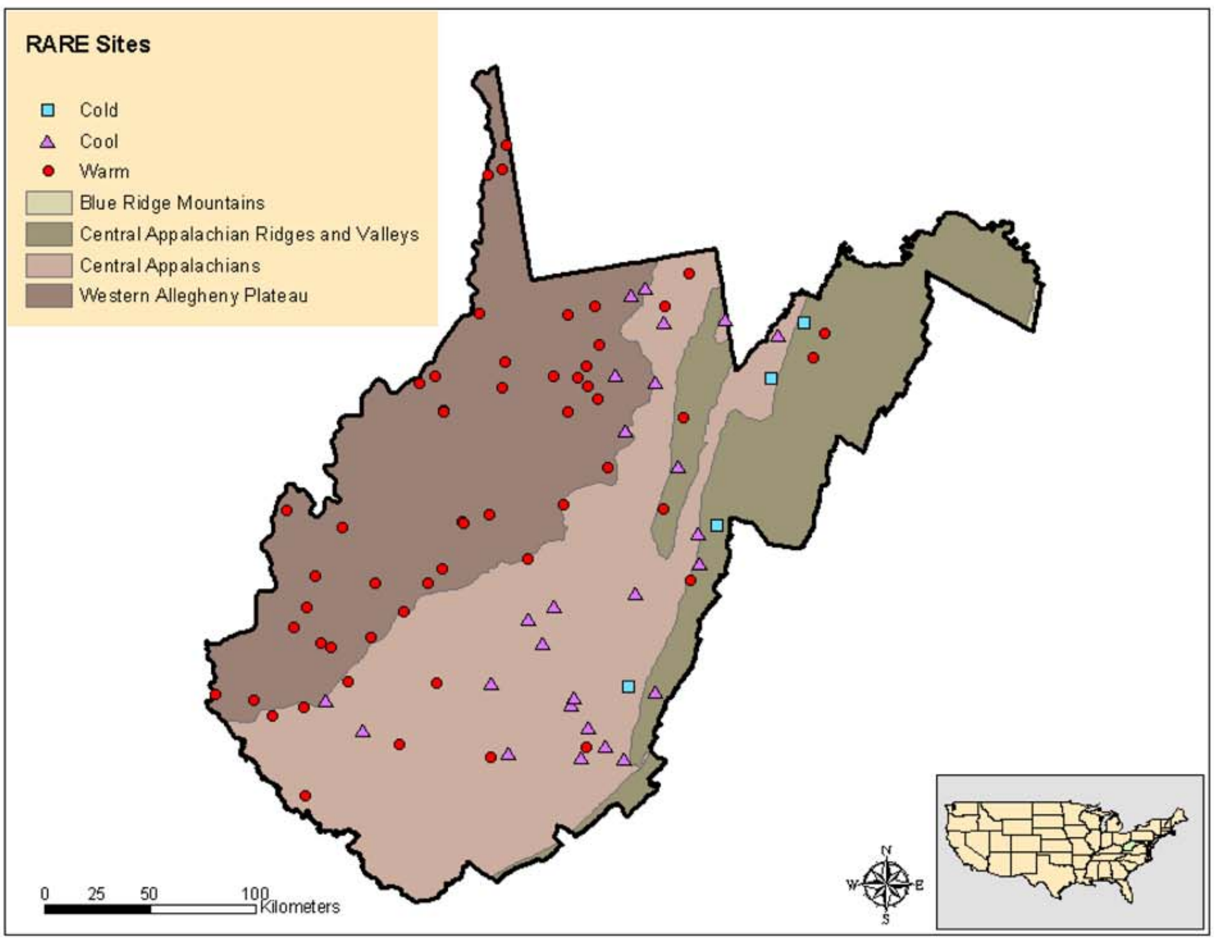
United States Environmental Protection Agency

Development of a Fish Index of Biotic Integrity to Assess the Condition of West Virginia Streams:

Technical Support Document



EPA/600/R-06/010
February 16, 2006

Development of a Fish Index of Biotic Integrity to Assess the Condition of West Virginia Streams:

Technical Support Document

David M. Walters
U.S. EPA Office of Research and Development

United States Environmental Protection Agency

Office of Research and Development
National Exposure Research Laboratory
26 West Martin Luther King Jr. Blvd.
Cincinnati, OH 45268

NOTICE

The research described in this document has been funded by the United States Environmental Protection Agency REMAP and RARE programs. It has been subjected to Agency peer and administrative review and approved for publication as an EPA document.

Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

The correct citation for this document is:

Walters, D.M. 2006. Development of a Fish Index of Biotic Integrity to Assess the Condition of West Virginia Streams: Technical Support Document. EPA/600/R-06/010. U.S. Environmental Protection Agency, Cincinnati, Ohio.

ACKNOWLEDGEMENTS

This report relied on the generous support of numerous individuals. Frank McCormick (U.S. Forest Service) provided data analysis and technical support. Dan Cincotta (West Virginia DNR) led crews to collect fish and other stream data. Alicia Shelton (SoBran Inc.) provided database support and analysis of stream temperature data. Naomi Detenbeck (U.S. EPA) provided land cover data. Lou Reynolds (U.S. EPA) and Karen Blocksom (U.S. EPA) provided reviews and helpful comments to improve the report.

Table of Contents

List of Figures

List of Tables

List of Appendices

Introduction

The landscape, streams, and fish assemblages of the Mid-Atlantic region have endured a long history of human impacts. Streams in the highlands have been subjected to stresses from acid deposition, mining, logging, agriculture, and urban development (Raitz et al. 1984; Whitney 1994; Jones et al. 1997). Agriculture and clear-cutting of highland and valley forests have exacerbated soil erosion and sedimentation (USDA 1996). Active and abandoned coal mining resulted in mine drainage that affected approximately 4,000 km of streams (Herlihy et al. 1990; USEPA 1995). Extensive areas of the Ridge, Blue Ridge, and Appalachian Plateau ecoregions have poorly buffered soils and steep slopes, rendering these highland streams particularly susceptible to acid precipitation (Herlihy et al. 1993).

To accurately assess and manage the biotic integrity of aquatic ecosystems, a comprehensive inventory of the biotic resources should be conducted, the current conditions of streams should be determined, and the impacts of stressors (e.g., acid deposition, stream disturbances, mine drainage, agricultural runoff, erosion, and domestic and industrial pollution) should be evaluated (Kazyak et al. 1994). Although biomonitoring over the years has investigated many different assemblages for use as indicators of water quality (e.g., Karr 1981, Lazorchak et al. 1998), recent USEPA guidance documents recommend that fish and macroinvertebrate community analyses be adopted in state water evaluation programs (US EPA 2002).

The West Virginia Division of Natural Resources (WVDNR) has conducted fishery studies in wadeable streams since the early 1950s, focusing primarily on the status of game fish populations. Survey data collected by WVDNR, including gamefish availability, standing crop, and recruitment, are used to indirectly assess the ecological condition of stream resources, but West Virginia has not initiated the use of fishes in their statewide assessment program.

Fish species exhibit diverse morphological, ecological, and behavioral adaptations to their natural habitat and, thus, are particularly effective indicators of the condition of aquatic systems (Karr et al. 1986; Fausch et al. 1990; Simon and Lyons 1995). Human disturbance of streams and landscapes alters key attributes of aquatic ecosystems: water quality, habitat structure, hydrological regime, energy flow, and biological interactions (Karr and Dudley 1981). The index of biological integrity (IBI) was developed to assess the condition of water bodies by direct evaluation of biological attributes (Karr et al. 1986). The IBI is a composite index that integrates structural, ecological, trophic, and reproductive attributes of fish assemblages at multiple levels of organization (Fausch et al. 1990). Originally developed for assessment of Midwestern U.S. warmwater streams, it has been modified for use in other regions and waters (Simon and Lyons 1995; Lyons et al. 1996; Hughes et al. 1998; McCormick et al. 2001). Several authors have argued that the IBI must be modified when it is applied in different ecoregions (Fausch et al. 1984; Miller et al. 1988). In the Mid-Atlantic region, researchers have developed IBIs for specific ecoregions (Scott and Hall 1997; Roth et al. 1998; Smogor and Angermeier 1999) or applied it to specific systems (Leonard and Orth 1986).

