

## Research Article

# Modeling Aquatic Macroinvertebrate Richness Using Landscape Attributes

**Marcia S. Meixler and Mark B. Bain**

*Department of Ecology, Evolution and Natural Resources, Rutgers University, 14 College Farm Road, New Brunswick, NJ 08901, USA*

Correspondence should be addressed to Marcia S. Meixler; [meixler@aesop.rutgers.edu](mailto:meixler@aesop.rutgers.edu)

Received 4 November 2014; Revised 24 December 2014; Accepted 30 December 2014

Academic Editor: Jean-Guy Godin

Copyright © 2015 M. S. Meixler and M. B. Bain. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We used a rapid, repeatable, and inexpensive geographic information system (GIS) approach to predict aquatic macroinvertebrate family richness using the landscape attributes stream gradient, riparian forest cover, and water quality. Stream segments in the Allegheny River basin were classified into eight habitat classes using these three landscape attributes. Biological databases linking macroinvertebrate families with habitat classes were developed using life habits, feeding guilds, and water quality preferences and tolerances for each family. The biological databases provided a link between fauna and habitat enabling estimation of family composition in each habitat class and hence richness predictions for each stream segment. No difference was detected between field collected and modeled predictions of macroinvertebrate families in a paired *t*-test. Further, predicted stream gradient, riparian forest cover, and total phosphorus, total nitrogen, and suspended sediment classifications matched observed classifications much more often than by chance alone. High gradient streams with forested riparian zones and good water quality were predicted to have the greatest macroinvertebrate family richness and changes in water quality were predicted to have the greatest impact on richness. Our findings indicate that our model can provide meaningful landscape scale macroinvertebrate family richness predictions from widely available data for use in focusing conservation planning efforts.

## 1. Introduction

The value of high biodiversity areas for resistance to disturbance [1], increased productivity [2], resource utilization [3–6], and rare species protection [7] has been well documented. However, biodiversity has declined over the last few decades [8–10] prompting responses from many global and regional institutions [11]. Internationally, the Convention on Biological Diversity (CBD) and European Union (EU) set targets to stem the current biodiversity loss by 2010; however, those targets have not been achieved [12]. Recent assessments show that biodiversity continues to be under serious pressure and that the policy response, though sometimes regionally successful, is not currently sufficient to stop the decline [12]. Regionally, many conservation groups have set targets for species and habitat restoration and protection with the goal of maintaining or increasing biodiversity (e.g., [13–15]). Further, community level planning, as opposed to single species conservation, has begun to gain traction as a viable alternative in the biodiversity arena [16, 17]. Thus, there is a clear need for models that can identify the locations of high biodiversity

in watersheds, compare aquatic richness distributions among regions, and provide watershed-scale information useful for informing conservation decisions.

Progress toward meeting these goals has been slowed by vast scientific challenges, especially in aquatic environments. Watersheds can span large land areas, encompass a connected range of streams sizes, and integrate natural and altered habitat properties, making aquatic environments particularly difficult to model. Further, to be most useful, aquatic biodiversity assessment models need to be rapidly developed, cost-efficient, and accurate [18–20]. Predictive models based on landscape variables derived using geographic information systems (GIS) are thus useful [21] for directing management and conservation efforts by natural resource planners [18, 22] especially in areas where field data collections have not been completed or are difficult to perform [23–25].

The use of GIS models to predict diversity of biotic communities has been well documented for terrestrial environments (e.g., [16, 26, 27]) and is starting to receive more attention in aquatic environments [28–30] though historically far less attention has been paid to the development

TABLE 1: Habitat classification criteria.

	Categories	High	Low
Stream gradient	Criteria	>1.5/1000 slope	≤1.5/1000 slope
Riparian forest cover	Categories	Closed	Open
	Criteria	>50% forest cover in 30 m buffer	≤50% forest cover in 30 m buffer or >3000 km <sup>2</sup> drainage area
Water quality	Categories	Suitable for life support	Biologically stressed
	Criteria	Stream segment is above EPA criteria thresholds for all three pollutants	Stream segment is below EPA criteria threshold for any of the three pollutants

of models for diversity prediction in aquatic communities [31, 32]. Most of the existing aquatic classification efforts have extensive data requirements (e.g., [33, 34]), are hierarchical (e.g., [35, 36]), use data that are difficult to obtain (e.g., [37]), or are based on only a single landscape attribute (e.g., [38–40]). Further, few studies link life history preferences of aquatic organisms to habitat classes for predictive modeling [17, 29].

Benthic macroinvertebrates account for much of the biodiversity in stream ecosystems as they are diverse in terms of number of species and possess numerous functional roles in aquatic environments [21, 41, 42]. Further, they have been positively correlated with other measures of biodiversity [21, 43] making them an ideal group for stream biotic richness prediction as an indicator of overall biodiversity. Family-level bioassessment data have been used quite extensively for conservation planning and have been shown to be well correlated with species- and genus-level data [44–47]. Family-level bioassessment data also have advantages in terms of cost reduction and minimal expertise needed for taxonomic identification.

Thus, the aim of this study was to develop and test a rapid, repeatable, and inexpensive GIS model for macroinvertebrate family richness prediction in riverine environments using commonly available digital data. We discuss the accuracy, strengths, and weaknesses of the model and identify some of the broad implications of using this model for identification of areas for further study and possible future conservation efforts.

## 2. Methods

**2.1. Model Design.** We developed and tested our macroinvertebrate family richness model in the upper Allegheny River basin in western New York State. The upper Allegheny River basin comprises approximately 4870 km<sup>2</sup> of first to fourth order streams north of the Pennsylvania-New York State line (Figure 1) in Chautauqua, Cattaraugus, and Allegany counties. This region provides an ideal study location as biotic diversity is known to be high and digital data are readily available for model development. In addition, the region supports a variety of land uses and land cover types including agricultural farming (crop and dairy, 25%), residential and urban development (1.5%), primary and secondary growth forest (67%), and wetlands and lakes (6%, as measured from the National Land Cover Database (NLCD) 2011).

Our macroinvertebrate family richness model was developed by predicting habitat classes for stream segments and

then defining habitat-family relations to link biota with habitat classes. We could thus predict richness in each stream segment (Figure 2). The first step in our modeling process was to define stream segments as a stream or river reach from tributary confluence to tributary confluence on the United States Environmental Protection Agency River Reach File Version 3.0 at the 1:100,000 scale (USEPA RF3; [48]). Then, we categorized each stream segment as one of eight habitat classes using three landscape attributes: stream gradient (a surrogate for substrate), riparian forest cover, and water quality. These attributes were chosen for their influence on macroinvertebrate community composition, availability of data, and ability to link to biota.

**2.2. Habitat Classification.** Habitat classification was performed by predicting stream gradient, riparian forest cover, and water quality for each stream segment in our study area. These predictions were then used to classify each stream segment into one of eight habitat classes.

The first of the three landscape attributes to be modeled was stream gradient. Stream gradient was used to classify stream segments into two groups likely to differ in substrate composition: (1) coarse substrate, defined as gravel, pebble, cobble, and boulders, present in high slope streams, and (2) fine substrate, defined as sand, silt, and clay, found in low slope streams. The division between high and low gradient streams was obtained by plotting stream gradient and dominant substrate values from 50 sites throughout the United States [49]. Most fine substrate sites had slopes less than 1.5/1000 m while coarse substrate sites generally had higher slopes (Table 1). A one-tailed sign test of matched field collected dominant substrate and stream gradient data ( $N = 38$ ) from the Allegheny River basin verified the high likelihood that the 1.5/1000 m cutoff was appropriate for delineating stream substrate based on gradient ( $P = 0.0002$ ). Since the data used to determine this cutoff were drawn from a national study [49], we support the broad use of this cutoff for future studies without further field verification.

Gradient-substrate relations were directed at identifying stream segments with coarse substrate important in providing attachment sites and microconditions for many aquatic macroinvertebrates [50]. Substrate dominated by fine sediment is often unstable habitat and is known to support a reduced density and diversity of macroinvertebrate taxa [50]. However, some species prefer such habitats and community structure may thus differ drastically between habitats of coarse and fine particles [51, 52].

