NCEE Working Paper

Valuing Aquatic Ecosystem Health at a National Scale: Modeling Biological Indicators Across Space and Time

Ryan Hill, Chris Moore, Jessie Doyle, Scott G. Leibowitz, Paul Ringold, and Brenda Rashleigh

Working Paper 20-04
November, 2020

U.S. Environmental Protection Agency
National Center for Environmental Economics
https://www.epa.gov/environmental-economics

NCEE
NATIONAL CENTER FOR ENVIRONMENTAL ECONOMICS

VALUING AQUATIC ECOSYSTEM HEALTH AT A NATIONAL SCALE: MODELING BIOLOGICAL INDICATORS ACROSS SPACE AND TIME

Ryan Hill, Chris Moore, Jessie Doyle, Scott G. Leibowitz, Paul Ringold, Brenda Rashleigh

ABSTRACT: The US Environmental Protection Agency (EPA) uses a water quality index (WQI) to estimate the benefits of proposed regulations. However, the existing WQI focuses mainly on metrics related to human use values, such as recreation, and fails to capture aspects important to nonuse values of aquatic ecosystems, such as existence values. Here, we identify an appropriate index of biological health for use in stated preference (SP) surveys that seek to quantify the nonuse value of streams and lakes anywhere within the conterminous US (CONUS). We used a literature review and focus groups to evaluate two aquatic indices that are regularly reported by the EPA's National Aquatic Resources Surveys: (1) multimetric indices (MMIs) and (2) the observed-to-expected ratio of taxonomic composition (O/E). Focus group participants had difficulty interpreting the meaning of a hypothetical 5-point change in MMI values on a 100-point scale in response to changes in water or habitat quality. This difficulty arose because a 5-point change can occur due to many unique combinations of the individual metrics that compose an MMI. In contrast, participants found it easier to interpret loss in native taxa (O/E) as an index of biological condition. We chose the O/E index because of this superior interpretability when assessed against MMIs. In addition to index selection, we modeled and interpolated the values of O/E to 1.1 million stream segments and 297,071 lakes across the CONUS to provide data for SP studies at any scope or scale, from local watersheds to the entire lower 48 states. As part of this effort, we also modeled and interpolated the areas of streams

(m^2) to place them in the same unit as lakes to describe the quantity of resources affected by policy scenarios. Focus groups found comparisons of management scenarios easier to interpret when aquatic resources were placed into the same units and especially when presented as percentages of area. Finally, we discuss future work to link O/E with water quality and habitat models that will allow us to forecast changes in the metric resulting from regulatory action.

KEYWORDS: Existence value, ecosystem health, stated preference, random forest modeling

JEL CODES: Q51, Q57

DISCLAIMER

VALUING AQUATIC ECOSYSTEM HEALTH AT A NATIONAL SCALE: MODELING BIOLOGICAL INDICATORS ACROSS SPACE AND TIME

Ryan Hill[1], Chris Moore[2†], Jessie Doyle[3], Scott G. Leibowitz[1], Paul Ringold[1,] Brenda Rashleigh[4]

[1] US EPA, Center for Public Health and Environmental Assessment, Pacific Ecological Systems Division, Corvallis, Oregon

[2] US EPA, National Center for Environmental Economics, Washington, DC

[3] Oak Ridge Institute for Science and Education Participant c/o US EPA, CPHEA, PESD, Corvallis, Oregon

[4] US EPA, Office of Research and Development, Narragansett, Rhode Island

[†] Corresponding author: moore.chris@epa.gov

1. Introduction

The 1972 US Clean Water Act (CWA; 33 U.S.C. §1251 et seq.) mandates the maintenance and restoration of the chemical, physical, and biological integrity of US waters. The US Environmental Protection Agency (EPA) often quantifies the outcome of CWA regulations on aquatic resources with a water quality index (WQI; Brown et al. 1970). The WQI characterizes the suitability of surface waters for a range of human uses such as boating, fishing, and swimming with a unidimensional indicator comprised of multiple water quality parameters. Of these parameters, chemical indicators (e.g., nitrogen, phosphorous, and dissolve oxygen) are typically used, but physical (e.g., total suspended solids) and biological (e.g., fecal coliform and biological oxygen demand) indicators of water quality can also be used (Walsh and Wheeler 2013). The WQI is the primary metric used by the EPA to value water quality changes in monetary terms.

Over the past several decades, benefits analysis of significant surface water quality regulations have typically used meta-analyses of stated preference (SP) studies that estimate economic value from responses to survey questions about peoples' willingness to pay for (or accept) hypothetical changes in environmental quality (Johnston et al. 2017). The studies included in these meta-analyses either directly used the WQI in SP surveys or valued changes that were converted to WQI by EPA for inclusion in meta-analyses (USEPA 2015). However, the EPA's reliance on the WQI to estimate the benefit of proposed policies fails to recognize a recent shift in how the Agency analyzes CWA regulations. Specifically, estimated benefits have expanded from primarily human health and recreational uses to include impacts on biological integrity, such as the composition of species within an ecosystem (Griffiths et al. 2012), for which the WQI can be a poor indicator. Thus, contemporary benefit-cost analysis of CWA regulations requires a metric that captures changes in ecological integrity beyond those that affect use values, such as recreation.

Society's value for environmental quality that is not driven by human uses is known as "existence value" (Madariaga and McConnell 1987, Larson 1993). Existence value is a component of nonuse value[1] that is derived from the satisfaction that people get from stewardship of the environment even if they will never use the resource in question (see Crowards 1997). Importantly, values held by nonusers of a resource can be small per household but equal or exceed use-based value when aggregated across many households (Moore et al. 2018). Thus, failure to account for this source of benefits may underestimate total benefits of CWA regulations. The only way to capture existence value is through SP studies, but even the best designed SP surveys can be cognitively challenging for respondents (Johnson and Mathews

[1] Other sources of nonuse value could include bequest and option values depending on how it is defined. See Larson (1993) for a discussion.

2001) and identifying indicators that avoid conflation or confusion between use and nonuse values is a major challenge.

To more accurately capture the existence portion of total economic value, we are pursuing a dual-index approach to complement the WQI with a second indicator appropriate for SP surveys. Such a survey should transparently represent the WQI as an indicator of suitability for recreational activities and the complementary indicator would convey the biological integrity of aquatic resources. The purpose of the dual-index survey design is to collect valuation data that captures existence value and allow a more accurate estimation of total economic value by forecasting changes in two distinct dimensions of water quality. A suitable complementary metric for use in the SP surveys should be salient to survey participants, capture existence value of freshwater ecosystems, and avoid conflation with the use values captured by the WQI. Our search is limited to existing measures of biological condition that are routinely collected at a national scale by the EPA (i.e., USEPA 2009, 2016). We also limited our search to indicators of streams and lakes to match current work within EPA to apply scenario-based water quality modeling to a broad swath of these ecosystems within the US. We describe our process for selecting a metric of biological condition by subjecting candidate metrics to focus group research and evaluating them based on the Schultz et al. (2012) criteria of measurability, interpretability, applicability, and comprehensiveness.

How to describe the *quantity* of surface water being improved is a less nuanced detail of SP study design, yet as important to benefits estimation as water quality. The measure must be applicable to large geographic areas and multiple waterbody types. It should be easily understood by survey respondents and have a consistent interpretation across respondents and researchers. The SP literature contains various examples of surface water quantity metrics (e.g., Van Houtven et al. 2014 use of percent of lakes in a state) but, to our knowledge, there has not been a systematic comparison of alternatives, especially when multiple types of waterbodies are presented to survey respondents (e.g., lakes and streams). We explored this feature of SP study design with focus groups and identify a single measure of quantity that can be applied to both streams and lakes which substantially reduced the cognitive burden of focus group participants when considering the quality and quantity of differing waterbody types.

The current condition of nearby waters can influence the willingness of survey respondents to pay for incremental improvements in water quality, often in a nonlinear fashion (Newbold et al. 2018). EPA's recent analyses of CWA regulations estimate the willingness of representative households in each census block to pay for water quality changes within a specified radius (e.g., USEPA 2015). To do so nationally, while accounting for geographic variation in nearby biological condition, requires data on each stream segment and lake in the conterminous US (CONUS) to match with the locations of survey participants. However, existing observational datasets of biological condition collected by academic or government institutions are only available for a small subset of waterbodies. To account for geographic variation in biological condition and improve national benefits analysis, we have undertaken a modeling effort that will allow our selected metric of biological condition to be integrated into the geographic scope of EPA's current valuation paradigm. Our current modeling effort seeks to spatially interpolate observed values of biological condition to unsampled streams and lakes and impart the coverage needed for a national analysis of CWA regulations. A parallel effort also seeks to model and interpolate our selected water quantity metric so that condition estimates from streams and lakes can be placed into the same units. Upon completion, we plan to publish these interpolations for use in other analyses, including non-market valuation efforts. A publicly available national dataset of this kind could benefit resource economists by providing a consistent measure for comparing the nonuse benefits and costs of proposed policies among studies. Further, a dataset of interpolated conditions at this fine spatial scale is flexible and would allow resource economists to aggregate the measure of biological condition to any geographic resolution appropriate for their application (e.g., political or natural boundaries).

The remainder of the paper is as follows. Section 2 describes how data on aquatic biological condition are collected by the EPA and the candidate metrics we considered for nonuse valuation. In Section 3, we describe our process and criteria for selecting metrics of both aquatic resource condition and quantity for use in a SP survey, including the outcome of focus groups that helped guide this selection. Section 4 describes our initial methods and results for spatially interpolating observations of aquatic biological condition and quantity. In Section 5, we discuss future work, including possible approaches for linking biological condition to scenarios of water quality and habitat condition to complement WQI scenarios in analyses of CWA regulations, and Section 6 concludes.

2. Available Data & Metrics on Aquatic Biological Condition

The EPA's National Aquatic Resource Surveys (NARS) program provides a unique set of spatially extensive data, including ecological indices, that we can test with focus groups as potential metrics to complement the WQI. The NARS program was designed to provide assessments of the Nation's aquatic resources, including lakes, streams, wetlands, and coastal waters (Shapiro et al. 2008). The program does so through spatially balanced sampling of aquatic resources that allows the EPA to compute summaries and make inferences of water quality status at both regional and national scales (Olsen et al. 1999, Stevens and Olsen 2004). As a part of NARS, about 2000 rivers and streams (henceforth streams) and 1200 lakes (Figure 1) are sampled every five years in collaboration with State and Tribal partners. These respective sampling efforts are called the National Rivers and Streams Assessment (NRSA) and the National Lakes Assessment (NLA). The target population of the NRSA is all CONUS streams with flowing water during the sample period (April-September). Likewise, permanent lakes, reservoirs, and ponds greater than one hectare in surface area make up the NLA target population (USEPA 2017). During both surveys, physical, chemical, and biological data are collected at each sample location, including counts of fishes (NRSA), diatoms (NRSA), benthic macroinvertebrates (NRSA and NLA), and zooplankton (NLA). From these biological data, ecological indicators are developed and made available to the public as reports and data.[2]

Among candidate metrics collected by NARS, we chose to test quantitative measures of biological condition that can be estimated on continuous rather than nominal or ordinal scales. Early in the selection process, we ruled out qualitative descriptions of biological condition (e.g., poor, fair, and good) because they are problematic when applied to regulatory analysis. For example, if the valuation study only quantifies willingness to pay for water quality improvements that cross a threshold from one category to another, it is not obvious how to estimate benefits for improvements that do not cross one of these thresholds. Changes in pollutant levels or other ecosystem stressors were also ruled out early because survey respondents would be forced to speculate about the eventual outcome of such changes on the

[2] https://www.epa.gov/national-aquatic-resource-surveys

features of biological condition that they care about. Finally, pollutants may be redundant with indicators included in the WQI, which could lead respondents to conflate the selected metric with direct human use values.

Initial filtering of candidates within the NARS left us with two potential companion metrics to the WQI: (1) multimetric indices (MMI; Karr 1981, 1991, 1999) and (2) observed-to-expected (O/E) taxonomic composition (also called RIVPACS; Moss et al. 1987, Hawkins et al. 2000). We considered these metrics because both are available for nationwide use in non-market valuation since they are routinely collected and reported by NARS for several thousand streams and lakes across the CONUS. Additionally, both indicators are measured on continuous scales and designed to directly assess the biological condition of aquatic ecosystems, in contrast to indirect measures such as pollutants or the presence of aquatic stressors. Finally, both have undergone considerable development and testing over the last 40 years (Buss et al. 2014) and are accepted and widely used in state, regional, and national monitoring programs (e.g., Mazor et al. 2016, Larson et al. 2019). The detailed mechanics of MMIs and O/E approaches are beyond the scope of this paper, though a brief summary can help clarify the relative advantages of each when assessed with focus groups. Our description of each is tailored to their specific applications within the NARS, but there are numerous variations on these approaches available to practitioners.

An MMI is designed to be a comprehensive index of ecological health through the aggregation of several individual metrics that are calculated from taxonomic data (e.g., fish counts or occurrences; Karr 1999). Dozens of potential metrics are available for inclusion in an MMI, each designed to capture specific aspects of the biological composition or function of aquatic organisms. For the NRSA MMI, metrics were selected to represent six categories of biological information: (1) habitat preferences, (2) taxonomic richness, (3) pollution tolerance of collected taxa, (4) taxonomic composition , (5) taxonomic evenness/diversity , and (6) feeding groups (Stoddard et al. 2008). To choose final metrics for the NRSA MMI, candidate metrics were subjected to a series of tests, including reproducibility and responsiveness. Only one metric was chosen to represent each category. Values of the selected metrics across the six categories were rescaled and summed to produce an aggregated score of biological condition (see Stoddard et al. 2008 for details).

