicting diatom communities at the genus level for the rapid biological assessment of rivers: Prediction of river diatoms. *Freshw. Biol* 41, 317–331. 10.1046/j.1365-2427.1999.00433.x.
- Coste M, Boutry S, Tison-Rosebery J, Delmas F, 2009. Improvements of the Biological Diatom Index (BDI): Description and efficiency of the new version (BDI-2006). *Ecol. Indic* 9, 621–650. 10.1016/j.ecolind.2008.06.003.
- Coste M, 1982. Étude des méthodes biologiques d'appréciation quantitative de la qualité des eaux. Rapport Cemagref QE Lyon-AF Bassin Rhône Méditerranée Corse
- Dixit SS, Smol JP, Kingston JC, Charles DF, 1992. Diatoms: powerful indicators of environmental change. *Environ. Sci. Technol* 26, 22–33. 10.1021/es00025a002.
- European Union, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. The European Parliament and the Council of the European Union. Official Journal of the European Communities L 327/1:1–72.
- Feio MJ, Almeida SFP, Craveiro SC, Calado AJ, 2007. Diatoms and macroinvertebrates provide consistent and complementary information on environmental quality. *Fundam. Appl. Limnol. Arch. Für Hydrobiol* 169, 247–258. 10.1127/1863-9135/2007/0169-0247.
- Guiry MD, 2012. How many species of algae are there? *J. Phycol* 48, 1057–1063. 10.1111/j.1529-8817.2012.01222.x. [PubMed: 27011267]

- Hajnal É, Padisák J, 2008. Analysis of long-term ecological status of Lake Balaton based on the ALMOBAL phytoplankton database. *Hydrobiologia* 599, 227–237. 10.1007/s10750-007-9207-x.
- Hering D, Johnson RK, Kramm S, Schmutz S, Szoszkiewicz K, Verdonschot PFM, 2006. Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: a comparative metric-based analysis of organism response to stress. *Freshw. Biol* 51, 1757–1785. 10.1111/j.1365-2427.2006.01610.x.
- Herlihy AT, Paulsen SG, Sickle JV, Stoddard JL, Hawkins CP, Yuan LL, 2008. Striving for consistency in a national assessment: the challenges of applying a reference-condition approach at a continental scale. *J. North Am. Benthol. Soc* 27, 860–877. 10.1899/08-081.1.
- Herlihy AT, Sifneos JC, Hughes RM, Peck DV, Mitchell RM, 2020. The relation of lotic fish and benthic macroinvertebrate condition indices to environmental factors across the conterminous USA. *Ecol. Indic* 112, 105958 10.1016/j.ecolind.2019.105958.
- Hill BH, Herlihy AT, Kaufmann PR, Stevenson RJ, McCormick FH, Johnson CB, 2000. Use of periphyton assemblage data as an index of biotic integrity. *J. North Am. Benthol. Soc* 19, 50–67. 10.2307/1468281.
- Hill BH, Stevenson RJ, Pan Y, Herlihy AT, Kaufmann PR, Johnson CB, 2001. Comparison of correlations between environmental characteristics and stream diatom assemblages characterized at genus and species levels. *J. North Am. Benthol. Soc* 20, 299–310. 10.2307/1468324.
- Hill BH, Herlihy AT, Kaufmann PR, DeCelles SJ, Vander Borgh MA, 2003. Assessment of streams of the eastern United States using a periphyton index of biotic integrity. *Ecol. Indic* 2, 325–338. 10.1016/S1470-160X(02)00062-6.
- Kahlert M, Albert R-L, Anttila E-L, Bengtsson R, Bigler C, Eskola T, Gälman V, Gottschalk S, Herlitz E, Jarlman A, Kasperoviciene J, Kokoci ski M, Luup H, Miettinen J, Paunksnyte I, Piirsoo K, Quintana I, Raunio J, Sandell B, Simola H, Sundberg I, Vilbaste S, Weckström J, 2009. Harmonization is more important than experience—results of the first Nordic-Baltic diatom intercalibration exercise 2007 (stream monitoring). *J. Appl. Phycol* 21, 471–482. 10.1007/s10811-008-9394-5.
- Kahlert M, Kelly M, Albert R-L, Almeida SFP, Bešta T, Blanco S, Coste M, Denys L, Ector L, Fránková M, Hlúbiková D, Ivanov P, Kennedy B, Marvan P, Mertens A, Miettinen J, Picinska-Faltynowicz J, Rosebery J, Tornés E, Vilbaste S, Vogel A, 2012. Identification versus counting protocols as sources of uncertainty in diatom-based ecological status assessments. *Hydrobiologia* 695, 109–124. 10.1007/s10750-012-1115-z.
- Kahlert M, Ács É, Almeida SFP, Blanco S, Dreßler M, Ector L, Karjalainen SM, Liess A, Mertens A, van der Wal J, Vilbaste S, Werner P, 2016. Quality assurance of diatom counts in Europe: towards harmonized datasets. *Hydrobiologia* 772, 1–14. 10.1007/s10750-016-2651-8.
- Kaufmann PR, Levine P, Robison EG, Seeliger C, Peck DV, 1999. Quantifying physical habitat in wadable streams U.S. Environmental Protection Agency. EPA/620/R_99/003. Washington, D.C.
- Kaufmann PR, Hughes RM, Paulsen SG, Peck DV, Seeliger CW, Weber MH, Mitchell RM, In press. Physical habitat in conterminous US streams and rivers, Part 1: Geoclimatic controls and anthropogenic alteration. *Ecological Indicators*, XXX (XXXX) 109046.
- Kaufmann PR, Faustini JM, Larsen DP, Shirazi MA, 2008. A roughness-corrected index of relative bed stability for regional stream surveys. *Geomorphology* 99, 150–170. 10.1016/j.geomorph.2007.10.007.
- Keck F, Rimet F, Franc A, Bouchez A, 2016. Phylogenetic signal in diatom ecology: perspectives for aquatic ecosystems biomonitoring. *Ecol. Appl* 26, 861–872. 10.1890/14-1966. [PubMed: 27411256]
- Kelly M, Bennett C, Coste M, Delgado C, Delmas F, Denys L, Ector L, Fauville C, Ferréol M, Golub M, Jarlman A, Kahlert M, Lucey J, NiChatháin B, Pardo I, Pfister P, Picinska-Faltynowicz J, Rosebery J, Schranz C, Schaumburg J, van Dam H, Vilbaste S, 2009a. A comparison of national approaches to setting ecological status boundaries in phytobenthos assessment for the European Water Framework Directive: results of an intercalibration exercise. *Hydrobiologia* 621, 169–182. 10.1007/s10750-008-9641-4.
- Kelly M, King L, NiChatháin B, 2009b. The conceptual basis of ecological status assessments using diatoms. *Biol. Environ. Proc. R. Ir. Acad* 109, 175–189. 10.3318/BIOE.2009.109.3.175.