Over the last decade, a number of stream surveys and indicator development studies have been conducted in the Mid-Atlantic Region. McCormick et al. (2001) developed a regional index of biotic integrity for the assessment of Mid-Atlantic Highland wadeable streams. Data for the Mid-Atlantic Highlands Assessment (MAHA) were collected from 1993-1996 as part of an

Environmental Monitoring and Assessment Program (EMAP) study. In 1997-1998, the EMAP study was expanded to include sites on the Coastal Plain of the eastern US and incorporate non-wadeable streams as part of the Mid-Atlantic Integrated Assessment (MAIA). In 2000, USEPA Region 3 initiated a Regional-EMAP (REMAP) and a Regional Applied Research Effort (RARE) project in West Virginia (WV) in cooperation with the Office of Research and Development and West Virginia Department of Natural Resources (Cincotta et al. 2001).

The primary purpose of this report is to document the development of a fish IBI for wadeable streams in WV and to determine the applicability of the IBI in streams with different thermal regimes. The IBI will be developed using fish data collected at EMAP sites from 1993-1998. The RARE sites are used as an independent data set to test the robustness of the IBI across stream temperature regimes (i.e., cold, cool, and warm water streams).

Methods

Study Area and Survey Design

Omernik (1987) and Woods et al. (1996) identified three primary ecoregions (Ridge and Valley, Western Allegheny Plateau, and Central Appalachian Plateau) in West Virginia (a small segment of the Blue Ridge lays in the eastern panhandle). The Ridge and Valley province consists of roughly parallel northeast-southwest trending ridges and valleys that have a variety of widths, heights, and geologic materials. The Western Allegheny Plateau is characterized by rounded hills separated by narrow valleys. The Central Appalachian ecoregion is primarily a high, dissected, rugged plateau; its terrain, cool climate, and infertile soils limit agriculture, resulting in a mostly forested land cover. Extensive mixed mesophytic forests and mixed oak forests typically remain on the upland terrain. Agriculture (dairy, livestock, and general farms) and residential developments are concentrated in the valleys. Bituminous coal mines are common, and have caused the siltation and acidification of streams.

Stream sites for the EMAP projects (hereafter referred to as “EMAP sites”) were selected using a randomized systematic design with a spatial component (Overton et al. 1991; Herlihy et al. 2000). The sample population of streams in the region was delineated from digitized USGS topographic maps (1:100,000 scale). Sample probabilities were set so that roughly equal numbers of first-, second-, and third-order streams would be selected. Details on the sampling framework are provided by Davis and Scott (2000) and McCormick et al. (2001). EMAP surveys included the Mid-Atlantic Highlands Assessment (MAHA) sampled 1993-1996 and the Mid-Atlantic Integrated Assessment (MAIA) sampled 1997-1998. Several streams in WV were also sampled under a Regional EMAP (REMAP) study affiliated with the MAHA study, although habitat sampling was less extensive for the REMAP sites. The total number of sites sampled in West Virginia under these programs was 96 with 12 revisits (n = 108 samples).

RARE sites were selected using a stratified, random design. Sites were stratified by hydrologic regime and land use (Detenbeck et al. 2004). Hydrologic regime was based on watershed storage and main channel length, whereas land use was characterized as either high- or low-intensity. High- and low- intensity designations were based on literature values for thresholds of land-use activity at which clear degradation in biological or chemical condition was observed. The land-use based classes were developed for the predominant land uses (e.g., agricultural, urban/residential, and mining activities) in each ecoregion.

Watersheds in the Potomac River basin were not included in the survey. A total of 119 sites were selected and sampled in 2001-2002. Site selection processes are detailed in Detenbeck et al. (in review). Stream sites surveyed in West Virginia for the EMAP/REMAP and RARE projects are shown in Figure 1.

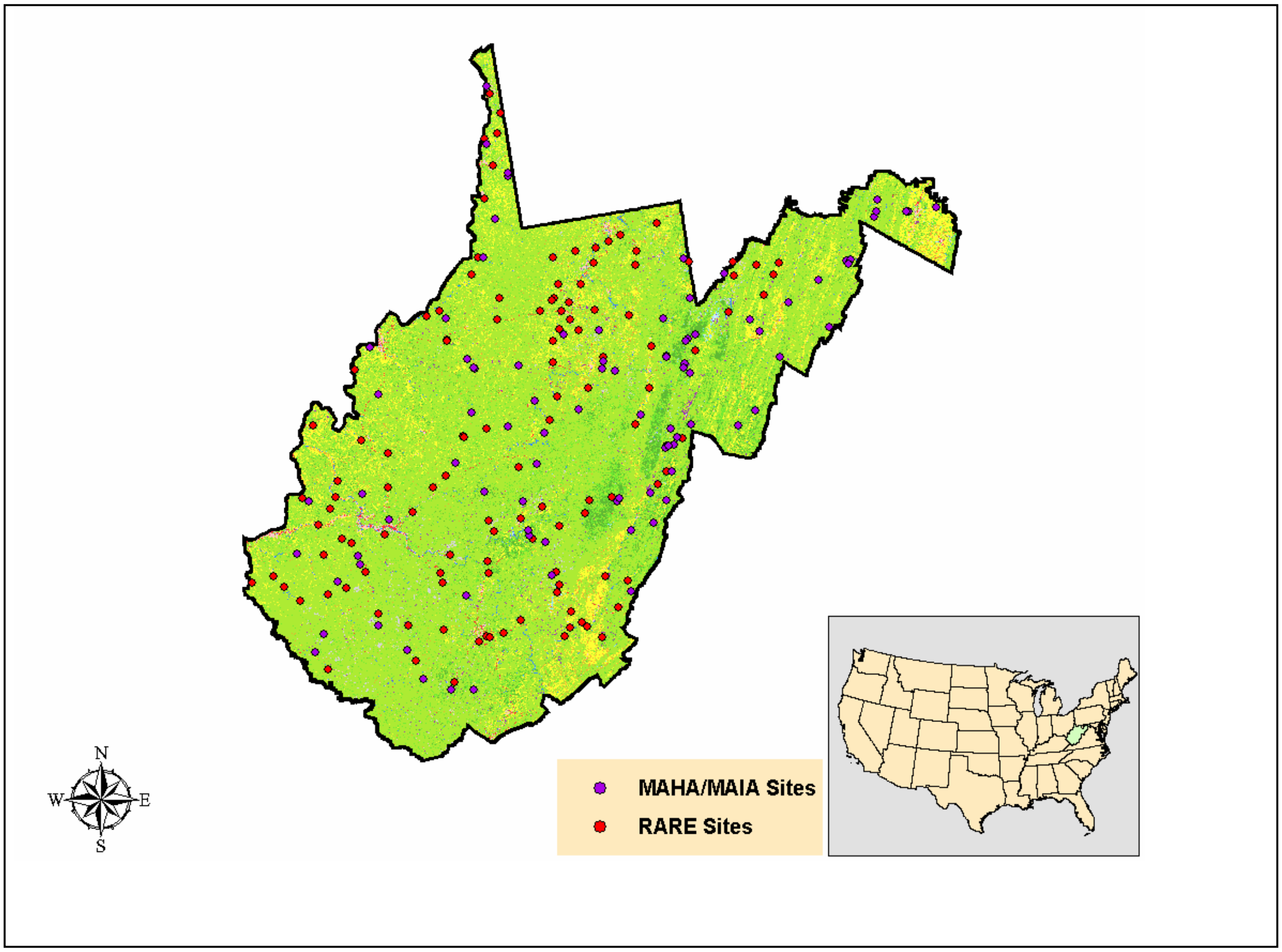


Figure 1. Site collection locales for West Virginia 1993-2002. RARE (Regional Applied Research Effort) sites were sampled in 2001 and 2002. MAHA (Mid-Atlantic Highlands Assessment)/MAIA (Mid-Atlantic Integrated Assessment) sites were sampled from 1993-98 as part of the EMAP program.