In contrast to an MMI, an O/E index quantifies the loss of aquatic taxa due to human-related stressors (Hawkins 2006). It does so by comparing the list of taxa observed (O) at an assessed site to the taxonomic composition that would be expected (E) in the absence of such stressors (Moss et al. 1987). To estimate E, several steps must be taken. First, a set of sample sites are identified to represent a regional benchmark against which assessed sites can be compared. These sites are typically selected to be minimally disturbed by human activity (also called “reference condition”; see definition below). Next, the selected sites are then clustered based on their observed taxonomic composition so that sites with many shared taxa are grouped together. These groupings represent the types of biological assemblages that would be expected in high quality streams. Next, the biological groups are then related to environmental variables (e.g., watershed area, soils, topography, and climate) with statistical modeling techniques, such as discriminant functions, that can estimate the probability of a new site’s membership in each of the identified biological groups based on its physical setting. In this way, a biological expectation can be made for a new site being assessed based on its physical watershed characteristics. Finally, these estimated group probabilities are used with taxon-specific occurrence frequencies to estimate the capture probabilities of each taxon and derive E. For example, if a taxon occurs frequently within a biological cluster of sites (e.g., occurs within 0.95 of sites within that group) and a new assessed site is estimated by the model to have a high probability of belonging to that group of sites (e.g., probability of class membership = 0.90), then the probability of capturing this taxon in this sample can be estimated by multiplying its frequency of occurrence in this biological group by the probability of the new site belonging to that group (i.e., 0.95 x 0.9 = 0.85). Additional calculations adjust for the possibility of site membership in more than one biological class (see Moss et al. 1987, Hawkins et al. 2000 for details). Once adjusted, E for the assessed site is the sum of individual taxa capture probabilities and O is the sum of observed taxa (richness). Ideally, the ratio O/E at sites with no-to-minimal human disturbance is 1, meaning all taxa that would be expected to occur under such conditions were observed. However, some variation above or below 1 is common due to sampling, laboratory, or modeling errors, as well as difficulties in identifying high-quality sites to set biological expectations. However, this variation can be quantified as model precision and accounted for when making assessments. Deviation from 1 at an assessed site indicates the loss of taxa from an ecosystem (e.g., O/E = 0.5 implies 50% of taxa have been lost) since E is ideally derived from a model of sites with minimal human-related disturbance (Hawkins 2006).

Critical to the formulation of both MMI and O/E indices in NARS is the concept of reference condition. Reference condition refers to the benchmark against which assessed sites are compared (Hawkins et al. 2010). In the case of MMIs, scores at reference-condition and “worst-case” sites are compared to identify metrics that are sensitive to ecological impairment

(Stoddard et al. 2008). For O/E, E is modeled directly with reference-condition sites to establish an ecological benchmark against which O can be compared (Moss et al. 1987, Hawkins et al. 2000). However, confusion about the meaning of reference condition is common even among environmental professionals (Stoddard et al. 2006), and the development and application of the concept has received much attention over the last several decades (Hawkins et al. 2010). Recognizing the need for a consistent definition, Stoddard et al. (2006) distinguished among several concepts for which the term "reference condition" was often used but which have substantially different meanings, including historical, pristine, minimally disturbed, and least disturbed conditions. Given that monitoring data do not exist for historical conditions (e.g., pre-intensive agriculture) and that few truly pristine or minimally disturbed areas exist, biological assessment programs typically use sites in least-disturbed condition (i.e., best available among candidate sites) that can deviate substantially from natural conditions, especially in heavily impacted regions. For example, to retain enough reference sites for statistical comparisons, Herlihy et al. (2008) were forced to use substantially higher criteria for nitrogen and phosphorous to select reference sites in the Temperate and Southern Plains of the US compared with other ecoregions. This means that "good condition" is a relative measure, and that absolute condition cannot be compared between regions using these metrics. While SP survey respondents need not understand the mechanics of how reference sites are selected, it is critical that respondents, and research economists alike, understand the concept of reference condition. An understanding of the benchmark against which current conditions are compared is also important to prevent incorrect interpretation as being truly undisturbed because little is known of how differing interpretations ecological benchmarks might affect SP survey outcomes.

3. Selecting an Index of Aquatic Biological Condition and Units of Quantity

Aquatic Biological Condition

In searching for a metric to complement the WQI, the way we collect and use the valuation data has implications for characteristics that an appropriate metric must possess. Schultz et al. (2012) examined the representation of ecological outcomes in SP studies and identified four criteria that metrics should meet for valid benefit estimation, which we used to compare MMI and O/E indices. Ecological indicators used in an SP survey should be *measurable*. Schultz et al. (2012) describe a measurable index as one that has “a clearly stated relationship to ecological data or model results.” Subjective descriptions of outcomes such as “good” or “poor” often fail to meet this criterion. *Interpretability* is another necessary feature of ecological indicators that ensures different values of the metric have consistent meanings to survey respondents, subject experts, and resource managers. Thirdly, an indicator with *applicability* will be relevant to the management scenario that is the subject of the survey. For an indicator to be applicable, survey respondents must understand and have a clearly defined preference in its outcome under these management scenarios. An applicable indicator will aid in scenario acceptance by the SP survey respondent and is required for the SP survey results to be relevant to the benefit estimation effort. Finally, the *comprehensiveness* of an indicator reflects the degree to which all direct and indirect ecosystem impacts are described by the metric(s) and understood by the respondent.

We began assessing the measurability of MMI and O/E indices by examining the literature. The ability of these indicators to assess the biological condition of aquatic ecosystems has been rigorously compared and refined over the last several decades (Hawkins et al. 2010, Buss et al. 2014). From this work, all aspects of their measurability (e.g., accuracy, precision, sensitivity and specificity to stressors) improved substantially. Improvements came principally through refinements in field and laboratory protocols (e.g., Vinson and Hawkins 1996, Ostermiller and Hawkins 2004), statistical and other analytical approaches (e.g., Van Sickle et al. 2005, Cao et al. 2007, Van Sickle et al. 2007), as well as refinement and better application of such concepts as reference condition (e.g., Hughes et al. 1986, Stoddard et al. 2006, Ode et al. 2016). Due to these improvements, MMI and O/E approaches can produce comparable regional assessments of aquatic condition (Hawkins 2006, Stribling et al. 2008). Therefore, we found no substantial difference in measurability between the candidate indicators for use in SP surveys.

MMIs, by definition, are designed to be comprehensive measures of biological condition through the inclusion of carefully selected life history and behavioral traits (e.g., Stoddard et al. 2008). The inclusion of these traits makes them more comprehensive than O/E indices. However, despite this comprehensiveness, MMIs may not capture the biotic response to all stressors. For example, an O/E index could detect the displacement of native taxa by invading taxa whereas an MMI may be insensitive to this stressor if invading taxa fill the ecological roles of the taxa they displace (Collier 2009 as cited in Mazor et al. 2016). Thus, the comprehensiveness of an index does not guarantee its sensitivity in all cases and was not, therefore, the primary criterion we used to select a complementary metric to the WQI.

We could not assess the interpretability or applicability of MMI or O/E indices by examining the literature alone. To our knowledge, no study has compared the ability of the public to understand or form preferences of ecological outcomes based on these indicators. Therefore, the EPA conducted a series of focus groups to evaluate MMI and O/E indices in a SP setting. A total of ten focus groups with eight to ten participants each were conducted in Washington DC; Chicago, IL; and Phoenix, AZ. Locations were chosen to work with participants that have a variety of experiences with, and interpretations of, water resource issues. Participants were selected to include equal numbers of men and women with a minimum of a high school diploma and to roughly match the general population with regards to race and income. We adopted an *emergent design* structure for the study, with early focus groups following a conversational format to identify dominant themes when considering environmental quality in aquatic environments (Morgan et al. 2008). As the study progressed, the discussions became more structured and considered topics such as how to convey scientific information to the general public and how to describe changes in our candidate metrics of biological condition.

Focus group participants struggled with their interpretation of MMIs. Specifically, focus group participants had difficulty understanding how a 5-point increase in such an index would be different from, say, a 10-point increase. An increase in the value is clearly an improvement, but how much of an improvement and exactly what features of biological condition would improve remained elusive. Our focus group results suggested this was due, in part, to the aggregation of various metrics into a single index. When presented with an improvement in an MMI, respondents were unable to describe what these specific changes meant because MMIs are comprised of multiple sub-metrics that can vary independently of one another so that many unique combinations of scores can produce the same value. Asking participants to keep track of the individual metrics that compose the MMI was too cognitively burdensome. Any measure that is designed to be comprehensive may suffer from this interpretability problem. In contrast, focus group participants more easily explained how a 5- or 10-point change in O/E related to loss of taxa. This may be because taxa are a salient "currency" of ecosystems (Zachos 2016) that are conceptually familiar to the public. So, while O/E may capture fewer dimensions of ecological health, it is one that focus group participants were already familiar with and could more easily understand.

Applicability requires that the survey respondent be able to connect the index to the value of interest and form an opinion on its outcome under various management scenarios (Schultz et al. 2012). While participants found no major differences in applicability between MMI and O/E indices themselves, the focus groups were valuable in informing how to present index information to survey respondents. When describing an index to focus group participants, we found that it was important to emphasize its reliance on macroinvertebrates or plankton, rather than fish. Without that emphasis, participants tended to focus on higher order taxa, such as fish and birds, which have a more direct link to use values through activities, such as recreational hunting, fishing, and wildlife viewing. Although some anglers may make an indirect link between the biological condition of lower order taxa and improved fishing and hunting, we found that the use of these taxa allowed most respondents to consider and form opinions on the outcome of biological condition separately from recreational uses when considering the valuation task.

Based on focus group results, we chose to pursue O/E as the measure of aquatic biological condition that will complement the WQI. On balance, we found little evidence for substantial differences in measurability and applicability of MMIs and O/E indices. MMIs are more comprehensive but it was unclear how this comprehensiveness translated to a superior index for nonuse valuation in an SP survey. Ultimately, the success of the SP survey will depend on the ability of survey respondents to correctly interpret changes in the selected biological index presented to them. Due to this importance, we chose to weight interpretability above other criteria and the greater ability of focus group participants to interpret O/E indices over MMIs drove our final decision on index choice. Finally, by considering taxonomic levels other than fish (i.e., macroinvertebrates in streams and plankton in lakes), focus group participants were better able to draw a direct link between the biological index and existence value, thereby minimizing conflation and confusion with the WQI.

Aquatic Ecosystem Quantity

Stated preference studies may seek to quantify the willingness of the public to pay for improvements in the biological condition in both streams and lakes. When using SP to value water quality changes over large geographic regions, especially those with which respondents may be unfamiliar, the survey must describe the quantity of water or size of waterbodies that would improve under a given scenario. Doing so will require that these waterbodies be placed in the same units to facilitate their presentation to survey participants.

Focus groups found that strictly visual representations did not include enough information to be interpreted consistently. Maps showing the land area over which improvements would occur did not convey the number or size of waterbodies included in that area. For example, a given area in the northeast US will generally have many more lakes and streams than the same area in the arid southwest. Further, we found that maps of water features could present drastically different pictures depending on how they were created and were too difficult for focus group participants to consistently interpret as a meaningful quantity measure. Recognizing that lotic and lentic ecosystems could be valued differently by respondents (Johnston et al. 2017), we tested using linear metrics for streams and areal metrics for lakes, consistent with the NARS program reporting on the status of these resources. This proved to be too cognitively burdensome, because it effectively doubles the number of attributes respondents must consider when evaluating scenarios. Based on these findings, we narrowed the universe of metrics down to numeric descriptors that can place lakes and streams into like units to facilitate their combined presentation to survey respondents.

One such measure that is currently used by EPA to transfer values from a meta-analysis of valuation studies (Corona et al. *forthcoming*) is shoreline length. Shoreline length measures the perimeter of freshwater lakes and reservoirs and the length of both banks of a stream. Satisfying the measurability and comprehensiveness criteria, this measure was received more favorably by focus group participants than visual representations. Some participants, however, pointed out that a narrow length of stream could have more shoreline length than a large lake. Others pointed out that, depending on how islands are treated in the calculation, lakes with more islands would have larger values using this metric, which led them to reject the measure as a meaningful indication of quantity.

A second measure that may be robust to these issues identified by the focus group is water surface area. Focus group participants tended to agree that surface area is a more accurate reflection of quantity when assessing water quality improvements than the other candidates. The performance of this metric improved further when focus group participants were told what percent of the total surface area in the CONUS the affected area represents. Based on these findings, we selected area (to be expressed as square miles to survey participants) as the common unit for expressing O/E values in streams and lakes due to the flexibility of also being able to express them as a percent of the total surface area within a region or the CONUS.

Having chosen metrics for aquatic biological health and the amount of water experiencing an improvement, we can move into Phase 2 of the project to augment the observational NARS data with modeled values for sites that are not in the NARS sample. Generally, survey respondents' willingness to pay for a given improvement will depend on current conditions of aquatic resources and we assume that they are aware of water quality conditions in local waterbodies. As such, we need current values for O/E in lakes and stream reaches in the CONUS to match with survey participants and their census blocks. Below, we describe our efforts to model and provide these measures at a geographic scope and granularity that is appropriate for integration in the current EPA paradigm for valuation of changes in surface water quality.

4. Interpolating Resource O/E and Quantity

Modeling O/E

For modeling and interpolation, we used O/E scores of stream benthic macroinvertebrates from the 2013-14 NRSA and lake plankton from the 2007 NLA as the dependent variables (see Appendix S1 for data and model QA/QC procedures). These years were used to develop our modeling and interpolation methodology, but later model iterations could include additional survey years. Although we modeled different taxa in streams and lakes, O/E is a unitless measure of ecological completeness that has a consistent meaning regardless of the taxonomic groups used in its derivation. Streams and lakes were modeled separately because some independent variables differed between them (see StreamCat and LakeCat descriptions below). Further, a model was created for each waterbody type for each of the three NARS ecoregions for a total of six models (Figure 1; Eastern Highlands: EHIGH; Plains and Lowlands: PLNLOW; and Western Mountains: WMTNS). We created regional models because these are the original regions used to develop the O/E assessments (Yuan et al. 2008) and Hill et al. (2017) found that creating regional models improved estimates of biological condition due to differences in the quality of reference sites among them; an issue that was not solved by including region as an independent variable in the models.