- Kelly MG, Whitton BA, 1995. The Trophic Diatom Index: a new index for monitoring eutrophication in rivers. *J. Appl. Phycol* 7, 433–444. 10.1007/BF00003802.
- Kilgour BW, Somers KM, Matthews DE, 1998. Using the normal range as a criterion for ecological significance in environmental monitoring and assessment. *Écoscience* 5, 542–550. 10.1080/11956860.1998.11682485.
- Kociolek JP, Spaulding SA, 2000. Freshwater diatom biogeography. *Nova Hedwig* 71, 223–241. 10.1127/nova/71/2000/223.
- Lange K, Townsend CR, Matthaei CD, 2016. A trait-based framework for stream algal communities. *Ecol. Evol* 6, 23–36. 10.1002/ece3.1822. [PubMed: 26811771]
- Lavoie I, Campeau S, Grenier M, Dillon PJ, 2006. A diatom-based index for the biological assessment of eastern Canadian rivers: an application of correspondence analysis (CA). *Can. J. Fish. Aquat. Sci* 63, 1793–1811. 10.1139/f06-084.
- Lavoie I, Lento J, Morin A, 2010. Inadequacy of size distributions of stream benthic diatoms for environmental monitoring. *J. North Am. Benthol. Soc* 29, 586–601. 10.1899/09-062.1.
- Lee SS, Bishop IW, Spaulding SA, Mitchell RM, Yuan LL, 2019. Taxonomic harmonization may reveal a stronger association between diatom assemblages and total phosphorus in large datasets. *Ecol. Indic* 102, 166–174. 10.1016/j.ecolind.2019.01.061. [PubMed: 32802000]
- Leira M, Chen G, Dalton C, Irvine K, Taylor D, 2009. Patterns in freshwater diatom taxonomic distinctness along an eutrophication gradient. *Freshw. Biol* 54, 1–14. 10.1111/j.1365-2427.2008.02086.x.
- Leitão RP, Zuanon J, Villéger S, Williams SE, Baraloto C, Fortunel C, Mendonça FP, Mouillot D, 2016. Rare species contribute disproportionately to the functional structure of species assemblages. *Proc. R. Soc. B Biol. Sci* 283, 20160084. 10.1098/rspb.2016.0084.
- Magee TK, Blocksom KA, Fennessy MS, 2019. A national-scale vegetation multimetric index (VMMI) as an indicator of wetland condition across the conterminous United States. *Environ. Monit. Assess* 191, 322. 10.1007/s10661-019-7324-4. [PubMed: 31222469]
- Mazor RD, Rehn AC, Ode PR, Engeln M, Schiff KC, Stein ED, Gillett DJ, Herbst DB, Hawkins CP, 2016. Bioassessment in complex environments: designing an index for consistent meaning in different settings. *Freshw. Sci* 35, 249–271. 10.1086/684130.
- Omernik JM, 1987. Ecoregions of the Conterminous United States. *Ann. Assoc. Am. Geogr* 77, 118–125. 10.1111/j.1467-8306.1987.tb00149.x.
- Pan Y, Stevenson RJ, Hill BH, Herlihy AT, Collins GB, 1996. Using Diatoms as Indicators of Ecological Conditions in Lotic Systems: A Regional Assessment. *J. North Am. Benthol. Soc* 15, 481–495. 10.2307/1467800.
- Passy SI, 2007a. Diatom ecological guilds display distinct and predictable behavior along nutrient and disturbance gradients in running waters. *Aquat. Bot* 86, 171–178. 10.1016/j.aquabot.2006.09.018.
- Passy SI, 2007b. Community analysis in stream biomonitoring: what we measure and what we don't. *Environ. Monit. Assess* 127, 409–417. 10.1007/s10661-006-9290-x. [PubMed: 16964529]
- Ponader KC, Potapova MG, 2007. Diatoms from the genus *Achnantheidium* in flowing waters of the Appalachian Mountains (North America): Ecology, distribution and taxonomic notes. *Limnologia* 37, 227–241. 10.1016/j.limno.2007.01.004.
- Potapova M, Charles DF, 2003. Distribution of benthic diatoms in US rivers in relation to conductivity and ionic composition. *Freshw. Biol* 48, 1311–1328.
- Potapova M, Charles DF, 2007. Diatom metrics for monitoring eutrophication in rivers of the United States. *Ecol. Indic* 7, 48–70. 10.1016/j.ecolind.2005.10.001.
- Potapova M, Hamilton PB, 2007. Morphological and ecological variation within the *Achnantheidium minutissimum* (Bacillariophyceae) species complex. *J. Phycol* 43, 561–575. 10.1111/j.1529-8817.2007.00332.x.
- Pouřková A, Špačková J, Kelly MG, Duchoslav M, Mann DG, 2008. Ecological variation within *Sellaphora* species complexes (Bacillariophyceae): specialists or generalists? *Hydrobiologia* 614, 373–386. 10.1007/s10750-008-9521-y.
- Prechtl J, 2004. Intracellular Spheroid Bodies of *Rhopalodia gibba* Have Nitrogen-Fixing Apparatus of Cyanobacterial Origin. *Mol. Biol. Evol* 21, 1477–1481. 10.1093/molbev/msh086. [PubMed: 14963089]