Integrating datasets: quantifying seasonal, regional and methodological biases — These analyses draw on three large, complex datasets that differed in some degree in terms of sampling methodology and scale. In an effort to minimize bias, the data sets were compared to identify potential sources of seasonal, sample gear, and regional bias.

Most MAHA data (1993 - 1996) were collected during the spring, whereas the majority of MAIA data (1997 - 1998) were collected during the summer. During 1995 and 1996 MAHA crews made 29 revisits to 16 sites (region-wide) during the summer, in anticipation of the MAIA summer sampling. Likewise, MAIA crews sampled 23 sites (region-wide) in both spring and summer 1997 and 1998. However, six samples collected in spring 1997 show much lower richness than samples collected at the same sites again during summer 1997 and 1998 (i.e., richness > 6 species and richness only 17-44% of subsequent samples). These samples suggest data quality problems may exist and were treated as outliers and excluded from the seasonal analysis. Fish species richness between seasons was compared using bivariate plots to identify any seasonal bias.

Fish sampling in the MAHA survey was originally restricted to wadeable streams with basin areas < 500 km^2 in the upland ecoregions (Central Appalachian and Western Allegheny plateaus, Blue Ridge and Ridge & Valleys) of the Mid-Atlantic Highlands. The MAIA survey was extended to the Piedmont and Coastal Plain, neither of which is in West Virginia. McCormick et al. (2001) found no significant differences in unscored metric values across ecoregions or basins for the nine metrics in their IBI. They thus combined all site data for their analyses. RARE sites were also wadeable streams draining < 500 km^2. However, sampling was confined to West Virginia, so the spatial extent of sampling was much smaller than the EMAP surveys. To test the assumption that the IBI of McCormick et al. (2001) was as representative of the condition of streams in West Virginia as it was of the Mid-Atlantic region, cumulative distribution functions (CDFs) of IBI scores of all MAHA and MAIA sites were compared to CDFs of the scores from only the West Virginia sites for those projects.

Fish collection methods varied between the EMAP and RARE surveys, so data from a subset of sites sampled were compared using both techniques to determine if sampling method biased metric values. As part of the 1997 MAIA project, WVDNR sampled 10 sites using both the parallel wire and backpack electrofishing methods. Sampling events were at least two weeks apart. Raw metric scores were calculated for each of the nine metrics and the MAHA IBI developed by McCormick et al. (2001). Scores derived from the parallel wire and backpack methods were directly compared using regression analysis to identify potential method bias.

<u>*Environmental Variables*</u>

Environmental variables were collected for EMAP and RARE sites at the basin and reach scales. These data were collected to serve as stressor gradients for evaluating fish metrics and IBI responsiveness. Environmental variables fall into four main categories: water chemistry, physical habitat and riparian condition, basin and reach landscape characteristics, and temperature.

Water Chemistry — Water chemistry variables were derived from two primary sources, in-situ measurements and water samples collected for laboratory analysis. In-situ measures included dissolved oxygen (DO), pH, conductivity, and temperature. These data were collected using standard field equipment and methods outlined in Plafkin et al. (1989). Water samples for laboratory analysis were collected using standard EMAP protocols (Lazorchak et al. 1998). Laboratory analysis of water samples followed standard US EPA methods (Davis and Scott 2000: Cincotta et al. 2001). Analytes included nutrients, major anions and cations, alkalinity, suspended solids, and heavy metals. Chemistry variables were collected at a single visit to each site.

Physical Habitat and Riparian Condition — Stream habitat and riparian condition were measured using methods provided by Kaufmann and Robison (1998). Major elements measured included channel dimensions, channel gradient, channel substrate, habitat complexity and cover, riparian vegetation, anthropogenic alterations, and channel riparian alterations. Measurements were taken along the stream thalweg and 11 cross-sectional transects evenly spaced within the reach. This type of habitat survey is laborious (requiring about 3 h per site) and was conducted at a subset of sites (n = 67 EMAP sites and n = 104 RARE sites). A qualitative habitat survey (Barbour et al. 1999) was also conducted at all EMAP sites and 104 RARE sites.

Basin and Reach Landscape Characteristics — Basin and reach landscape variables were compiled from a variety of spatially extensive digital coverages. Basins were described in terms of morphometry (e.g., size, shape, and topography) and land cover. EMAP site basins were characterized using digital topographic coverages, Landsat Thematic Mapper (TM) data from 1991-1993, and aerial photographs (Herlihy et al. 1998). Reach-scale attributes such as elevation and gradient were also tabulated for survey sites. Similar watershed attributes were measured for RARE sites. Variables, GIS databases, and methods are detailed in Detenbeck et al. (in review). Land use for RARE basins were determined from the National Landcover Database (NLCD), updated for surface mining.

Temperature — Temperature was collected once at EMAP and RARE sites at the time fishes were sampled. Additionally, temperature loggers (StowAway® TidbiT®; Onset Corporation) were deployed at RARE sites between late May and mid July and were retrieved in the fall (September-October). Temperature was recorded hourly. Retrieved temperature data were plotted to identify and remove any data that were recorded either prior to deployment or after removal from the stream. Weekly maximum and minimum temperature was calculated for all sites. Weekly maxima were averaged for all sites to demonstrate summer trends in temperature across the state. These trends were used to identify the timing of peak summer temperatures. Presumably, this is the part of the summer when fishes would be most stressed by high stream temperature. Henceforth, the period is referred to as "summer". Summer temperature data were used to calculate mean weekly temperature (the average of weekly maximum and minimum temperatures; Wehrly et al. 2003). Based on these data streams were classified into three temperature categories (cold <19˚ C; cool 19-22˚ C; and warm >22˚ C) (Wehrly et al. 2003).

<u>*Fish Assemblages*</u>

Fishes were sampled at EMAP sites using a combination of backpack electrofishing and seining (McCormick and Hughes 1998). Fishes were sampled in a single pass and reach length was scaled to 40 times mean stream width. Minimum and maximum lengths of 150 and 500 m, respectively, were used. RARE sites were sampled using electric parallel wire technique (Holton and Sullivan 1954; Cincotta et al. 2001). Fish were sampled in a single pass over a reach of 160 m.

Metric Selection and IBI Development

Database Management — The fish assemblage data collected in 1993-1996 as part of the EMAP Mid-Atlantic Highlands Assessment (MAHA) were used to test fish assemblage metrics for responsiveness to anthropogenic stressors (n = 45 sites plus 4 revisits). From those data, an IBI was developed for the state of West Virginia. Data collected in 1997 and 1998 by the Mid-Atlantic Integrated Assessment were then used to validate the metrics and IBI (n = 43 sites plus 8 revisits). Finally, data from WVDNR's RARE sites were used to calculate IBI scores from the 2000-2001 project years. Data from the MAHA and MAIA surveys had been entered into the EMAP database management system and subjected to data entry quality assurance, including verification of species identification based on museum vouchers. Data from the REMAP/RARE project were entered by WVDNR and compiled using EMAP Surface Waters Information Management (SWIM) protocols.

Ecological Attributes for Metric Development — Fish ecological characteristics (e.g., spawning guild and tolerance level) were originally compiled for the MAHA study (McCormick et al. 2001). Species characterizations published in McCormick et al. (2001) were based largely on descriptions in Jenkins and Burkhead (1994), with occasional reference to Trautman (1981) and Pflieger (1975). Several fish species that were collected during the MAIA and RARE surveys were not included in the original MAHA database or had not been completely or correctly characterized. New characteristics were added to the file and re-evaluated the criteria, consistency and inclusion of several original MAHA metrics (i.e., macro-omnivores, invertivore-piscivores, tolerant, intolerant, benthic invertivore, benthic habitat, clean substrate spawner). Taxonomic data were updated, erroneous species identifications were corrected, and native/alien designations were revised in both MAIA and MAHA datasets. This file (*FISHCHAR*) is available electronically with the supplementary material provided with this report.