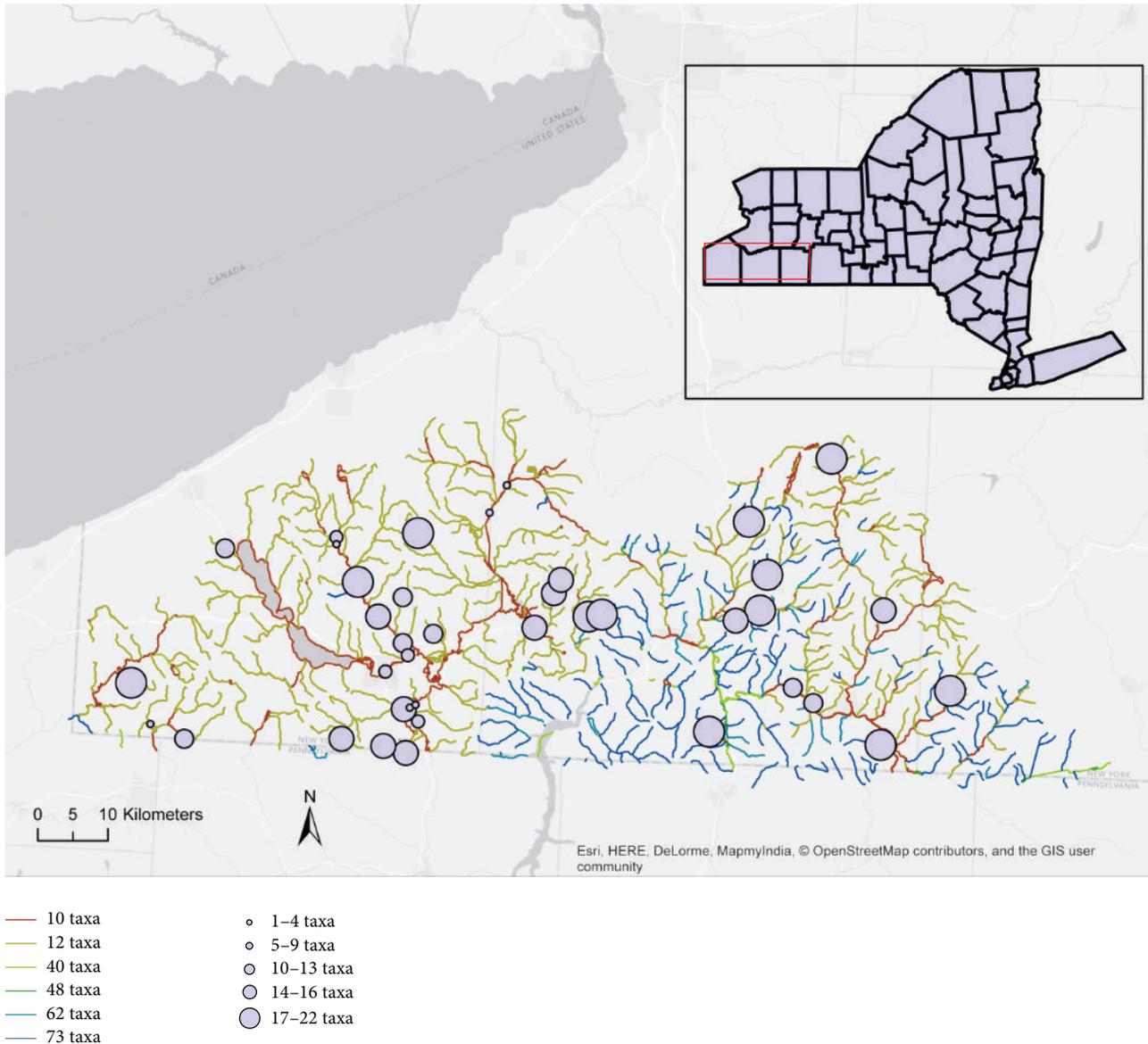


FIGURE 1: Predicted (streams) and field collected (dots) macroinvertebrate family richness in the upper Allegheny River basin of western New York.

Stream gradient classification was calculated in a GIS by obtaining elevations for the furthest upstream and downstream points of each stream segment from digital elevation models (DEMs) and dividing by the stream length between the two points. Stream gradient values were grouped into “high” or “low” categories using the 1.5/1000 m slope classification criterion. Computations were performed on DEMs at the 1:24,000 scale since they provided the most precise measurements of elevation for our study area.

The second of the three landscape attributes to be modeled was riparian forest cover. Riparian forest cover is considered important as a source of coarse particulate organic matter and as an impediment to in-stream primary production by providing shade [50]. Small to midsize streams without closed canopies can be considered degraded because lack of dense riparian vegetation is often associated with

higher runoff, bank destabilization, and human land use such as urban development, agriculture, and grazing [50]. We classified riparian forest cover in a GIS by calculating the percentage of forested land (NLCD 2001) in a 30-meter buffer [53, 54] surrounding each stream segment. Streams with >50% canopy cover were classified as “closed” canopy and streams with ≤50% canopy cover were classified as “open” canopy ([55]; Table 1). Large size streams, with total cumulative drainage area of >3000 km<sup>2</sup> [56], would be too wide for riparian vegetation to reduce primary productivity and thus were also classified as “open” canopy.

The last of the three landscape attributes to be modeled was nonpoint source (NPS) pollution. NPS pollution caused by agricultural land use is a primary source of stream impairment in the United States, and elevated sedimentation is a principal pollutant causing stream degradation [57].

TABLE 2: Eight habitat classifications and the number of sites sampled in each.

Habitat type	Stream gradient	Riparian forest cover	Water quality	Number of sampled sites
HCB	High	Closed	Stressed	10
HCG	High	Closed	Suitable	3
HOB	High	Open	Stressed	5
HOG	High	Open	Suitable	0
LCB	Low	Closed	Stressed	6
LCG	Low	Closed	Suitable	1
LOB	Low	Open	Stressed	13
LOG	Low	Open	Suitable	1

H = high gradient, L = low gradient, C = closed canopy, O = open canopy, G = water quality suitable for life support, and B = water quality biologically stressed.

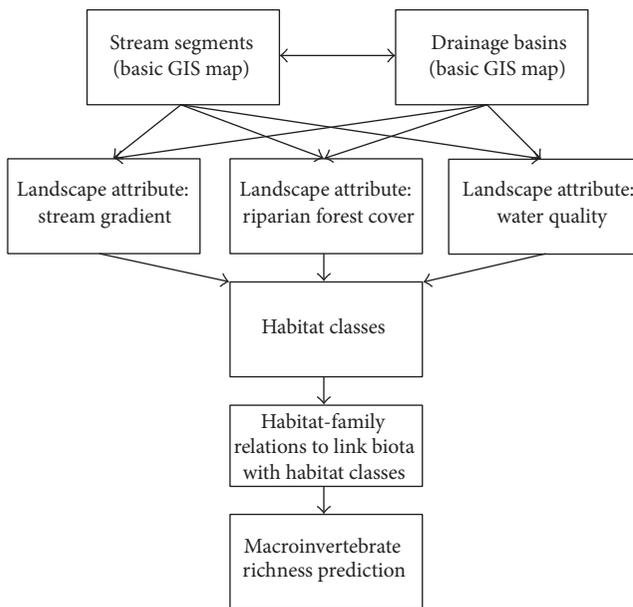


FIGURE 2: GIS model algorithm.