- Prygiel J, Coste M, 1998. Mise au point de l'Indice Biologique Diatomée, un indice diatomique pratique applicable au réseau hydrographique français. *L'Eau, l'industrie, les nuisances* 40–45.
- R Development Core Team, 2017. *R: A language and environment for statistical computing* R Foundation for Statistical Computing. Vienna, Austria.
- Riato L, Della Bella V, Leira M, Taylor JC, Oberholster PJ, 2017. A diatom functional-based approach to assess changing environmental conditions in temporary depressional wetlands. *Ecol. Indic* 78, 205–213. 10.1016/j.ecolind.2017.03.018.
- Rimet F, Bouchez A, 2011b. Taxonomic resolution of river diatoms: structuring impact of environmental parameters and implications for biomonitoring. in: 7th Symposium for European Freshwater Sciences, Catalan Institute for Water Research (ICRA) . Girona, Spain pp.2. fihal-02806457.
- Rimet F, Bouchez A, 2019. Life-forms, cell-sizes and ecological guilds of diatoms in European rivers doi:10.15454/XLQ40G.
- Rimet F, Bouchez A, 2011a. Use of diatom life-forms and ecological guilds to assess pesticide contamination in rivers: lotic mesocosm approaches. *Ecol. Indic* 11, 489–499.
- Rimet F, Bouchez A, 2012a. Life-forms, cell-sizes and ecological guilds of diatoms in European rivers. *Knowl. Manag. Aquat. Ecosyst* 01 10.1051/kmae/2012018.
- Rimet F, Bouchez A, 2012b. Biomonitoring river diatoms: Implications of taxonomic resolution. *Ecol. Indic* 15, 92–99. 10.1016/j.ecolind.2011.09.014.
- Rumeau A, Coste M, 1988. Initiation à la systématique des diatomées d'eau douce pour l'utilisation pratique d'un indice diatomique générique. *Bull. Fr. Peche Piscic* 309, 1–69.
- Smol JP, Stoermer EF (Eds.), 2010. *The diatoms: applications for the environmental and earth sciences*, 2nd ed. Cambridge University Press, Cambridge; New York.
- Soininen J, Jamoneau A, Rosebery J, Passy SI, 2016. Global patterns and drivers of species and trait composition in diatoms: Global compositional patterns in stream diatoms. *Glob. Ecol. Biogeogr* 25, 940–950. 10.1111/geb.12452.
- Stenger-Kovács C, Buczkó K, Hajnal É, Padišák J, 2007. Epiphytic, littoral diatoms as bioindicators of shallow lake trophic status: Trophic Diatom Index for Lakes (TDIL) developed in Hungary. *Hydrobiologia* 589, 141–154. 10.1007/s10750-007-0729-z.
- Stenger-Kovács C, Lengyel E, Crossetti LO, Üveges V, Padišák J, 2013. Diatom ecological guilds as indicators of temporally changing stressors and disturbances in the small Torna-stream, Hungary. *Ecol. Indic* 24, 138–147. 10.1016/j.ecolind.2012.06.003.
- Stenger-Kovács C, Körmendi K, Lengyel E, Abonyi A, Hajnal É, Szabó B, Buczkó K, Padišák J, 2018. Expanding the trait-based concept of benthic diatoms: Development of trait- and species-based indices for conductivity as the master variable of ecological status in continental saline lakes. *Ecol. Indic* 95, 63–74. 10.1016/j.ecolind.2018.07.026.
- Stevenson RJ, Zalack JT, Wolin J, 2013. A multimetric index of lake diatom condition based on surface-sediment assemblages. *Freshw. Sci* 32, 1005–1025. 10.1899/12-183.1.
- Stoddard JL, Herlihy AT, Peck DV, Hughes RM, Whittier TR, Tarquinio E, 2008. A process for creating multimetric indices for large-scale aquatic surveys. *J. North Am. Benthol. Soc* 27, 878–891. 10.1899/08-053.1.
- Straile D, Jochimsen MC, Kümmerlin R, 2013. The use of long-term monitoring data for studies of planktonic diversity: a cautionary tale from two Swiss lakes. *Freshw. Biol* 58, 1292–1301. 10.1111/fwb.12118.
- Tapolczai K, Bouchez A, Stenger-Kovács C, Padišák J, Rimet F, 2016. Trait-based ecological classifications for benthic algae: review and perspectives. *Hydrobiologia* 10.1007/s10750-016-2736-4.
- Tapolczai K, Bouchez A, Stenger-Kovács C, Padišák J, Rimet F, 2017. Taxonomy- or trait-based ecological assessment for tropical rivers? Case study on benthic diatoms in Mayotte island (France, Indian Ocean). *Sci. Total Environ* 607–608, 1293–1303. 10.1016/j.scitotenv.2017.07.093.
- Tyree MA, Bishop IW, Hawkins CP, Mitchell R, Spaulding SA, 2020. Reduction of taxonomic bias in diatom species data. *Limnol. Oceanogr. Methods* 18, 271–279. 10.1002/lom3.10350. [PubMed: 34025297]