Minimal Disturbance Criteria and Disturbance Scores — Expectations for biological metrics and indices typically are based on conditions at minimally-disturbed locations (reference sites; Hughes 1995). Eleven measures of disturbance (i.e., stressors) were used to score both reference sites and highly-disturbed sites (Table 1). Landscape (catchment) variables included: % catchment as agriculture; % catchment as urban; and % catchment as agriculture, urban & mined. Water chemistry variables included: chloride, ammonia, sulfate, acid neutralizing capacity, nitrate, total nitrogen and total phosphorus. Disturbance designations were made for each disturbance indicator with values above the appropriate criterion level. Sites were evaluated for each stressor and scored a 1 for exceeding minimally disturbed criteria or a 3 for exceeding the highly disturbed criteria, except acid neutralizing capacity, for which higher values are more desirable. If sites characteristics were below the minimum disturbance criterion, they received a disturbance score of 0 for that variable. Cumulative scores based on the designation of disturbance classifications were determined by the sum of scores for each disturbance variable. Sites with 0-1 points were classified as "reference", sites with 2-6 points were classified as intermediate and sites with >6 points were classified as highly disturbed. .

Candidate Metric Evaluation and Selection — In general, the positive metrics (those expected to increase with better conditions) excluded introduced species, whereas negative metrics included all species. Sixty-eight candidate metrics were developed in four categories: taxonomic, trophic, reproductive and tolerance (variable names and descriptions are presented in Appendix I). The 13 metrics that failed the range test and the two metrics that failed the signal to noise test in McCormick et al. (2001) were not included. The candidate metrics list included "non-tolerant" versions (with the tolerant species removed) for most of the positive metrics.

Table 1. Disturbance classification thresholds for EMAP sites (1993-1998).

Variable	Highly Disturbed	Minimally Disturbed
Agriculture (% basin)	45	15
Agriculture + urban + mining (% basin)	50	15
Urban (% basin)	3.5	0.2
Chloride (µeq/l)	900	≤ 100
Ammonia (µeq/l)	8	2
Nitrate (µeq/l)	100	3
Total Nitrogen (µg/l)	>750	≤ 750
Total Phosphorus (µg/l)	150	≤ 20
Sulfate (µeq/l)	1,000	≤ 400
Acid Neutralizing Capacity (µeq/l)	<50	≥ 50
Rapid Habitat (mean score)	<12	≥ 16

Candidate metrics were screened with four successive "filters" following the approaches described in Hughes et al. (1998) and McCormick et al. (2001). In the evaluation process, each metric was examined for its scoring range, variability, responsiveness, and redundancy. Metrics were rejected if they failed a range test (applied only to richness metrics; rejected if raw value ranges between 0 and 2 species) or a signal to noise test (ratio<3, where signal was the variance among sites and noise was the variance among repeat visits (Kaufmann et al. 1999). The range test was only applied to richness metrics.

To determine if the candidate metric was responsive to human disturbance, Spearman correlations and bivariate plots (Hughes et al. 1998) were used to test the responsiveness of the remaining candidate metrics to physical habitat structure and water quality (pH; sulfate concentration; total nitrogen concentration; total phosphorus concentration; chloride concentration; percent sands and fine substrate; relative bed substrate stability; density of large woody debris; fish cover; indices of riparian and channel disturbance; and indices of channel, riparian, and watershed quality). Metrics were plotted against two aggregate measures of human disturbance. First, metric sensitivity was analyzed in streams classified into different disturbance classes (i.e., reference, mixed, nutrients, and mine) based on stream chemistry (Herlihy 1990, Davis and Scott 2000). Second, metric response was analyzed in streams classified using Bryce Condition Class, an index that uses watershed and local stressors (e.g., in-stream sediment and habitat, basin forest cover, etc) to evaluate human disturbance (Bryce et al. 1999).

Scatterplots of the metric values versus stream size (log_{10} watershed area), were visually assessed with sites coded by membership in the "reference" sites, highly disturbed sites, or intermediate disturbance sites (all other sites). These comparisons were used to determine if metrics were biased for stream size across the disturbance categories. Metrics were retained for which a majority of reference sites had better values than did the majority of disturbed sites. Pearson's Product Moment correlation was used to test for redundancy among metrics. Only one metric out of each correlated pair ($r \geq 0.75$) was retained. All statistical analyses were conducted in PC-SAS for Windows, release 8.02 (SAS 2001).

Adjustments for watershed area — Some metrics were correlated with watershed size. These were normalized for a watershed size of 100 km^2 following the approach described by Urquhart (1982). The regression equation of the metrics with watershed area (log_{10} watershed area in km^2) for the reference sites was calculated. That reference regression equation was then applied to all sites, and their residuals were calculated. Next, the expected value for reference data at a standardized watershed area of 100 km^2 was determined, and this constant was applied to residuals. This resulted in all observations having non-negative values.

Metric and IBI scoring — Metric scoring followed McCormick et al. (2001). Metrics were scored on a continuous scale from 0-10 based on the distribution of scores from sites in the calibration data set. IBI scores were calculated by taking the sum of the nine metrics scores and multiplying the sum by 1.11 to give index scores that ranged from 0-100.

Responsiveness of the IBI — The responsiveness of the IBI to stressors was evaluated by plotting it against chemical and physical habitat variables as well as four aggregate measures that represented general disturbance gradients among watersheds. Sites were classified as minimally-disturbed, intermediate, and highly disturbed using a combination of Rapid Bioassessment Habitat (RBP) variables (Barbour et al. 1999) and chemistry variables (see McCormick et al. (2001) for details on classification). Streams were also classified into different disturbance classes (i.e., reference, mixed, nutrients, and mine) based on stream chemistry (Herlihy 1990, Davis and Scott 2000). A third aggregate measure was the continuous disturbance scores derived from 11 stressors listed in Table 1 (see Methods on Minimal Disturbance Criteria). Finally, Bryce Condition Class, an index that uses watershed and local stressors (e.g., in-stream sediment and habitat, basin forest cover, etc) was calculated to evaluate human disturbance (Bryce et al. 1999).

Relationship between IBI and stream temperature — Relationships between IBI and stream temperature regime were assessed using three methods. First, box plots were used to illustrate the distribution of cold, cool, and warm water sites among reference, intermediate, and disturbed sites. Temperature was only recorded once at these EMAP sites, so thermal categories were assigned based on the presence or absence of cool and coldwater taxa such as sculpins and salmonids. Second, regression analysis was used to compare RARE site IBI scores with temperature recorded in the field at the time fishes were sampled. Finally, RARE sites were categorized as cold, cool, and warm based on continuous summer temperature data. Box plots were used to identify trends in metrics and IBI scores among these temperature classes.

Results

Comparing seasonal, regional, and methodological differences among datasets — Plots of species richness at these seasonal calibration sites for spring versus summer showed little seasonal bias (Figure 2). Likewise, a CDF plot for the population of upland streams in the spring (MAHA) and in the summer (MAIA) were virtually identical (Figure 3). These analyses suggest that seasonal bias is minimal at the site (Figure 2) or regional scale (Figure 3). These lines of evidence support the use of West Virginia MAHA data for metric screening and West Virginia MAIA data for evaluating the responsiveness of the IBI to disturbance.