Agricultural land use, in addition to forestry and urbanization, affects sediment supply and runoff and alters the rates of surface water flow into adjacent and downstream water bodies. Each of these effects, in turn, can threaten biotic populations in aquatic systems [58–60]. In our model, water quality was classified using an adaptation of a GIS nonpoint source pollution runoff model originally developed by Adamus and Bergman [61]. We used inputs of land cover (EROS Data Center 1991–1993), soils (STATSGO 1994), average annual rainfall (Northeast Regional Climate Center 1961–1990), runoff coefficients [61], and pollutant concentrations [61] to determine the cumulative annual pollutant loading of total phosphorous (TP), total nitrogen (TN), and suspended sediment (SS) to each stream segment from its drainage basin. We adapted the model [62] to compare concentrations to the allowable USEPA pollutant criteria thresholds for the study area (ecoregion 7, subcoregion 61) for TP and TN, which are 0.03563 and 1 mg/L, respectively [63], and the strictest SS 30-day average in warm water streams (90 mg/L; [64]). A stream segment was classified as acceptable for each pollutant if its estimate was below the pollution criteria; otherwise the

stream segment was considered substandard. If at least two out of the three pollutants were considered within acceptable levels for a single stream segment, the reach was classified as “suitable for life support.” Otherwise, the stream segment was classified as “biologically stressed” (Table 1).

Once modeling of the three landscape attributes was complete, each stream segment was then classified into one of eight habitat classes given its classifications for stream gradient (high/low), riparian forest cover (open/closed), and water quality (suitable for life support/biologically stressed; Table 2). Stream segments along lake shores, reservoir margins, wetlands, or the state border were classified as undesignated. Some additional stream segments failed to receive classification due to errors in the digitized drainage basin and hydrology data layers resulting in multiple stream segments per drainage basin or stream segments without drainage basins. All undesignated stream segments were dropped from the analysis; however, they were retained in the GIS system and maps to maintain continuity of watershed connections. A total of 1016 (80%) stream segments received classifications.

**2.3. Defining Habitat-Family Relationships.** Each macroinvertebrate family in the Allegheny River watershed was associated with one or more of the same eight habitat classes using their preferences and tolerances for life habits, feeding guild, and water quality (Table 4). Preference and tolerance information was used to link macroinvertebrate families with the same eight habitat classes (defined in the habitat classification section) and thus enable us to predict family richness for each habitat class. First, bioassessment surveys completed in 1981 and 1989–1990 by the New York State Department of Environmental Conservation, Division of Water [65, 66], were used to determine the seventy-three macroinvertebrate families (mostly aquatic insects) present in the study area.

Next, life habit preferences (burrower, climber, clinger, sprawler, swimmer, and not specific) were used to link macroinvertebrate families with substrate, one of the components in our habitat classification model (represented in our model by stream gradient [51, 52]). In some rare cases several substrate preferences were listed for a single macroinvertebrate family in the literature. Our model could not accommodate multiple preferences for a single family, and thus we made the assumption that the first listed preference was dominant and used that for classification.

Feeding guilds (collector filterer, collector gatherer, scraper, shredder, and predator) were used to provide information on macroinvertebrate preferences for leaves and detritus originating from riparian forest cover, another component of our habitat classification model [65]. Finally, family-level tolerances for degraded water quality were used to link macroinvertebrate families to the water quality component of our habitat classification model. We used the categories intolerant and tolerant based on Hilsenhoff 1988 [44]. Taxa with values of eight or higher were considered tolerant. When taxa were not assigned a family tolerance by Hilsenhoff 1988 [44], an appropriate genus or species tolerance [65] was used.

Once preferences and tolerances were determined for each macroinvertebrate family, we grouped the families in terms of habitat classes (Table 4). We would expect that all functional feeding groups would be present (shredders, scrapers, filterers, gatherers, and predators) in closed canopy, high gradient streams. We would not expect the fauna of streams with closed canopies but low gradients to commonly include the scraper guild. Open canopy streams of high gradient would be expected to contain all feeding guilds except the shredders, while open canopy, low gradient streams would contain only gatherers, filterers, and predators. These habitat associations were further refined by eliminating families considered to be clingers from each of the low gradient sites (Table 4). Streams with water quality suitable for life support were expected to contain macroinvertebrates both tolerant and intolerant of water quality degradation, while biologically stressed water quality was only expected to contain macroinvertebrates tolerant to water quality degradation.

**2.4. Predicting Richness.** Using the preference and tolerance information for stream gradient, riparian forest cover, and water quality, each macroinvertebrate family was classified into one of eight habitat classes for macroinvertebrate richness prediction (Table 4). A single macroinvertebrate family can be classified into several habitat classes based on its preferences and tolerances. The number of families per habitat class was then tallied and stream segments with high or low predicted macroinvertebrate family richness were identified (Figure 3).

**2.5. Observed Data Collection.** A survey of 39 sites in the upper Allegheny River basin was completed between late May and mid-August 1998 during baseflow conditions. Summer samples are used routinely for stream bioassessments in New York State [65]. Sites were originally chosen using stratified random sampling to maintain an equal number of sites in each habitat class present in the study area. However, several sites chosen randomly were inaccessible, located in wetlands or were dry. Such sites were replaced with more accessible locations of the same habitat class where possible. Some sites were randomly located along the same channel. Site visitation order was randomized. For one habitat class (high gradient streams with open canopies and water quality suitable for life support) we failed to find suitable sites for sampling; therefore, this category is not represented in the field data collections (Table 2). High gradient streams with open canopies either are in poor condition with low water

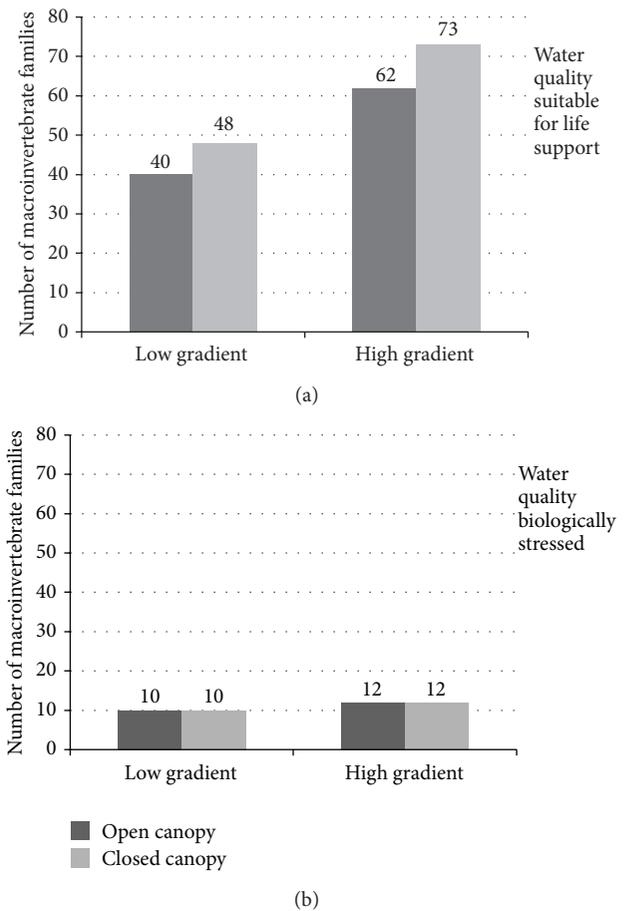


FIGURE 3: Predicted macroinvertebrate family richness for the eight habitat classes based on stream gradient, riparian forest cover, and water quality.

quality (e.g., mowed banks) or are very large. Both conditions reduce the number of sites available for sampling in this particular habitat class.

Latitude, longitude, and elevation measurements were obtained at the upstream and downstream ends of each stream segment using a global positioning system (GPS) unit with an antenna on a 3.6 m pole. One hundred data points were taken at each site when possible. In a GIS, the latitude and longitude coordinates were graphed and the distance along the stream between upstream and downstream data points was measured (i.e., run). Then, the difference in upstream and downstream elevations (i.e., rise) was divided by the estimated distance between these points (run) in GIS to calculate observed stream gradient for each site.

Eight substrate measurements were taken at equal intervals along each of the three transects in pool, riffle, and run habitats to obtain 24 measurements at each of the 39 sites. Substrate was randomly chosen at each of the eight locations along each transect and measured across the intermediate axis with a millimeter ruler and then classified into one of five categories based on size: boulder (>256 mm), cobble (65–256 mm), pebble (17–64 mm), gravel (2–16 mm), and fine sediment such as silt, clay, or sand (<2 mm). The dominant substrate was identified from the 24 measurements and the

site was classified as having coarse (gravel, pebble, cobble, or boulder) or fine (sand, silt, or clay) substrate.

