Regional Frameworks and Candidate Metrics for Assessing Biotic Integrity in Mid-Atlantic Highland Streams

PAUL L. ANGERMEIER*

U.S. Geological Survey, Virginia Cooperative Fish and Wildlife Research Unit,[1]
Department of Fisheries and Wildlife Sciences,
Virginia Polytechnic Institute and State University,
Blacksburg, Virginia 24061-0321, USA

ROY A. SMOGOR[2]

Department of Fisheries and Wildlife Sciences,
Virginia Polytechnic Institute and State University,
Blacksburg, Virginia 24061-0321, USA

JAY R. STAUFFER

School of Forestry, Pennsylvania State University,
University Park, Pennsylvania 16802, USA

Abstract.—Regionally stratified biological criteria are being used increasingly to assess stream quality. We used multivariate analysis of variance and canonical analysis to examine the utility of two regional frameworks (basins and ecoregions) and 14 candidate metrics of local fish assemblages for assessing the biotic integrity of streams in the mid-Atlantic highlands (montane areas from Pennsylvania to Virginia). In particular, we determined (1) how metrics varied naturally among basins and ecoregions and (2) which metrics varied most consistently with site quality. We also examined the ability of preliminary multimetric indices (MMIs) to distinguish site quality. Metrics varied meaningfully among both basins and ecoregions, but most metrics differed more among basins. The basin effect was especially strong for taxonomic metrics (e.g., number of species [TOTSP]), which reflected the influence of zoogeography on fish community composition. Few metrics differed strongly among both basins and ecoregions. Collectively, metrics distinguished among high-, medium-, and low-quality sites within most regions, but the discriminative ability of individual metrics differed by region. The number of darter or sculpin species (DOSSP) was the only metric related to site quality both in most basins and most ecoregions. Metric differences among site-quality classes were more consistent with a priori expectations within basins than within ecoregions. In each of five regions, we built an MMI from the most discriminative metrics. Only DOSSP and the proportional abundance of tolerants were included in all five MMIs. All MMIs included taxonomic and reproductive (e.g., proportional abundance of simple lithophils, excluding tolerants) metrics, but not all included trophic metrics (e.g., proportional abundance of invertivores). Multimetric indices distinguished between high- and low-quality sites in each region tested, but they usually did not do so to a greater degree than did taxonomic metrics alone. Among the metrics included in MMIs, TOTSP was most consistently related to site quality. Our findings indicate that both basins and ecoregions provide useful frameworks for regionalizing biotic assessments based on fishes and that metric utility may vary considerably among regions even when regions are environmentally similar. To enhance MMI performance, we encourage an increased reliance on region-specific empirical relations in the development of metrics and scoring criteria.

A main goal of environmental assessment is to measure human impacts on ecosystems. Useful as-

* Corresponding author: biota@vt.edu

[1] The Unit is jointly sponsored by the U.S. Geological Survey, the Virginia Polytechnic Institute and State University, and the Virginia Department of Game and Inland Fisheries.

[2] Present address: Illinois Department of Natural Resources, Watershed Management Section, 600 North Grand Avenue West, Springfield, Illinois 62702, USA.

sessment protocols must distinguish between the variation in environmental conditions that occur naturally from place to place and the variation caused by human impacts. Regionalization is a critical component of current approaches to assessing and regulating the quality of aquatic ecosystems. Ecologically defined regions help managers to stratify the natural spatial variation among aquatic ecosystems as well as the variation in ecosystem response to human impacts. Region-specific assessment criteria provide a workable

compromise between the untenable extremes of uniform nationwide criteria and unique criteria for each water body (Hughes et al. 1990). Adoption of regional frameworks has catalyzed recent progress in using biota to assess directly the biological integrity of water resources (Karr 1991). Resource managers and regulators generally agree that ecological regions exist, but there is considerable debate regarding how to define these regions (Omernik 1995; Omernik and Bailey 1997).

The aquatic ecoregions of the U.S. Environmental Protection Agency (US EPA) have been used widely in the context of assessing water quality (Hughes et al. 1990, 1994). These ecoregions are defined on the basis of soil, physiography, natural terrestrial vegetation, and land use, the relative importance of which varies among suites of ecoregions. In some areas (e.g., the mid-Atlantic highlands; Figure 1), ecoregion boundaries largely coincide with physiographic boundaries. Although ecoregions are not defined on the basis of aquatic biota, several studies have demonstrated that variation in aquatic biota among ecoregions often exceeds variation within ecoregions. Most such studies have examined variation in fish-species composition (Hawkes et al. 1986; Larsen et al. 1986; Rohm et al. 1987; Lyons 1989), but a few have also examined periphyton or macroinvertebrate assemblages (Whittier et al. 1988; Barbour et al. 1996; Pan et al. 1996). Most of these studies used data from a single state, even though ecoregions typically transcend state boundaries.

Regional frameworks other than ecoregions may more precisely stratify geographic variation in fish assemblages. For example, variation in fish-species composition is closely tied to the phylogenetic history of drainage basins, including speciation and dispersal events (Hocutt and Wiley 1986; Mayden 1992). This relationship is especially relevant given the widespread use of fish assemblages as bioassessment tools (Southerland and Stribling 1995; US EPA 1996). Direct comparisons of the utility of basins versus that of ecoregions as regional frameworks for biotic assessment (Hughes et al. 1987; Smogor 1996) indicate that no single framework will be most useful in every situation. Omernik and Bailey (1997) argue that the two frameworks are complementary and that their relative utility is often a function of the spatial extent associated with the question being asked. Additional study is needed to discern the circumstances and spatial extents for which each regional framework is more useful. Herein
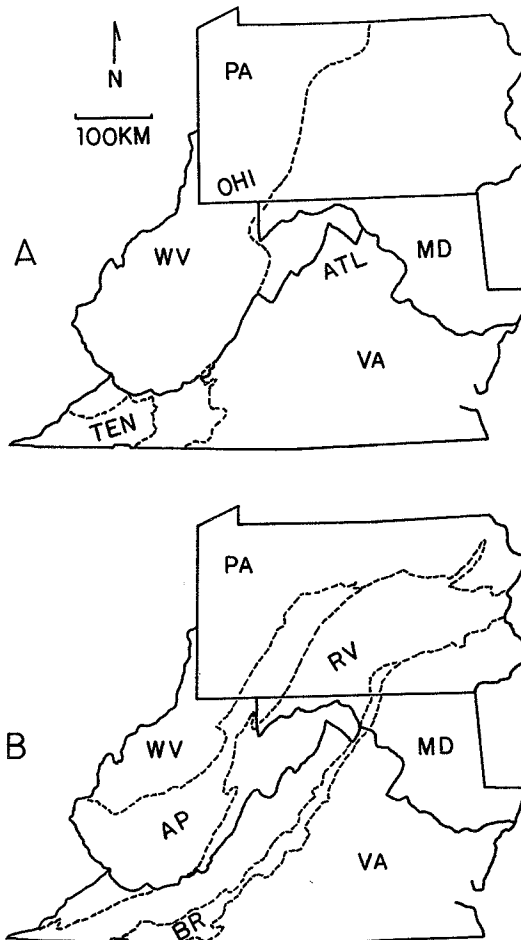

FIGURE 1.—Map of the mid-Atlantic highlands showing boundaries of four states, three drainage basins (A), and three ecoregions (B). The states Maryland (MD), Pennsylvania (PA), Virginia (VA), and West Virginia (WV) are delineated by solid lines. The Atlantic slope (ATL), Ohio River (OHI), and Tennessee River (TEN) basins and the Blue Ridge (BR), Ridge and Valley (RV), and Appalachian Plateau (AP) ecoregions are delineated by dashed lines.

we use "region" generically, and we treat ecoregions and basins as types of regions.

The increasing use of biological indicators has improved assessments of environmental quality relative to assessments based solely on physicochemical indicators. In particular, the use of biologically based multimetric indices (MMIs) has grown rapidly over the last two decades. Multimetric indices typically integrate information on many attributes of a biotic community (one attribute per metric) into a numerical index that is scaled to reflect the ecological health of the com-

munity. The index of biotic integrity (IBI), the first MMI to gain wide acceptance (Karr et al. 1986; Karr 1991), originally used data on fish communities (i.e., local assemblages) to assess the ecological health of streams in the midwestern United States. Later applications of this approach led to MMI development for streams in other regions of North America (Miller et al. 1988; Steedman 1988; Lyons et al. 1995) and in other continents (reviewed in Hughes and Oberdorff 1999). The approach has also been applied to macroinvertebrates (Kerans and Karr 1994; Barbour et al. 1996) and periphyton (Rosen 1995). Continued development of MMIs for new regions is needed to help resource managers and regulators accurately assess the ecological health of water bodies.

A key strength of MMIs as assessment tools is their broad ecological basis, which has two dimensions. First, many MMIs integrate information from multiple levels of biotic organization by including metrics that summarize the health of individual organisms, populations, and the entire community. Second, some metrics describe the taxonomic composition of the community (e.g., number of species), whereas others describe functional composition (e.g., proportional abundance of carnivores). This integrative ecological foundation enables MMIs to detect a broader array of human impacts than can physicochemical measures of water quality, including impacts on flow regime, habitat structure, and biotic interactions (US EPA 1990; Yoder and Rankin 1998).

The utility of MMIs is greatest when they are coupled with a regional framework so that the metrics and their scoring criteria reflect region-specific attributes of natural biotic communities and region-specific responses of those communities to human impacts. For example, dozens of metrics have been substituted for the original IBI metrics as investigators have applied the approach to different ecosystems (Simon and Lyons 1995). Such flexibility enhances the ability of MMIs to accurately reflect biotic responses to environmental degradation. The most common types of community-level metrics reflect taxonomic composition (e.g., number of species), trophic habits (e.g., proportional abundance of omnivores), reproductive habits (e.g., proportional abundance of lithophilic spawners), and tolerance to human impact (e.g., proportional abundance of tolerant species). Region-specific adaptations of MMIs typically have been based on expert knowledge or intuition, but recently developed protocols increasingly rely on empirical relations to select metrics and to derive scoring criteria (Barbour et al. 1995; Barbour et al. 1996; Hughes et al. 1998).