- USEPA, 2006. U.S. Environmental Protection Agency. Wadeable Streams Assessment: a collaborative survey of the nation's streams EPA 841-B-06-002. Washington, DC.
- USEPA, 2008. U.S. Environmental Protection Agency. National Rivers and Streams Assessment: Laboratory Methods Manual EPA 841-B-07-010. Washington, DC.
- USEPA, 2013a. U.S. Environmental Protection Agency. National Rivers and Streams Assessment 2013–2014: Field Operations Manual – Non-Wadeable EPA-841-B-12-009a. Washington, DC.
- USEPA, 2013b. U.S. Environmental Protection Agency. National Rivers and Streams Assessment 2013–2014: Field Operations Manual – Wadeable EPA-841-B-12-009b. Washington, DC.
- USEPA, 2016a. U.S. Environmental Protection Agency. National rivers and streams assessment 2008–2009: technical report EPA/841/R-16/008. Washington, D.C.
- USEPA, 2016b. U.S. Environmental Protection Agency. National Aquatic Resource Surveys. [National Rivers and Streams Assessment 2008–2009] (data and metadata files) Available from U.S. EPA web page: <https://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys>.
- USEPA, 2016c. National Lakes Assessment 2012: A Collaborative Survey of Lakes in the United States EPA 841-R-16-113. Washington, DC.
- USEPA, 2020a. U.S. Environmental Protection Agency. National Rivers and Streams Assessment 2013–2014: A Collaborative Survey EPA 841-R-19-001. Washington, DC.
- USEPA, 2020b. U.S. Environmental Protection Agency. National Aquatic Resource Surveys. [National Rivers and Streams Assessment 2013–2014] (data and metadata files) Available from U.S. EPA web page: <https://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys>.
- Van Sickle J, 2010. Correlated Metrics Yield Multimetric Indices with Inferior Performance. *Trans. Am. Fish. Soc* 139, 1802–1817.
- Verberk WCEP, van Noordwijk CGE, Hildrew AG, 2013. Delivering on a promise: integrating species traits to transform descriptive community ecology into a predictive science. *Freshw. Sci* 32, 531–547. 10.1899/12-092.1.
- Werner P, Adler S, Dreßler M, 2016. Effects of counting variances on water quality assessments: implications from four benthic diatom samples, each counted by 40 diatomists. *J. Appl. Phycol* 28, 2287–2297. 10.1007/s10811-015-0760-9.
- Wetzel CE, Ector L, Van de Vijver B, Compère P, Mann DG, 2015. Morphology, typification and critical analysis of some ecologically important small naviculoid species (Bacillariophyta). *Fottea* 15, 203–234. 10.5507/fot.2015.020.
- Whitton BA, Kelly MG, 1995. Use of algae and other plants for monitoring rivers. *Austral Ecol* 20, 45–56. 10.1111/j.1442-9993.1995.tb00521.x.
- Zampella RA, Laidig KJ, Lowe RL, 2007. Distribution of Diatoms in Relation to Land Use and pH in Blackwater Coastal Plain Streams. *Environ. Manage* 39, 369–384. 10.1007/s00267-006-0041-0. [PubMed: 17219257]