The extent of riparian forest cover (closed or open) was assessed visually in the field and then augmented using imagery for the same locations in a GIS (GeoEye, 2002, 41 cm resolution). We did not rely solely on field assessment for this metric because not all portions of the riparian zone of a stream were visible from access points in the field. Fortunately, the entire riparian zones were visible using imagery in a GIS thereby allowing us to make a more accurate assessment of whether stream segments had closed or open canopies. Using information both from the field and from a GIS, the entire portion of each of the 39 stream segments was identified and visually surveyed and the amount of forest cover was approximated [67]. Stream segments with less than or equal to 50% forest cover in the riparian zones on both sides of the river were classified as open, as were rivers that were obviously large, and those with greater than 50% were classified as closed. All observed classifications were performed by the same observer to maintain uniformity in responses.

Water chemistry measurements for TP, TN, and SS were taken at the downstream end of each site in riffle, pool, and run habitats, where applicable, between July 27 and July 30. This time period was chosen to take advantage of conditions when the nitrogen content was at its lowest point and water was the clearest [50]. Three water samples of 250 ml were obtained for suspended sediment measurements at each of the sites and stored in a cooler with ice. After the field day was completed, the samples were pumped through preweighed filters (cellulose nitrate filter membranes; 45 microns) and dried in an oven at 103–105°C for one hour and then in a dessicator for 24 hours, after which the filters were weighed again. Three additional water samples of 100 ml were taken from each of the sites for total dissolved nitrogen and total dissolved phosphorous measurements. These were stored in a freezer until processing was completed at a lab ten months later.

Kick net sampling for macroinvertebrates was performed in three riffles chosen randomly from the bottom, middle, and top of each site [46, 68, 69]. Sampling was performed in the section of the riffle with the fastest flow. If riffles were not present, runs or pools were sampled instead. A three-minute kick sample was collected using D-frame sweep nets of mesh size 0.5 mm [65]. Macroinvertebrate organisms were collected and stored in Nalgene bottles with 70% ethyl alcohol. If the collection revealed low numbers of macroinvertebrates, then sampling was repeated until a single collection of 100 organisms was obtained. In the laboratory, the sample was transferred to and distributed homogeneously over the bottom of a gridded enamel pan. A small amount of the sample (approximately a tablespoon) was randomly removed with a spatula and placed in a petri dish containing 70% ethyl alcohol. This portion of the sample was examined under a stereomicroscope and the organisms were sorted to family, placed in vials containing 70% ethyl alcohol, and counted until we reached 100 organisms [65]. For the samples from stream segments with low macroinvertebrate abundance, all 100 organisms in the collection were retained

and sorted to family. The summed taxa of the three samples from each stream segment were tallied to obtain observed family richness.

**2.6. Data Analysis.** Observed field data were tested against model predictions for macroinvertebrate family richness and all landscape attributes. Statistical analysis of landscape attributes was necessary as successful family richness prediction rests on the ability of the model to accurately represent its component parameters. Observed and predicted macroinvertebrate family richness were assessed using the paired *t*-test. Predicted and observed stream gradient, riparian forest cover, and water quality classifications were compared using the sign test, a nonparametric test with matched samples. Matching observed and predicted classifications were tagged with one sign and mismatched classifications were assigned the inverse sign. The sign test provides a probability that the obtained ratio of matches and mismatches differs from the outcome of random results. That is, the probability result from a sign test indicates the likelihood that the results were due to random variation rather than the hypothesized effect. For example, model predictions of water quality and field collected water samples could be compared for a set of stream segments. If predictions and field samples for paired locations were both of good quality, or both of poor quality, that would indicate a match. If predictions were of good quality but field samples were of poor quality (or vice versa), that would indicate a mismatch. The sign test determines if matches occur more frequently than by chance alone; thus, if a match between the predictions of water quality and field collected samples is strong, we can have confidence in our model predictions. The results of the sign test then are used to judge confidence in our findings (Graphpad Software Inc., San Diego, CA USA). Statistical significance was assumed at the five percent level of probability ( $P \leq 0.05$ ).

### 3. Results

No difference was detected between the observed and predicted number of macroinvertebrate families in a paired *t*-test ( $T = 1.5318$ ,  $P = 0.1339$ ; Table 3). The habitat type expected to have the greatest macroinvertebrate family richness (73 taxa) was clearly closed canopy, high gradient streams with water quality suitable for life support. These relatively uncommon (13.3%) stream segments averaged 3.3 km in length, somewhat longer than stream segments experiencing some form of degradation (2.6 km). Stream segments in this habitat class were primarily found in the central and eastern parts of the study area with very rare instances in the more agriculturally developed western portion of the study area.

Our linked habitat-macroinvertebrate family richness predictions indicate marked declines with water quality degradation (Figure 3). However, changes in stream gradient and riparian forest cover appear to have little effect on the relative richness of macroinvertebrate families. On average across the habitat classes, 62% of the families were predicted to be lost with water quality degradation. Low gradient streams were associated with 15% fewer families than high gradient streams and open riparian zones were associated

TABLE 3: Accuracy of macroinvertebrate family richness predictions and landscape attributes.

	Paired <i>t</i> -test	<i>P</i>	Number of matches
Macroinvertebrate family richness	1.532	0.134	
Stream gradient		0.012	27
Riparian forest cover		0.005	28
Water quality (total nitrogen)		<0.0001	31
Water quality (total phosphorus)		0.027	26
Water quality (suspended sediment)		<0.0001	39

Paired *t*-test = *T* value, *P* = *P* value, and number of matches = matches using the sign test. Sign test was used for attributes where only *P* value is present. Results were termed significant for  $P < 0.05$ .

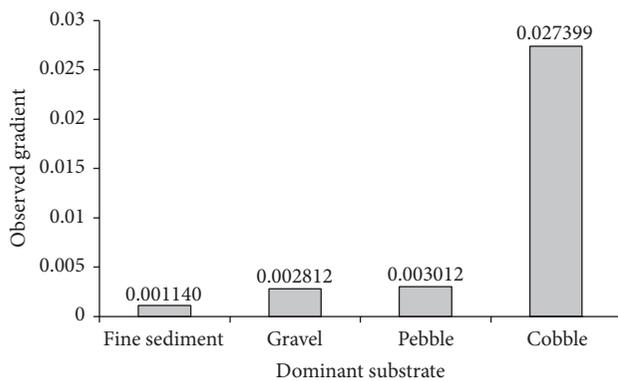


FIGURE 4: Median observed stream gradient measurements for four field collected dominant substrate classes.

with 4% fewer families than closed canopy riparian stream segments in our predictions.

Approximately 50% of the streams segments in the upper Allegheny River basin were predicted to have high stream gradient. High gradient reaches were widely distributed throughout the watershed though concentrated in low order streams. Predicted and observed stream segments, classified as high or low gradient, were compared using the sign test [62]. Mismatches between observed and predicted high and low gradient categories were uncommon (12 out of 39; 31%) and much less likely than by chance alone ( $P = 0.012$ ; Table 3).

A graph (Figure 4) of observed dominant substrate (fine sediment, gravel, pebble, and cobble) and observed gradient values (without missing data points) shows a slight increase in median gradient between fine sediment (median = 0.0011,  $n = 7$ ) and gravel (median = 0.0028,  $n = 7$ ) and an even smaller increase between gravel and pebble (median 0.0030,  $n = 18$ ). However, a large increase exists between gradient medians of pebble and cobble (median 0.0274,  $n = 2$ ). This indicates that the classification criterion we used (1.5/1000 m) was suitable for separating stream segments with a dominant substrate of fine sediment from those with coarser substrate (e.g., cobble; Figure 4).

In terms of riparian forest cover, the model predicted that approximately 36% of the stream segments in the watershed would have closed canopy and that they would be fairly well distributed throughout the study area. Predicted and observed riparian forest cover were also compared using

the sign test. Mismatches between predicted and observed riparian forest cover classifications as open or closed canopy were uncommon (11 out of 39; 28%) and much less likely than by chance alone ( $P \leq 0.005$ ; Table 3).