As independent variables in the models, we used watershed metrics to characterize the natural and anthropogenic setting of each stream or lake watershed (see Appendix S2 for a list of the predictor variables used to model O/E). To interpolate to unsampled streams, the same set of independent watershed variables must be available for model calibration and application to unsampled locations (e.g., Maloney et al. 2018). For example, Hill et al. (2017) interpolated probabilities of MMI class membership (i.e., good vs. poor condition) to 1.1 million stream segments by building and applying models with the EPA StreamCat dataset, which we also used here. StreamCat is an extensive dataset of landscape metrics for about 2.65 million stream segments and their associated watersheds within the CONUS (Hill et al. 2016). These data characterize both natural (e.g., soils, geology, and climate) and anthropogenic (e.g., urbanization and agriculture) landscape features within watersheds and have been frequently used as covariates for model development and application in recent years (e.g., Beck et al.

2019, Guillon et al. 2020, Maloney et al. 2020). However, our current objective to interpolate O/E scores to unsampled streams and lakes differs from Hill et al. (2017) in at least two ways. First, Hill et al. (2017) modeled MMI response classes (i.e., good vs. poor), whereas O/E scores must be modeled as a continuous measure to estimate benefits for improvements that do not cross a class threshold. Second, Hill et al. (2017) only interpolated conditions of streams, while our current objective also includes lakes. The subsequent development of the EPA LakeCat dataset (Hill et al. 2018) allows the application of these approaches to lentic ecosystems as well. However, some differences between StreamCat and LakeCat metrics precluded the development of a model of lakes and streams together.

We used random forests to separately model O/E scores of streams and lakes. Random forest modeling is a non-parametric machine learning technique that builds many individual decision trees (Breiman et al. 1984) from randomized subsets of the original data (Breiman 2001). Use of many randomized subsets of the data stabilizes model results and improves final model performance as measured with samples that are excluded from tree construction (also called "out-of-bag" samples; Cutler et al. 2007). To construct individual trees, splits of the response variable across values of the independent variables are tested. Optimal splits are identified as those that minimize the sums of squares of the response variable after splitting. Once individual trees are constructed, statistical predictions (i.e., spatial interpolations in our case) can then be made when independent variables (i.e., StreamCat and LakeCat) at unsampled locations are used as new input to the random forest trees. During model development, we did no variable selection because it is unnecessary for random forests when the principal purpose of the model is interpolation rather than interpretation (Fox et al. 2017).

Advantages of using random forest are that it makes no assumptions about the normality or independence of input variables, requires very little tuning, and captures non-linear relationships and interactions which other modeling techniques may not. These advantages are important because ecological data often violate statistical assumptions and exhibit non-linear interactions that can be accounted for with random forest. Furthermore, random forest models often outperform other modeling techniques (Prasad et al. 2006), especially when compared against linear techniques (Cutler et al. 2007).

The models explained 25-30% of the variation in stream O/E scores and 13-36% of the variation in lake O/E scores (pseudo r-squared; Table 1). Model root mean squared errors (RMSE) were 0.25-0.27 (streams) and 0.22-0.25 (lakes) out of observed O/E scores of 0-1.6. Despite explaining a low percentage of the variation in O/E scores, model residual errors showed no spatial biases (i.e., clusters of over or under predictions) when plotted as maps (not shown here). Further, residual autocorrelation in NARS samples has been shown to be negligible due to the distance between sites and most sites have non-overlapping watersheds (Fox et al. 2020). Maps of interpolated O/E scores showed distinct shifts in values at ecoregion boundaries (Figure 2); a behavior that has been observed in previous model interpolations and attributed to differences in reference site quality among regions (Hill et al. 2017). For example, marked transitions are visible in interpolated lake O/E scores between the WMTNS and PLNLOW regions in the states of Montana and Wyoming (Figure 2C) and in streams between EHIGH and PLNLOW (Figure 2D). In addition to regional differences within lakes and streams, patterns of observed (cf. Figure 2A and B) and interpolated (cf. Figure 2C and D) O/E scores differed substantially between lakes and streams across the CONUS. This difference between stream and lake O/E scores was most apparent in Montana and Wyoming where lake O/E scores were substantially lower than those observed in streams.

Modeling Width

Consistent with focus group findings, we sought to place both streams and lakes into areal units. Doing so will help to describe the amount of water being improved for each waterbody type (stream vs. lake) under a hypothetical policy. Further, areas can provide a weight for averaging ecological measures to policy-relevant regions for analysis of SP results. An area for each lake is available from the waterbodies layer of the National Hydrography Dataset Plus Version 2 (McKay et al. 2012). Unfortunately, the hydrography dataset only contains lengths for streams. To estimate area, we first modeled widths which we then multiplied by stream lengths to produce estimates of stream area. We modeled stream widths from field measurements of the 2008-09 and 2013-14 NRSA. Within these data, two width measurements were available: bankfull and wetted widths. We chose to model wetted width because it better replicates summertime low flow conditions when biological samples were taken, rather than the near-flood conditions of bankfull width which typically occur at other times of the year. Wetted width was measured from water edge to edge at the time of sampling. As with O/E, we used random forest modeling with StreamCat data as independent predictors (Appendix S2). However, we found that regional models of width did not improve performance and a single CONUS model was constructed. Using this model, we multiplied estimated widths by the stream lengths of the hydrography data to calculate area for each stream segment in the CONUS.

The model explained a high proportion of the variation in measured stream widths (R^2 = 0.83) and did not exhibit spatial biases (mapped model residuals not shown here). Patterns of interpolated widths followed those expected for major river networks of the United States (Figure 3). Recognizing that survey respondents may value streams and lakes differently, we found it more informative to compare the relative areas of streams and lakes across hydrologic boundaries. By area, streams were the dominant water resource in 1.06 million local catchments (the finest spatial grain produced by this study) whereas lakes were dominant in only 312,739 catchments (Figure 5A). However, lakes progressively became the dominant waterbody feature within regions when areas were aggregated to increasingly coarse geographic units (Figure 5B-D).

5. Discussion

In this study, we outlined our strategy to improve the accuracy and completeness of aquatic resource valuation with a second indicator that can complement the WQI (Figure 1). The EPA's historical reliance on the WQI for benefit-cost analyses of CWA regulations failed to capture the full suite of values the public derives from freshwater resources. Existence values are an important and possibly large (in aggregate) source of benefits that should be accounted for in benefit-cost analyses prepared to help to inform CWA policies. However, estimating the existence value of aquatic resources presents a major challenge for research economists, in part, because of potential confusion and conflation with the use-based values that are the primary focus of the WQI. A further challenge is estimating this complementary metric at a geographic scope that is relevant to resource valuation for national policy. Our results advance the state of valuation in at least two important ways. First, based on feedback from focus groups, we identified an biological metric that met criteria for valid benefit estimation (Schultz et al. 2012) yet reduces conflation with the use-based WQI when collecting SP data and when forecasting aquatic conditions under counterfactual scenarios. Second, we developed preliminary models as a proof-of-concept for interpolating values of both the quality and quantity of stream and lake ecological conditions nationally. Here, we consider these advancements as well as further work that is needed to improve our model of biological integrity and link it to other EPA models of water quality.

Selecting a Complementary Metric

A critical step towards valuation of aquatic resources was identifying a complementary yet distinct metric to the WQI. We were able to evaluate MMIs and O/E indices for their measurability and comprehensiveness with existing literature because the merits of each were vigorously debated near the turn of the millennium (Hawkins et al. 2010). This debate improved both metrics greatly, and we found little evidence within the literature of differences in their measurability when compared as regional assessments (e.g., Hawkins 2006). While similar in measurability, MMIs are generally accepted as comprehensive indicators of ecological health, which is not the case for the O/E. However, it was unclear how this comprehensiveness translated into a clear benefit because of tradeoffs in sensitivity and specificity to environmental stressors that can exist between MMIs and O/E indices (Mazor et al. 2016).

Focus groups were invaluable for assessing MMI and O/E indices and refining how researchers should present biological condition to SP survey respondents. It is critical for an SP survey that respondents be able to interpret the meaning of specific changes in an index value to produce accurate results, and on balance, interpretability outweighed considerations of index comprehensiveness. Further, an index must be interpretable for it to be applicable. Index applicability depends on the ability of a respondent to draw a direct link to a benefit they value and to form an opinion on its outcome under management scenarios; tasks that would be impossible if a metric is not interpretable. Focus groups also refined our understanding of how different taxonomic groups in SP surveys may influence the interpretation of indices for nonuse valuation. The use of non-fish taxa, such as macroinvertebrates and plankton, helped respondents consider ecosystems and management scenarios independently of their impacts on recreation. This finding is important as the analysis of surface water regulations expands beyond human health and direct use values. For example, due to their diversity and responsiveness to stressors, benthic macroinvertebrates are one of the most commonly used taxonomic groups in freshwater bioassessments worldwide (Chessman et al. 2007, Buss et al. 2014), making them good candidates for nonuse benefits studies globally.

Modeling Resource Condition and Quantity

The models of O/E and stream width illustrate how ecological health and quantity can be spatially interpolated for use in resource valuation. As with similar EPA efforts (Hill et al. 2017), the finalized interpolations will be made available for download by practitioners and researchers outside of the EPA for use in resource planning and valuation. Coupled with resource valuation functions and with ecological production functions, the interpolations of ecological health could provide valuable insight into the economic benefits offered by competing management and conservation scenarios. For example, the Center for Large Landscape Conservation recently used interpolated probabilities of good MMI condition published by the EPA (Hill et al. 2017) in a tool[3] designed to help prioritize conservation within the Missouri Headwaters Basin. Tools such as this could benefit greatly from valuation of competing conservation scenarios that this multi-phase project may provide.

The fine spatial grain and near-continental extent of these interpolations has several advantages for SP survey analysis. First, this framework is flexible. Analysis of streams and lakes in tandem or in isolation is possible because each resource type was modeled separately. Second, interpolations can be spatially partitioned or aggregated to fit a variety of needs, such as valuation of resources within a political or management boundary or to match other EPA models and tools used for aquatic resource assessment and valuation, such as HAWQS and BenSPLASH (Corona et al. *forthcoming*). However, as illustrated above, care must be taken to understand how the dominance of resource type might vary as interpolated values are aggregated. Aggregation to coarser resolutions may weight lakes over streams due to their areal dominance when equal valuation may be desired. Yet the flexibility of the framework will allow users to return and examine data at their original resolution to understand how aggregation might affect analytical results.

Currently, the models explain a low proportion of the variation in O/E, which could occur for several reasons. First, models with low r-squared values are common within ecology and our results may be in line with or better than many ecological models (Møller and Jennions 2002). Second, values of E in the ratios are from previous modeling that include model error. Thus, future work could seek to improve interpolations by returning to the original taxonomic data to develop models of O and E separately. Finally, although values of O/E ranged between 0

[3] https://uppermissouriheadwaters.org/

and 1.6 due to field, laboratory, and modeling errors, O/E values >1 are not necessary to represent taxonomic completeness for a SP survey. We will explore whether limiting the range of O/E values to between 0 and 1 for modeling can improve out-of-sample fit. In addition to these improvements, we will also explore methods to minimize the observed transitions at ecoregional borders and explore other modeling techniques to improve interpolations. Despite the current limitations, this study shows promise for providing interpolated estimates of ecological health and quantity to complement the WQI.

Linking Biological Integrity with Water Quality

Ultimately, the purpose of this research is to provide resource economists with the data and tools they need to estimate changes in value in response to proposed management. It is unlikely that the models described here can provide this capability. Rather, their purpose is to interpolate current conditions to unsampled lakes and streams and help account for geographic variation in valuation of benefits. We achieved this through a machine learning algorithm. Further, we made direct relationships between both natural and human-related watershed features with biological condition. The large number of watershed variables (Appendix S2) and their use to directly model O/E make it difficult to infer the mechanistic pathways by which proposed policies would affect biological condition through improvements in water quality or habitat condition.

To make the transition to a policy valuation tool, resource economists will need the means to forecast biological condition under baseline and regulatory scenarios. However, substantial work remains to build such forecasting capability. First, explicit linkages must be made between human-related watershed activity, instream water or habitat quality, and O/E scores (i.e., watershed stressors → *in situ* water or habitat quality → *in situ* O/E scores). For example, to infer potential sources of impairment in Nevada, USA streams, Vander Laan et al. (2013) related macroinvertebrate O/E values to measurements of heavy metals as well as O/E ratios of total dissolved solids (TDS; measured as electrical conductivity). Values of TDS O/E were largely driven by agriculture, mining, and urbanization, thereby providing an indirect assessment of how major land uses influence biological condition through water quality. Second, we must use modeling approaches that can describe mechanistic linkages between management and biological condition. Schmidt et al. (2019) used structural equation modeling to link land use in Midwestern watersheds to instream habitat structure, temperature, and pesticide and nutrient concentrations. These parameters were, in turn, linked to algal and benthic macroinvertebrate MMIs. The use of structural equation modeling in this way could disentangle mechanistic pathways between human-related landscape stressors, *in situ* water and habitat quality, and O/E scores (Grace et al. 2010). However, doing so will require a model framework that sufficiently describes the hypothesized pathways between each watershed stressor and water quality parameter and, ultimately, O/E (Grace and Bollen 2005). Defining these pathways is a significant challenge given the many direct and indirect ways human activity can affect aquatic biota. Finally, the models described in this paper and in the examples above are spatial models, that is, the relationships explain differences among sample locations rather than changes due to management over time. Although there is some evidence to support space-for-time substitution in ecological models (Blois et al. 2013), such assumptions should be explicitly tested before being used for decision making.

Process-based models could provide management scenarios for some water quality parameters. This capability exists for the WQI through water quality assessment tools such as SWAT[4] (Santhi et al. 2006) and SPARROW[5] (Smith et al. 1997). Such models allow a researcher to enter location-specific changes in pollutant loadings and/or management practices and recover the downstream water quality values needed to calculate the WQI (e.g., dissolved oxygen, total suspended solids). However, work would still be needed to link these water quality outputs with outcomes of O/E. A further limitation is that process-based models would be difficult to implement broadly at a national scale, potentially limiting their development to specific watershed and, hence, their utility under the geographic scope of EPA's current valuation paradigm. Finally, despite the process-based nature of these water quality models, the linkages to O/E values would still need to be made through spatial models due to the relative rarity of long-term biological datasets. Where they do exist, long-term biological datasets are often restricted to specific regions, making inference to other parts of the country difficult. Thus, a framework that comprehensively links management practices with biological condition for benefits valuations will be a major research challenge in the coming years.