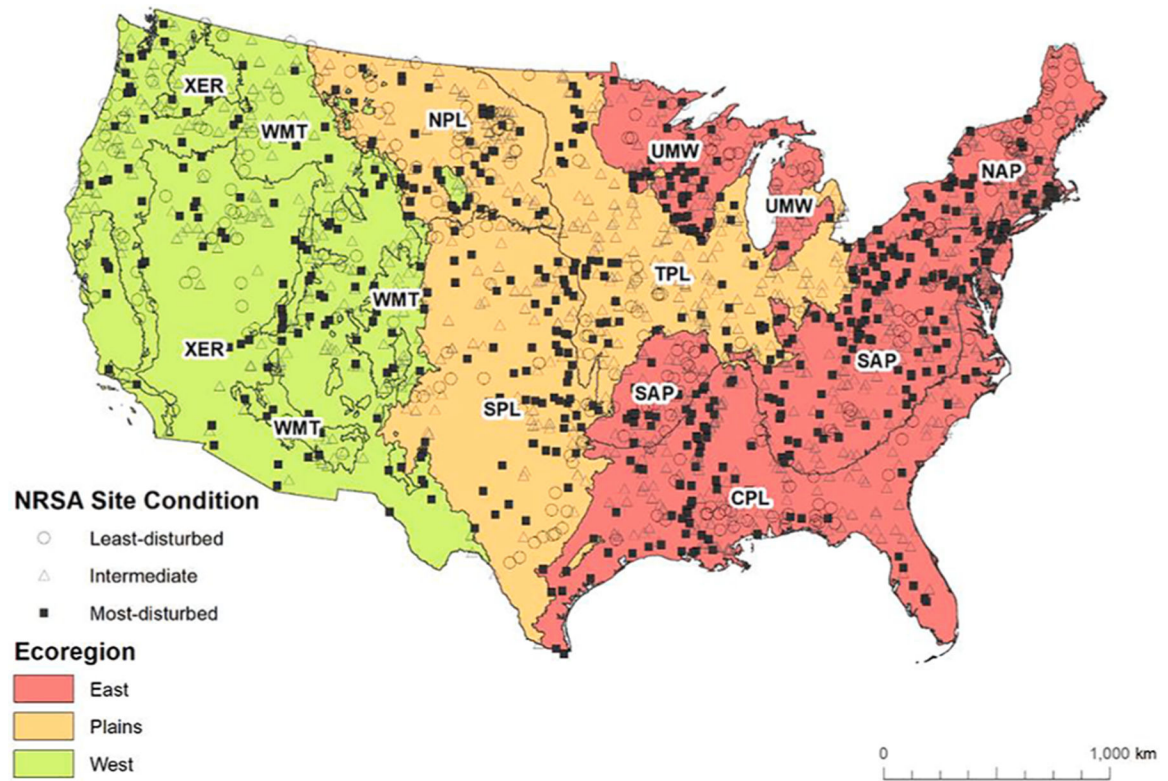


Fig. 1. Location of 2008–2009 sample sites within three large ecoregions used by NRSAs – the East, Plains and West. The 9 finer-scale ecoregions across all three regions are Northern Appalachians (NAP), Southern Appalachians (SAP), Coastal Plains (CPL), Upper Midwest (UMW), Temperate Plains (TPL), Northern Plains (NPL), Southern Plains (SPL), Western Mountains (WMT) and Xeric (XER).

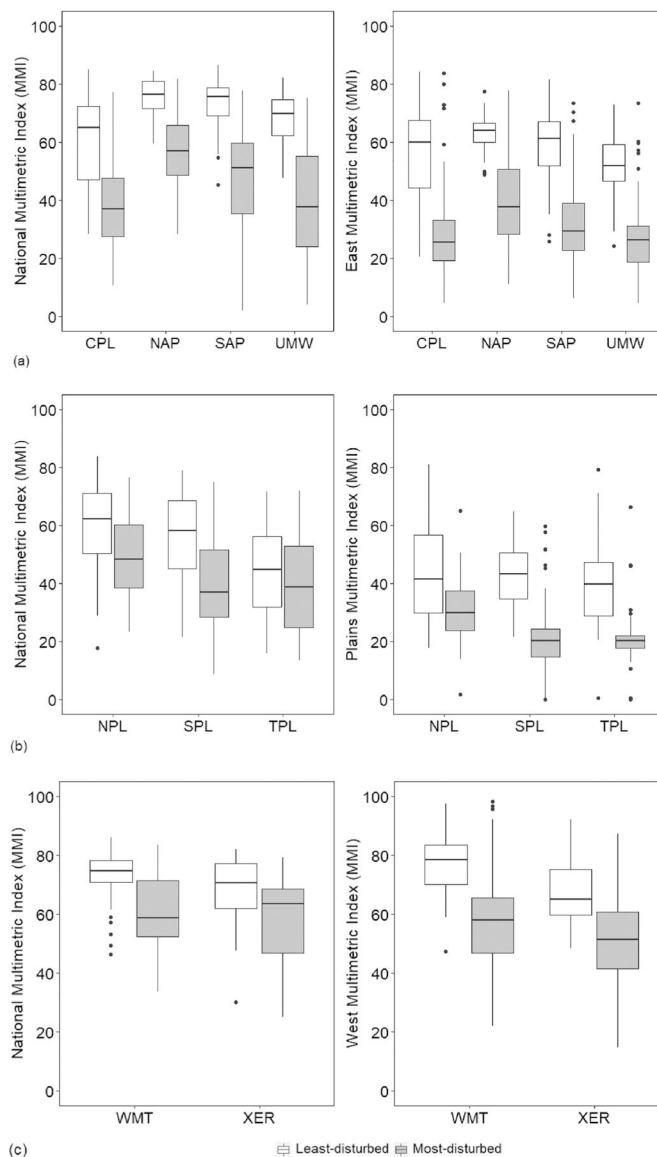


Fig. 2. Boxplots comparing the discrimination between least- and most-disturbed site scores for the national MMI and ecoregion-specific MMIs within each of the component ecoregions (a) the East, (b) the Plains, and (c) the West, based on the calibration data (2008–2009 NRSA). Refer to Table 1 for a description of the x-axis ecoregion acronyms. Horizontal lines represent the medians, boxes represent the interquartile ranges (25th and 75th percentiles), whiskers represent 5th and 95th percentiles, and the dots represent outliers.