Approximately 67% of the stream segments in the study area were predicted to be biologically stressed because of water quality degradation. Not surprisingly, the majority of degraded water quality stream segments were located in the portion of the watershed where agricultural and urban land uses were most concentrated (western half). The high quality stream segments were largely located in an area protected by the New York State Park system. Thus, water quality degradation appeared more clustered, regional, and prevalent than the riparian forest cover and gradient classification distributions. Predicted TN classifications matched observed classifications for all but 8 stream segments (79%). Predicted and observed TP classifications matched for 26 of the 39 stream segments (67%) and SS classifications matched for all stream segments (100%). The probability of obtaining 31 and 39 matching classifications out of 39 comparisons was <0.0001 in both cases (Table 3), indicating that the high rates of TN and SS matches were significant and highly confident results. The probability of obtaining 26 matching classifications out of 39 stream segment comparisons was 0.027 (Table 3), indicating that the rate of TP matches was slightly lower but still a significant result.

#### 4. Discussion

Our study was aimed at predicting macroinvertebrate family richness in riverine environments using the landscape attributes stream gradient, riparian forest cover, and water quality to define habitat classes for stream segments. Then we used macroinvertebrate life habit and feeding guild preferences and water quality tolerances to link biota with habitat classes. We could thus predict macroinvertebrate family richness for each stream segment. Our predictions were tested with a survey of macroinvertebrate family richness, stream gradient, riparian forest cover openness, and water quality sampling. Results of our techniques appear encouraging. Predicted and observed macroinvertebrate families were found to be not significantly different. Our findings indicate that our model can provide landscape scale macroinvertebrate family richness predictions from widely available data given reasonable time and resources. Only slight modifications would be necessary to use our model to predict richness



TABLE 4: Continued.

Family	Feeding guild	Water quality tolerance	Habitat preferences	HCB	HCG	HOB	HOG	LCB	LCG	LOB	LOG
<i>Peltoperlidae</i>	sh	I	clg		X						
<i>Perlidae</i>	pr	I	clg		X		X				
<i>Perlodidae</i>	pr	I	clg		X		X				
<i>Philopotamidae</i>	cf	I	clg		X		X				
<i>Phryganeidae</i>	sh	I	clb		X				X		
<i>Physidae</i>	cg	T	all	X	X	X	X	X	X	X	X
<i>Pisidiidae</i>	cf	I	bur		X		X		X		X
<i>Polycentropodidae</i>	cf	I	clg		X		X				
<i>Polymitarcyidae</i>	cg	I	bur		X		X		X		X
<i>Potamanthidae</i>	cg	I	sp		X		X		X		X
<i>Psephenidae</i>	sc	I	clg		X		X				
<i>Psychomyiidae</i>	cg	I	clg		X		X				
<i>Pteronarcyidae</i>	sh	I	clg		X						
<i>Pyrilidae</i>	sh	I	clb		X				X		
<i>Rhyacophilidae</i>	pr	I	clg		X		X				
<i>Saldidae</i>	pr	I	clb		X		X		X		X
<i>Sialidae</i>	pr	I	bur		X		X		X		X
<i>Simuliidae</i>	cf	I	clg		X		X				
<i>Siphonuridae</i>	cg	I	sw		X		X		X		X
<i>Sphaeriidae</i>	cf	I	bur		X		X		X		X
<i>Tabanidae</i>	pr	I	sp		X		X		X		X
<i>Talitridae</i>	cg	T	bur	X	X	X	X	X	X	X	X
<i>Tetrastemmatidae</i>	pr	T	all	X	X	X	X	X	X	X	X
<i>Tipulidae</i>	sh	I	bur		X				X		
<i>Tricorythidae</i>	cg	I	sp		X		X		X		X
<i>Tubificidae</i>	cg	T	bur	X	X	X	X	X	X	X	X

H = high gradient, L = low gradient, C = closed canopy, O = open canopy, G = water quality suitable for life support, and B = water quality biologically stressed; cf = collector filterer, cg = collector gatherer, sc = scraper, sh = shredder, pr = predator; I = intolerant, T = tolerant; bur = burrower, clb = climber, clg = clinger, sp = sprawler, sw = swimmer, and all = all life habits represented.

in new areas. Chiefly, one would need to only research the life habit, feeding guild, and water quality preferences and tolerances of new organisms, obtain landscape scale data to identify habitat classes for the new location, and procure pollution thresholds for the region.

One of the foremost contributions of this study is our research defining habitat-family relationships. Though the concept of linking organisms to habitats has been well documented [16], few studies exist that link life history preferences of aquatic organisms to habitat classes for predictive modeling [17, 29]. In our study, we found that high gradient streams with forested riparian zones and water quality suitable for life support were predicted to have the greatest macroinvertebrate family richness and changes in water quality were predicted to have the greatest impact on family richness. For streams with biologically stressed water quality, there was no difference between the expected numbers of families in closed canopy streams and open canopy streams. Similarly, the difference between low gradient and high gradient streams was minimal (10 versus 12 families; *Bithyniidae* and *Hydrobiidae* missing).

In our classification of macroinvertebrate families into habitat classes, we made the assumption that the fauna of streams with closed canopies but low gradients would

not commonly include the scraper guild. In our testing of streams with closed canopies and low gradients (7 sites), most of the samples did not include scraper families with the exception of *Elmidae* and *Heptageniidae* which were found in all but one of the sites. In the literature, some *Elmidae* species are considered gatherers [65] so it is possible that this family should be reclassified upon further investigation. *Heptageniidae* species are almost all classified as scrapers [65]; therefore, this family is properly categorized. *Heptageniidae* and possibly also *Elmidae* may be quite ubiquitous in the study area and found in habitats not normally expected to include scraper families.

We also made the assumption that open canopy streams of high gradient would be expected to contain all feeding guilds except shredders. In our testing of streams with open canopies and high gradients (5 sites), most of the samples included the shredder family *Tipulidae*. In the literature, *Tipulidae* species are equally distributed between predator, gatherer, and shredder feeding guilds [65]. Therefore, this family may be reclassified as predator or gatherer upon further investigation.

The assumption that open canopy, low gradient streams would contain only gatherers, filterers, and predators was violated at 12 out of the 14 sites sampled in this set of

habitat classes. Most sites had 1–3 families that were not gatherers, filterers, or predators, most commonly *Elmidae* and *Heptageniidae* (both classified as scrapers) and *Tipulidae* (classified as shredder). Similarly, we investigated the assumption that low gradient sites were not expected to include clingers. Quite a few common species of clingers were found throughout many of the 21 low gradient sites sampled for macroinvertebrate families in the study area (e.g., *Elmidae* and *Heptageniidae*). As noted previously, these families are quite ubiquitous and are likely to be found in a broad range of habitats [50]. Further, it is possible that one or more of these families are misclassified and further study should be undertaken to determine if the classifications in our study are correct. Overall, the families *Elmidae* and *Heptageniidae* are the two families accounting for most of the deviations from successful macroinvertebrate family habitat classification.

Streams with water quality suitable for life support were expected to contain macroinvertebrates both tolerant and intolerant to water quality degradation, while biologically stressed streams were only expected to contain macroinvertebrate families tolerant to water quality degradation. In a survey of the sampled sites, many sites with biologically stressed water quality contained families considered to be intolerant (e.g., *Caenidae*, *Corixidae*, *Ephemerellidae*, *Ephemeridae*, *Heptageniidae*, *Hydropsychidae*, *Perlidae*, *Polymitarcyidae*, *Potamanthidae*, *Sialidae*, and *Siphonuridae*). Many of these families legitimately indicate good water quality (e.g., *Perlidae*); however, some may deserve a closer look in terms of classification. Overall, 84% of the families predicted to exist in the study area were classified as intolerant of poor water quality. This is the factor that largely defines the big difference in number of families predicted as good versus poor water quality sites.

Overall, the assumptions made in the prediction of macroinvertebrate families into habitat classes seem valid with a few notable exceptions. Our linked macroinvertebrate family-habitat class relationships form the basis for our richness predictions and those predictions were found to be not statistically different from observed richness in the field.