In this study we developed preliminary region-specific MMIs to assess streams in the mid-Atlantic highlands of North America, an area targeted by the US EPA for development of a wide array of environmental assessment tools (e.g., Pan et al. 1996; Jones et al. 1997). Leonard and Orth (1986) developed an IBI for a small portion of this area, but appropriate regions have not been established, nor have metrics with area-wide utility been identified. To be useful in this large area, assessment tools must meet two important challenges. First, they must be sensitive to many common impact types, including acid precipitation, agriculture, coal mining, forestry, and urbanization (Jones et al. 1997). Second, because the area is zoogeographically complex and encompasses parts of two ichthyogeographic provinces (Mississippi and Atlantic Coast; Moyle and Cech 1988), assessment tools must accommodate substantial natural geographic variation in fish communities. Our analyses address four main objectives: (1) to determine how metrics vary naturally (i.e., at least-impacted sites) among basins and ecoregions; (2) to compare the utility of basins and ecoregions as regional frameworks for using MMIs; (3) to determine which metrics best reflect site quality in each potential bioassessment region; and (4) to illustrate the potential ability of regional MMIs to reflect site quality. Although we do not provide fully validated MMIs, we do recommend candidate metrics and identify research needs for further development of these useful indices.

Study Area

The mid-Atlantic highlands area includes the montane portions of Pennsylvania, Maryland, West Virginia, and Virginia (Figure 1). It comprises six ecoregions (Omernik et al. 1992): Blue Ridge Mountains (Blue Ridge), Central Appalachian Ridges and Valleys (Ridge and Valley), Central Appalachian Plateau (Appalachian Plateau), Northern Appalachian Plateau and Uplands, North Central Appalachians, and Western Allegheny Plateau. Most streams in this area drain into the Atlantic Ocean, but some streams in West Virginia and Virginia drain into the Ohio River basin, and streams in extreme southwest Virginia drain into the Tennessee River basin. Our data represent only the first three ecoregions listed above but all three major basins.

TABLE 1.—Numbers of sites for combinations of basin, ecoregion, and site quality (H = high, M = medium, L = low).

Ecoregion	Atlantic			Ohio			Tennessee			Total		
	H	M	L	H	M	L	H	M	L	H	M	L
Blue Ridge	10	3	1	5	1	1	0	1	0	15	5	2
Ridge and Valley	60	62	18	5	4	5	5	11	4	70	77	27
Appalachian Plateau	0	21	1	6	6	11	2	4	5	8	31	17
Total	70	86	20	16	11	17	7	16	9	93	113	46

Methods

Data sets.—The fish collections used in our analyses came from two sources and included site-specific data on species composition, proportional abundance, and habitat and water quality (details are available from P. L. Angermeier, Angermeier et al. 1998). Collections from Pennsylvania and West Virginia were made by J. R. Stauffer and associates from 1980 to 1995 during various surveys in third- to sixth-order streams. Fish were collected with seines or backpack electrofishers from April through October. Collectors attempted to capture all species present and typically continued sampling a site until no new species were captured. Collections from Virginia were made by P. L. Angermeier and associates from 1987 to 1992 during two surveys in third- to sixth-order streams. Virginia fishes were collected from June to September with a standardized electric-seine protocol designed to capture all species present (Angermeier et al. 1991; Smogor 1996). Sample sites were 70–200 m long; the longer sites were located in larger streams. Because low numbers of fish in samples may cause errors in estimating metrics (Fore et al. 1994), we eliminated from both data sets all samples (about 20) that comprised fewer than 50 fish.

The 252 sites retained for analyses were distributed haphazardly among basins and ecoregions (Table 1). Sites were assigned to basins on the basis of known drainage patterns and to ecoregions on the basis of a digitized coverage of ecoregion boundaries (provided by J. M. Omernik; US EPA, Corvallis, Oregon). Because coldwater fish communities are generally restricted to headwater streams (first- to third-order) in the mid-Atlantic highlands, our findings do not apply to these communities.

Each site was assigned retrospectively to a quality class (high, medium, or low), largely on the basis of field observations recorded by the collectors. Site quality was assessed primarily by considering riparian and instream habitat features, adjacent land use, and observed point sources of pollution. High-quality sites appeared to have natural channel morphology and substrate composition, with well-vegetated riparian zones and stable banks. If the streams drained forest, there was no evidence of recent clear-cutting. Some high-quality streams drained pasture but were fenced to prohibit access by livestock. Some low-quality sites were associated with severe impacts, including acid-mine drainage, recent channelization, heavy livestock use, substantial point sources of pollution, and clear-cutting to the streambank. Medium-quality sites exhibited features that were intermediate between those of high- and low-quality sites. Smogor (1996) and Smogor and Angermeier (1999a, in press) provide additional details regarding assignments of Virginia sites to quality classes.

Species attributes.—Each of the 134 fish species in our samples was assigned to one category for each of nine descriptors of native status, taxonomy, trophic and reproductive habits, and tolerance for human impacts (details in Angermeier et al. 1998). Assignments, based on Jenkins and Burkhead (1994), generally followed Angermeier (1995) and Smogor (1996) and were used to compute metrics for each sample (Table 2). Native versus nonnative status of each species in each major river drainage was determined from Jenkins and Burkhead (1994) and from Stauffer et al. (1995). We recognized five major drainages from the Atlantic basin (Potomac, Rappahannock, James, Roanoke, and Pee Dee), three major drainages from the Ohio basin (Monongahela, Kanawha, and Big Sandy), and two major drainages from the Tennessee basin (Clinch and Holston). Minnow, sucker, sunfish, and sculpin species included all species in the families Cyprinidae, Catostomidae, Centrarchidae, and Cottidae, respectively. Darter species included all species in the subfamily Etheostomatini. Assignment to tolerance categories followed previous IBI applications (e.g., Karr et al. 1986; Lyons 1992) and was based on documented changes in overall abundance or geographic range attributable to human impacts. Tolerant and intol-

TABLE 2.—Metrics selected for bioassessment of stream-fish communities and expected responses to human impact. Acronyms used in the text are in parentheses. "Proportional abundance" refers to the number of individual fish of the named species' category relative to the total number in the sample. The tolerance metric TOLPR was grouped with reproductive metrics in analyses.

Metric	Response
Taxonomic	
Number of species (TOTSP)	Negative
Number of nonnative species (INTSP)	Positive
Number of minnow species (MINSP)	Negative
Number of sucker species (SUCSP)	Negative
Number of sunfish species (SUNSP)	Negative
Number of darter or sculpin species (DOSSP)	Negative
Tolerance	
Proportional abundance of tolerants (TOLPR)	Positive
Trophic	
Proportional abundance of nonbenthic generalist feeders (GENPR)	Positive
Proportional abundance of invertivores (INVPR)	Negative
Proportional abundance of benthic specialist invertivores (BSIPR)	Negative
Proportional abundance of carnivores[a] (CARPR)	Negative
Reproductive	
Number of late-maturing species (LAMSP)	Negative
Proportional abundance of nonsimple nonlithophils (NNLPR)	Positive
Proportional abundance of simple lithophils, excluding tolerants (SLTPR)	Negative

[a] All carnivores and piscivores are specialist feeders.

erant species included only those exhibiting the strongest positive and negative responses, respectively, to human impacts.

Feeding ecology was characterized by use of the water column, number of food types eaten, and trophic guild. We designated two water column uses, benthic and nonbenthic. Benthic species are largely restricted to the stream bottom and have specialized adaptations, such as enlarged pectoral fins or inferior mouth position; nonbenthic species are less specialized and feed throughout the water column. We recognized four major food types: (1) detritus, (2) algae or vascular plants, (3) invertebrates, excluding crayfish, and (4) fish (including fish blood) or crayfish. Generalist feeders included species in which three or more food types are major diet components for adults; specialist feeders typically eat a single food type as adults. We assigned species to one of five trophic guilds on the basis of the primary food type of adults. Guilds reflected common gradations from herbivory to piscivory: (1) detritivore–herbivore, (2) herbivore–invertivore, (3) invertivore, (4) invertivore–piscivore, and (5) piscivore. "Carnivores" included species in guilds 4 or 5.

Reproductive behavior was characterized by age

at maturity, spawning substrate, and parental care. We designated two age-at-maturity classes, early and late. Females of late-maturing species typically do not reproduce before 3 years of age. We distinguished between species that require unsilted mineral substrates (e.g., sand, pebble, and boulder) to spawn (i.e., lithophils) and species that are less specialized (nonlithophils). We also distinguished between species that build nests or use cavities in which to spawn or that guard their eggs or young (i.e., nonsimple spawners) and species that do neither (simple spawners).

Metrics.—We examined four types of metrics: taxonomic, tolerance, trophic, and reproductive. Metrics reflecting the health of individual fish were not considered, because those data were not available. Metrics reflecting population health were not examined, because collection methods were not standardized among fish samples. Initially we considered 62 potential metrics (26 taxonomic, 4 tolerance, 18 trophic, and 14 reproductive) based on previous IBI applications (Miller et al. 1988; Simon and Lyons 1995) and personal experience. These metrics reflected the typical effects of human impacts and were easily obtained from field samples. The initial list was reduced to 14 metrics (Table 2) by eliminating those metrics that were strongly correlated with others (i.e., with Pearson's $r \geq 0.90$ or $r \leq -0.90$; Tabachnick and Fidell 1996) or that exhibited very low variance (e.g., zeroes at most sites). Elimination of strongly correlated metrics was necessary, because excessive intercorrelation of variables can destabilize the inversion of the correlation matrix and confound interpretation of the multivariate statistics that we intended to use (Tabachnick and Fidell 1996). All remaining metrics had pairwise correlation coefficients with absolute values of 0.90 or less at high-quality sites within each basin and ecoregion. The single exception to this pattern was a Pearson's r of 0.94 between numbers of species (TOTSP) and minnow species (MINSP) in the Appalachian Plateau, which we allowed because it occurred in only one region and because the default diagnostics in the SAS software identified no collinearity problems during the analyses. The screening process eliminated all but one tolerance metric, proportional abundance of tolerants (TOLPR), which we grouped with reproductive metrics in further analyses.

We limited the metrics in each of the three remaining groups (taxonomic, trophic, and reproductive) to a maximum of six, because our smallest sample size for high-quality sites in any basin or

ecoregion was seven. We sought to avoid ratios of number of metrics to sample size that were 1 or greater, because such ratios can excessively bias inferential tests in multivariate analyses (Tabachnick and Fidell 1996). Even so, in our analysis of MMIs (see Multimetric indices section below), metrics slightly outnumbered samples from a site-quality class in two regions, thereby potentially biasing our inferential tests.