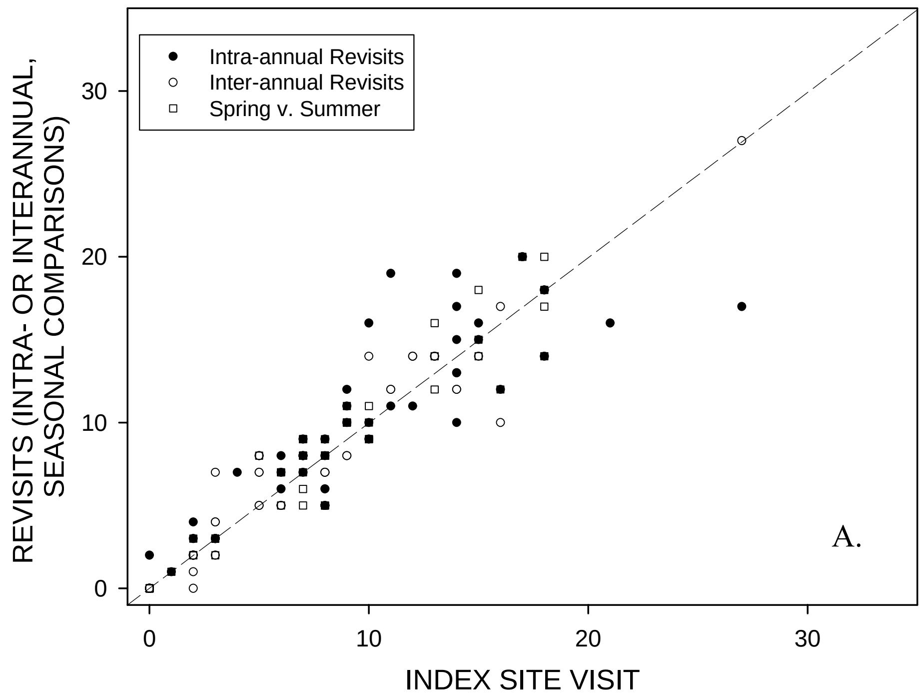
The CDFs of all MAHA and MAIA sites along with the subset of MAHA and MAIA sites from West Virginia sites for those projects were calculated to test the assumption that the IBI of McCormick et al. (2001) was equally representative of stream condition at the state and regional scales. Strong similarity among the plots suggests that the spectrum of stressors and fish assemblage responses found in the region at large were represented in West Virginia (Figure 4). As such, the approach of McCormick et al. (2001) was applied for metric and IBI calculation of the RARE sites.

The comparison of raw metric scores from 10 sites sampled with the parallel wire and backpack electrofishing methods showed strong agreement between methods for seven of the nine metrics (Figure 5). The proportion of invertivore-piscivores (higher for the backpack method) and clean gravel spawners (higher for the parallel wire method) differed between methods but the net effect of these differences on the resulting IBI was negligible.

IBI Development and Testing

Metric selection — After excluding the 13 metrics from McCormick et al. (2001) that failed the range test, no additional metrics were excluded based on this criterion. Two metrics (NTROPH, PNEST) failed the signal to noise test. The results of the redundancy and responsiveness tests are shown in Table 2. Metrics failed the responsiveness test if they did not show clear separation between reference and disturbed sites. The results of this analysis produced a list of metrics identical to that of McCormick et al. (2001). With the revisions to the species list and the modifications of the assemblage characteristics, several metrics were revised, but not significantly changed. All of these metrics were significantly correlated with some measures of water quality and habitat (Table 3 and electronic supplementary table HABITAT CORRELATION MATRIX.xls). Many of the positive richness metrics (those expected to increase with better conditions) were positively correlated with chemical and habitat stressor variables. This was likely due to the positive relationship between watershed area and many stressor variables (electronic supplementary table, HABITAT CORRELATION MATRIX.xls). In many cases, fish richness metrics increase as a function of stream size.

Non-tolerant versions of the clean substrate spawner and piscivore/invertivore metrics were selected whereas McCormick et al. (2001) kept all species. The definition of non-tolerant species for the cyprinid richness and benthic habitat species richness metrics (McCormick et al. 2001) was modified by excluding all tolerant species rather than one or two species originally excluded.



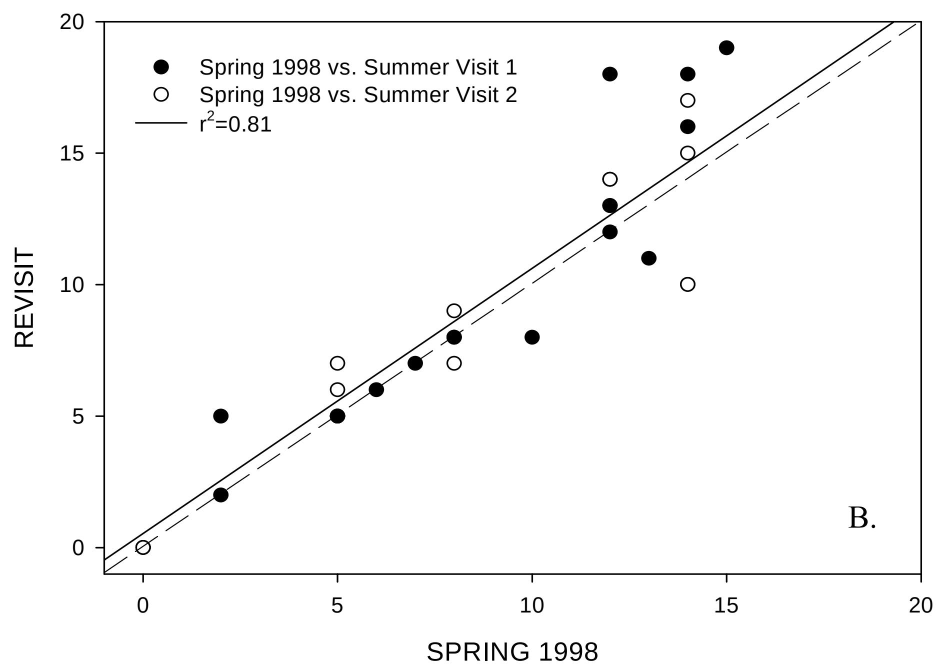


Figure 2. A. Intra- and inter-annual, and seasonal comparisons of native species richness. Most of the 68 intra-annual revisits occurred within the same season (Spring, 1993-1994; Summer, 1995-1998). In 1995 and 1996, 16 sites sampled in the spring were revisited in the summer of 1995 and 1996. Five 1993-1994 spring sites were re-sampled during summer 1997-1998 as part of the Mid-Atlantic Integrated Assessment (MAIA). Diagonal line represents 1:1 relationship. B. Intra-annual comparisons of 23 sites sampled in 1998 for MAIA. The dashed line represents 1:1 relationship and the solid line shows the linear regression for the dataset.

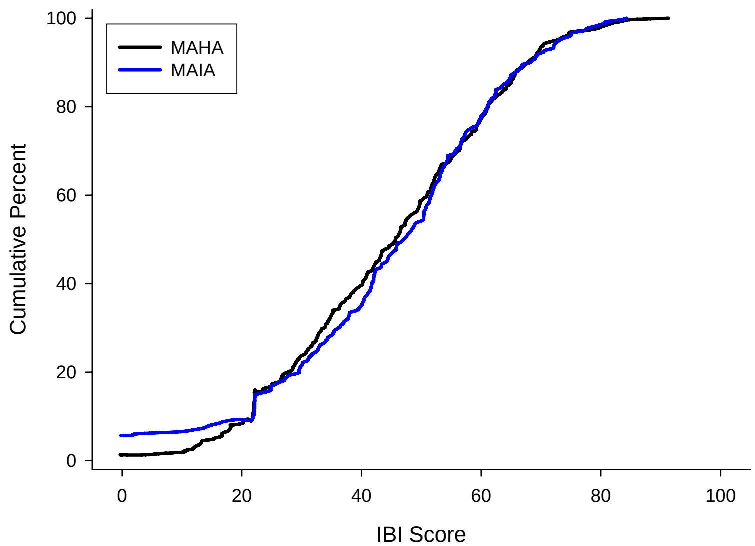


Figure 3. Cumulative distribution functions (CDFs) of index of biotic integrity (IBI) scores for the Mid-Atlantic Highlands Assessment (MAHA) and the Mid-Atlantic Integrated Assessment (MAIA). Indices were calculated using the IBI of McCormick et al. (2001).