6. Concluding Remarks

We reported on the results of the first two phases of a multi-phase project to improve the way EPA values aquatic resources: (1) identify a metric that satisfies important criteria for use in stated preference valuation and (2) interpolate observational data to achieve spatial resolution required to estimate the valuation equation. Focus groups were critical in evaluating several

[4] https://swat.tamu.edu/

[5] https://www.usgs.gov/mission-areas/water-resources/science/sparrow-modeling-estimating-nutrient-sediment-and-dissolved?qt-science_center_objects=0#qt-science_center_objects

indices of aquatic biological condition and exposing a tension between index comprehensiveness and interpretability. In the end, concerns of interpretability outweighed those of comprehensiveness, and we found that O/E, while not perfect, satisfies the criteria best and suits our specific purpose of valuing water quality changes throughout the CONUS. We also reported on modeling to interpolate stream macroinvertebrate O/E values to streams across the CONUS. Our hope is that spatially explicit maps of ecological health can improve nonuse benefits estimation based on SP surveys. Although additional refinements are needed to improve model performance, a nationally consistent dataset of ecological condition could facilitate the acquisition of such data by research economists and help to improve comparability among studies of nonuse benefits.

Acknowledgments

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 between the U.S. Department of Energy and EPA. The views expressed in this journal article are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.

Literature Cited

Beck, M. W., R. D. Mazor, S. Johnson, K. Wisenbaker, J. Westfall, P. R. Ode, R. Hill, C. Loflen, M. Sutula, and E. D. Stein. 2019. Prioritizing management goals for stream biological integrity within the developed landscape context. Freshwater Science **38**:883-898.

Blois, J. L., J. W. Williams, M. C. Fitzpatrick, S. T. Jackson, and S. Ferrier. 2013. Space can substitute for time in predicting climate-change effects on biodiversity. Proceedings of the National Academy of Sciences **110**:9374.

Breiman, L. 2001. Random Forests. Machine Learning **45**:5-32.

Breiman, L., J. Friedman, R. Olshen, and C. Stone. 1984. Classification and Regression Trees. Chapman & Hall/CRC.

Brown, R. M., N. I. McClelland, R. A. Deininger, and R. G. Tozer. 1970. A Water Quality Index-Do we Dare? Water Sewage Works **11**:339-343.

Buss, D. F., D. M. Carlisle, T.-S. Chon, J. Culp, J. S. Harding, H. E. Keizer-Vlek, W. A. Robinson, S. Strachan, C. Thirion, and R. M. Hughes. 2014. Stream biomonitoring using macroinvertebrates around the globe: a comparison of large-scale programs. Environmental Monitoring and Assessment **187**:4132.

Cao, Y., C. P. Hawkins, J. Olson, and M. A. Kosterman. 2007. Modeling natural environmental gradients improves the accuracy and precision of diatom-based indicators for Idaho streams. Journal of the North American Benthological Society **26**:566-585.

Chessman, B., S. Williams, and C. Besley. 2007. Bioassessment of streams with macroinvertebrates: effect of sampled habitat and taxonomic resolution. Journal of the North American Benthological Society **26**:546-565.

Collier, K. J. 2009. Linking multimetric and multivariate approaches to assess the ecological condition of streams. Environmental Monitoring and Assessment **157**:113-124.

Corona, J., T. Doley, C. Griffiths, M. Massey, C. Moore, S. Muela, B. Rashleigh, W. Wheeler, S. D. Whitlock, and J. Hewitt. *forthcoming*. An Integrated Assessment Model for Valuing Water Quality Changes in the U.S. Land Economics.

Crowards, T. O. M. 1997. Nonuse Values and the Environment: Economic and Ethical Motivations. Environmental Values **6**:143-167.

Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. Random forests for classification in ecology. Ecology **88**:2783-2792.

Fox, E. W., R. A. Hill, S. G. Leibowitz, A. R. Olsen, D. J. Thornbrugh, and M. H. Weber. 2017. Assessing the accuracy and stability of variable selection methods for random forest modeling in ecology. Environmental Monitoring and Assessment **189**:316.

Fox, E. W., J. M. Ver Hoef, and A. R. Olsen. 2020. Comparing spatial regression to random forests for large environmental data sets. Plos One **15**:e0229509.

Grace, J. B., T. M. Anderson, H. Olff, and S. M. Scheiner. 2010. On the specification of structural equation models for ecological systems. Ecological Monographs **80**:67-87.

Grace, J. B., and K. A. Bollen. 2005. Interpreting the Results from Multiple Regression and Structural Equation Models. Bulletin of the Ecological Society of America **86**:283-295.

Griffiths, C., H. Klemick, M. Massey, C. Moore, S. Newbold, D. Simpson, P. Walsh, and W. Wheeler. 2012. U.S. Environmental Protection Agency Valuation of Surface Water Quality Improvements. Review of Environmental Economics and Policy **6**:130-146.

Guillon, H., C. F. Byrne, B. A. Lane, S. Sandoval Solis, and G. B. Pasternack. 2020. Machine Learning Predicts Reach-Scale Channel Types From Coarse-Scale Geospatial Data in a Large River Basin. Water Resources Research **56**:e2019WR026691.

Hawkins, C. P. 2006. Quantifying biological integrity by taxonomic completeness: its utility in regional and global assessments. Ecological Applications **16**:1277-1294.

Hawkins, C. P., R. H. Norris, J. N. Hogue, and J. W. Feminella. 2000. Development and evaluation of predictive models for measuring the biological integrety of streams. Ecological Applications **10**:1456-1477.

Hawkins, C. P., J. R. Olson, and R. A. Hill. 2010. The reference condition: predicting benchmarks for ecological and water-quality assessments. Journal of the North American Benthological Society **29**:312-343.

Herlihy, A., S. G. Paulsen, J. Van Sickle, J. L. Stoddard, C. P. Hawkins, and L. L. Yuan. 2008. Striving for consistency in a national assessment: the challenges of applying a reference-condition approach at a continental scale. Journal of the North American Benthological Society **27**:860-877.

Hill, R. A., E. W. Fox, S. G. Leibowitz, A. R. Olsen, D. J. Thornbrugh, and M. H. Weber. 2017. Predictive Mapping of the Biotic Condition of Conterminous-USA Rivers and Streams. Ecological Applications **27**:2397-2415.

Hill, R. A., M. H. Weber, R. M. Debbout, S. G. Leibowitz, and A. R. Olsen. 2018. The Lake-Catchment (LakeCat) Dataset: characterizing landscape features for lake basins within the conterminous USA. Freshwater Science **37**:208-221.

Hill, R. A., M. H. Weber, S. G. Leibowitz, A. R. Olsen, and D. J. Thornbrugh. 2016. The Stream-Catchment (StreamCat) Dataset: A Database of Watershed Metrics for the Conterminous United States. Journal of the American Water Resources Association (JAWRA) **52**:120-128.

Hughes, R. M., D. P. Larsen, and J. M. Omernik. 1986. Regional reference sites: a method for assessing stream potentials. Environmental Management **10**:629-635.

Johnson, F. R., and K. E. Mathews. 2001. Sources and Effects of Utility-Theoretic Inconsistency in Stated-Preference Surveys. American Journal of Agricultural Economics **83**:1328-1333.

Johnston, R. J., K. J. Boyle, W. Adamowicz, J. Bennett, R. Brouwer, T. A. Cameron, W. M. Hanemann, N. Hanley, M. Ryan, R. Scarpa, R. Tourangeau, and C. A. Vossler. 2017. Contemporary Guidance for Stated Preference Studies. Journal of the Association of Environmental and Resource Economists **4**:319-405.

Karr, J. R. 1981. Assessment of biotic integrity using fish communities. Fisheries **66**:21-27.

Karr, J. R. 1991. Biological integrity: a long-neglected aspect of water resource management. Ecological Applications **1**:66-84.

Karr, J. R. 1999. Defining and measuring river health. Freshwater Biology **41**:221-234.

Larson, C., G. Merritt, J. Janisch, J. Lemmon, M. Rosewood-Thurman, B. Engeness, S. Polkowske, and G. Onwumere. 2019. The first statewide stream macroinvertebrate bioassessment in Washington State with a relative risk and attributable risk analysis for multiple stressors. Ecological Indicators **102**:175-185.

Larson, D. M. 1993. On Measuring Existence Value. Land Economics **69**:377-388.

Liaw, A., and M. Wiener. 2002. Classification and Regression by randomForest. R News **2**:18-22.

Madariaga, B., and K. E. McConnell. 1987. Exploring existence value. Water Resources Research **23**:936-942.

Maloney, K. O., K. P. Krause, C. Buchanan, L. E. Hay, G. J. McCabe, Z. M. Smith, T. L. Sohl, and J. A. Young. 2020. Disentangling the potential effects of land-use and climate change on stream conditions. Global Change Biology **26**:2251-2269.

Maloney, K. O., Z. M. Smith, C. Buchanan, A. Nagel, and J. A. Young. 2018. Predicting biological conditions for small headwater streams in the Chesapeake Bay watershed. Freshwater Science **37**:795-809.

Mazor, R. D., A. C. Rehn, P. R. Ode, M. Engeln, K. C. Schiff, E. D. Stein, D. J. Gillett, D. B. Herbst, and C. P. Hawkins. 2016. Bioassessment in complex environments: designing an index for consistent meaning in different settings. Freshwater Science **35**:249-271.

McKay, L., T. Bondelid, T. Dewald, J. Johnston, R. Moore, and A. Reah. 2012. NHDPlus Version 2: User Guide. (Available from: http://www.horizon-systems.com/NHDPlus/NHDPlusV2_home.php).

Møller, A. P., and M. D. Jennions. 2002. How Much Variance Can Be Explained by Ecologists and Evolutionary Biologists? Oecologia **132**:492-500.

Moore, C., D. Guignet, C. Dockins, K. B. Maguire, and N. B. Simon. 2018. Valuing Ecological Improvements in the Chesapeake Bay and the Importance of Ancillary Benefits. Journal of Benefit-Cost Analysis **9**:1-26.

Morgan, D., C. Fellows, and H. Guevara. 2008. Emergent Approaches to Focus Group Research. Pages 189-205 *in* S. Nagy Hesse-Biber and P. Leavy, editors. Handbook of emergent methods. Guilford Publications, New York, NY.

Moss, D., M. T. Furse, J. F. Wright, and P. D. Armitage. 1987. The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. Freshwater Biology **17**:41-52.

Newbold, S. C., P. J. Walsh, D. M. Massey, and J. Hewitt. 2018. Using structural restrictions to achieve theoretical consistency in benefit transfers. Environmental and Resource Economics **69**:529-553.

Ode, P. R., A. C. Rehn, R. D. Mazor, K. C. Schiff, E. D. Stein, J. T. May, L. R. Brown, D. B. Herbst, D. Gillett, K. Lunde, and C. P. Hawkins. 2016. Evaluating the adequacy of a reference-site pool for ecological assessments in environmentally complex regions. Freshwater Science **35**:237-248.

Olsen, A. R., J. Sedransk, D. Edwards, C. A. Gotway, W. Liggett, S. Rathbun, K. H. Reckhow, and L. J. Yyoung. 1999. Statistical Issues for Monitoring Ecological and Natural Resources in the United States. Environmental Monitoring and Assessment **54**:1-45.

Ostermiller, J. D., and C. P. Hawkins. 2004. Effects of sampling error on bioassessments of stream ecosystems: application to RIVPACS-type models. Journal of the North American Benthological Society **23**:363-382.

Prasad, A. M., L. R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. Ecosystems **9**:181-199.

Santhi, C., R. Srinivasan, J. G. Arnold, and J. R. Williams. 2006. A modeling approach to evaluate the impacts of water quality management plans implemented in a watershed in Texas. Environmental Modelling & Software **21**:1141-1157.

Schmidt, T. S., P. C. Van Metre, and D. M. Carlisle. 2019. Linking the Agricultural Landscape of the Midwest to Stream Health with Structural Equation Modeling. Environmental Science & Technology **53**:452-462.

Schultz, E. T., R. J. Johnston, K. Segerson, and E. Y. Besedin. 2012. Integrating Ecology and Economics for Restoration: Using Ecological Indicators in Valuation of Ecosystem Services. Restoration Ecology **20**:304-310.

Shapiro, M. H., S. M. Holdsworth, and S. G. Paulsen. 2008. The need to assess the condition of aquatic resources in the US. Journal of the North American Benthological Society **27**:808-811, 804.

Smith, R. A., G. E. Schwarz, and R. B. Alexander. 1997. Regional interpretation of water-quality monitoring data. Water Resources Research **33**:2781-2798.

Stevens, D. L., and A. R. Olsen. 2004. Spatially Balanced Sampling of Natural Resources. Journal of the American Statistical Association **99**:262-278.

Stoddard, J. L., A. T. Herlihy, D. V. Peck, R. M. Hughes, T. R. Whittier, and E. Tarquinio. 2008. A process for creating multimetric indices for large-scale aquatic surveys. Journal of the North American Benthological Society **27**:878-891.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson, and R. H. Norris. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications **16**:1267-1276.

Stribling, J. B., B. K. Jessup, and D. L. Feldman. 2008. Precision of benthic macroinvertebrate indicators of stream condition in Montana. Journal of the North American Benthological Society **27**:58-67.

USEPA (U.S. Environmental Protection Agency) Office of Water. 2015. Benefit and Cost Analysis for the Effluent Limitations Guidelines and Standards for the Steam Electric Power Generating Point Source Category (EPA-821-R-15-005). Washington, DC.

USEPA (US Environmental Protection Agency) Office of Water and Office of Research and Development. 2009. National Lakes Assessment 2007: A Collaborative Survey (EPA/841/R-09/001). Washington, DC.

USEPA (US Environmental Protection Agency) Office of Water and Office of Research and Development. 2016. National Rivers and Streams Assessment 2008-2009: A Collaborative Survey (EPA/841/R-16/007). Washington, DC.

USEPA (US Environmental Protection Agency) Office of Water and Office of Research and Development. 2017. National Lakes Assessment 2012: Technical Report (EPA841-R-16-114). Washington, DC.