The behavior of proportional abundance metrics expected to be inversely related to human impact (e.g., proportional abundance of invertivores) may be masked by including tolerant species in their computation, because the proportional abundance of tolerants typically increases with impact. We assessed this potential shortcoming by examining metric versions that excluded tolerants (e.g., proportional abundance of invertivores, excluding tolerant species). We inspected box plots (not shown) of selected metrics and their tolerants-excluded counterparts to discern obvious differences in distribution or pattern among site-quality classes. No such differences were apparent, except for the proportional abundance of simple lithophilic spawners, which was the only tolerants-excluded metric selected for further analysis (Table 2). Excluding tolerants from this metric increased the separation of values between high- and low-quality sites and made relations among site-quality classes more consistent with the expectation that we would find greater values at higher quality sites.

Analyzing metric variation.—For each metric group, we used one-way multivariate analysis of variance (MANOVA) and canonical analysis (descriptive discriminant function analysis) to examine how metrics at high-quality sites varied among basins and among ecoregions. We focused on high-quality sites in these analyses because we were interested in natural interregional variation in fish communities. However, we could not test simultaneously for basin and ecoregion "effects", because some basin–ecoregion combinations were represented inadequately in the data sets. We used the same two analytical methods to examine how metrics varied among site-quality classes within each region (e.g., basin). Sample sizes were too small to allow these analyses in a few regions (e.g., only two low-quality sites in the Blue Ridge ecoregion).

To interpret MANOVAs and canonical analyses, we used measures analogous to the coefficient of determination (R^2) in regression. As a measure of overall explanatory power in MANOVA, we used Pillai's trace (Pt; SAS Institute 1990) divided by

one less than the number of statistical groups (s). We used the canonical r^2 (can. r^2) to represent the proportion of generalized variance in metrics that is attributable to differences among statistical groups (i.e., regions or site-quality classes) for individual canonical functions. Pt divided by s is the average can. r^2 across all canonical functions.

Confident interpretation of MANOVAs and canonical analyses requires that one meet assumptions of multivariate normality and homogeneity of variance–covariance structure. We did not test for multivariate normality, because no conclusive test was readily available. However, we increased the likelihood of normality and homogeneity by transforming metrics to meet analogous univariate assumptions (Tabachnick and Fidell 1996) prior to analyses. For species counts we used a square-root transformation ($[x + 0.5]^{0.5}$); for proportions we used an arcsine transformation (arcsine $[x^{0.5}]$; Sokal and Rohlf 1981). Because we found few outliers and acceptable symmetry in distributions of MANOVA residuals of the transformed dependent variables, we believe that our interpretations are not confounded by bias due to unmet assumptions of normality.

Multimetric indices.—We developed a MMI for each of the five regions in which at least one metric group (i.e., taxonomic, trophic, or reproductive metrics) varied meaningfully and reliably among site-quality classes. We used two criteria to make this judgment: (1) for any metric group, can. $r^2 \geq 0.10$ and $P \leq 0.10$ for the first canonical function, which represents the maximal differences in metrics among site-quality classes; and (2) the first canonical axis represented a site-quality gradient (i.e., mean canonical scores of medium-quality sites were between mean scores of low- and high-quality sites.

In each MMI we included metrics that varied meaningfully and reliably among site-quality classes. We selected metrics on the basis of three criteria: (1) each metric belonged to a group shown to differ among site-quality classes in the region (i.e., can. $r^2 > 0.10$ and $P < 0.10$); (2) each metric had a large (≤ -0.45 or ≥ 0.45) canonical structure coefficient, which is the Pearson's r between metric values and canonical scores; and (3) the relationship between each metric and site quality was consistent with the expected response to human impacts (Table 2). The third criterion was assessed by examining signs of structure coefficients and box plots (not shown) of metric values by site-quality class. Each MMI included metrics from two or three metric groups. Canonical analyses of

TABLE 3.—Results of one-way multivariate analyses of variance in taxonomic, trophic, and reproductive metrics from high-quality sites. Separate analyses were performed on variance among Atlantic (ATL), Ohio (OHI), and Tennessee (TEN) basins and among Blue Ridge (BR), Ridge and Valley (RV), and Appalachian Plateau (AP) ecoregions. Values of Pt/s[a] are also shown for contrasts between pairs of basins and ecoregions.

Metric	df	Pt/s	Approximate F	P	Pt/s for contrasts
Variance among basins					
Taxonomic	12, 172	0.33	6.91	0.0001	ATL versus OHI: 0.14
					ATL versus TEN: 0.51
					OHI versus TEN: 0.48
Trophic	8, 176	0.17	4.37	0.0001	ATL versus OHI: 0.20
					ATL versus TEN: 0.18
					OHI versus TEN: 0.08
Reproductive	8, 176	0.14	3.70	0.0005	ATL versus OHI: 0.21
					ATL versus TEN: 0.15
					OHI versus TEN: 0.02
Variance among ecoregions					
Taxonomic	12, 172	0.15	2.46	0.0055	BR versus RV: 0.19
					BR versus AP: 0.20
					RV versus AP: 0.10
Trophic	8, 176	0.14	3.63	0.0006	BR versus RV: 0.24
					BR versus AP: 0.04
					RV versus AP: 0.08
Reproductive	8, 176	0.15	4.03	0.0002	BR versus RV: 0.29
					BR versus AP: 0.10
					RV versus AP: 0.04

[a] Pt/s = Pillai's trace (Pt) divided by one less than the number of statistical groups (s).

MMI metrics, when compared to analyses of metric groups separately, enabled us to judge if combinations of metrics from different groups were more closely related to site quality than combinations from a single group.

Data values for each metric selected for a MMI were scored with a 3, 2, or 1, following a protocol similar to that of Barbour et al. (1996). A score of 3 represented the least-impacted condition; 1 represented the most-impacted condition. For metrics that were positively related to site quality (e.g., number of species), the lowest quartile (25th percentile) of the distribution of metric values from high-quality sites was used as the minimum value for scoring a 3. The minimum metric value for scoring a 2 was one-half the value of the 25th percentile; lower metric values were scored a 1. For metrics negatively related to site quality (e.g., proportional abundance of tolerants), we used an analogous mirror-image process focused on the 75th percentile. The MMI score for each site was the sum of the metric scores; thus, a MMI with five metrics potentially ranged from 5 to 15.

In many applications of the IBI, scoring criteria for species-richness metrics are adjusted for stream size (Simon and Lyons 1995). In general, the need for such adjustments should be evaluated in a region-specific manner (Smogor and Anger-

meier 1999a). Because we were unable to acquire adequate stream-size information for sites outside Virginia, we did not attempt to adjust scoring criteria to reflect the potential influences of stream size.

For each region, we used a Kruskal–Wallis test (SAS Institute 1990) to assess the relation between MMI scores and site quality. We also used canonical analyses to examine relations between component metrics and site quality.

Results

Metric Variation among Basins and Ecoregions

For high-quality sites, omnibus MANOVAs indicated that all three groups of metrics varied meaningfully (all Pt/s values > 0.13) and reliably (all P values < 0.006) among some basins and some ecoregions (Table 3). Trophic and reproductive metrics differed as much among basins as they did among ecoregions, but taxonomic metrics differed more among basins than they did among ecoregions. All metric groups differed to a similar degree among ecoregions, but basin effects were stronger for taxonomic metrics than for trophic or reproductive metrics (Table 3). Contrasts among basins indicated that the Tennessee was the most taxonomically distinctive, whereas the Atlantic

was the most trophically and reproductively distinctive. Contrasts among ecoregions indicated that the Blue Ridge was the most taxonomically and reproductively distinctive, whereas the Ridge and Valley was most trophically distinctive (Table 3).

Canonical analyses helped explain MANOVA patterns by revealing which metrics in each group differed most among particular regions. We report results for only the first canonical axes, because in only one case (variation in taxonomic metrics among basins) did subsequent axes at least 10% of the variance in metrics (our criterion for meaningful variation). The greatest variation was observed for taxonomic metrics among basins, especially number of darter or sculpin species (DOSSP) and TOTSP (Table 4). This pattern reflected the greater species richness at sites in the Tennessee. Trophic differences among basins were most closely related to variation in proportional abundances of nonbenthic generalist feeders (GENPR) and invertivores (INVPR) (Table 4), which were highest in the Atlantic. Reproductive differences among basins were most closely related to variation in proportional abundances of simple lithaphils, excluding tolerants (SLTPR) and nonsimple nonlithaphils (NNLPR) and in TOLPR (Table 4). The latter two metrics were highest in the Atlantic, whereas SLTPR was lowest there.

The metrics that varied most among ecoregions generally were not the same as those that varied most among basins. Taxonomic differences among ecoregions were largely due to variations in numbers of sucker species (SUCSP) and nonnative species (INTSP; Table 4); these matrices were highest in the Blue Ridge. Trophic differences among ecoregions were largely due to variation in GENPR and INVPR (Table 4); these metrics were highest in the Ridge and Valley. Reproductive differences among ecoregions were most closely related to variation in SLTPR, number of late-maturing species (LAMSP), and TOLPR (Table 4); SLTPR and LAMSP were highest in the Blue Ridge, whereas TOLPR was lowest there. In summary, some metrics from each of the three groups differed meaningfully among basins and among ecoregions, but the particular metrics distinguishing basins were usually different from those distinguishing ecoregions. Only the trophic metrics GENPR and INVPR and the reproductive metrics SLTPR and TOLPR were among the strongest contributors to both interbasin and interecoregion differences for their respective metric groups.

For high-quality sites, we also examined metric

TABLE 4.—Results of canonical analyses of taxonomic, trophic, and reproductive metrics from high-quality sites. Separate analyses examined metric differences among Atlantic (ATL), Ohio (OHI), and Tennessee (TEN) basins and among Blue Ridge (BR), Ridge and Valley (RV), and Appalachian Plateau (AP) ecoregions. Metric acronyms are defined in Table 2. Canonical r^2 is the proportion of generalized variance in metrics explained by basin or ecoregion. Structure coefficients reflect correlations between metrics and canonical scores. Region means reflect positions of sites, grouped by region, along each canonical axis.