Prediction accuracy from our model was high across all three landscape attributes on which the habitat classification was based. Stream gradient predictions matched observed gradient data more than what would be expected by chance alone indicating that measurements from 1:24,000 scale DEM digital data can accurately be used to estimate stream gradient. Stream gradient acts as a surrogate for substrate by separating organisms which favor sand, silt, and clay (low gradient streams) from those which favor cobble, pebble, and boulders (high gradient streams [49]). We would expect low gradient stream segments to have more fine sediment which would fill interstitial spaces and possibly reduce oxygen absorption by macroinvertebrates leading to lower taxa richness [50]. Our plot of observed median gradients against dominant substrate in the upper Allegheny River basin lent support for the use of the 1.5/1000 m classification criterion to separate sites with dominant fine sediment substrate from those with dominant coarse substrate (gravel, pebble, cobble, or boulder).

Categorical riparian forest cover predictions demonstrated greater than chance agreement with observed forest cover. Open canopy streams have lower inputs of detritus, a vital food source for some macroinvertebrates, leading to lower taxa richness in those water bodies [50]. We obtained this favorable result despite several mitigating factors. First, the digital land use maps used in our model were composed of pixels which represent an area of 30 meters by 30 meters. The area inside the pixel is designated as a single land use type (i.e., forest and agriculture) despite the fact that several land use types may actually be present in the area represented by the pixel [23]. Thus, narrow forested riparian zones next to streams may have been overlooked at this scale in the predicted data. Further, observed riparian forest cover characterizations were somewhat subjective and difficult to determine for an entire stream segment from confluence to confluence. Despite these potential difficulties, our modeling methods appear to yield accurate results for riparian forest cover prediction.

Our adapted nonpoint source pollution load screening model was designed to predict whether TP, TN, and SS concentrations were higher than USEPA quality criteria (see Meixler and Bain 2010a for further explanation and validation of this model [62]). Organic pollution has been closely linked with low taxa richness [50]. Our model was able to accurately match observed water quality classifications for all parameters, despite some limitations. The coarse, simple GIS approach of the model was intended for annual prediction of parameters and could not accurately reflect subtle changes in pollutant concentrations. The field samples, collected over one short time interval at baseflow conditions, were not a thorough test of annual water quality conditions since grab samples reflect instantaneous conditions that are quickly and easily changed by agricultural manipulations, rainfall events, and biological uptake [70]. In baseflow conditions, much of the water in the channel comes from groundwater sources [50]. Pollutants associated with runoff from various land uses are more likely to be deposited in streams following high rainfall events. Therefore, multiple field samplings taken at times of high runoff volume may more accurately represent annual water quality levels. Additional error may have been caused by misclassifications in the creation of the digital data [18] and inappropriate runoff coefficients and pollutant concentration values for western New York State. Despite these sources of potential error, our model predictions closely matched observed water quality classifications.

The primary advantages of our model are its minimal data requirements, rapid data preprocessing, useful output scale, and broad applicability in similar aquatic environments [17]. Our model only needs digital land use, soils, DEM, and rainfall data and a list of known macroinvertebrates to be applied to new regions, though we recommend that model developers seek out local family-level macroinvertebrate tolerances and preferences and local runoff coefficients/pollutant concentrations (for water quality prediction) for new areas. Further, the procedures presented here are largely independent of region unless the hydrology (e.g., desert) or dominant land cover type (e.g., plains) is markedly different.

Several limitations of the model and sampling protocol should be acknowledged. Building a classification structure for a model requires the use of absolute boundaries on what were often true continua [71]. The habitat classes described here were merely representations of a continuum of habitat patterns. Further, interpretations of habitat features were limited by the scale of the available maps. Fine scale land use, soils, and DEM maps may better define the true composition of the region and improve prediction quality. The automated nature of our GIS classification system enables us to address such limitations through parameter modifications to accommodate a variety of specialized circumstances such as differences in regional area, the development of highly specialized models for key individual taxa, and a variety of other small-scale applications. Future work may test the accuracy of the model for these specialized scenarios.

Sources of variability in our observed macroinvertebrate family data collections stem from misidentification of uncommon macroinvertebrates, gear and field technician inefficiency in macroinvertebrate capture, and unevenness of habitat class sampling. Sources of variability in our macroinvertebrate family predictions stem from inexact model parameterization and from challenges in classifying macroinvertebrates with complex life strategies into discrete habitat classes. In two sites, observed and predicted family richness values were exactly equal. In fifteen sites predicted values were greater than observed values. The sources of variability in our observed data collection would likely result in lower family richness estimates (i.e., misidentification of uncommon taxa and inefficiency of macroinvertebrate capture) and could possibly obscure biases in the macroinvertebrate family predictions (i.e., uneven habitat class sampling coverage). Thus, with improvements in observed data sampling techniques we would expect our observed values to increase and our observed and predicted richness correlation to become stronger.

Although an attempt was made to choose sites using stratified random sampling, many fewer sites existed with water quality suitable for life support and many of those had access restrictions. Thus, few candidate streams existed in which to sample sites with good water quality and these categories were thus not as well represented among the field collected samples. Our paired *t*-test showed significant similarities across all sampled habitat classes between predicted and observed macroinvertebrate family richness. However, if the model were to have a bias toward lower macroinvertebrate family richness predictions, our observed richness values may not be sufficiently robust or broad enough to detect this bias. We recommend that further research in the future not only sample macroinvertebrates more evenly across all habitat classes but also sample in more sites overall.

Finally, our samples were collected in the summer between the months of late May and mid-August [65]. This time period may have influenced the composition of our macroinvertebrate community for several reasons. First, many macroinvertebrates are less detectable in the summer months due to the timing of their life cycles when they are in early life stages, emerging or diapausing [51]. The long period of sampling may have resulted in some taxa having

been collected early in the summer while others, at different sites, were missed late in the summer although they occurred at these sites earlier in the year. These factors may have led to biases in the results, as spatial variation among streams may be related to seasonal differences in macroinvertebrate communities among sites. Macroinvertebrate taxa are also likely to move among stream segments during their life histories and are capable of occupying a broad range of habitat [50]. A further source of error may have resulted from our sampling focus on riffle habitats in streams at the 1:100,000 scale leading to an underestimation of richness in stream segments that are largely pool habitat. Future studies should test model predictions using samples from a variety of in-stream habitats and should assess the model's ability to predict macroinvertebrate richness at finer scales.

Effective conservation of biodiversity in aquatic communities requires the identification and protection of key landscapes and communities [31]. To do so, ecologists and landscape planners need to design protocols to assess the health of the biotic community and develop strategies for conservation. The habitat classification approach presented here is one of many such methods grouping stream habitat into classes in an effort to identify homologous regions with similar attributes (e.g., [71, 72]). However, few classification systems go further to link faunal assemblages and thus biotic richness with aquatic habitats developed in a GIS. It is clear that the modeling procedures presented here have considerable potential to predict macroinvertebrate family richness at the watershed scale. Planners can use information from this model to locate areas of high macroinvertebrate family richness in which to focus efforts for further study or find stream segments which might be improved through best management practices, stream side buffers, or creation of wetlands [73].

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This study was supported by the United States Geological Survey under cooperative agreements no. 1434-HQ-97-RU-01553 and RWO no. 40 and benefited from research performed in the study area under a grant from the Nature Conservancy. Special thanks go to Greg Galbreath for compiling the macroinvertebrate data in Table 4. The authors also wish to thank Jordan Gass and Andrew Koo for their fieldwork assistance and Magdeline Laba, Steve Smith, and Thomas Hopper for their GIS assistance in support of this project.

## References

- [1] C. S. Elton, *The Ecology of Invasions by Animals and Plants*, Methuen Publishing, London, UK, 1958.
- [2] D. Tilman, D. Wedin, and J. Knops, "Productivity and sustainability influenced by biodiversity in grassland ecosystems," *Nature*, vol. 379, no. 6567, pp. 718–720, 1996.