Van Houtven, G., C. Mansfield, D. J. Phaneuf, R. von Haefen, B. Milstead, M. A. Kenney, and K. H. Reckhow. 2014. Combining expert elicitation and stated preference methods to value ecosystem services from improved lake water quality. Ecological Economics **99**:40-52.

Van Sickle, J., C. P. Hawkins, D. P. Larsen, and A. T. Herlihy. 2005. A null model for the expected macroinvertebrate assemblages in streams. Journal of the North American Benthological Society **24**:178-191.

Van Sickle, J., D. P. Larsen, and C. P. Hawkins. 2007. Exclusion of rare taxa affects performance of the O/E index in bioassessments. Journal of the North American Benthological Society **26**:319-331.

Vander Laan, J. J., C. P. Hawkins, J. R. Olson, and R. A. Hill. 2013. Linking land use, in-stream stressors, and biological condition to infer causes of regional ecological impairment in streams. Freshwater Science **32**:801-820.

Vinson, M. R., and C. P. Hawkins. 1996. Effects of Sampling Area and Subsampling Procedure on Comparisons of Taxa Richness among Streams. Journal of the North American Benthological Society **15**:392-399.

Walsh, P. J., and W. J. Wheeler. 2013. Water quality indices and benefit-cost analysis. Journal of Benefit-Cost Analysis **4**:81-105.

Yuan, L. L., C. P. Hawkins, and J. V. Sickle. 2008. Effects of regionalization decisions on an O/E index for the US national assessment. Journal of the North American Benthological Society **27**:892-905, 814.

Zachos, F. E. 2016. Species Concepts in Biology: Historical Development, Theoretical Foundations and Practical Relevance. Springer.

Tables and Figures

Table 1. Performances of preliminary random forest models of stream (NRSA 2013-14) and lake (NLA 2007) O/E scores. (see Figure 1 for regional boundaries). RMSE=Root Mean Square Error. Pseudo-R^2 as defined in Liaw and Wiener (2002).

Model	Regions	Psuedo-R^2	RMSE
NRSA 2013-14	Continental US	28	0.26
	Eastern highlands	25	0.26
	Plains and lowlands	25	0.27
	Western Mountains	30	0.25
NLA 2007	Continental US	30	0.23
	Eastern highlands	13	0.23
	Plains and lowlands	36	0.22
	Western Mountains	18	0.25

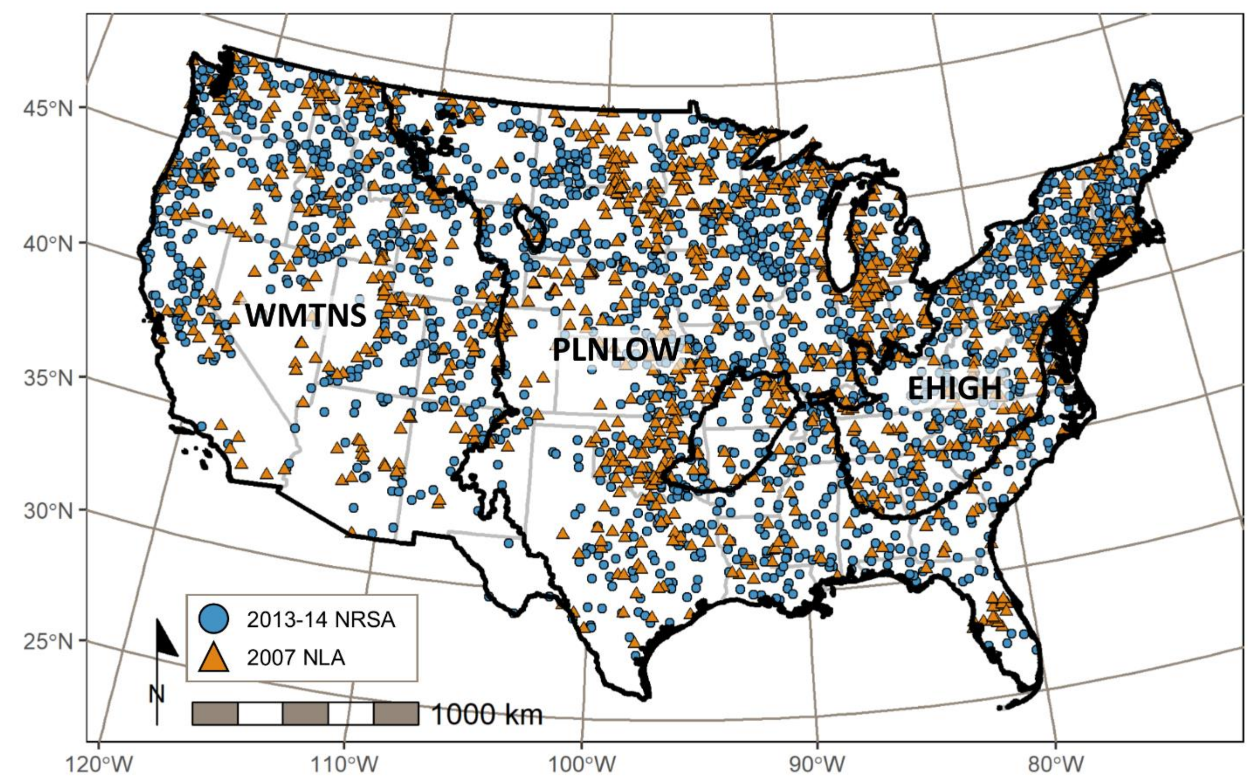

Figure 1. 2013-14 National Rivers and Streams Assessment (NRSA) and 2007 National Lakes Assessment (NLA) sample sites. The USEPA conducts surveys that include new sample locations on a 5-year cycle. Ecoregion abbreviations: Western Mountains (WMTNS), Plains and Lowlands (PLNLOW), and Eastern Highlands (EHIGH).

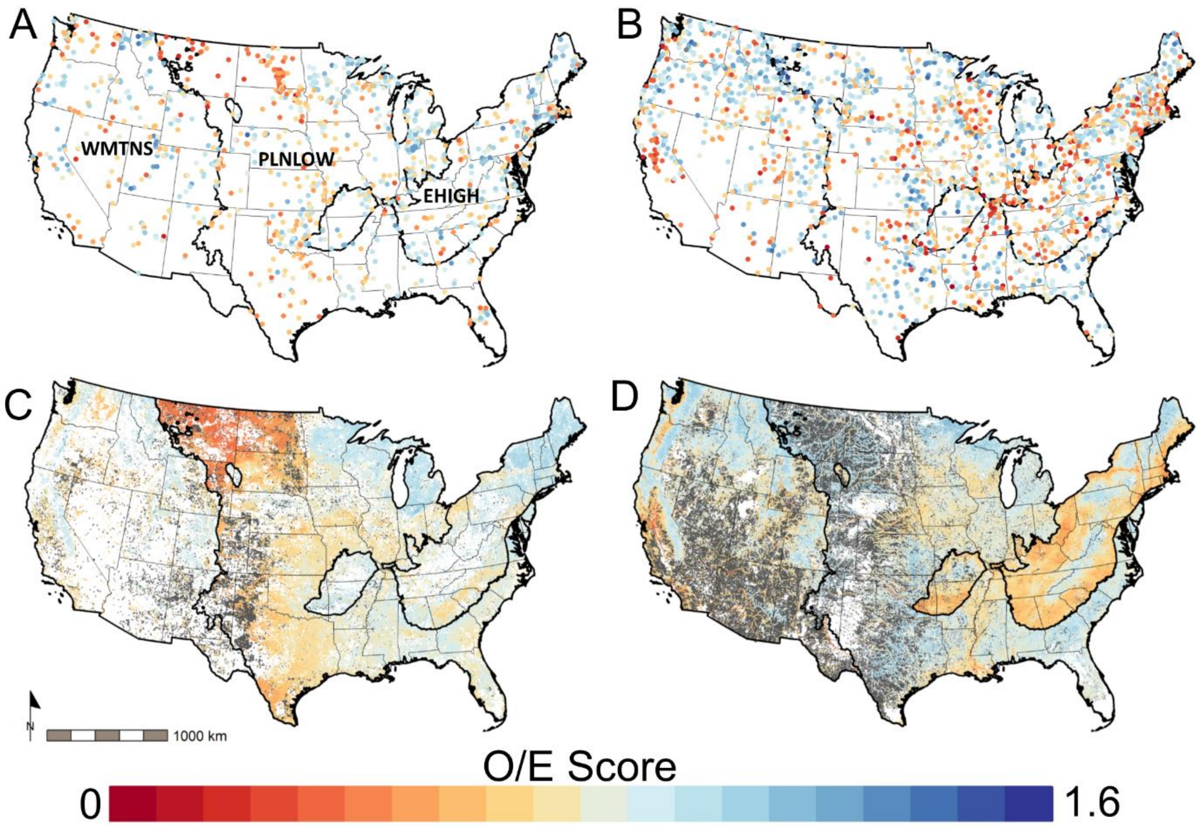

Figure 2. Distribution of sample sites with O/E scores from the (A) 2007 National Lakes Assessment (NLA) and (B) 2013-14 National Rivers and Streams Assessment (NRSA) and model interpolated O/E scores for (C) lakes and (D) streams. Separate O/E assessments for streams and lakes were developed within three ecoregions: Western Mountains (WMTNS), Plains and Lowlands (PLNLOW), and Eastern Highlands (EHIGH). Dark grey areas in (C) and (D) represent lakes or stream that are outside of the sampling frame of the EPA NLA or NRSA, respectively, and were excluded from model interpolation.

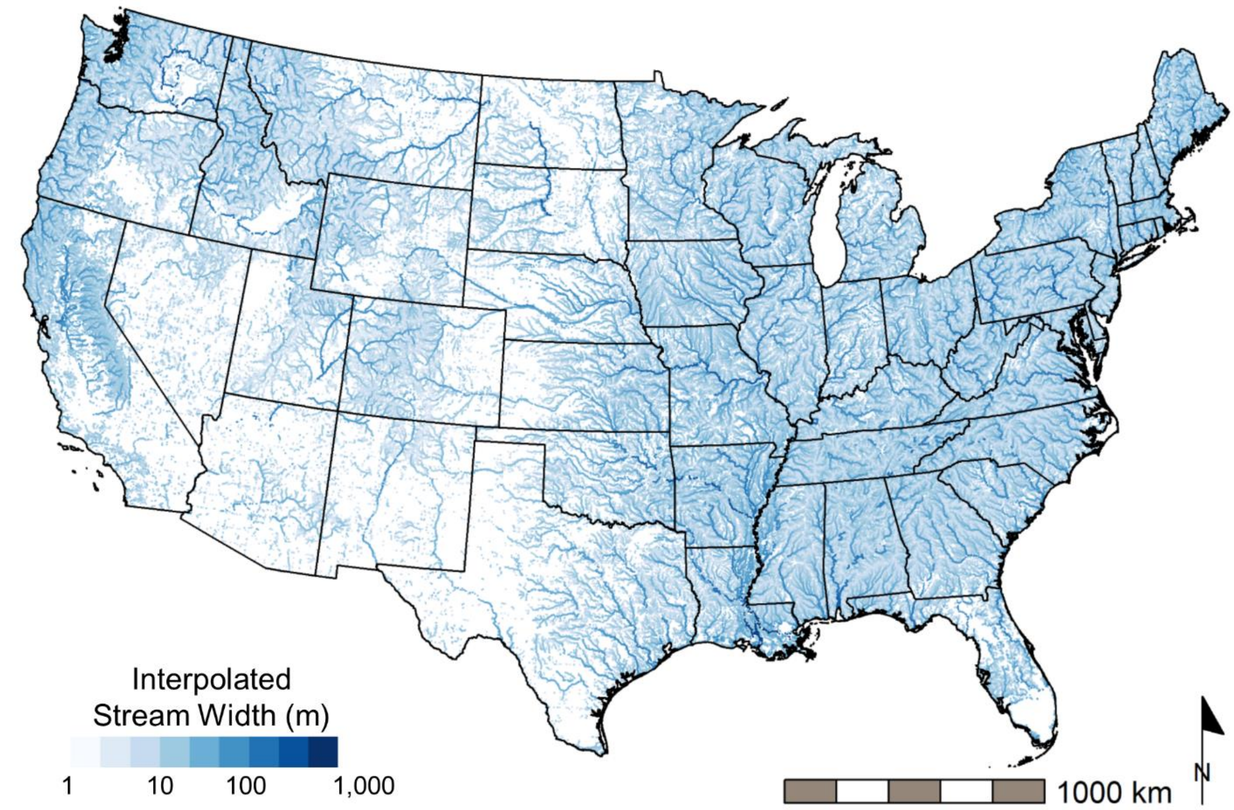

Figure 3. Stream widths (meters) interpolated to 1.1 million stream and river segments of the National Hydrography Dataset (version 2).