Metric or statistic	Basins, axis 1		Ecoregions, axis 1
Taxonomic			
Canonical r^2	0.52		0.22
P	0.0001		0.0048
df	12, 170		12, 170
Structure coefficients			
TOTSP	−0.46		0.23
MINSP	−0.32		0.30
SUCSP	−0.21		0.37
SUNSP	−0.12		−0.07
DOSSP	−0.72		−0.05
INTSP	0.20		0.37
Region means			
ATL	0.24	BR	1.10
OHI	0.50	RV	−0.13
TEN	−3.56	AP	−0.91
Trophic			
Canonical r^2	0.26		0.26
P	0.0001		0.0004
df	8, 174		8, 174
Structure coefficients			
GENPR	0.57		0.54
INVPR	0.54		0.53
BSIPR	−0.13		−0.04
CARPR	−0.11		0.03
Region means			
ATL	0.33	BR	−1.19
OHI	−0.89	RV	0.32
TEN	−1.29	AP	−0.61
Reproductive			
Canonical r^2	0.27		0.29
P	0.0003		0.0001
df	8, 174		8, 174
Structure coefficients			
LAMSP	−0.38		0.64
TOLPR	0.60		−0.62
NNLPR	0.69		−0.51
SLTPR	−0.69		0.66
Region means			
ATL	0.34	BR	1.39
OHI	−0.99	RV	−0.32
TEN	−1.17	AP	0.15

differences among basin–ecoregion combinations, but in this pursuit we were constrained by small sample sizes for some combinations. Because metrics differed little between Ridge and Valley and Appalachian Plateau (see contrasts in Table 3), we

pooled sites in these two ecoregions. Even so, only four combinations had sufficient numbers of high-quality sites to allow valid multivariate comparisons: Atlantic–Blue Ridge; Atlantic–Ridge and Valley–Appalachian Plateau; Ohio–Ridge and Valley–Appalachian Plateau; and Tennessee–Ridge and Valley–Appalachian Plateau. One-way MANOVAs indicated that metrics differed among basin–ecoregion combinations to a degree very similar to that observed among basins alone (Pt/s = 0.26, 0.15, and 0.15 for taxonomic, trophic, and reproductive metrics, respectively; all P values = 0.0001; Table 3). As was the case for basins, taxonomic metrics differed more among basin–ecoregion combinations than did trophic or reproductive metrics.

Metric Variation among Site-Quality Classes within Regions

Site quality within basins.—Metric differences among site-quality classes were inconsistent among metric groups and among basins. Results from MANOVA (not shown) indicated that none of the three metric groups varied meaningfully with site quality in the Atlantic (all Pt/s values = 0.04; all P values > 0.08). Overall, metrics differed most among site-quality classes in the Tennessee, and the greatest differences occurred between high- and low-quality sites.

Canonical analyses clarified relations between individual metrics and site quality within basins. We report results for only the first canonical axes, because subsequent axes did not explain a statistically significant ($P < 0.05$) portion of the metric variance within basins or ecoregions. In the Ohio, six metrics were strongly correlated with canonical scores (i.e., structure coefficients ≤ -0.45 or ≥ 0.45); TOTSP, INVPR, and SLTPR were most closely related to site quality in their respective metric groups (Table 5). In the Tennessee, eight metrics were strongly correlated with canonical scores; DOSSP, GENPR, and LAMSP were most closely related to site quality in their respective metric groups (Table 5). Although the metrics exhibiting the strongest relations with site quality differed among basins, five metrics (TOTSP, MINSP, DOSSP, TOLPR, and SLTPR) were related with site quality in both the Ohio and Tennessee basins.

For each basin, we plotted sites along the first canonical axes to illustrate differences in metrics among site-quality classes. We present only the plot for the Tennessee basin (Figure 2), which showed the greatest separation among site-quality classes for all three metric groups. For all metric groups in the Ohio and Tennessee, class means were ordered as expected (i.e., medium between high and low; Table 5). All groups, except trophic metrics in the Ohio, were related to site quality as expected; contrary to expectations, INVPR in the Ohio was inversely related to site quality (Table 5). Thus, five canonical axes represented interpretable site-quality gradients within basins.

Site quality within ecoregions.—As was the case for basins, metric differences among site-quality classes were inconsistent among metric groups and among ecoregions. Results from MANOVA (not shown) indicated that no metric group differed reliably with site quality in the Blue Ridge (all P values > 0.52), and reproductive metrics differed unreliably with site quality in the Ridge and Valley (Pt/s = 0.04; P = 0.06) and the Appalachian Plateau (Pt/s = 0.10; P = 0.20). Overall, metrics differed most among site-quality classes in the Appalachian Plateau, and the greatest differences typically occurred between medium- and low-quality sites.

Canonical analyses clarified relation between individual metrics and site quality within ecoregions. In the Ridge and Valley, four taxonomic and two trophic metrics differed meaningfully and reliably with site quality; MINSP and INVPR were most closely related to site quality in their respective metric groups (Table 5). Taxonomic metrics in the Ridge and Valley exhibited patterns similar to those seen in the Ohio and Tennessee basins. In the Appalachian Plateau, only one taxonomic and two trophic metrics differed among site-quality classes; DOSSP and proportional abundance of benthic specialist invertivores (BSIPR) were most closely related to site quality in their respective metric groups (Table 5). The metrics exhibiting the strongest relations with site quality differed among ecoregions, and only two metrics (DOSSP and GENPR) were related with site quality in both the Ridge and Valley and Appalachian Plateau ecoregions. Number of darter or sculpin species was the only metric related to site quality in most basins and in most ecoregions.

For each ecoregion, we plotted sites along the first canonical axes to illustrate differences in metrics among site-quality classes (plots not shown). For taxonomic and trophic metrics, only the class means in the Appalachian Plateau were ordered as expected (i.e., medium between high and low) and related to site quality as expected (Table 5). Thus, only two canonical axes represented interpretable site-quality gradients within ecoregions.

TABLE 5.—Results of canonical analyses of taxonomic, trophic, and reproductive metrics from all sites. Separate analyses examined metric differences among site-quality classes within each of three basins: Atlantic (ATL), Ohio (OHI), and Tennessee (TEN), and each of three ecoregions: Blue Ridge (BR), Ridge and Valley (RV), and Appalachian Plateau (AP). Metric acronyms are explained in Table 2. Canonical r^2 is the proportion of generalized variance in metrics explained by site-quality class. Structure coefficients reflect correlations between metrics and canonical scores; coefficients are shown for only those axes with $P < 0.10$. Class means reflect positions of sites, grouped by quality class, along each canonical axis.

Metric or statistic	ATL, axis 1	OHI, axis 1	TEN, axis 1	BR, axis 1	RV, axis 1	AP, axis 1
Taxonomic						
Canonical r^2	0.07	0.39	0.69	0.34	0.11	0.36
P	0.2475	0.0257	0.0007	0.7314	0.0060	0.0066
df	12, 336	12, 72	12, 48	12, 28	12, 332	12, 96
Structure coefficients						
TOTSP		0.73	0.81		0.76	0.15
MINSP		0.64	0.57		0.83	0.08
SUCSP		0.42	0.24		0.14	0.16
SUNSP		0.31	0.23		0.50	−0.17
DOSSP		0.57	0.86		0.68	0.71
INTSP		0.44	0.03		0.36	0.15
Quality-class means						
High	0.20	0.75	2.08	−0.06	−0.27	0.65
Medium	0.01	0.37	0.16	0.87	0.21	0.44
Low	−0.73	−0.95	−1.90	−1.76	0.11	−1.10
Trophic						
Canonical r^2	0.05	0.28	0.38	0.18	0.10	0.23
P	0.0852	0.0934	0.0767	0.7136	0.0089	0.0086
df	8, 340	8, 76	8, 52	8, 32	8, 336	8, 100
Structure coefficients						
GENPR	0.68	0.15	0.70		0.77	0.51
INVPR	−0.53	0.90	0.05		−0.78	−0.12
BSIPR	0.19	0.06	−0.63		−0.39	−0.80
CARPR	0.42	−0.05	0.03		0.24	0.14
Quality-class means						
High	−0.06	−0.51	−1.02	−0.01	−0.25	−0.38
Medium	0.17	−0.43	−0.14	−0.47	0.36	−0.35
Low	−0.53	0.76	1.05	1.24	−0.40	0.81
Reproductive						
Canonical r^2	0.06	0.36	0.49	0.32	0.08	0.11
P	0.0894	0.0038	0.0087	0.5065	0.0617	0.2068
df	8, 340	8, 76	8, 52	8, 32	8, 336	8, 100
Structure coefficients						
LAMSP	−0.06	0.37	0.86		0.23	
TOLPR	0.67	−0.67	−0.50		−0.52	
NNLPR	−0.70	−0.41	−0.02		0.78	
SLTPR	0.03	0.90	0.66		0.29	
Quality-class means						
High	−0.16	0.72	1.37	0.25	0.13	−0.40
Medium	−0.03	0.33	0.12	0.05	0.11	−0.17
Low	0.69	−0.89	−1.28	−2.00	−0.66	0.51

Metric variation with site quality in basins was more pronounced and more consistent with expectations than was variation in ecoregions. The can. r^2 values for basins typically exceeded those for ecoregions (Table 5), and more canonical axes represented meaningful and reliable site-quality gradients for basins than for ecoregions (five of nine axes versus two of nine axes; Table 5).