- [3] R. H. MacArthur and E. O. Wilson, *The Theory of Island Biogeography*, Princeton University Press, Princeton, NJ, USA, 1967.
- [4] D. Tilman, *Resource Competition and Community Structure*, Monographs in Population Biology, Princeton University Press, Princeton, NJ, USA, 1982.
- [5] D. Tilman, "Community invasibility, recruitment limitation, and grassland biodiversity," *Ecology*, vol. 78, no. 1, pp. 81–92, 1997.
- [6] S. L. Pimm, *The Balance of Nature?* University of Chicago Press, Chicago, Ill, USA, 1991.
- [7] D. H. Wright and J. H. Reeves, "On the meaning and measurement of nestedness of species assemblages," *Oecologia*, vol. 92, no. 3, pp. 416–428, 1992.
- [8] P. R. Ehrlich and E. O. Wilson, "Biodiversity studies: science and policy," *Science*, vol. 253, no. 5021, pp. 758–762, 1991.
- [9] R. Dirzo, H. S. Young, M. Galetti, G. Ceballos, N. J. B. Isaac, and B. Collen, "Defaunation in the Anthropocene," *Science*, vol. 345, no. 6195, pp. 401–406, 2014.
- [10] P. J. Seddon, C. J. Griffiths, P. S. Soorae, and D. P. Armstrong, "Reversing defaunation: restoring species in a changing world," *Science*, vol. 345, no. 6195, pp. 406–412, 2014.
- [11] G. Mace, H. Masundire, and J. Baillie, "Biodiversity," in *Millennium Ecosystem Assessment, 2005. Current State and Trends: Findings of the Condition and Trends Working Group, Ecosystems and Human Well-Being*, vol. 1, Island Press, Washington, DC, USA, 2005.
- [12] European Environmental Agency (EEA), *Progress towards the European 2010 Biodiversity Target*, Schultz Grafisk, Copenhagen, Denmark, 2009.
- [13] C. Groves, *Drafting a Conservation Blueprint. A Practitioner's Guide to Planning for Biodiversity*, Island Press, Washington, DC, USA, 2003.
- [14] R. L. Pressey, R. M. Cowling, and M. Rouget, "Formulating conservation targets for biodiversity pattern and process in the Cape Floristic Region, South Africa," *Biological Conservation*, vol. 112, no. 1-2, pp. 99–127, 2003.
- [15] D. Klein, "Results of New York customer consultations," in *Lake Ontario Improvement Opportunity Assessment Project*, The Natural Heritage Institute, Ed., Great Lakes Protection Fund, Evanston, Ill, USA, 2008.
- [16] J. M. Scott, F. Davis, B. Csuti et al., "Gap Analysis: a geographic approach to protection of biological diversity," *Wildlife Monographs*, vol. 123, 1, p. 41, 1993.
- [17] M. S. Meixler and M. B. Bain, "A GIS framework for fish habitat prediction at the river basin scale," *International Journal of Ecology*, vol. 2012, Article ID 146073, 10 pages, 2012.
- [18] K. Tucker, S. P. Rushton, R. A. Sanderson, E. B. Martin, and J. Blaiklock, "Modelling bird distributions—a combined GIS and Bayesian rule-based approach," *Landscape Ecology*, vol. 12, no. 2, pp. 77–93, 1997.
- [19] R. W. Tiner, "Remotely-sensed indicators for monitoring the general condition of 'natural habitat' in watersheds: an application for Delaware's Nanticoke River watershed," *Ecological Indicators*, vol. 4, no. 4, pp. 227–243, 2004.
- [20] M. N. Snyder, S. J. Goetz, and R. K. Wright, "Stream health rankings predicted by satellite derived land cover metrics," *Journal of the American Water Resources Association*, vol. 41, no. 3, pp. 659–677, 2005.
- [21] J. Heino, H. Mykrä, and J. Kotanen, "Weak relationships between landscape characteristics and multiple facets of stream macroinvertebrate biodiversity in a boreal drainage basin," *Landscape Ecology*, vol. 23, no. 4, pp. 417–426, 2008.
- [22] T. W. Norton and J. E. Williams, "Habitat modelling and simulation for nature conservation: a need to deal systematically with uncertainty," *Mathematics and Computers in Simulation*, vol. 33, no. 5-6, pp. 379–384, 1992.
- [23] Z. Naveh and A. Lieberman, *Landscape Ecology: Theory and Application*, Springer, New York, NY, USA, 1994.
- [24] L. A. Bojorquez-Tapia, I. Azuara, E. Ezcurra, and O. Flores-Villela, "Identifying conservation priorities in Mexico through geographic information systems and modeling," *Ecological Applications*, vol. 5, no. 1, pp. 215–231, 1995.
- [25] M. S. Meixler, M. B. Bain, and M. Todd Walter, "Predicting barrier passage and habitat suitability for migratory fish species," *Ecological Modelling*, vol. 220, no. 20, pp. 2782–2791, 2009.
- [26] A. S. L. Rodrigues, H. R. Akçakaya, S. J. Andelman et al., "Global gap analysis: priority regions for expanding the global protected-area network," *BioScience*, vol. 54, no. 12, pp. 1092–1100, 2004.
- [27] S. M. McMahon, S. P. Harrison, W. S. Armbruster et al., "Improving assessment and modelling of climate change impacts on global terrestrial biodiversity," *Trends in Ecology & Evolution*, vol. 26, no. 5, pp. 249–259, 2011.
- [28] H. Mykrä, J. Aroviita, J. Kotanen, H. Hämäläinen, and T. Muotka, "Predicting the stream macroinvertebrate fauna across regional scales: influence of geographical extent on model performance," *Journal of the North American Benthological Society*, vol. 27, no. 3, pp. 705–716, 2008.
- [29] N. Péru and S. Dolédec, "From compositional to functional biodiversity metrics in bioassessment: a case study using stream macroinvertebrate communities," *Ecological Indicators*, vol. 10, no. 5, pp. 1025–1036, 2010.
- [30] I. R. Waite, L. R. Brown, J. G. Kennen et al., "Comparison of watershed disturbance predictive models for stream benthic macroinvertebrates for three distinct ecoregions in western US," *Ecological Indicators*, vol. 10, no. 6, pp. 1125–1136, 2010.
- [31] P. L. Angermeier and M. R. Winston, "Characterizing fish community diversity across Virginia landscapes: prerequisite for conservation," *Ecological Applications*, vol. 9, no. 1, pp. 335–349, 1999.
- [32] D. L. Strayer and D. Dudgeon, "Freshwater biodiversity conservation: recent progress and future challenges," *Journal of the North American Benthological Society*, vol. 29, no. 1, pp. 344–358, 2010.
- [33] S. S. Wall, C. R. Berry Jr., C. M. Blausey, J. A. Jenks, and C. J. Kopplin, "Fish-habitat modeling for gap analysis to conserve the endangered Topeka shiner (*Notropis topeka*)," *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 61, no. 6, pp. 954–973, 2004.
- [34] S. P. Sowa, G. Annis, M. E. Morey, and D. D. Diamond, "A gap analysis and comprehensive conservation strategy for riverine ecosystems of Missouri," *Ecological Monographs*, vol. 77, no. 3, pp. 301–334, 2007.
- [35] P. L. Angermeier and I. J. Schlosser, "Conserving aquatic biodiversity: beyond species and populations," *American Fisheries Society Symposium*, vol. 17, pp. 402–414, 1995.
- [36] J. V. Higgins, M. T. Bryer, M. L. Khoury, and T. W. Fitzhugh, "A freshwater classification approach for biodiversity conservation planning," *Conservation Biology*, vol. 19, no. 2, pp. 432–445, 2005.