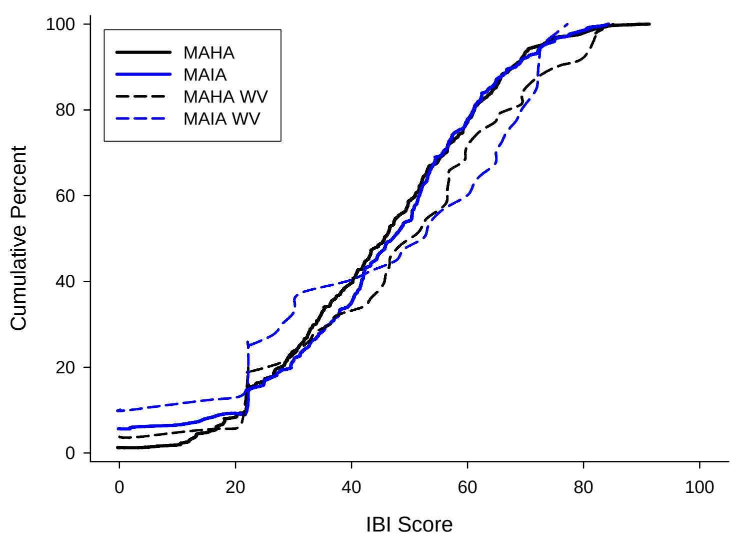


Figure 4. Cumulative distribution functions (CDFs) of index of biotic integrity (IBI) scores for the Mid-Atlantic Highlands Assessment (MAHA) and the Mid-Atlantic Integrated Assessment (MAIA). Indices were calculated using the IBI of McCormick et al. (2001). Region-wide results (solid lines) and state-specific results (dashed lines) are similar.

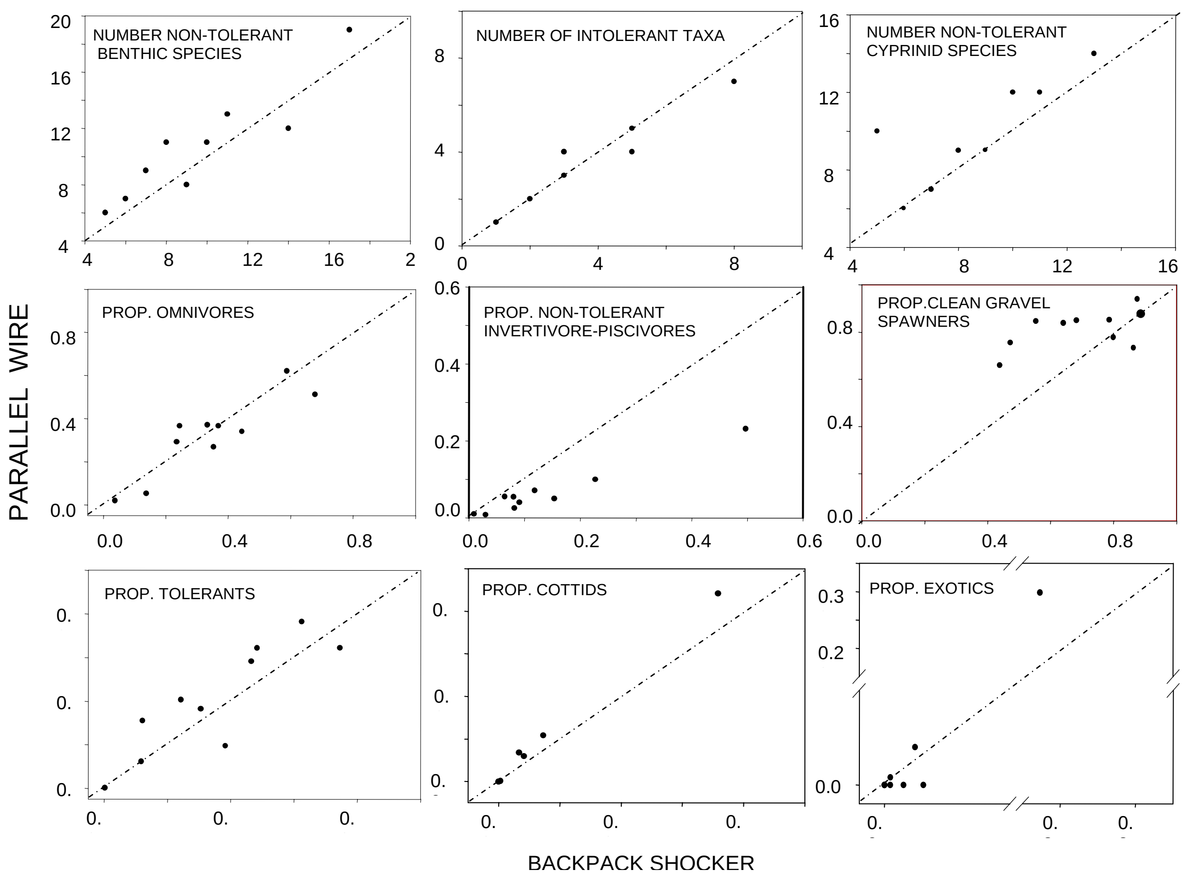


Figure 5. Comparison of metric results from EMAP backpack and WVDNR parallel wire electrofishing methods. The dashed line represents the 1:1 relationship. The proportion of non-tolerant invertivore-piscivores and the proportion of non-tolerant clean gravel spawning individuals show poor agreement compared with other metrics.

Table 2. Results of metric evaluation process. Metric definitions are provided in Appendix I. Original metrics refer to those selected by McCormick et al. (2001) for the MAHA IBI.

Failed responsiveness test	Failed redundancy test	Original IBI metrics	Retained metrics for WV
NUMNATSP	PCOLD	NSCYPR2	NSCYPR_NONTOL
NFAM	PBCLN	NSBENT2	NSBENT_HAB_NONTOL
NSDART	PTREPRO	NSINTOL	NSINTOL
NSICTA	NSP	PGRAVEL	PCLNSPWNR_NONTOL
NSCENT_NONTOL	NUMSPEC	PCOTTID	PCOTTID
NSCATO2	NSCENT	PTOLE	PTOLE
NSCATO_NONTOL	NSCATO	PMACRO	PMACRO
NSBENT_HAB	PCYPR	PPISCINV2	PPISCINV_NONTOL
NSCOLU	PCYPTL	PEXOT	PEXOT
NSBENT_INV	PINTOL		
NSBENT_INV_NONTOL	PPISC		
NSEXOT	PBENT_INV_NONTOL		
NSTOLE	PFISHBUG		
NSCLNSPWNR	PPISCINV		
NSCLNSPWNR_NONTOL			
NREPROS			
PCATO			
PCATO_NONTOL			
PCENT			
PCENT_NONTOL			
PBCST			
PATNG			
PNTGU			
PCGBU			
PCLNSPWNR			
PCOLD2			
PMICRO			
PMICRO2			
PBENT_INV			
PCARN			
PHERB			
PINVERT			
POMNI_H			

Table 3. Spearman rank correlation coefficients (r > 0.2; P <0.01) between EMAP fish assemblage metric variables and chemical and physical habitat variables. The data are from the MAHA survey (n = 45 sites for chemical and land cover data; n = 26 sites for physical habitat data. Full table of correlations of fish assemblage metrics with physical habitat variables is incorporated as a hyperlink to an Excel file to conserve space. Stressor acronym definitions are provided in Appendix III.