Site quality within basin–ecoregions.—We also examined metric variation among site-quality classes within each of two basin–ecoregion combinations (Atlantic–Ridge and Valley–Appalachian Plateau and Ohio–Ridge and Valley–Appalachian Plateau). Combinations including the Blue Ridge ecoregion lacked enough low-quality sites to include in the analysis. We also excluded the Tennessee–Ridge and Valley–Appalachian Plateau combination because it included only two fewer
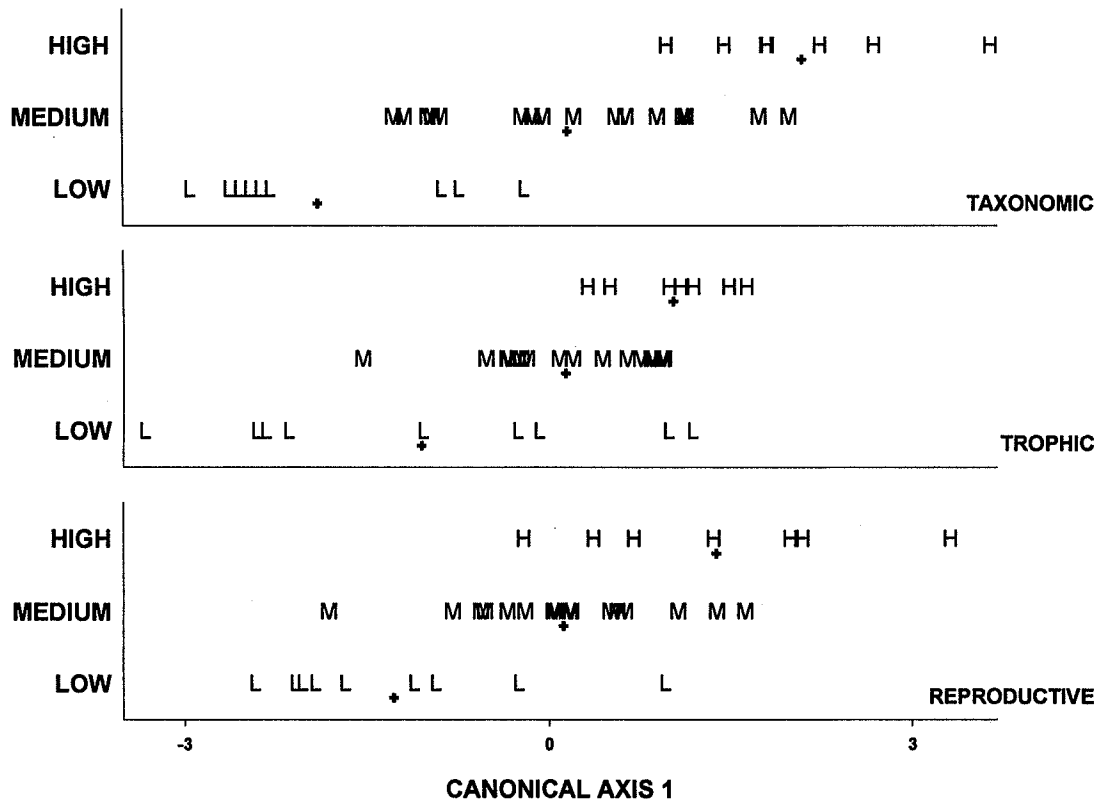
FIGURE 2.—Distributions of sites in the Tennessee River basin along the first canonical axes, representing taxonomic, trophic, and reproductive metrics. High-, medium-, and low-quality sites (H, M, and L, respectively) are plotted separately. Means of quality classes are indicated with pluses just below the rows of plotted points.

sites than the Tennessee basin. One-way MANO-VAs and canonical analyses (not shown) indicated that metric variation among site-quality classes within combinations largely reflected metric variation within basins. In particular, no metric group differed meaningfully in the Atlantic–Ridge and Valley–Appalachian Plateau (all can. r^2 values $\leq$ 0.06), and patterns of metric variation in the Ohio–Ridge and Valley–Appalachian Plateau were similar to, but weaker than, patterns in the Ohio (can. r^2 = 0.39, 0.25, and 0.26 for taxonomic, trophic, and reproductive metrics, respectively; P = 0.083, 0.203, and 0.051, respectively; Table 5). Thus, subdividing basins into basin–ecoregion units did not increase the metric variance explained by site quality, nor did it reveal new patterns of metric variation. However, given that we could not adequately examine most basin–ecoregion combinations as a result of the unevenly distributed samples, our assessment of the utility of such combinations is incomplete.

Multimetric Indices

We built MMIs for five regions: two basins (Ohio and Tennessee), two ecoregions (Ridge and Valley and Appalachian Plateau), and a basin–ecoregion combination (Ohio–Ridge and Valley–Appalachian Plateau). We relied on canonical analyses and univariate box plots to select component metrics for MMIs (see Methods—Multimetric indices). However, we slightly relaxed our initial criteria for metric selection in order to allow reproductive metrics to be represented in all MMIs. Three exceptions seemed reasonable to ensure that the MMIs had a broad ecological basis. For the Ridge and Valley MMI, we included TOLPR, even though variation in reproductive metrics (as a group) among site-quality classes was low (can. r^2 = 0.08; Table 5). Similarly, for the Appalachian Plateau MMI, we included TOLPR and SLTPR, even though reproductive metrics (as a group) differed unreliably with site quality (P = 0.21; Table 5). Box plots showed that in all three exceptions,

TABLE 6.—Scoring criteria for selected metrics in each of five regions: Ohio (OHI) and Tennessee (TEN) basins, Ridge and Valley (RV) and Appalachian Plateau (AP) ecoregions, and Ohio–Ridge and Valley–Appalachian Plateau (ORVAP) basin–ecoregion combination. Acronyms for metrics are given in Table 2. Greater scores reflect higher site quality.

Region	Metric	Score = 1	Score = 2	Score = 3
OHI	TOTSP	0–6	7–12	≥13
	MINSP	0–3	4–6	≥7
	DOSSP	0	1	≥2
	TOLPR	0.65–1.00	0.30–0.64	0.00–0.29
	SLTPR	0.00–0.07	0.08–0.15	0.16–1.00
TEN	TOTSP	0–10	11–21	≥22
	MINSP	0–4	5–8	≥9
	DOSSP	0–2	3–5	≥6
	GENPR	0.50–1.00	0.006–0.49	0.000–0.005
	BSIPR	0.00–0.11	0.12–0.23	0.24–1.00
	LAMSP	0–1	2	≥3
	TOLPR	0.60–1.00	0.21–0.59	0.00–0.20
	SLTPR	0.00–0.11	0.12–0.23	0.24–1.00
RV	TOTSP	0–4	5–9	≥10
	MINSP	0–2	3–4	≥5
	SUNSP	0	1	≥2
	DOSSP	0	1	≥2
	TOLPR	0.79–1.00	0.57–0.78	0.00–0.56
AP	DOSSP	0	1	≥2
	GENPR	0.58–1.00	0.16–0.57	0.00–0.15
	BSIPR	0.00–0.03	0.04–0.07	0.08–1.00
	TOLPR	0.77–1.00	0.53–0.76	0.00–0.52
	SLTPR	0.00–0.05	0.06–0.12	0.13–1.00
ORVAP	TOTSP	0–5	6–10	≥11
	MINSP	0–2	3–4	≥5
	SUNSP	0	1	≥2
	DOSSP	0	1	≥2
	TOLPR	0.72–1.00	0.43–0.71	0.00–0.42
	SLTPR	0.00–0.07	0.08–0.15	0.16–1.00

the metric varied, as expected, from high- to low-quality sites in the region. Finally, for the Ohio–Ridge and Valley–Appalachian Plateau MMI, we included TOLPR and SLTPR, even though the order of site-quality class means for reproductive metrics was slightly incongruous (i.e., the mean for medium-quality sites slightly exceeded the mean for high-quality sites).

Metric composition of the MMIs differed by region (Table 6). Only DOSSP and TOLPR were selected in all five regions, but three metrics (TOTSP, MINSP, and SLTPR) were selected in four regions. All five MMIs included at least one taxonomic and one reproductive metric, but trophic metrics were included inconsistently (Table 6). Only the MMI for the Tennessee basin included multiple metrics from all three groups. Two taxonomic metrics (SUCSP and INTSP), two trophic metrics (INVPR and proportional abundance of carnivores; CARPR), and one reproductive metric (NNLPR) were never selected for a MMI.

Of those included in MMIs, several metrics differed only slightly among site-quality classes within regions. For example, in four of the five MMIs

that included DOSSP and in both MMIs that included number of sunfish species (SUNSP), the criteria for receiving a score of 1 versus a score of 3 differed by only two species (Table 6). The metric LAMSP exhibited a similarly narrow range of variation in the Tennessee basin (Table 6). Among proportional metrics, SLTPR varied least with site quality; in three of the four MMIs in which it was included, the criteria for receiving a score of 1 versus a score of 3 differed by less than 0.10 (Table 6). Results for GENPR were unusual in that the values needed to score a 3 were very low (≤0.005) in the Tennessee (Table 6). Additional study would be needed to verify whether estimates of these metrics would be precise enough for them to consistently distinguish sites of differing quality.

Values of MMIs varied meaningfully and reliably (all P values < 0.011 for Kruskal–Wallis tests) among site-quality classes in each region. Medium-quality sites differed from low-quality sites in each region, but they were distinct from high-quality sites only in the Tennessee basin (Figure 3). The two MMIs for basins showed the most
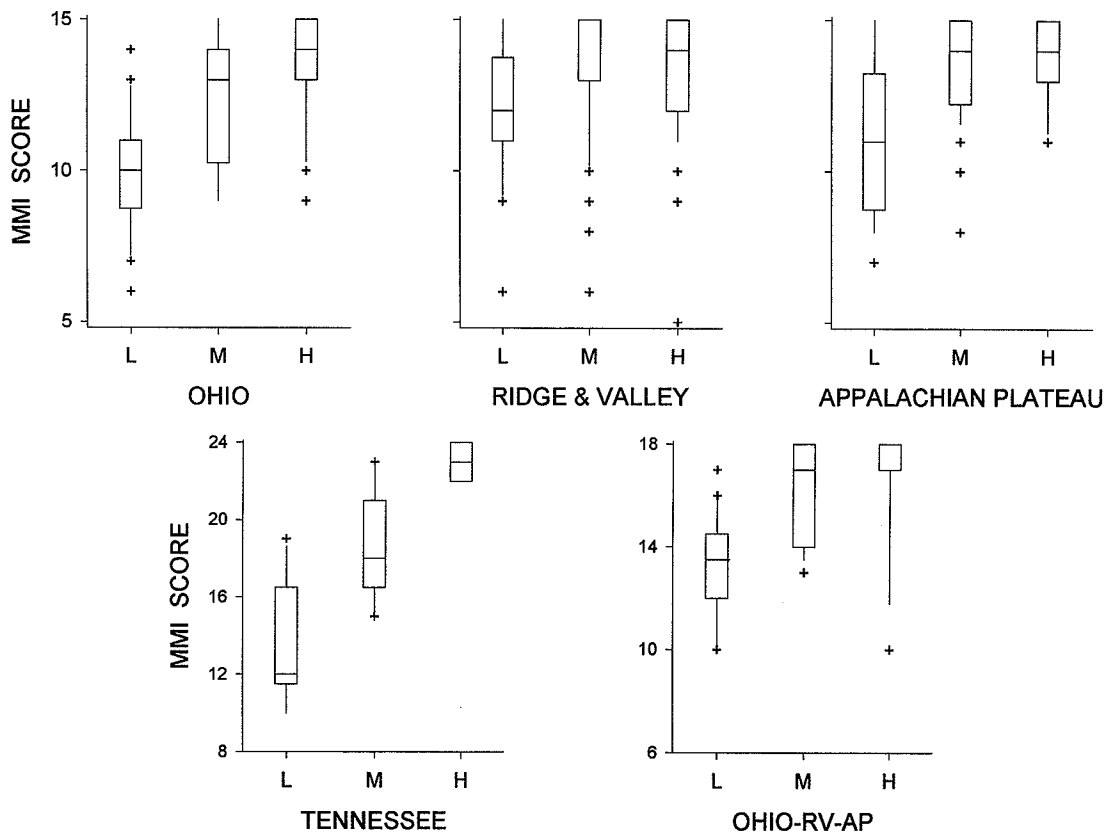
FIGURE 3.—Box plots of multimetric index (MMI) scores for sites in five regions of the mid-Atlantic highlands. High-, medium-, and low-quality sites (H, M, and L, respectively) are plotted separately. Boxes encompass the 25th to 75th percentiles. Horizontal lines inside the boxes represent medians (50th percentiles), but in two cases the median coincides with the 25th percentile. Vertical lines extending from the boxes encompass the 10th to 90th percentiles. Pluses represent more extreme outliers. Numbers on each y-axis reflect the minimum and maximum values possible for MMI scores.

discrimination among site-quality classes, whereas the two MMIs for ecoregions showed the least discrimination (Figure 3). In neither ecoregion did median MMI scores for high-quality sites exceed those for medium-quality sites.