- [37] R. Céréghino, Y.-S. Park, A. Compin, and S. Lek, "Predicting the species richness of aquatic insects in streams using a limited number of environmental variables," *Journal of the North American Benthological Society*, vol. 22, no. 3, pp. 442–456, 2003.
- [38] F. B. Lottspeich, "Watersheds as the basic ecosystem: this conceptual framework provides a basis for a natural classification system," *Journal of the American Water Resources Association*, vol. 16, no. 4, pp. 581–586, 1980.
- [39] R. J. Naiman, D. G. Lonzarich, T. J. Beechie et al., "General principles of classification and the assessment of conservation potential in rivers," in *River Conservation and Management*, P. J. Boon, P. Calow, and G. E. Petts, Eds., pp. 93–124, John Wiley & Sons, Chichester, UK, 1992.
- [40] L. P. Aadland, "Stream habitat types: their fish assemblages and relationship to flow," *North American Journal of Fisheries Management*, vol. 13, no. 4, pp. 790–806, 1993.
- [41] J. B. Wallace and J. R. Webster, "The role of macroinvertebrates in stream ecosystem function," *Annual Review of Entomology*, vol. 41, pp. 115–139, 1996.
- [42] A. P. Covich, M. A. Palmer, and T. A. Crowl, "The role of benthic invertebrate species in freshwater ecosystems," *BioScience*, vol. 49, pp. 119–127, 1999.
- [43] S. Gascón, D. Boix, and J. Sala, "Are different biodiversity metrics related to the same factors? A case study from Mediterranean wetlands," *Biological Conservation*, vol. 142, no. 11, pp. 2602–2612, 2009.
- [44] W. L. Hilsenhoff, "Rapid field assessment of organic pollution with a family-level biotic index," *Journal of the North American Benthological Society*, vol. 7, no. 1, pp. 65–68, 1988.
- [45] B. C. Chessman, K. M. Trayler, and J. A. Davis, "Family- and species-level biotic indices for macroinvertebrates of wetlands on the Swan Coastal Plain, Western Australia," *Marine & Freshwater Research*, vol. 53, no. 5, pp. 919–930, 2002.
- [46] G. J. Pond, M. E. Passmore, F. A. Borsuk, L. Reynolds, and C. J. Rose, "Downstream effects of mountaintop coal mining: comparing biological conditions using family- and genus-level macroinvertebrate bioassessment tools," *Journal of the North American Benthological Society*, vol. 27, no. 3, pp. 717–737, 2008.
- [47] E. Turak, R. Marchant, L. A. Barmuta, J. Davis, S. Choy, and L. Metzeling, "River conservation in a changing world: invertebrate diversity and spatial prioritisation in south-eastern coastal Australia," *Marine and Freshwater Research*, vol. 62, no. 3, pp. 300–311, 2011.
- [48] R. C. Horn, S. A. Hanson, and L. D. McKay, *History of the US EPA's River Reach File: A National Hydrographic Database Available for ARC/INFO Applications*, Office of Wetlands, Oceans, and Watersheds, Office of Water, US Environmental Protection Agency, Washington, DC, USA, 1994.
- [49] H. H. Barnes Jr., "Roughness characteristics of natural channels," U. S. Geological Survey Water Supply Paper 1849, U. S. Geological Survey, Washington, DC, USA, 1967.
- [50] J. D. Allan and M. M. Castillo, *Stream Ecology: Structure and Function of Running Waters*, Springer, Amsterdam, The Netherlands, 1997.
- [51] D. G. Smith, *Pennak's Freshwater Invertebrates of the United States*, John Wiley & Sons, New York, NY, USA, 4th edition, 2001.
- [52] R. W. Merritt, K. W. Cummins, and M. B. Berg, *An Introduction to the Aquatic Insects of North America*, Kendall/Hunt Publishing, Dubuque, Iowa, USA, 4th edition, 2008.
- [53] A. J. Castelle, A. W. Johnson, and C. Conolly, "Wetland and stream buffer size requirements—a review," *Journal of Environmental Quality*, vol. 23, no. 5, pp. 878–882, 1994.
- [54] N. E. Roth, J. D. Allan, and D. L. Erickson, "Landscape influences on stream biotic integrity assessed at multiple spatial scales," *Landscape Ecology*, vol. 11, no. 3, pp. 141–156, 1996.
- [55] M. Munn, J. Frey, and A. Tesoriero, "The influence of nutrients and physical habitat in regulating algal biomass in agricultural streams," *Environmental Management*, vol. 45, no. 3, pp. 603–615, 2010.
- [56] J. D. Milliman and K. L. Farnsworth, *River Discharge to the Coastal Ocean: A Global Synthesis*, Cambridge University Press, New York, NY, USA, 2011.
- [57] United States Environmental Protection Agency, "National water quality inventory: report to congress—2002 reporting cycle," Tech. Rep. EPA 841-R-07-001, Office of Water, US Environmental Protection Agency, 2007.
- [58] D. E. Blockstein, "An aquatic perspective on U.S. biodiversity policy," *Fisheries*, vol. 17, no. 3, pp. 26–30, 1992.
- [59] M. A. Cairns and R. T. Lackey, "Biodiversity and management of natural resources: the issues," *Fisheries*, vol. 17, no. 3, pp. 6–10, 1992.
- [60] R. M. Hughes and R. F. Noss, "Biological diversity and biological integrity: current concerns for lakes and streams," *Fisheries*, vol. 17, no. 3, pp. 11–19, 1992.
- [61] C. L. Adamus and M. J. Bergman, "Estimating nonpoint source pollution loads with GIS screening model," *Water Resources Bulletin*, vol. 31, no. 4, pp. 647–655, 1995.
- [62] M. S. Meixler and M. B. Bain, "A water quality model for regional stream assessment and conservation strategy development," *Environmental Management*, vol. 45, no. 4, pp. 868–880, 2010.
- [63] United States Environmental Protection Agency, "Ambient water quality criteria recommendation: rivers and streams," EPA 822-b-00-018, US Environmental Protection Agency, Washington, DC, USA, 2000.
- [64] W. Swietlik, *Developing Water Quality Criteria for Suspended and Bedded Sediments (SABS): Potential Approaches*, Environmental Protection Agency, Office of Water, Washington, DC, USA, 2003.
- [65] R. W. Bode, M. A. Novak, and L. E. Abele, *Quality Assurance Work Plan for Biological Stream Monitoring in New York State*, Department of Environmental Conservation, Division of Water, New York, NY, USA, 1991.
- [66] R. W. Bode, *Twenty Year Trends in Water Quality of Rivers and Streams in New York State Based on Macroinvertebrate Data 1972–1992*, Division of Water, New York State Department of Environmental Conservation, Albany, NY, USA, 1993.
- [67] M. S. Meixler and M. B. Bain, "Landscape scale assessment of stream channel and riparian habitat restoration needs," *Landscape and Ecological Engineering*, vol. 6, no. 2, pp. 235–245, 2010.
- [68] M. T. Barbour, J. Gerritsen, B. D. Snyder et al., *Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish*, U.S. Environmental Protection Agency, Office of Water, Washington, DC, USA, 2nd edition, 1999, EPA 841-B-99-002.
- [69] C. Riley, S. Inamdar, and C. Pennuto, "Use of benthic macroinvertebrate indices to assess aquatic health in a mixed-landuse watershed," *Journal of Freshwater Ecology*, vol. 22, no. 4, pp. 539–551, 2007.

- [70] C. M. Rohm, J. W. Giese, and C. C. Bennett, "Evaluation of an aquatic ecoregion classification of streams in Arkansas," *Journal of Freshwater Ecology*, vol. 4, no. 1, pp. 127–140, 1987.
- [71] P. S. Seelbach, M. J. Wiley, J. C. Kotanchik et al., "A landscape-based ecological classification system for river valley segments in lower Michigan (MI-VSEC version 1.0)," Research Report 2036, Fisheries Division, State of Michigan Department of Natural Resources, Lansing, Mich, USA, 1997.
- [72] J. R. Maxwell, C. J. Edwards, M. E. Jensen et al., "A hierarchical framework of aquatic ecological units in North America (nearctic zone)," General Technical Report NC-176, United States Department of Agriculture, Forest Service, North Central Forest Experiment Station, St. Paul, Minn, USA, 1995.
- [73] J. D. Stednick, "The influence of riparian/wetland systems on surface water quality," in *Restoration, Creation and Management of Wetland and Riparian Ecosystems in the American West*, K. M. Mutz, D. J. Cooper, M. L. Scott, and L. K. Miller, Eds., Symposium of the Rocky Mountain Chapter of the Society of Wetland Scientists, Denver, Colo, USA, 1998.