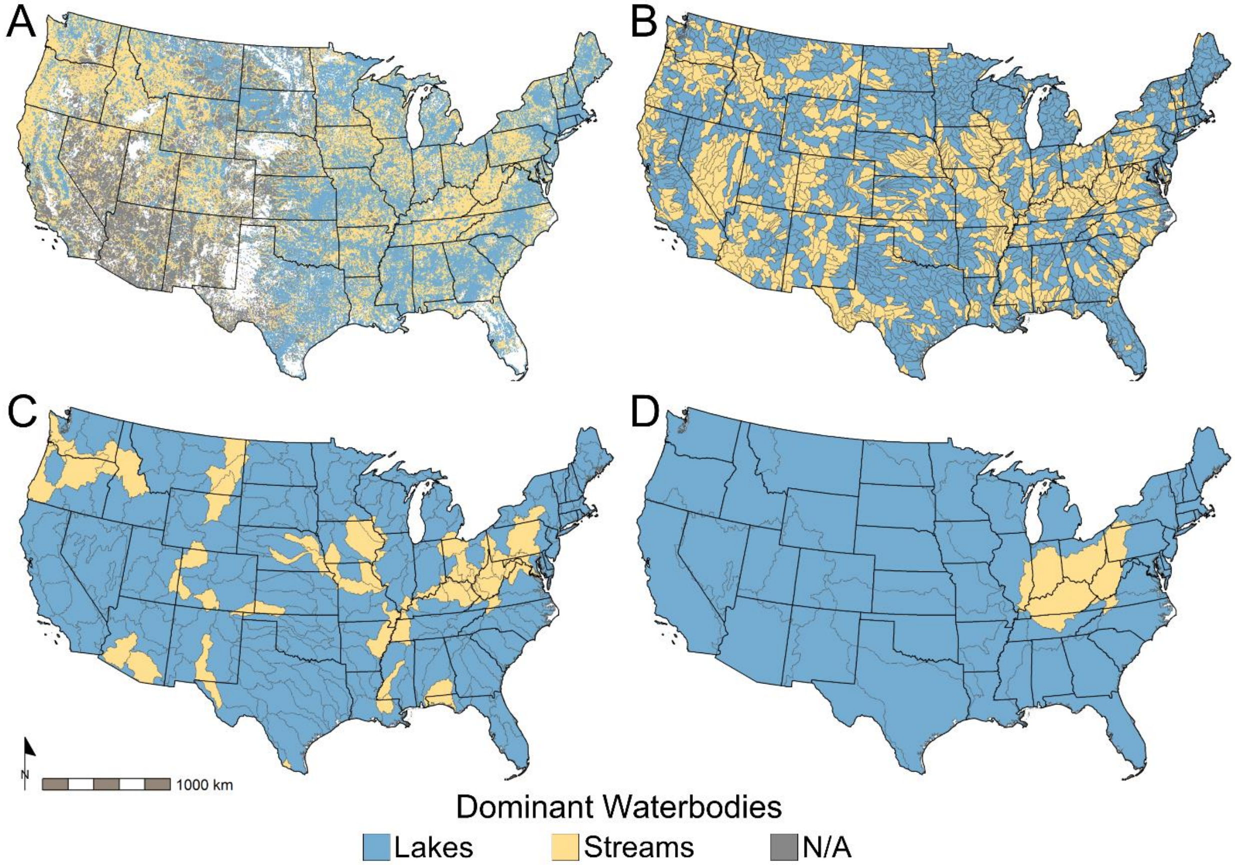

Figure 4. Dominant waterbody type by area within (A) NHDPlus catchments, (B) 8-digit HUCs, (C) 4-digit HUCs, and (D) 2-digit HUCs. Note that map (A) visually overrepresents the dominance of lakes in NHDPlus catchments because lakes are plotted on top of streams. Streams were dominant in 1.06 million catchments compared with 312,739 dominated by lakes. Grey areas in (A) and labeled “N/A” represent catchments that were excluded from model interpolations because they lacked lakes or streams within the NLA or NRSA sampling frames.

Appendix S1. QA/QC

O/E Values and Stream Widths

Data on the biological condition of streams and lakes and stream widths were obtained from the USEPA National Aquatics Resources Survey (NARS) program. This EPA program uses well established field, lab, and analytical procedures with documented QA/QC procedures. Data and complete documentation are available publicly via the NARS website: https://www.epa.gov/national-aquatic-resource-surveys. Data used in this analysis from the 2013/14 National Rivers and Streams Assessment are not yet available online, but were collected and analyzed with methods reported in the 2008/09 NRSA Technical Appendix (USEPA 2016).

StreamCat and LakeCat Data

This analysis used existing data from the StreamCat and LakeCat datasets (Hill et al. 2016, Hill et al. 2018). These datasets are available to the public via: https://www.epa.gov/national-aquatic-resource-surveys/streamcat-dataset and https://www.epa.gov/national-aquatic-resource-surveys/lakecat-dataset. These data were developed under EPA-approved Quality Assurance Project Plans. All data were processed using Python (including scripted ArcGIS tools) and R code to create national GIS layers if not already in that form and then summarized to National Hydrography Dataset version 2 catchments (McKay et al. 2012). The processing was documented in Python or R scripts so that it is repeatable and available for review for QA purposes.

All existing catchment attribute data were evaluated to ensure that each existing data source used was downloaded completely and without corruption of coordinates or attributes, including the following:

- Projections: Verify the geographic projections of all Landscape Layers (LLs) and that each existing data source has the correct input coordinate system information; e.g., convert to USGS Albers Equal Area Conic Projection.
- Units: Check the units of each LL with metadata. Convert to SI.
- Data distribution: Examine the range and distribution of values of each LL. Is the range realistic? Does the histogram have an odd distribution or outliers?
- Visual inspection: Visually inspect each LL. Are there strange gaps, edges, or other anomalous features within the raster?
- No Data: Verify how no-data values are represented in each LL. Are no-data locations represented by 0s or other numeric values (e.g., -9999)?
- Combinations: Check values of all second-order LLs, i.e., those that are based on manipulations/combinations of raw LLs.

To ensure that subsequent processing was based on the correct input data, LLs were assessed against QA steps listed above and investigate data quality concerns.

Models

Statistical models were developed using NARS National Rivers and Streams Assessment (NRSA) and National Lake Assessment (NLA) data. All NARS data are collected under an approved NARS QA plan (see above). Statistical model assessments were completed using out-of-bag data from the random forest models. Data used as model input was reviewed by comparing the results for consistency with the source data, completing range checks for each variable, and looking for multivariate outliers. All statistical models will be developed using existing R or Python software and packages. The performance of the statistical spatial predictions based on these models will be assessed using standard statistical methods, e.g., r-squared values and root mean squared error.

Literature Cited

Hill, R. A., M. H. Weber, R. M. Debbout, S. G. Leibowitz, and A. R. Olsen. 2018. The Lake-Catchment (LakeCat) Dataset: characterizing landscape features for lake basins within the conterminous USA. Freshwater Science **37**:208-221.

Hill, R. A., M. H. Weber, S. G. Leibowitz, A. R. Olsen, and D. J. Thornbrugh. 2016. The Stream-Catchment (StreamCat) Dataset: A Database of Watershed Metrics for the Conterminous United States. Journal of the American Water Resources Association (JAWRA) **52**:120-128.

McKay, L., T. Bondelid, T. Dewald, J. Johnston, R. Moore, and A. Reah. 2012. NHDPlus Version 2: User Guide. (Available from: http://www.horizon-systems.com/NHDPlus/NHDPlusV2_home.php).

USEPA (US Environmental Protection Agency) Office of Water and Office of Research and Development. 2016. National Rivers and Streams Assessment 2008-2009 Technical Report (EPA/841/R-16/008). Washington, DC.

Appendix S1. Names and descriptions of StreamCat and LakeCat predictor variables used in random forest models of observed-to-expected taxonomic composition models. Mean values (SD) of predictor variables are provided for the three ecoregions within the conterminous US used for model development: EHIGH = Eastern Highlands, PLNLOW = Plains and Lowlands, and WMTNS = Western Mountains. Some variables are available for streams and lakes whereas others are only available for streams as indicated by the O/E Model column. A subset of StreamCat variables were also used to model wetted widths (WW) of streams; Y (yes) or N (no).

		STREAMS				**LAKES**				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
AgKffactWs	Mean soil erodibility (Kf) factor (unitless) of soils within watershed on agricultural land. The Kf factor is used in the Universal Soil Loss Equation (USLE) and represents a relative index of susceptibility of bare, cultivated soil to particle detachment and transport by rainfall.	0.06 (0.06)	0.11 (0.1)	0.01 (0.03)		0.04 (0.05)	0.09 (0.09)	0.01 (0.04)	Streams, Lakes	Y
Al2O3Ws	Mean % of lithological aluminum oxide (Al2O3) content in surface or near surface geology within watershed	9.89 (3.73)	9.51 (3.21)	11.63 (2.95)		10.58 (3.95)	10.29 (3.32)	11.45 (2.76)	Streams, Lakes	N
BFIWs	Baseflow is the component of streamflow that can be attributed to ground-water discharge into streams. The Baseflow Index (BFI) is the ratio of baseflow to total flow, expressed as a percentage, within watershed.	44.04 (10.89)	41.08 (17.94)	63.54 (11.7)		44.39 (11.66)	42.68 (19.65)	61.54 (14.05)	Streams, Lakes	Y

		STREAMS				LAKES				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
CanalDensWs	Density of NHDPlus line features classified as canal, ditch, or pipeline within the upstream watershed (km/ square km)	0 (0.01)	0.02 (0.06)	0.02 (0.05)		0 (0)	0.02 (0.12)	0.04 (0.15)	Streams, Lakes	Y
CaOWs	Mean % of lithological calcium oxide (CaO) content in surface or near surface geology within watershed	7.57 (7.58)	8.99 (8.38)	6.67 (4.49)		6.72 (7.04)	6.2 (8.22)	5.43 (4.2)	Streams, Lakes	N
CatAreaSqKm	Catchment area (square km) at NHDPlus stream segment outlet, i.e., at the most downstream location of the vector line segment	2.66 (3.7)	4.94 (11.52)	6.39 (49.45)		0 (0)	0 (0)	0 (0)	Streams	N
CatAreaSqKmRp100	Catchment area (square km) within a 100-m buffer of NHD streams	0.33 (0.33)	0.5 (0.89)	0.48 (0.93)		0 (0)	0 (0)	0 (0)	Streams	N
CBNFWs	Mean rate of biological nitrogen fixation from the cultivation of crops in kg N/ha/yr, within watershed	2.64 (4.47)	10.82 (14.1)	0.81 (2.23)		1.49 (2.45)	8.07 (10.63)	1.19 (4.83)	Streams, Lakes	Y
ClayWs	Mean % clay content of soils (STATSGO) within watershed	20.1 (10.68)	24.66 (9.24)	19.09 (6.93)		19.67 (11.57)	22.28 (11.22)	18.04 (8.87)	Streams, Lakes	Y
CoalMineDensWs	Density of coal mines sites within watershed (mines/square km)	0.08 (0.39)	0.05 (0.32)	0.01 (0.05)		0.04 (0.18)	0.01 (0.14)	0 (0)	Streams, Lakes	Y
CompStrgthWs	Mean lithological uniaxial compressive strength (megaPascals) content in surface or near surface geology within watershed	103.75 (35.45)	50.13 (39.63)	100.15 (33.81)		112.58 (41.18)	34.96 (43.53)	84.88 (44.73)	Streams, Lakes	Y
DamDensWs	Density of georeferenced dams within watershed (dams/ square km) based on the National Inventory of Dams (https://catalog.data.gov/dataset/national-inventory-of-dams)	0.02 (0.06)	0.02 (0.04)	0 (0.01)		0.11 (0.26)	0.05 (0.15)	0.05 (0.24)	Streams, Lakes	Y

		STREAMS				LAKES				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
DamNIDStorWs	Total possible volume of all reservoirs (NID_STORA in NID) per unit area of watershed (cubic meters/square km) based on the National Inventory of Dams (https://catalog.data.gov/dataset/national-inventory-of-dams)	75705.09 (152446.23)	57233.92 (231998.68)	46878.1 (113821.07)		147727.6 (217310.06)	124020.98 (460555.93)	173688.76 (461114.03)	Streams, Lakes	Y
DamNrmStorWs	Normal (most common) volume of all reservoirs (NORM_STORA in NID) per unit area of watershed (cubic meters/square km) based on the National Inventory of Dams (https://catalog.data.gov/dataset/national-inventory-of-dams)	42348.84 (98735.73)	23234.15 (48809.22)	41264.11 (102988.36)		83213.61 (134198.72)	67317.06 (330127.2)	148554.58 (431422.12)	Streams, Lakes	Y
ElevWs	Mean watershed elevation (m)	374.75 (189.5)	569.82 (532.49)	1847.16 (728.15)		352 (210.25)	454.31 (397.9)	1776.78 (925.33)	Streams, Lakes	Y
Fe2O3Ws	Mean % of lithological ferric oxide (Fe_2O_3) content in surface or near surface geology within watershed	3.86 (1.76)	6.49 (4.55)	6.66 (3.13)		4.18 (2.57)	8.77 (6.89)	8.44 (5.14)	Streams, Lakes	N
FertWs	Mean rate of synthetic nitrogen fertilizer application to agricultural land in kg N/ha/yr, within watershed	5.62 (8.22)	21.18 (23.23)	1.86 (6.46)		3.71 (7.58)	16.82 (17.76)	1.78 (8.96)	Streams, Lakes	Y
HUDen2010Ws	Mean housing unit density (housing units/square km) within watershed	31.57 (68.64)	15.17 (61.08)	3.64 (14.93)		38.17 (97.5)	33.7 (102.63)	17.68 (95.4)	Streams, Lakes	Y
HUDen2010WsRp100	Mean housing unit density (housing units/square km) within watershed and within a 100-m buffer of NHD stream lines	30.62 (68.24)	13.43 (51.26)	3.59 (15.97)		0 (0)	0 (0)	0 (0)	Streams	Y

		STREAMS				LAKES				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
HydrlCondWs	Mean lithological hydraulic conductivity (micrometers per second) content in surface or near surface geology within watershed	0.93 (4.01)	23.16 (39.23)	12.68 (23.79)		1.93 (13.81)	39.01 (69.12)	24.2 (50.01)	Streams, Lakes	Y
InorgNWetDep_2008Ws	Annual gradient map of precipitation-weighted mean deposition for inorganic nitrogen wet deposition from nitrate and ammonium for 2008 in kg of N/ha/yr, within watershed	4.21 (0.89)	3.92 (1.53)	1.75 (0.81)		4.17 (1)	4.05 (1.4)	1.75 (1)	Streams, Lakes	N
K2OWs	Mean % of lithological potassium oxide (K2O) content in surface or near surface geology within watershed	1.99 (0.8)	1.83 (0.55)	2.1 (0.65)		2.15 (0.98)	1.92 (0.6)	2.08 (0.64)	Streams, Lakes	N
KffactWs	Mean soil erodibility (Kf) factor (unitless) of soils within watershed. The Kf factor is used in the Universal Soil Loss Equation (USLE) and represents a relative index of susceptibility of bare, cultivated soil to particle detachment and transport by rainfall.	0.3 (0.06)	0.29 (0.07)	0.27 (0.07)		0.29 (0.07)	0.27 (0.09)	0.28 (0.08)	Streams, Lakes	Y
ManureWs	Mean rate of manure application to agricultural land from confined animal feeding operations in kg N/ha/yr, within watershed	3.82 (10.05)	2.67 (5.25)	0.17 (0.96)		2.39 (7.4)	1.73 (4.89)	0.19 (0.79)	Streams, Lakes	Y
MgOWs	Mean % of lithological magnesium oxide (MgO) content in surface or near surface geology within watershed	2.61 (2.14)	2.88 (2.33)	3.02 (1.57)		2.76 (2.66)	2.03 (2.01)	2.53 (1.78)	Streams, Lakes	N
MineDensWs	Density of mine sites within watershed (mines/square km)	0 (0)	0 (0)	0 (0.01)		0 (0.02)	0 (0.02)	0 (0)	Streams, Lakes	Y