Our MMI metrics, which included metrics from multiple metric groups, reliably discriminated among site-quality classes, but they usually did so no more than did taxonomic metrics alone. For all five regions with MMIs, MMI metrics varied with site quality to a similar degree as taxonomic metrics (compare Tables 5 and 7). In only one region (the Ohio–Ridge and Valley–Appalachian Plateau) were MMI metrics more strongly related to site quality than were taxonomic metrics (can. $r^2 = 0.46$ versus 0.39). This pattern reflects the consistently weaker relations that trophic and reproductive metrics exhibited with site quality, compared to relations exhibited by taxonomic metrics (Table 5).

In each region, the metrics most closely related to the first canonical axis in the analysis of the MMI metrics (Table 7) were also most closely related to the first canonical axis in the analyses of individual metric groups (Table 5). Among metrics included in MMIs, TOTSP was most consistently and strongly related to site quality. Structure coefficients for TOTSP were among the highest in every region except the Appalachian Plateau (Table 7), where TOTSP was not included in the MMI. Structure coefficients for DOSSP were among the highest in three regions (Tennessee, Ridge and Valley, and Appalachian Plateau).

Discussion

Basins Versus Ecoregions as Regional Frameworks

The primary goal of defining bioassessment regions is to stratify geographic variation in natural

TABLE 7.—Results of canonical analyses of the metrics included in the multimetric index for each of five regions. Region acronyms are explained in Table 6. Separate analyses examined metric differences among three site-quality classes within each region. Metric acronyms are explained in Table 2. Canonical r^2, structure coefficients, and class means are explained in Table 5. Empty cells indicate that a metric was not included in the region's index.

Statistic or metric	OHI, axis 1	TEN, axis 1	RV, axis 1	AP, axis 1	ORVAP, axis 1
Canonical r^2	0.38	0.69	0.11	0.28	0.46
P	0.0142	0.0045	0.0048	0.0094	0.0208
df	10, 74	16, 44	10, 334	10, 98	12, 58
Structure coefficients					
Taxonomic metrics					
TOTSP	0.74	0.80	0.76		0.58
MINSP	0.66	0.56	0.83		0.42
SUNSP			0.48		0.45
DOSSP	0.58	0.84	0.68	0.89	0.46
Trophic metrics					
GENPR		−0.35		−0.39	
BSIPR		0.32		0.67	
Reproductive metrics					
LAMSP		0.56			
TOLPR	−0.63	−0.32	−0.39	−0.35	−0.44
SLTPR	0.86	0.43		0.51	0.55
Quality-class means					
High	0.79	2.18	0.08	0.67	0.90
Medium	0.29	0.11	0.20	0.32	0.62
Low	−0.93	−1.90	−0.80	−0.90	−1.00

biotic conditions and in ecosystem response to human impacts. Properly defined regions (strata) enhance the precision and accuracy of assessments by minimizing the confounding effects of geographic variation. Both drainage basins and aquatic ecoregions provide reasonable frameworks for regionalizing bioassessments because they represent regions with distinctive fish assemblages (Table 3). In the mid-Atlantic highlands, ecoregions correspond closely with physiographic provinces, which strongly influence fish distribution (Jenkins and Burkhead 1994; Stauffer et al. 1995). However, drainage basins more clearly reflected the inherent geographic variation in the fish-community attributes we analyzed, primarily because of pronounced interbasin variation in taxonomic metrics. In particular, values for DOSSP and TOTSP were especially large in the Tennessee basin, which is a center of fish diversification in North America (McAllister et al. 1986). Smogor and Angermeier (in press) found that taxonomic metrics were also more variable than were trophic and reproductive metrics among physiographic regions of Virginia. If bioassessment regions of the mid-Atlantic highlands are based on ecoregions, scoring criteria for taxonomic, trophic, and reproductive metrics would differ to a similar degree among regions; if regions are based on basins, differences in scoring criteria among regions would predominantly be associated with taxonomic metrics and

would reflect historical differences in fish speciation and dispersal. Nevertheless, scoring criteria for a few metrics (e.g., GENPR, INVPR, and SLTPR) would exhibit interregional differences regardless of which framework was used (Table 4).

Interregional variation in taxonomic metrics was only loosely related to interregional variation in functional (i.e., trophic and reproductive) metrics. We expect interregional variation in functional metrics to generally reflect variation in environmental constraints, such as productivity or bottom type, and we expect variation in taxonomic metrics to reflect variation in either environmental constraints or phylogenetic history. That is, two streams may exhibit differences in taxonomic metrics because of different zoogeographic settings rather than because of different environmental conditions. For fishes and other aquatic fauna that possess a limited ability to disperse over land, zoogeographic boundaries usually coincide with basin boundaries (Hocutt and Wiley 1986; Mayden 1992). The failure of functional metrics to vary as much among basins as do taxonomic metrics suggests that the additional species in speciose communities (e.g., in the Tennessee basin) are often functionally similar, at least with respect to our coarse descriptions of function. However, taxonomic and functional metrics may covary if functionally distinctive species have limited geographic distributions due to dispersal barriers. For ex-

ample, of the 22 species of carnivores represented in our data, the number that are native to Virginia portions of the mid-Atlantic highlands is highly variable, even within basins (Jenkins and Burkhead 1994). Seven carnivores are native to the Roanoke drainage and one to the Pee Dee drainage (both in the Atlantic basin); three carnivores are native to the Kanawha and seven to the Big Sandy (both in the Ohio basin). Notably, the Roanoke, Pee Dee, and Kanawha adjoin each other. Such geographic variation in taxonomic composition could cause geographic variation in potential ranges of proportional abundance metrics (e.g., CARPR) and could reduce the ability of these metrics to detect human impacts in regions that transcend zoogeographic boundaries (e.g., some ecoregions).

Interrelations between taxonomic and functional metrics dictate caution when one is transferring scoring criteria for functional metrics among regions, even when the regions are environmentally similar, as is the case in the mid-Atlantic highlands. The supposition that geographic variation in functional metrics only reflects environmental variation may not be valid. Although functional metrics often vary less among adjoining regions than do taxonomic metrics (Smogor and Angermeier, in press; this study), interregional differences in functional metrics can be substantial (Table 4), especially when regions are zoogeographically distinctive.

The relative ability of basins and ecoregions to stratify geographic variation in biotic attributes and human impacts likely depends on the geographic extent of the analysis (Omernik and Bailey 1997). Geographic extent influences the range of observed variation (in attributes and impacts) and the juxtaposition of the areas in which the extremes in variation occur (i.e., whether the extremes occur in the same region). Our findings illustrate that interecoregion differences in metrics are likely small or difficult to discern when the considered ecoregions are similar (e.g., all highlands). Conversely, interbasin differences in metrics are likely obscured if individual basins drain composites of very dissimilar ecoregions (e.g., lowlands and highlands), thereby inflating metric variation within basins. In a statewide analysis of Virginia fish communities, Smogor and Angermeier (in press) found that physiography accounted for more variation in community metrics than did river drainage or ecoregion. Several drainages in that analysis drained both lowland (Coastal Plain) and highland physiography, which exhibited the greatest inter-

regional differences in taxonomic, trophic, and reproductive metrics. The interbasin differences in taxonomic metrics observed herein largely reflected that the mid-Atlantic highlands span two distinct ichthyogeographic provinces (Moyle and Cech 1988). Areas that are similar in size to these highlands but that are confined to a single ichthyogeographic province (e.g., much of the midwestern United States) are less likely to exhibit such large interbasin differences in taxonomic metrics.

Neither basins nor ecoregions provide an ideal regional framework for biotic assessments based on fish metrics. Basins may not correspond with geographic differences in functional attributes associated with key environmental conditions; ecoregions may not correspond with geographic differences in taxonomic composition associated with zoogeography. Intuitively, we expect frameworks that incorporate both types of region to most effectively minimize confounding effects of geographic variation in fish metrics, thereby enhancing precision and accuracy of assessments. For example, each drainage–physiography combination in Virginia supports a distinctive type of fish community (Angermeier and Winston 1999); some of this geographic variation is likely expressed in metrics that are relevant to bioassessments. Although the basin–ecoregion combinations examined herein explained no more variation in fish metrics than did basins alone (Table 4), our analyses were constrained by small sample sizes and lack of information for some combinations. We believe that such combinations warrant further examination in the mid-Atlantic highlands and in other areas in which bioassessment tools are being developed.

Metric Utility in Assessing Stream Quality

The metrics most closely related to site quality differed among basins, among ecoregions, and between basins and ecoregions. No metric group was related strongly to site quality in the Atlantic basin or the Blue Ridge ecoregion. Only five metrics (TOTSP, MINSP, DOSSP, TOLPR, and SLTPR) were related to site quality in both the Ohio and Tennessee basins (Table 5). Only two metrics (DOSSP and GENPR) were related to site quality in both the Ridge and Valley and Appalachian Plateau ecoregions (Table 5). These lists also illustrate the inconsistent utility of the three metric groups. The list for basins lacks a trophic metric; the list for ecoregions lacks a reproductive metric. In their analysis of Virginia fish communities, Smogor and

Angermeier (1999b) also found considerable interregional variation in relations between metrics and site quality. Based on our analyses in regions in which metrics were related to site quality, five metrics seem to be broadly useful in the mid-Atlantic highlands: TOTSP, MINSP, DOSSP, TOLPR, and SLTPR (Table 6). However DOSSP and SLTPR require further study to determine if they can be estimated precisely enough to reliably detect differences in site quality.