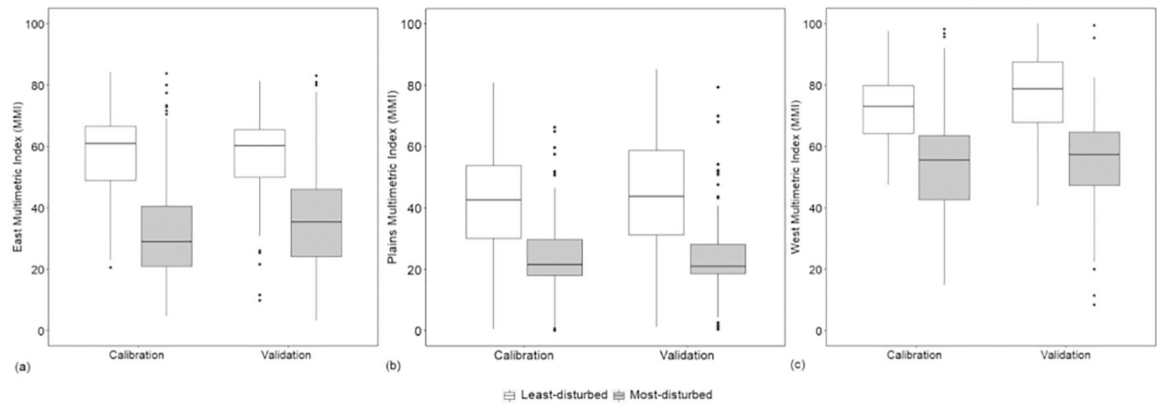


Fig. 3. Boxplots comparing (a) East MMI, (b) Plains MMI, and (c) West MMI scores between least- and most-disturbed sites for the validation data (2013–2014 NRSA) relative to the calibration data (2008–2009 NRSA). Horizontal lines represent the medians, boxes represent the interquartile ranges (25th and 75th percentiles), whiskers represent 5th and 95th percentiles, and the dots represent outliers.