	NO3	NTL	PTL	TSS	TURB	AG_TOT	MINE_TOT	URB_TOT	DIS-TOT	PCT_GF	PCT_SA	PCT_SAFN	XEMBED	W1_HALL
NUMNATSP						0.48		0.32	0.47				0.36	
NATNSP				-0.39				0.26			0.29			0.42
NUMSPEC				-0.41				0.28		-0.28				0.38
NSP				-0.42				0.25		-0.31	0.28			0.33
NSEXOT	-0.26													0.43
NATIVFAM				-0.35		0.36		0.34	0.36				0.27	
NSCATO				-0.31		0.32			0.31					
NSCATO_NONTOL														0.31
NSSUCK							0.33	0.38		-0.42			0.31	
NSCENT							0.30	0.44	0.28	-0.27	0.29		0.39	
NSCENT_NONTOL				-0.35		0.33		0.31	0.33		0.28			
NSBASS														0.39
NSMINN						0.33								
NSCYPR						-0.39			-0.42			-0.50	-0.50	
NSCYPR_NONTOL								0.35			0.31		0.32	0.42
NSCYPR2			-0.35			-0.38			-0.37					
NSDART			0.36			0.38			0.38					
NSICTA														0.33
NSBENT2			-0.26	-0.58	-0.42									
NSBHAB	-0.45	-0.39											0.34	0.34
NSBENT_HAB	-0.45	-0.37						0.37			0.36		0.41	0.40
NSBENT_HAB_NONTOL						0.36		0.36	0.36	-0.29				0.41
NSCOLU									0.33				0.39	
NSINTOL					-0.38		-0.26					-0.30		
NSSENS														
NSTOLE				-0.37		0.30		0.29	0.28					0.45
NTROPH														
NSBENT_INV	-0.39	-0.34									0.34		0.34	0.40
NSBENT_INV_NONTOL														0.39
NSCLNSPWNR						0.34		0.41	0.41				0.38	
NSCLNSPWNR_NONTOL	-0.38	-0.35												
NREPROS				-0.34										0.43
PCATO			-0.30	-0.42	-0.38					-0.38				
PCATO_NONTOL												0.39	0.40	
PCENT						0.35		0.28	0.36					

	NO3	NTL	PTL	TSS	TURB	AG_ TOT	MINE_ TOT	URB_ TOT	DIS- TOT	PCT_GF	PCT_SA	PCT_SAFN	XEMBE D	W1_HALL
PCENT_NONTOL														0.33
PCOTTID						0.36		0.34	0.39	-0.29			0.33	
PCYPR														
PCYPR_NONTOL												0.34		0.38
PATNG								-0.12						
PBCLN														0.37
PBCST			0.29	0.37						0.34				
PCGBU						0.25								0.32
PNEST														
PNTGU														
PBENTSP										-0.33				
PGRAVEL			0.36			0.38			0.37					
PCLNSPWNR	-0.36													
PCLNSPWNR_NONTOL										0.37				
PCOLD1														
PCOLD2										0.37				
PHIDE				-0.58	-0.41					-0.28				
PHIDE_NONTOL														0.34
PEXOT									0.26					
PNIS														
PINTOL														
PTOLE	0.33													
PTREPRO							0.26	0.34		-0.31	0.32	0.29	0.31	
PBENT														
PBENT_INV														0.29
PBENT_INV_NONTOL	-0.30						0.27				0.31			
PCARN														
PHERB	-0.36	-0.32												
PINSE												0.35		
PINVERT							0.27			-0.39				
PMACOMNI							0.38	0.45		-0.32		0.32	0.39	
PMACRO											0.27			0.39
PMICRO	0.34													
PMICRO2								0.37						
POMNI_H														
PPISC			0.39											
PPISCINV														
PPISCINV_NONTOL														0.34
PPISCINV2														

Metric and IBI Scoring — Raw scores for richness metrics (intolerant, cyprinid, and benthic habitat species) increased with watershed size and required calibration. Raw scores for the proportion metrics were unrelated to watershed size. These metrics were scored somewhat differently than McCormick et al. (2001; Table 4). Positive metrics were scored based on the 50th percentile of reference sites and 10th percentile of disturbed sites. Negative metrics were scored based on the 50th percentile of reference site and 90th percentile of disturbed sites. The numbers values were rounded to the nearest 10 to simplify scoring. For example, the 50th percentile of % Cottid individuals was 7% and was rounded to 10% for metric scoring. The scored metrics were plotted against watershed size and observed no association with size for any metric. Scored metrics were summed and calibrated to a range of 0 - 100, with the following exceptions. The IBI was set to 0 for sites with watershed areas >2 km^2 and < 10 individuals; the IBI was not calculated for sites with watershed areas $\leq$ 2 km^2 and <10 individuals (McCormick et al. (2001).

Table 4. Revised West Virginia (WV) Streams IBI metric scoring criteria. UPPER = data value expected for assemblage in good ecological condition; values $\geq$UPPER were given metric score of 10. LOWER = data value expected for assemblage in poor condition; values $\leq$ LOWER were given metric score of 0. For negative metrics, values $\leq$ UPPER were scored as 10; value $\geq$ LWER were scored 0. Data values between LOWER and UPPER were rescaled to a range of 0 to 10. Richness metrics were calibrated to watershed area (LWSAREA = $\log_{10}$(watershed area (km^2) + 1)) based on data from the entire MAHA region (McCormick et al. 2001). Scored metric correlations with WV IBI scores in right column.

Metric	UPPER	LOWER	r
% Cottid Individuals	10	0	0.39
% Non-tolerant Clean Spawner Individuals	40	0	0.69
% Macro-omnivore Individuals (Negative)	0	20	0.29
% Tolerant Individuals (Negative)	30	100	0.73
% Non-native Individuals (Negative)	0	10	0.24
Intolerant Species	1+(3*LWSAREA)	0	0.74
Cyprinid Species	1+(4.5*LWSAREA)	0	0.55
% Non-tolerant Piscivore/ Invertivore Individuals	10	0	0.18
Non-tolerant Benthic Habitat Species	5.5*LWSAREA	0	0.69

Responsiveness of IBI to disturbance gradients — The revised IBI was sensitive to four measures of watershed condition. Minimally-disturbed sites had higher IBI scores than sites with intermediate or high levels of disturbance (Figure 6a.). Sites with minimal chemical disturbance had higher IBI scores than sites with mixed anthropogenic impacts, or sites with chemistry profiles indicating agricultural or mining impacts (Figure 6b.). IBI was negatively correlated with disturbance score, indicating that IBI declined with increasing disturbance (Figure 6c.). Likewise, the IBI showed predictable declines in response to Bryce condition class. Those sites that showed the most cumulative human impacts had the lowest scores (Figure 6d.). Disturbed sites tended to have more variable IBI scores than reference sites (Figure 6b and 6d).

The West Virginia IBI was not correlated with watershed size. In general, reference sites had higher scores than intermediate and highly disturbed sites (Bryce et al. 1999). The univariate distributions of reference site IBI scores were very similar for all three methods of selecting reference sites. The IBI was responsive to catchment and riparian disturbance, sedimentation and nutrients (Appendix II).

IBI scoring Criteria — The approach described in McCormick et al. (2001) was used to set narrative criteria based on the IBI. IBI scores exceeding the 75th percentile for the reference sites (IBI>81) were classified as having excellent biotic integrity. Scores between the 75th and 25th percentiles ($70 < IBI \leq 81$) were identified as having good biotic integrity. Scores between the 5th and 25th percentiles ($56 < IBI \leq 70$) were described as being in fair condition and sites with scores below the 5th percentile were judged to be poor condition.

<u>*Comparisons of IBI performance across thermal regimes*</u>

Temperature classification of RARE sites — Data for 84 of the 119 RARE streams (71%) were analyzed. Sites were excluded from the analysis because data loggers were lost, failed, or returned abnormally high temperatures (i.e., > 32˚ C) indicating that the instrument was not properly submerged throughout the deployment. Plots of mean maximum temperature indicated a peak in summer temperature in the weeks around July 23rd (Figure 7A). The weeks of July 16-Aug 13 were selected to represent "summer" because this represented the longest continuous block of significantly warmer weekly temperatures and because weekly data for all 84 sites were available for this period. Based on these summer data, most sites were warm water (n = 54; Figure 7B). Only four of the 84 sites met the criteria for cold water streams.