Some metrics were positively related with site quality in some regions but were negatively related in other regions. The trophic metric INVPR was negatively related to site quality in the Ohio basin (Table 5), a pattern opposite that exhibited in our other regions as well as in previous IBI applications. For montane sites in Virginia, Smogor and Angermeier (1999b) also found that trophic metrics exhibited more contradictory relations with site quality than did other metric groups. The unpredictability of relations between some metrics and site quality underscores the need to empirically validate expected relations in order to ensure the accuracy of bioassessments. Despite the broad utility of some metrics and the multimetric approach, metric relations with site quality should continue to be treated as testable hypotheses rather than as assumptions (Fausch et al. 1990; Fore et al. 1996).

Interregional variation in metric utility, which is common, probably has three main sources: region-specific natural ranges of variation in metrics, region-specific human alteration of ecosystems, and region-specific responses by ecosystems to alteration. The interregional differences in metric utility that we observed in the mid-Atlantic highlands suggest that metrics useful in one region may be ineffective in nearby regions. Previous research has also shown pronounced interregional differences in the ability of particular metrics to detect human impacts on streams (Angermeier and Karr 1986; Smogor and Angermeier 1999b). Nevertheless, TOTSP, DOSSP, TOLPR, and SLTPR (or variations) seem to be widely useful in upland warmwater streams of North America (Simon and Lyons 1995). Comparability of bioassessments is likely enhanced by using MMIs that share metrics.

Metric utility can be influenced by accounting for (or failing to account for) key intraregional variables, such as stream size, in the analysis of metric versus site-quality relations. The criteria for scoring taxonomic metrics in IBI applications are commonly stratified to reflect relations between species richness and stream size (i.e., different criteria are used for different-sized streams). Stream-size effects on fish metrics vary greatly among basins (Fausch et al. 1984; Smogor and Angermeier 1999a) and should be examined on a region-by-region basis. By not accounting for stream-size effects in our analyses, we could not determine how much metric variation (among regions or among site-quality classes) was attributable to differences in stream size. If stream-size distributions differed among regions or among site-quality classes, some of the metric variation that we attributed to interregion or interclass differences could be due to stream-size effects. Smogor and Angermeier (1999a) documented stream-size effects on three taxonomic metrics (including INTSP and SUNSP) and a functional metric (CARPR) for mountain streams in Virginia. Similar effects may apply to the entire mid-Atlantic highland area; they require thorough investigation before IBI protocols are implemented. Given that stream-size effects are most common for taxonomic metrics, examination of such effects would be especially likely to enhance the utility of taxonomic metrics in bioassessments in the mid-Atlantic highlands.

Only 10–40% of the variance in fish metrics was typically attributable to differences among site-quality classes (Tables 5, 7). Several factors may have obscured real patterns or created spurious ones. Our data were collected independently by several investigators who used a variety of sampling protocols and purposes. These inconsistencies may have introduced variation that obscured important patterns, despite the many sample sites (e.g., in the Atlantic basin). Sites were not chosen to represent the full range of stream quality or size in each region. Furthermore, sites were retrospectively assigned to quality classes, and background information (e.g., water quality, impact history, and watershed land use) was inadequate to precisely rank site quality in the context of regional conditions or to distinguish effects of stream size versus site quality. These are major shortcomings for developing bioassessment tools, which should be based largely on empirical differences observed across known impact gradients. A better approach to selecting useful metrics would be to use least-impacted sites and historical accounts to identify reference conditions in the regions of interest (Hughes 1995) and then to compare reference conditions with conditions at suites of sites that represent a wide range of impact levels. Our findings regarding metric utility should be validated by studies that directly compare biotic conditions across the full range of impact levels in each re-

gion. Such studies could reveal additional useful metrics and would help determine if metrics with small but consistent differences among site-quality classes (e.g., DOSSP, SUNSP, LAMSP, and SLTPR) are reliable and sensitive enough to be included in MMIs.

Development of MMIs

Each of our five preliminary MMIs distinguished between high- and low-quality sites, thereby illustrating their utility in assessing mid-Atlantic highlands streams. Although taxonomic metrics were more strongly related with site quality than were functional metrics in our regions, we doubt that this pattern can be generalized across all regions. Previous research has suggested that MMIs incorporating several types of metrics (e.g., taxonomic, trophic, reproductive, and tolerance metrics) perform better than do narrowly focused MMIs (Angermeier and Karr 1986; Karr 1993). Full development of MMIs for use in bioassessment programs will require rigorous validation of metrics, including some that we did not examine. For example, metrics describing the health of individual fish are sensitive to many types of toxic pollution (Hughes and Gammon 1987; Yoder and Rankin 1995a; Sanders et al. 1999) and warrant evaluation for use in the mid-Atlantic highlands. Also, closer examination of relations between reproductive habits of Appalachian fishes and population responses to human impacts (e.g., Jones et al. 1999) may reveal additional metrics that are useful in bioassessments.

A major advantage of MMIs is that metrics and scoring criteria can be tailored to match each region's bioassessment needs. Our approach to MMI development is unusual, in that we selected metrics largely on the basis of empirical relations. The suites of metrics used in new MMIs typically have been selected largely on the basis of expert judgment or have been borrowed from other applications. Only recently have protocols for MMI development emerged that rely largely on empirical relations (Barbour et al. 1995; Barbour et al. 1996; Hughes et al. 1998). Greater reliance on empirical relations will require large data sets but will allow users of the resulting MMIs to be more confident about the performance of region-specific indices and should enhance the accuracy of assessments, even if the relations are derived from suboptimal data sets such as ours. Nevertheless, we advise against building MMIs with only those metrics that exhibit strong empirical relations, because such indices may lack the ecological foundation needed

to detect a broad range of human impacts. Expert judgment should still play a key role in MMI development, especially in the early stages and in ensuring that MMIs are sensitive to many potential (but perhaps unrealized) impacts.

Our findings should be valuable to state agencies that are currently developing fish-based MMIs to assess the quality of cool- and warmwater streams in the mid-Atlantic highlands. Development of analogous MMIs for coldwater streams will likely require separate data sets and analyses; this has been accomplished in Wisconsin (Lyons et al. 1996). Ultimately, each state determines how much validation and refinement is invested in MMIs as well as the role that MMIs will play in bioassessment programs (Southerland and Stribling 1995; US EPA 1996). However, cooperation among states that share bioassessment regions should make MMI development more cost-effective and should enhance the comparability of the resulting assessments. Cooperation is probably essential for states that include very small portions of regions (e.g., Maryland's portion of the Blue Ridge ecoregion), a situation that may preclude acquiring sufficient sample sizes to allow meaningful statistical analyses. Investing in MMI development and implementation can greatly enhance a state's ability to assess and communicate the environmental quality of aquatic ecosystems for purposes of management and regulation (Yoder and Rankin 1995b, 1998). We encourage more states in the mid-Atlantic highlands and in other areas to integrate these adaptable and effective tools into their bioassessment programs.

Acknowledgments

References

Angermeier, P. L. 1995. Ecological attributes of extinction-prone species: loss of freshwater fishes of Virginia. Conservation Biology 9:143–158.

Angermeier, P. L., and J. R. Karr. 1986. Applying an index of biotic integrity based on stream-fish com-

munities: c
tation. Nor
agement 6:
Angermeier, P.
1998. An a
the integrit
mid-Atlant
mental Pro
Angermeier, P. l
An electric
North Ame
11:352–35
Angermeier, P.
terizing fis
landscapes
ical Applic
Barbour, M. T.,
borg, E. M
1996. A fr
ida stream
nal of the
15:185–21
Barbour, M. T.
Multimetri
and measu
W. S. Da
assessmen
planning a
Boca Rato
Fausch, K. D.,
application
stream fisl
ican Fishe
Fausch, K. D.,
1990. Fis
mental de
editor. Bic
ican Fish
Maryland.
Fore, L. S., J.
tistical pr
used to ev
of Fisheri
Fore, L. S., J.
sessing in
evaluating
North Am
Hawkes, C. L
Fish ecor
patterns a
Environm
Hocutt, C. H.,
geograph
Wiley, N
Hughes, R. M
tus by co
31–47 in
ological a
source pl
lishers, B
Hughes, R. M
changes i

munities: considerations in sampling and interpretation. North American Journal of Fisheries Management 6:418–429.

Angermeier, P. L., R. A. Smogor, and J. R. Stauffer. 1998. An analysis of regions and metrics to assess the integrity of fish communities in streams of the mid-Atlantic highlands. Report to U.S. Environmental Protection Agency, Corvallis, Oregon.

Angermeier, P. L., R. A. Smogor, and S. D. Steele. 1991. An electric seine for collecting fish in streams. North American Journal of Fisheries Management 11:352–357.

Angermeier, P. L., and M. R. Winston. 1999. Characterizing fish community diversity across Virginia landscapes: prerequisite for conservation. Ecological Applications 9:335–349.

Barbour, M. T., J. Gerritsen, G. E. Griffith, R. Frydenborg, E. McCarron, J. S. White, and M. L. Bastian. 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. Journal of the North American Benthological Society 15:185–211.

Barbour, M. T., J. B. Stribling, and J. R. Karr. 1995. Multimetric approach for establishing biocriteria and measuring biological condition. Pages 63–77 in W. S. Davis and T. P. Simon, editors. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Fausch, K. D., J. R. Karr, and P. R. Yant. 1984. Regional application of an index of biotic integrity based on stream fish communities. Transactions of the American Fisheries Society 113:39–55.

Fausch, K. D., J. Lyons, J. R. Karr, and P. L. Angermeier. 1990. Fish communities as indicators of environmental degradation. Pages 123–144 in S. M. Adams, editor. Biological indicators of stress in fish. American Fisheries Society, Symposium 8, Bethesda, Maryland.

Fore, L. S., J. R. Karr, and L. L. Conquest. 1994. Statistical properties of an index of biological integrity used to evaluate water resources. Canadian Journal of Fisheries and Aquatic Sciences 51:1077–1087.

Fore, L. S., J. R. Karr, and R. W. Wisseman. 1996. Assessing invertebrate responses to human activities: evaluating alternative approaches. Journal of the North American Benthological Society 15:212–231.

Hawkes, C. L., D. L. Miller, and W. G. Layher. 1986. Fish ecoregions of Kansas: stream fish assemblage patterns and associated environmental correlates. Environmental Biology of Fishes 17:267–279.

Hocutt, C. H., and E. O. Wiley, editors. 1986. The zoogeography of North American freshwater fishes. Wiley, New York.