		STREAMS				LAKES				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
MineDensWsRp100	Density of mine sites within watershed and within 100-m buffer of NHD stream lines (mines/square km)	0 (0)	0 (0.01)	0 (0)		0 (0)	0 (0)	0 (0)	Streams	Y
Na2OWs	Mean % of lithological sodium oxide (Na2O) content in surface or near surface geology within watershed	1.4 (1.06)	1.08 (0.56)	1.98 (0.96)		1.67 (1.18)	1.18 (0.62)	1.82 (0.95)	Streams, Lakes	N
NABD_DensWs	Density of georeferenced dams within watershed (dams/ square km) based on the National Anthropogenic Barrier Dataset (https://www.sciencebase.gov/catalog/item/56a7f9dce4b0b28f1184dabd)	0.02 (0.06)	0.01 (0.07)	0 (0.01)		0.06 (0.17)	0.02 (0.07)	0.02 (0.11)	Streams, Lakes	N
NABD_NIDStorWs	Total possible volume of all reservoirs (NID_STORA in NID) per unit area of watershed (cubic meters/square km) based on the National Anthropogenic Barrier Dataset (https://www.sciencebase.gov/catalog/item/56a7f9dce4b0b28f1184dabd)	80822.59 (194162.66)	51461.74 (117233.53)	43495.35 (115151.38)		125580.91 (209343.74)	51742.3 (224956.31)	1807737.48 (24226661.75)	Streams, Lakes	N
NABD_NrmStorWs	Normal (most common) volume of all reservoirs (NORM_STORA in NID) per unit area of watershed (cubic meters/square km) based on the National Anthropogenic Barrier Dataset (https://www.sciencebase.gov/catalog/item/56a7f9dce4b0b28f1184dabd)	48146.12 (137703.27)	22717.79 (51310.57)	38127.67 (105098.11)		73403.44 (134228.48)	23512.95 (103148.64)	1454387.3 (19293353.75)	Streams, Lakes	N

		STREAMS				LAKES				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
NH4_2008Ws	Annual gradient map of precipitation-weighted mean deposition for ammonium ion concentration wet deposition for 2008 in kg of NH4/ha/yr, within watershed	2.52 (0.63)	2.97 (1.25)	1.24 (0.61)		2.5 (0.73)	3.05 (1.16)	1.23 (0.75)	Streams, Lakes	N
NO3_2008Ws	Annual gradient map of precipitation-weighted mean deposition for nitrate ion concentration wet deposition for 2008 in kg of NO3/ha/yr, within watershed	10 (2.2)	7.17 (2.89)	3.49 (1.66)		9.9 (2.38)	7.45 (2.83)	3.51 (2.04)	Streams, Lakes	N
NPDESDensWs	Density of permitted NPDES (National Pollutant Discharge Elimination System) sites within watershed (sites/square km)	0 (0)	0 (0)	0 (0)		0 (0.01)	0 (0.01)	0 (0)	Streams, Lakes	N
NPDESDensWsRp100	Density of permitted NPDES (National Pollutant Discharge Elimination System) sites within watershed and within 100-m buffer of NHD stream lines (sites/square km)	0 (0.01)	0 (0.01)	0 (0)		0 (0)	0 (0)	0 (0)	Streams	N
NWs	Mean % of lithological nitrogen (N) content in surface or near surface geology within watershed	0.06 (0.06)	0.24 (0.23)	0.11 (0.11)		0.06 (0.06)	0.24 (0.22)	0.17 (0.16)	Streams, Lakes	N
OmWs	Mean organic matter content (% by weight) of soils (STATSGO) within watershed	1.5 (1.55)	2.39 (4.28)	1.13 (0.71)		1.81 (2.56)	2.75 (4.32)	1.1 (0.8)	Streams, Lakes	Y
P2O5Ws	Mean % of lithological phosphorous oxide (P2O5) content in surface or near surface geology within watershed	0.16 (0.08)	0.18 (0.11)	0.18 (0.06)		0.16 (0.16)	0.17 (0.16)	0.17 (0.07)	Streams, Lakes	N
PctAg2006Slp10Ws	% of watershed area classified as ag land cover (NLCD 2006 classes 81-82) occurring on slopes greater than or equal to 10%	3.57 (4.77)	1.15 (3.88)	0.26 (2.05)		2.36 (4.25)	0.5 (1.59)	0.19 (0.97)	Streams, Lakes	Y

		STREAMS				LAKES				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
PctAg2006Slp20Ws	% of watershed area classified as ag land cover (NLCD 2006 classes 81-82) occurring on slopes greater than or equal to 20%	0.53 (1.12)	0.07 (0.37)	0.06 (0.58)		0.35 (1.14)	0.01 (0.03)	0.03 (0.24)	Streams, Lakes	Y
PctAgWs	% of watershed area representing the sum of NLD PctCropWs and PctHayWs (NLCD 2006 or 2011)	17.06 (17.79)	34.76 (29.58)	2.56 (7.03)		11.38 (13.65)	29.4 (26.72)	2.81 (9.36)	Streams, Lakes	Y
PctAlkIntruVolWs	% of watershed area classified as lithology type: alkaline intrusive volcanic rock	0 (0)	0.02 (0.17)	0.78 (5.37)		0 (0)	0 (0)	0.34 (2.74)	Streams, Lakes	Y
PctAlluvCoastWs	% of watershed area classified as lithology type: alluvium and fine-textured coastal zone sediment	1.56 (4.89)	11.53 (23.82)	7.11 (12.28)		0.46 (2.87)	12.15 (29.89)	12.05 (25.78)	Streams, Lakes	Y
PctBlWs	% of watershed area classified as barren land cover (NLCD 2006 or 2011)	0.32 (0.93)	0.33 (1)	2.1 (4.41)		0.4 (1.59)	0.25 (0.93)	3.98 (11.32)	Streams, Lakes	Y
PctBlWsRp100	% of watershed area classified as barren land cover (NLCD 2006 or 2011) within a 100-m buffer of NHD streams	0.25 (0.61)	0.26 (0.75)	0.93 (2.31)		0 (0)	0 (0)	0 (0)	Streams	Y
PctCarbResidWs	% of watershed area classified as lithology type: carbonate residual material	6.28 (19.79)	2.8 (11.83)	4.87 (13.23)		5.8 (20.77)	1.04 (8.98)	3.7 (12.96)	Streams, Lakes	Y
PctCoastCrsWs	% of watershed area classified as lithology type: coastal zone sediment, coarse-textured	0 (0)	0.64 (6.96)	0 (0)		0 (0)	2.14 (13.43)	0.31 (3.27)	Streams, Lakes	Y
PctColluvSedWs	% of watershed area classified as lithology type: colluvial sediment	23.83 (38.4)	8.08 (21.26)	0.48 (2.84)		22.44 (39.64)	4.33 (18.05)	0.29 (2.69)	Streams, Lakes	Y
PctConifWsRp100	% of watershed area classified as evergreen forest land cover (NLCD 2001 class 42)	9.42 (10.5)	6.88 (12.48)	48.86 (27.31)		0 (0)	0 (0)	0 (0)	Streams	Y
PctCropWsRp100	% of watershed area classified as crop land use (NLCD 2006 or 2011)	3.72 (7.58)	20.44 (24.35)	1.44 (5.62)		0 (0)	0 (0)	0 (0)	Streams	Y

		STREAMS				LAKES				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
PctDecidWsRp100	% of watershed area classified as deciduous forest land cover (NLCD 2006 or 2011) within a 100-m buffer of NHD streams	42.08 (22.24)	14.63 (17.17)	3.58 (8.49)		0 (0)	0 (0)	0 (0)	Streams	Y
PctEolCrsWs	% of watershed area classified as lithology type: eolian sediment, coarse-textured (sand dunes)	0.04 (0.76)	2.77 (11.19)	0.26 (1.17)		0.12 (1.68)	5.59 (21.32)	0.95 (9.45)	Streams, Lakes	Y
PctEolFineWs	% of watershed area classified as lithology type: eolian sediment, fine-textured (glacial loess)	0 (0.1)	4.15 (16.15)	0.4 (3.32)		0 (0)	2.11 (11.94)	0.59 (6.81)	Streams, Lakes	Y
PctExtruVolWs	% of watershed area classified as lithology type: extrusive volcanic rock	0 (0.01)	0.01 (0.19)	3.43 (12.17)		0 (0)	0.01 (0.28)	6.04 (20.68)	Streams, Lakes	Y
PctForestWs	% of watershed area representing the sum of NLD PctDecidWs, PctMxFstWs, and PctConfirWs (NLCD 2006 or 2011)	62.91 (0.02)	21.32 (0.13)	52.23 (0.51)		63.65 (20.91)	17.63 (22.07)	45 (29.42)	Streams, Lakes	Y
PctFireLossSum2yrWs	% Forest loss to fire (fire perimeter) between (2009-2010 or 2006-2007) within watershed	0 (0.01)	0.01 (0.17)	0.14 (0.48)		0.01 (0)	0 (0)	1.63 (0)	Streams, Lakes	N
PctFireSum2yrWsRp100	% Forest loss to fire (fire perimeter) between (2009-2010 or 2006-2007) within watershed and within 100-m buffer of NHD stream lines	0 (22.1)	0.01 (21.75)	0.12 (25.77)		0 (0)	0 (0)	0 (0)	Streams	N
PctFrstLossSum5yrWs	% Average forest cover loss (Tree canopy cover change) for 5 years (2009 - 2013 or 2003 - 2007) within watershed	1.04 (1.83)	1.09 (2.72)	1.71 (4.65)		1.14 (2.05)	0.65 (2.16)	1.45 (2.85)	Streams, Lakes	Y
PctFrstLossSum5yrWsRp 100	% Average forest cover loss (Tree canopy cover change) for 5 years (2009 - 2013 or 2003 - 2007) within watershed and within 100-m buffer of NHD stream lines	0.74 (1.54)	0.78 (1.9)	1.35 (4.68)		0 (0)	0 (0)	0 (0)	Streams	Y

		STREAMS				LAKES				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
PctGlacLakeCrsWs	% of watershed area classified as lithology type: glacial outwash and glacial lake sediment, coarse-textured	5.06 (13.41)	6.63 (16.57)	0.94 (6.64)		5.39 (17.18)	11.73 (27.29)	2.85 (14.04)	Streams, Lakes	Y
PctGlacLakeFineWs	% of watershed area classified as lithology type: glacial lake sediment, fine-textured	1.59 (7.7)	1.84 (8.55)	0.24 (1.79)		0.69 (5.57)	0.45 (4.42)	1.76 (11.83)	Streams, Lakes	Y
PctGlacTilClayWs	% of watershed area classified as lithology type: glacial till, clayey	1.69 (10.07)	2.97 (13.72)	0 (0)		0.64 (4.99)	3.46 (16.8)	0 (0)	Streams, Lakes	Y
PctGlacTilCrsWs	% of watershed area classified as lithology type: glacial till, coarse-textured	23.08 (37.69)	0.82 (7.54)	3.02 (9.81)		28.84 (42.68)	1.82 (13.06)	5.66 (15.56)	Streams, Lakes	Y
PctGlacTilLoamWs	% of watershed area classified as lithology type: glacial till, loamy	16.68 (31.76)	23.46 (35)	0.5 (5.58)		13.42 (32.51)	33.9 (43.61)	3.11 (15.88)	Streams, Lakes	Y
PctGrsWs	% of watershed area classified as grassland/herbaceous land cover (NLCD 2006 or 2011)	2.04 (3.87)	20.88 (25.3)	9.81 (10.87)		1.98 (4.31)	20.26 (27.1)	8.5 (13.36)	Streams, Lakes	Y
PctGrsWsRp100	% of watershed area classified as grassland/herbaceous land cover (NLCD 2006 or 2011) within a 100-m buffer of NHD streams	1.76 (3.53)	21.06 (25.32)	8.39 (10.41)		0 (0)	0 (0)	0 (0)	Streams	Y
PctHayWsRp100	% of watershed area classified as hay land use (NLCD 2006 or 2011)	11.11 (12.63)	7.91 (10.69)	1.55 (2.9)		0 (0)	0 (0)	0 (0)	Streams	N
PctHbWetWs	% of watershed area classified as herbaceous wetland land cover (NLCD 2006 or 2011)	0.3 (0.46)	1.16 (2.56)	0.47 (1.34)		0.38 (0.96)	2.08 (3.58)	0.6 (1.57)	Streams, Lakes	Y
PctHbWetWsRp100	% of watershed area classified as herbaceous wetland land cover (NLCD 2006 or 2011) within a 100-m buffer of NHD streams	1.1 (1.69)	2.71 (5.55)	1.25 (2.89)		0 (0)	0 (0)	0 (0)	Streams	Y