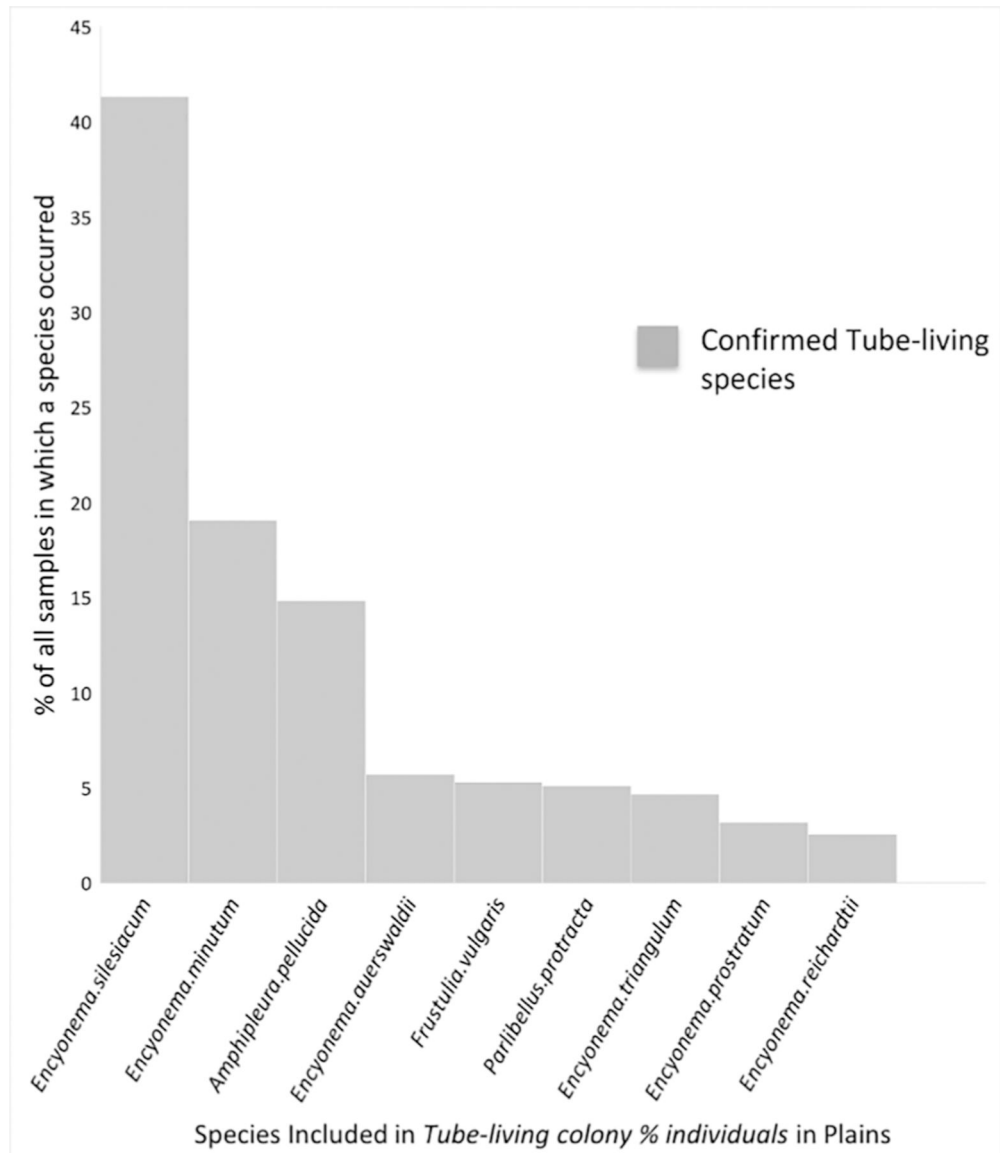


Fig. 4. Probabilities of occurrence for species included in the scoring of the Tube-living colony % individuals metric in the Plains based on 2008–2009 NRSA data. Only species present in 1% of sites are shown. Based on trait information at the species-level (Rimet and Bouchez, 2012a; 2019), species represented by the light grey bars were assigned the correct trait at the genus-level.

Table 1

Description of sample sites used to develop the MMIs based on the calibration dataset (2008–2009 NRSA) and evaluate the MMIs using the validation dataset (2013–2014 NRSA). Samples collected from these sites had at least 400 diatom valves.

MMI	Sites per ecoregion	Calibration dataset						Validation dataset					
		Total	LD	ID	MD	Revisits	Total	LD	ID	MD	Revisits		
National	All sites	2121	341	1167	613	165	1861	274	1017	570	154		
East	All Eastern sites	1075	154	594	327	96	942	114	515	313	98		
	Northern Appalachians (NAP)	242	44	133	65	31	298	34	183	81	31		
Plains	Southern Appalachians (SAP)	354	40	177	137	33	288	22	154	112	36		
	Coastal Plain (CPL)	297	39	183	75	20	220	36	112	72	25		
Plains	Upper Midwest (UMW)	182	31	101	50	12	136	22	66	48	6		
	All Plains sites	596	108	328	160	36	501	97	277	127	30		
West	Temperate Plains (TPL)	237	22	169	46	18	243	19	163	61	20		
	Northern Plains (NPL)	179	50	84	45	10	147	56	67	24	7		
West	Southern Plains (SPL)	180	36	75	69	8	111	22	47	42	3		
	All Western sites	450	79	245	126	33	418	63	225	130	26		
West	Western Mountains (WMT)	231	41	129	61	12	275	44	154	77	13		
	Xeric Region (XER)	219	38	116	65	21	143	19	71	53	13		

LD, least-disturbed sites; ID, intermediate disturbance; MD, most-disturbed sites.

Table 2

Assignment of traits to diatom taxa in the NRSA 2008–2009 and 2013–2014 datasets using genus-level taxonomy.

Traits	Definition of trait classification	Typical genus examples
Functional		
<i>Life-forms</i>		
Mobile	Free moving e.g., some species vertically migrate into the sediments to acquire nutrients or through sediments to reach light	<i>Achnanthes</i> , <i>Amphora</i> , <i>Nitzschia</i>
Tube-living colony	Species that form tubes inside which frustules can move	<i>Amphipleura</i> , <i>Encyonema</i> , <i>Frustulia</i>
Colonial	Species are attached by spines, stalks, or mucilage at their poles	<i>Aulacoseira</i> , <i>Eunotia</i> , <i>Tabellaria</i>
Non-colonial	Species are not attached; can be floating or free-moving	<i>Achnanthes</i> , <i>Navicula sensu lato</i> , <i>Nitzschia</i>
Filament colony	Species linked by spines	<i>Aulacoseira</i> , <i>Skeletonema</i>
Zig-zag colony	Pennate species connected by mucilage at their neighbor poles	<i>Diatoma</i> , <i>Tabellaria</i>
Stellate colony	Pennate species connected by mucilage at their opposed poles	<i>Asterionella</i> , <i>Staurosira</i>
Arbuscular colony	Species produce stalks at either pole; stalks diverge from each diatom frustule to form branching colonies	<i>Cymbella</i> , <i>Rhoicosphenia</i>
Rosette colony	Species attached to substrate by a short stalk at one pole; colonies look fan-shaped	<i>Fragilaria</i> , <i>Ulnaria</i>
Ribbon colony	Species attached to one another either by interlocking spines or by mucilage excretions on their valve face, forming long, ribbon-like colonies	<i>Eunotia</i> , <i>Fragilaria</i> , <i>Staurosira</i>
Pedunculate	Species grows upright to substrate, attached by a mucilage pad or stalk	<i>Achnanthes</i> , <i>Gomphonema</i> , <i>Reimeria</i>
Pad	Species grows upright to substrate, attached by a mucilage pad	<i>Fragilaria</i> , <i>Karayevia</i> , <i>Ulnaria</i>
Stalk	Species produce a stalk that sticks to the substrate. The stalk can be simple that is linked to one cell or can be branching stalks linked to several cells	<i>Gomphonema</i> , <i>Psammothidium</i> , <i>Suirella</i>
Adnate	Species grows parallel to substrate, attached by their valve face	<i>Amphora</i> , <i>Cocconeis</i> , <i>Rhopalodia</i>
<i>Ecological guilds</i>		
High-profile	Large species, or those that tend to form colonies such as filamentous, branched, chain-forming and tube-living	<i>Diatoma</i> , <i>Eunotia</i> , <i>Gomphonema</i>
Low-profile	Species of short stature, including prostrate, adnate, erect, and slow-moving species	<i>Achnantheidium</i> <i>Achnanthes</i> , <i>Amphora</i>
Motile	Fast-moving species	<i>Epithemia</i> , <i>Navicula sensu lato</i> , <i>Suirella</i>
Planktic	Solitary or colonial species that live in the water column or unattached to substrates	<i>Aulacoseira</i> , <i>Cyclotella</i> , <i>Stephanodiscus</i>
Physiological N-fixers	Atmospheric nitrogen-fixing species	<i>Epithemia</i> , <i>Rhopalodia</i>