Hughes, R. M. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31–47 in W. S. Davis and T. P. Simon, editors. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Hughes, R. M., and J. R. Gammon. 1987. Longitudinal changes in fish assemblages and water quality in the Willamette River, Oregon. Transactions of the American Fisheries Society 116:196–209.

Hughes, R. M., S. A. Heiskary, W. J. Matthews, and C. O. Yoder. 1994. Use of ecoregions in biological monitoring. Pages 125–151 in S. L. Loeb and A. Spacie, editors. Biological monitoring of aquatic systems. Lewis Publishers, Boca Raton, Florida.

Hughes, R. M., P. R. Kaufmann, A. T. Herlihy, T. M. Kincaid, L. Reynolds, and D. P. Larsen. 1998. A process for developing and evaluating indices of fish assemblage integrity. Canadian Journal of Fisheries and Aquatic Sciences 55:1618–1631.

Hughes, R. M., and T. Oberdorff. 1999. Applications of IBI concepts and metrics to waters outside the United States and Canada. Pages 79–93 in T. P. Simon, editor. Assessing the sustainability and biological integrity of water resources using fish communities. CRC Press, Boca Raton, Florida.

Hughes, R. M., E. Rexstad, and C. E. Bond. 1987. The relationship of aquatic ecoregions, river basins, and physiographic provinces to the ichthyogeographic regions of Oregon. Copeia 1987:423–432.

Hughes, R. M., T. R. Whittier, C. M. Rohm, and D. P. Larsen. 1990. A regional framework for establishing recovery criteria. Environmental Management 14:673–683.

Jenkins, R. E., and N. M. Burkhead. 1994. The freshwater fishes of Virginia. American Fisheries Society, Bethesda, Maryland.

Jones, E. B. D., III, G. S. Helfman, J. O. Harper, and P. V. Bolstad. 1999. Effects of riparian forest removal on fish assemblages in southern Appalachian streams. Conservation Biology 13:1454–1465.

Jones, K. B., K. H. Riitters, J. D. Wickham, R. D. Tankersley, Jr., R. V. O'Neill, D. J. Chaloud, E. R. Smith, and A. C. Neale. 1997. An ecological assessment of the United States Mid-Atlantic region: a landscape atlas. U.S. Environmental Protection Agency, EPA/600/R-97/130, Washington, D.C.

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

Karr, J. R. 1993. Defining and assessing ecological integrity: beyond water quality. Environmental Toxicology and Chemistry 12:1521–1531.

Karr, J. R., K. D. Fausch, P. L. Angermeier, P. R. Yant, and I. J. Schlosser. 1986. Assessing biological integrity in running waters: a method and its rationale. Illinois Natural History Survey, Special Publication 5, Champaign.

Kerans, B. L., and J. R. Karr. 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. Ecological Applications 4:768–785.

Larsen, D. P., J. M. Omernik, R. M. Hughes, C. M. Rohm, T. R. Whittier, A. J. Kinney, A. L. Gallant, and D. R. Dudley. 1986. Correspondence between spatial patterns in fish assemblages in Ohio streams and aquatic ecoregions. Environmental Management 10:815–828.

Leonard, P. M., and D. J. Orth. 1986. Application and testing of an index of biotic integrity in small, cool-

impacts. Expert
role in MMI de-
rly stages and in
to many potential

ole to state agen-
pping fish-based
cool- and warm-
c highlands. De-
s for coldwater
rate data sets and
ished in Wiscon-
ately, each state
n and refinement
e role that MMIs
ams (Southerland
6). However, co-
re bioassessment
velopment more
ce the compara-
nts. Cooperation
hat include very
Maryland's por-
), a situation that
t sample sizes to
alyses. Investing
lementation can
y to assess and
quality of aquat-
management and
995b, 1998). We
d-Atlantic high-
rate these adapt-
ir bioassessment

ts

der cooperative
of the Environ-
ent Program of
on Agency, Na-
l Research Lab-
We thank E. P.
R. M. Hughes,
, and an anony-
ments on earlier

ibutes of extinction-
fishes of Virginia.

986. Applying an
n stream-fish com-

water streams. Transactions of the American Fisheries Society 115:401–414.

Lyons, J. 1989. Correspondence between the distribution of fish assemblages in Wisconsin streams and Omernik's ecoregions. American Midland Naturalist 122:163–182.

Lyons, J. 1992. Using the index of biotic integrity (IBI) to measure environmental quality in warmwater streams of Wisconsin. U.S. Forest Service General Technical Report NC-149.

Lyons, J., S. Navarro-Perez, P. A. Cochran, E. Santana, and M. Guzman-Arroyo. 1995. Index of biotic integrity based on fish assemblages for the conservation of streams and rivers in west-central Mexico. Conservation Biology 9:569–584.

Lyons, J., L. Wang, and T. D. Simonson. 1996. Development and validation of an index of biotic integrity for coldwater streams in Wisconsin. North American Journal of Fisheries Management 16:241–256.

Mayden, R. L., editor. 1992. Systematics, historical ecology, and North American freshwater fishes. Stanford University Press, Stanford, California.

McAllister, D. E., S. P. Platania, F. W. Schueler, M. E. Baldwin, and D. S. Lee. 1986. Ichthyofaunal patterns on a geographic grid. Pages 17–51 in C. H. Hocutt and E. O. Wiley, editors. The zoogeography of North American freshwater fishes. Wiley, New York.

Miller, D. L., and thirteen coauthors. 1988. Regional applications of an index of biotic integrity for use in water resource management. Fisheries 13(5):12–20.

Moyle, P. B., and J. J. Cech, Jr. 1988. Fishes: an introduction to ichthyology. Prentice-Hall, Englewood Cliffs, New Jersey.

Omernik, J. M. 1995. Ecoregions: a spatial framework for environmental management. Pages 49–62 in W. S. Davis and T. P. Simon, editors. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Omernik, J. M., and R. G. Bailey. 1997. Distinguishing between watersheds and ecoregions. Journal of the American Water Resources Association 33:935–949.

Omernik, J. M., D. D. Brown, C. W. Kiilsgaard, and S. M. Pierson. 1992. Draft ecoregions and subecoregions of the Blue Ridge Mountains, central Appalachian ridges and valleys, and central Appalachians of EPA Region 3. U.S. Environmental Protection Agency, Corvallis, Oregon.

Pan, Y., R. J. Stevenson, B. H. Hill, A. T. Herlihy, and G. B.Collins. 1996. Using diatoms as indicators of ecological conditions in lotic systems: a regional assessment. Journal of the North American Benthological Society 15:481–495.

Rohm, C. M., J. W. Giese, and C. C. Bennett. 1987. Evaluation of an aquatic ecoregion classification of streams in Arkansas. Journal of Freshwater Ecology 4:127–140.

Rosen, B. H. 1995. Use of periphyton in the development of biocriteria. Pages 209–215 in W.S. Davis

and T. P. Simon, editors. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Sanders, R. E., R. J. Miltner, C. O. Yoder, and E. T. Rankin. 1999. The use of external deformities, erosion, lesions, and tumors (DELT anomalies) in fish assemblages for characterizing aquatic resources: a case study of seven Ohio streams. Pages 225–246 in T. P. Simon, editor. Assessing the sustainability and biological integrity of water resources using fish communities. CRC Press, Boca Raton, Florida.

SAS Institute. 1990. SAS/STAT user's guide: volume 2. SAS Institute, Cary, North Carolina.

Simon, T. P., and J. Lyons. 1995. Application of the index of biotic integrity to evaluate water resource integrity in freshwater ecosystems. Pages 245–262 in W. S. Davis and T. P. Simon, editors. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Smogor, R. A. 1996. Developing an index of biotic integrity (IBI) for warmwater wadeable streams in Virginia. Master's thesis. Virginia Polytechnic Institute and State University, Blacksburg.

Smogor, R. A., and P. L. Angermeier. 1999a. Effects of drainage basin and anthropogenic disturbance on relations between stream size and IBI metrics in Virginia. Pages 249–272 in T. P. Simon, editor. Assessing the sustainability and biological integrity of water resources using fish communities. CRC Press, Boca Raton, Florida.

Smogor, R. A., and P. L. Angermeier. 1999b. Relations between fish metrics and measures of anthropogenic disturbance in three IBI regions in Virginia. Pages 585–610 in T. P. Simon, editor. Assessing the sustainability and biological integrity of water resources using fish communities. CRC Press, Boca Raton, Florida.

Smogor, R. A., and P. L. Angermeier. In press. Determining a regional framework for assessing biotic integrity in Virginia streams. Transactions of the American Fisheries Society.

Sokal, R. R., and F. J. Rohlf. 1981. Biometry. Freeman, San Francisco.

Southerland, M. T., and J. B. Stribling. 1995. Status of biological criteria development and implementation. Pages 81–96 in W. S. Davis and T. P. Simon, editors. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Stauffer, J. R., Jr., J. M. Boltz, and L. R. White. 1995. The fishes of West Virginia. Proceedings of the Academy of Natural Sciences of Philadelphia 146:1–389.

Steedman, R. J. 1988. Modification and assessment of an index of biotic integrity to quantify stream quality in southern Ontario. Canadian Journal of Fisheries and Aquatic Sciences 45:492–501.

Tabachnick, B. G., and L. S. Fidell. 1996. Using multivariate statistics. Harper and Row, New York.

USEPA (United States Environmental Protection Agen-

cy). 1990
guidance
5-90-004,

USEPA (Unit
cy). 1996
programs
R-96-007

Whittier, T. R
Correspo
patterns i
Journal o
1278.

Yoder, C. O.,
teria pro

cy). 1990. Biological criteria: national program guidance for surface waters. USEPA, EPA-440/5-90-004, Washington, D.C.

USEPA (United States Environmental Protection Agency). 1996. Summary of state biological assessment programs for streams and rivers. US EPA, EPA 230-R-96-007, Washington, D.C.

Whittier, T. R., R. M. Hughes, and D. P. Larsen. 1988. Correspondence between ecoregions and spatial patterns in stream ecosystems in Oregon. Canadian Journal of Fisheries and Aquatic Sciences 45:1264–1278.

Yoder, C. O., and E. T. Rankin. 1995a. Biological criteria program development and implementation in Ohio. Pages 109–144 in W. S. Davis and T. P. Simon, editors. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Yoder, C. O., and E. T. Rankin. 1995b. Biological response signatures and the area of degradation value: new tools for interpreting multimetric data. Pages 263–286 in W. S. Davis and T. P. Simon, editors. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Yoder, C. O., and E. T. Rankin. 1998. The role of biological indicators in a state water quality management process. Environmental Monitoring and Assessment 51:61–88.