		STREAMS				LAKES				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
PctHydricWs	% of watershed area classified as lithology type: hydric, peat and muck	0.02 (0.15)	0.92 (5.8)	0 (0)		0.01 (0.07)	0.62 (5.97)	0 (0)	Streams, Lakes	
PctIceWs	% of watershed area classified as ice/snow land cover (NLCD 2006 or 2011)	0 (0)	0.02 (0.19)	0.45 (2.5)		0 (0)	0 (0)	0.28 (1.54)	Streams, Lakes	Y
PctIceWsRp100	% of watershed area classified as ice/snow land cover (NLCD 2006 or 2011) within a 100-m buffer of NHD streams	0 (0)	0.01 (0.2)	0.22 (2.08)		0 (0)	0 (0)	0 (0)	Streams	Y
PctImpWs	Mean imperviousness of anthropogenic surfaces (NLCD 2006 or 2011) within watershed	2.12 (4.43)	1.45 (4.18)	0.39 (1.03)		2.17 (4.49)	2.53 (6.55)	1.17 (4.34)	Streams, Lakes	Y
PctImpWsRp100	Mean imperviousness of anthropogenic surfaces (NLCD 2006 or 2011) within watershed and within a 100-m buffer of NHD stream lines	1.8 (3.43)	1.01 (2.96)	0.48 (1.26)		0 (0)	0 (0)	0 (0)	Streams	Y
PctMxFstWsRp100	% of watershed area classified as mixed deciduous/evergreen forest land cover (NLCD 2006 or 2011)	8.8 (10.98)	2.12 (5.25)	1.76 (5.58)		0 (0)	0 (0)	0 (0)	Streams	Y
PctNonAgIntrodManagVegWs	% Nonagriculture nonnative introduced or managed vegetation landcover type reclassed from LANDFIRE Existing Vegetation Type (EVT), within watershed	6.53 (15.44)	9.66 (18.19)	2.26 (4.38)		1.61 (5.91)	4.8 (10.75)	2.26 (6.6)	Streams, Lakes	Y
PctNonAgIntrodManagVegWsRp100	% Nonagriculture nonnative introduced or managed vegetation landcover type reclassed from LANDFIRE Existing Vegetation Type (EVT), within catchment and within 100-m buffer of NHD stream lines	5.52 (13.39)	7.99 (15.55)	2.56 (5.07)		0 (0)	0 (0)	0 (0)	Streams	Y

		STREAMS				LAKES				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
PctNonCarbResidWs	% of watershed area classified as lithology type: non-carbonate residual material	11.71 (27.03)	29.6 (38.1)	33.56 (34.41)		10.65 (26.03)	19.68 (36.73)	24.46 (34.47)	Streams, Lakes	Y
PctOwWs	% of watershed area classified as open water land cover (NLCD 2006 or 2011)	1.25 (1.93)	1.39 (2.63)	0.46 (0.92)		6.7 (6.78)	10.72 (12.57)	6.22 (9.52)	Streams, Lakes	Y
PctOwWsRp100	% of watershed area classified as open water land cover (NLCD 2006 or 2011) within a 100-m buffer of NHD streams	0.27 (0.48)	0.22 (0.44)	0.11 (0.34)		0 (0)	0 (0)	0 (0)	Streams	Y
PctSalLakeWs	% of watershed area classified as lithology type: saline like sediment	0 (0)	0 (0.03)	0.13 (1)		0 (0)	0 (0.01)	1.43 (9.6)	Streams, Lakes	Y
PctShrbWs	% of watershed area classified as shrub/scrub land cover (NLCD 2006 or 2011)	2.58 (3.52)	8.08 (14.34)	28.45 (23.03)		2 (4.88)	3.79 (10.33)	27.77 (26.42)	Streams, Lakes	Y
PctShrbWsRp100	% of watershed area classified as shrub/scrub land cover (NLCD 2006 or 2011) within a 100-m buffer of NHD streams	2.15 (3.54)	7.89 (14.69)	27.33 (23.4)		0 (0)	0 (0)	0 (0)	Streams	Y
PctSilicicWs	% of watershed area classified as lithology type: silicic residual material	7.7 (23.02)	3.06 (11.42)	42.72 (36.79)		11.27 (30.07)	0.36 (3.37)	35.77 (39.89)	Streams, Lakes	Y
PctUrbHiWsRp100	% of watershed area classified as developed, high-intensity land use (NLCD 2006 or 2011) within a 100-m buffer of NHD streams	0.21 (0.79)	0.11 (0.68)	0.02 (0.08)		0 (0)	0 (0)	0 (0)	Streams	Y
PctUrbLoWsRp100	% of watershed area classified as developed, low-intensity land use (NLCD 2006 or 2011) within a 100-m buffer of NHD streams	1.98 (3.56)	1.23 (3.76)	0.46 (1.4)		0 (0)	0 (0)	0 (0)	Streams	Y
PctUrbMdWsRp100	% of watershed area classified as developed, medium-intensity land use (NLCD 2006 or 2011) within a 100-m buffer of NHD streams	0.91 (2.33)	0.4 (1.83)	0.17 (1.09)		0 (0)	0 (0)	0 (0)	Streams	Y

		STREAMS				LAKES				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
PctUrbOpWs	% of watershed area classified as developed, open space land use (NLCD 2011 class 21)	5.94 (7.19)	3.99 (4.24)	0.99 (1.68)		6.56 (8.23)	5.31 (5.23)	1.74 (3.82)	Streams, Lakes	Y
PctUrbOpWsRp100	% of watershed area classified as developed, open space land use (NLCD 2011 class 21) within a 100-m buffer of NHD streams	6.47 (7.12)	3.65 (4.58)	1.47 (2.56)		0 (0)	0 (0)	0 (0)	Streams	Y
PctUrbWs	% of watershed area representing the sum of NLD PctUrbLoWs, PctUrbMdWs, and PctUrbHiWs (NLCD 2006 or 2011)	3.74 (8.19)	2.52 (7.63)	0.56 (1.94)		3.91 (8.42)	4.73 (12.34)	2.08 (8.57)	Streams, Lakes	Y
PctWaterWs	% of watershed area classified as lithology type: water	0.39 (1.23)	0.29 (1.06)	0.25 (0.76)		0.26 (1.15)	0.6 (3.99)	0.71 (3.15)	Streams, Lakes	Y
PctWdWetWs	% of watershed area classified as woody wetland land cover (NLCD 2006 or 2011)	3.49 (5.19)	5.14 (10.05)	0.59 (0.85)		3.04 (5.22)	5.83 (10.11)	1.03 (2.43)	Streams, Lakes	Y
PctWdWetWsRp100	% of watershed area classified as woody wetland land cover (NLCD 2006 or 2011) within a 100-m buffer of NHD streams	9.69 (12.39)	10.46 (15.42)	2.39 (3.75)		0 (0)	0 (0)	0 (0)	Streams	Y
PermWs	Mean permeability (cm/hour) of soils (STATSGO) within watershed	7.06 (4.23)	7.05 (6.23)	8.28 (5.27)		7.15 (4.76)	9.58 (9.67)	9.27 (6.91)	Streams, Lakes	Y
Pestic97Ws	Mean pesticide use (kg/km2) in yr. 1997 within watershed	19.61 (39.09)	58.34 (84.62)	22.99 (163.31)		13.33 (41.78)	76.66 (240.71)	18.86 (98.85)	Streams, Lakes	
PopDen2010Ws	Mean population density (people/square km) within watershed	74.95 (170.28)	36 (148.08)	8.24 (39.49)		86.44 (215.77)	75.83 (227.49)	35.45 (195.47)	Streams, Lakes	Y
PopDen2010WsRp100	Mean population density (people/square km) within watershed and within a 100-m buffer of NHD stream lines	72.67 (169.41)	31.82 (125.25)	7.88 (39.54)		0 (0)	0 (0)	0 (0)	Streams	Y
Precip0809Ws	PRISM climate data - Average mean precipitation (mm) within the watershed. Period: 2008 + 2009	1303.96 (203.42)	906.13 (386.73)	857.16 (559.28)		1340.88 (186.73)	904.58 (347.5)	798.95 (494.09)	Streams, Lakes	Y

		STREAMS				LAKES				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
Precip8110Ws	PRISM climate data - 30-year normal mean precipitation (mm): Annual period: 1981-2010 within the watershed	1188.78 (136.6)	824.15 (309.03)	950.96 (652.58)		1203.72 (137.62)	839.12 (289.69)	870.02 (542.6)	Streams, Lakes	Y
RckDepWs	Mean depth (cm) to bedrock of soils (STATSGO) within watershed	125.44 (18.02)	135.28 (22.47)	113.71 (21.65)		123.29 (21.7)	141.53 (20.46)	120.46 (25.3)	Streams, Lakes	Y
RdCrsSlpWtdWs	Density of roads-stream intersections (2010 Census Tiger Lines-NHD stream lines) multiplied by NHDPlusV21 slope within watershed (crossings*slope/square km)	50.83 (215.69)	237.06 (1347.74)	145.51 (538.94)		0 (0)	0 (0)	0 (0)	Streams	Y
RdCrsWs	Density of roads-stream intersections (2010 Census Tiger Lines-NHD stream lines) within watershed (crossings/square km)	0 (0)	0 (0)	0 (0)		0.34 (0.39)	0.2 (0.3)	0.25 (0.33)	Streams, Lakes	Y
RdDensWs	Density of roads (2010 Census Tiger Lines) within watershed (km/square km)	2.01 (1.46)	1.52 (1.17)	1 (0.74)		2.28 (1.76)	1.93 (1.76)	1.36 (1.6)	Streams, Lakes	Y
RdDensWsRp100	Density of roads (2010 Census Tiger Lines) within watershed and within a 100-m buffer of NHD stream lines (km/square km)	2.08 (1.44)	1.45 (1.07)	1.1 (0.78)		0 (0)	0 (0)	0 (0)	Streams	Y
RunoffWs	Mean runoff (mm) within watershed	563.15 (135.74)	219.22 (170.33)	439.96 (574.2)		563.15 (133.67)	221.33 (169.74)	406.37 (540.75)	Streams, Lakes	Y
SandWs	Mean % sand content of soils (STATSGO) within watershed	31.46 (11.79)	32.83 (18.9)	38.73 (11.52)		32.41 (12.45)	38.53 (23.86)	40.44 (13.69)	Streams, Lakes	Y
SiO2Ws	Mean % of lithological silicon dioxide (SiO2) content in surface or near surface geology within watershed	57.45 (12.21)	51.34 (12.87)	56.2 (8.71)		58.51 (12.57)	52.65 (13.47)	55.23 (9.55)	Streams, Lakes	N
SLOPE	NHD Flowline Slope	0.01 (0.02)	0 (0.01)	-20.26 (450.29)		0 (0)	0 (0)	0 (0)	Streams	Y

		STREAMS				LAKES				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
SN_2008Ws	Annual gradient map of precipitation-weighted mean deposition for average sulfur & nitrogen wet deposition for 2008 in kg of S+N/ha/yr, within watershed	594.55 (118.81)	448.37 (189.25)	179.33 (82.51)		589.58 (136.12)	468.2 (179.68)	178.41 (96.52)	Streams, Lakes	N
SuperfundDensWs	Density of Superfund sites within watershed (sites/square km)	0 (0.03)	0 (0.01)	0 (0)		0 (0.02)	0 (0.02)	0 (0)	Streams, Lakes	N
SuperfundDensWsRp100	Density of Superfund sites within watershed and within a 100-m buffer of NHD stream lines (sites/square km)	0 (0.02)	0 (0)	0 (0)		0 (0)	0 (0)	0 (0)	Streams	N
SWs	Mean % of lithological sulfur (S) content in surface or near surface geology within watershed	0.26 (0.61)	0.33 (0.65)	0.15 (0.34)		0.17 (0.23)	0.29 (1.05)	0.11 (0.17)	Streams, Lakes	Y
Tmax8110Ws	PRISM climate data - 30-year normal maximum temperature (°C): Annual period: 1981-2010 within the watershed	15.91 (3.79)	16.73 (4.91)	12.61 (3.71)		27.07 (4.63)	25.68 (7.97)	22.92 (6)	Streams, Lakes	Y
Tmean0809Ws	PRISM climate data - Average mean temperature (°C) within the watershed. Period: 2008 + 2009	9.8 (3.48)	9.69 (5.14)	6.2 (3.47)		9.92 (3.59)	10.01 (5.41)	6.48 (4.28)	Streams, Lakes	Y
Tmean8110Ws	PRISM climate data - 30-year normal mean temperature (°C): Annual period: 1981-2010 within the watershed	10.03 (3.57)	10.4 (4.83)	6.15 (3.47)		10.14 (3.69)	10.68 (5.01)	6.49 (4.36)	Streams, Lakes	Y
Tmin8110Ws	PRISM climate data - 30-year normal minimum temperature (°C): Annual period: 1981-2010 within the watershed	-5.85 (6.37)	-8.03 (7.92)	-7.67 (5.31)		-7.15 (6.01)	-4.5 (8.61)	-7.04 (6.73)	Streams, Lakes	Y
TRIDensWs	Density of TRI (Toxic Release Inventory) sites within watershed (sites/square km)	0.01 (0.06)	0.01 (0.03)	0 (0)		0.01 (0.03)	0.02 (0.12)	0 (0.01)	Streams, Lakes	N

		STREAMS				LAKES				
Predictor Variables	**Predictor Description**	**EHIGH**	**PLNLOW**	**WMTNS**		**EHIGH**	**PLNLOW**	**WMTNS**	**O/E Model**	**WW Model**
TRIDensWsRp100	Density of TRI (Toxic Release Inventory) sites within watershed and within a 100-m buffer of NHD stream lines (sites/square km)	0.01 (0.04)	0.01 (0.12)	0 (0)		0 (0)	0 (0)	0 (0)	Streams	N
WetIndexWs	Mean Composite Topographic Index (CTI) [Wetness Index] within watershed	751.55 (78.65)	868.05 (109.26)	688.42 (73.83)		759.04 (123.67)	829.87 (171.56)	732.53 (177.33)	Streams, Lakes	N
WsAreaSqKm	Watershed area (square km) at NHDPlus stream segment outlet, i.e., at the most downstream location of the vector line segment	4174.57 (16612.58)	58951.73 (297077.06)	15301.23 (59097.46)		671.48 (3491.53)	764.62 (5218.86)	2317.24 (14589.83)	Streams, Lakes	Y
WsAreaSqKmRp100	Watershed area (square km) within a 100-m buffer of NHD streams	670.48 (2829.15)	8978.64 (47091.48)	2195.9 (8446.51)		0 (0)	0 (0)	0 (0)	Streams	Y
WtDepWs	Mean seasonal water table depth (cm) of soils (STATSGO) within watershed	131.99 (35.71)	139.53 (42.37)	176.06 (10.18)		133.93 (35.7)	134.19 (39.01)	171.57 (21.93)	Streams, Lakes	Y