Table 3

Final metrics included in the best performing 4-metric national MMI, and 5-metric East MMI, Plains MMI, and West MMI based on the calibration dataset (2008–2009 NRSA). For each metric, positive or negative (+/–) expected responses to anthropogenic disturbances are shown based on distributions of metric scores between least- and most-disturbed sites, and existing literature.

Metric	National MMI	East MMI	Plains MMI	West MMI	Response
Traits					
Tube-living colony % individuals		X	X		–
High-profile % individuals	X	X			–
High-profile % taxa				X	–
Low-profile % individuals	X				+
Planktic % taxa	X			X	+
N-fixers % individuals			X		–
Motile-S1 % individuals	X				+
Genera					
<i>Achnanthydium</i> + <i>Achnanthes</i> % individuals		X	X	X	–
<i>Cyclotella</i> % individuals			X		+
<i>Cymbella</i> % taxa			X		–
<i>Navicula</i> % individuals		X			+
<i>Nitzschia</i> % individuals		X			+
<i>Melosira</i> % individuals				X	+
<i>Surirella</i> % individuals				X	+

Table 4

Performance statistics for the final national MMI, and the MMI for the East, Plains, and West regions, based on the calibration data (2008–2009 NRSA).

MMI	REDUNDANCY		SENSITIVITY		REPEATABILITY		PRECISION		RESPONSIVENESS	
	Max <i>r</i> among metrics	Mean <i>r</i> among metrics	% of MD sites identified as impacted	S/N	SD among LD sites	<i>t</i> -score between mean LD & MD sites				
National	0.42	0.21	24.4	7.8	14.8	16.4				
East	0.45	0.24	43.4	6.7	12.3	18.9				
Plains	0.48	0.20	10.8	5.7	13.7	10.0				
West	0.32	0.13	31	4.3	11.5	8.5				

r, Pearson correlation coefficient; LD, least-disturbed sites; MD, most-disturbed sites; S/N, signal/noise ratio; SD, standard deviation.

Table 5

t-scores to compare the discrimination between mean values of least-disturbed (LD) and most-disturbed (MD) sites for the national MMI and ecoregion-specific MMIs within each of the component ecoregions based on the calibration dataset. Refer to Table 1 for the number of LD and MD sites within each ecoregion applied in the *t*-tests. The higher the *t*-score the better the discrimination.

Region	<i>t</i> -scores			
	National MMI	East MMI	Plains MMI	West MMI
Northern Appalachians (NAP)	8.3	9.6	–	–
Southern Appalachians (SAP)	9.1	11	–	–
Coastal Plain (CPL)	6.7	7.9	–	–
Upper Midwest (UMW)	7.9	7.9	–	–
Temperate Plains (TPL)	1.6	–	3.3	–
Northern Plains (NPL)	3.3	–	4.2	–
Southern Plains (SPL)	4.9	–	8.3	–
Western Mountains (WMT)	5.1	–	–	5.8
Xeric Region (XER)	3.3	–	–	6.6

Table 6

Comparison of responsiveness between the diatom MMIs developed for the East, Plains, and West regions in this study, and existing NRSA and NLA MMIs within each of the three regions. Performance is based on the responsiveness of the MMI using *t*-scores to compare mean values between least-disturbed (LD) and most-disturbed (MD) sites; the higher the *t*-score, the greater the MMI responsiveness. Note, the range of *t*-scores represents the minimum and maximum *t*-scores associated with each component ecoregion for a given region, while a single *t*-score represents all sites across all ecoregions for a region.

RESPONSIVENESS (<i>t</i>-score between mean LD & MD sites)						
MMI Type	River and Stream Diatoms	Stream Macroinvertebrates	Stream Fish	Lake Macroinvertebrates	Lake Zooplankton	
Region						
East	18	10.1–14.7	7.7–15.4	7.3–8.5		4.9–5.8
Plains	9.6	5.7–9.7	5.9–9.8	6.0		4.5
West	8.5	14.7–15.7	7.6–8.7	9.4		6.3