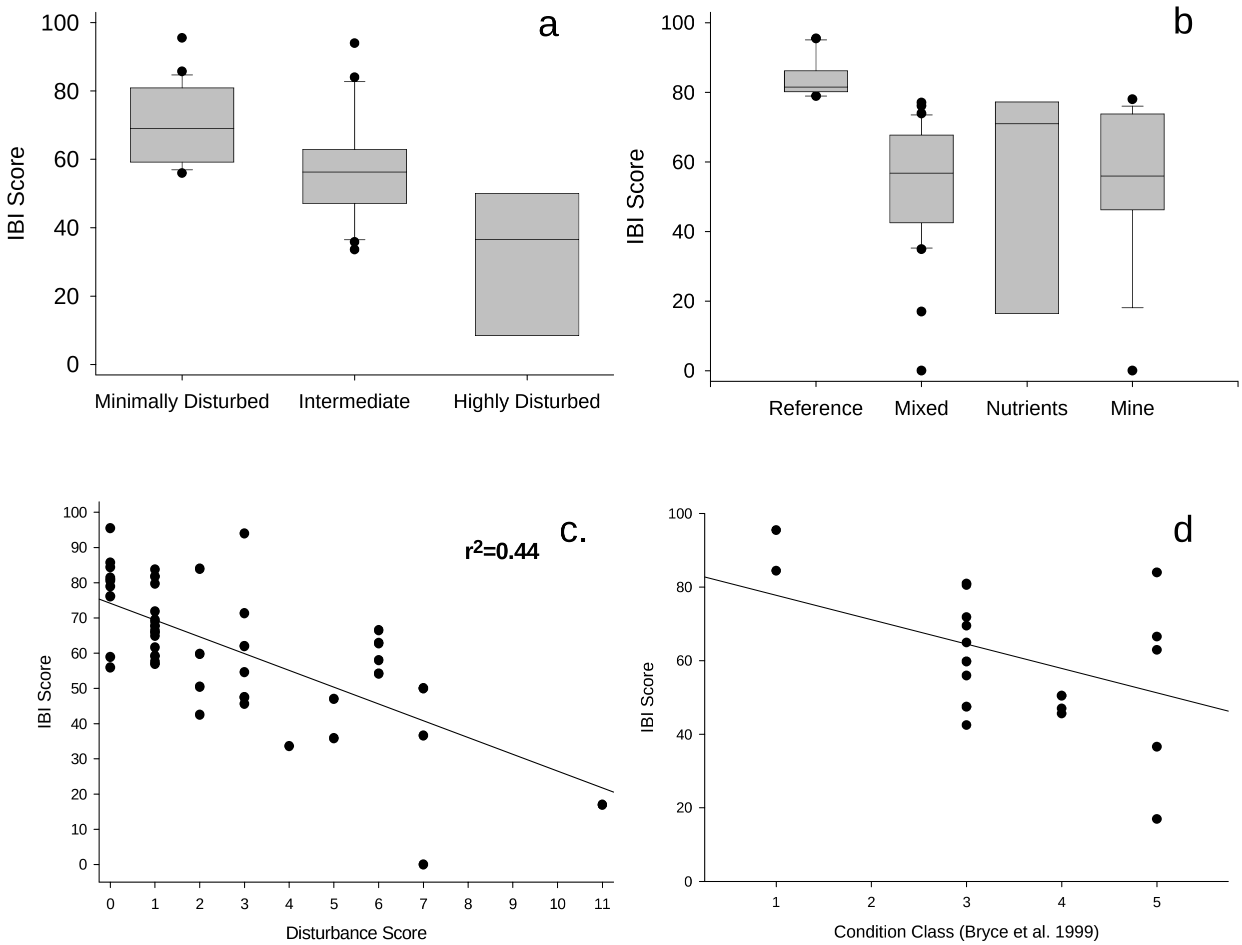


Figure 6. Responsiveness of IBI scores to measures of anthropogenic disturbance (MAIA data, n = 43 sites plus 8 revisits). Disturbance was determined by (a.) classification of sites using RBP and chemistry variables (McCormick et al. 2001), (b.) chemical classification (Scott and Davis 2000), (c.) cumulative values for eleven disturbance criteria (see Methods for details on disturbance calculations), (d.) condition class scores of Bryce et al. (1999). Higher condition class scores correspond with higher levels of disturbance. Because of the data requirements for calculation of condition class, fewer sites were classified.

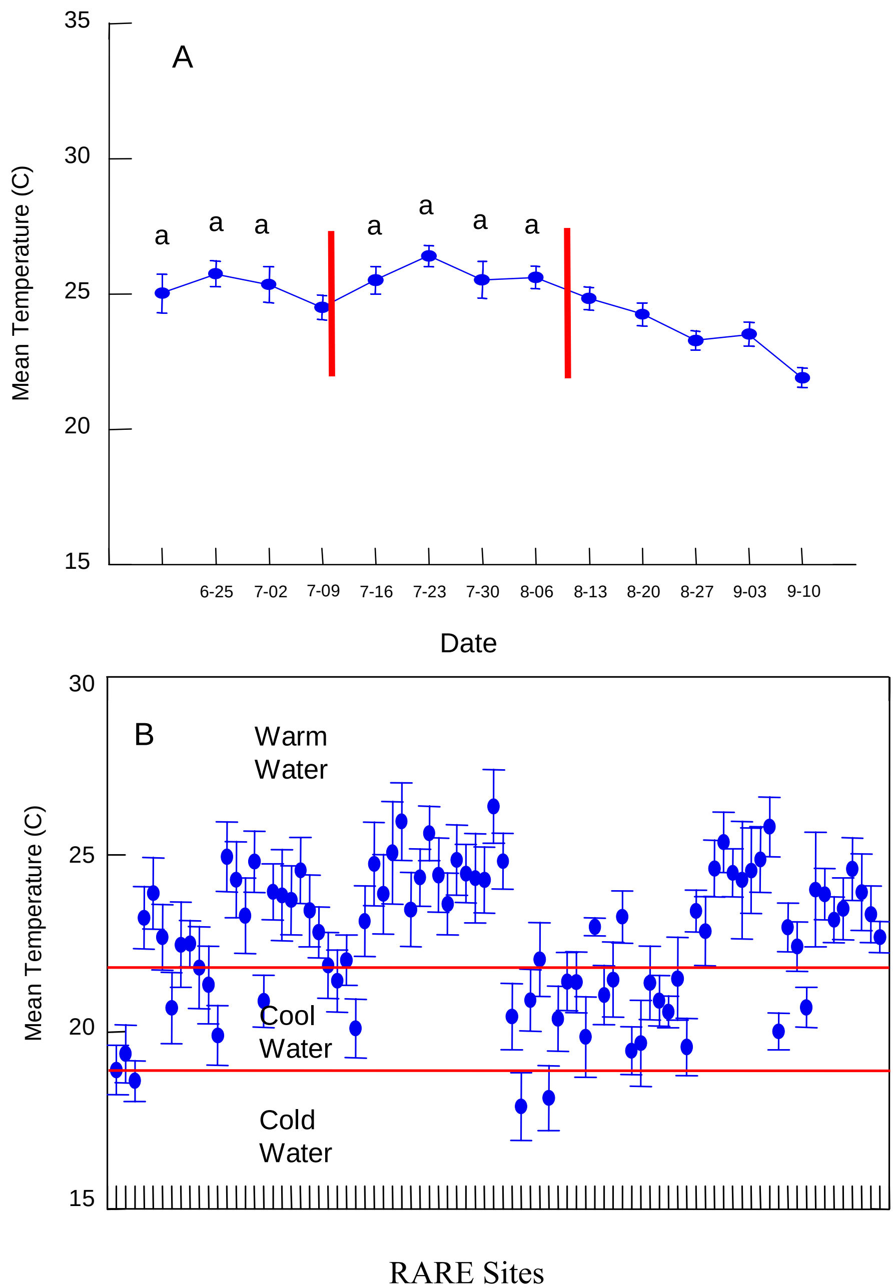


Figure 7. **(A)** Means of weekly maximum temperatures for all streams with temperature loggers. (n = 84 sites total; n = 986 total weekly observations). Deployment and retrieval of temperature loggers was staggered throughout the season, so observations for a given week vary from 33-44 sites. Temperature data are from 2001 (41 sites) and 2002 (43 sites). Bars represent a 95% confidence interval on the observations and weeks designated with "a" are significantly warmer than other weeks ($p < 0.05$). **(B)** Mean summer temperature. Means are calculated from weekly maximum and minimum temperatures for the weeks of July 16- Aug 12. Lines indicate temperature class designations for streams in Michigan (Wehrly et al. 2003). Based on these criteria, n = 4 cold, 26 cool, and 54 warm streams.

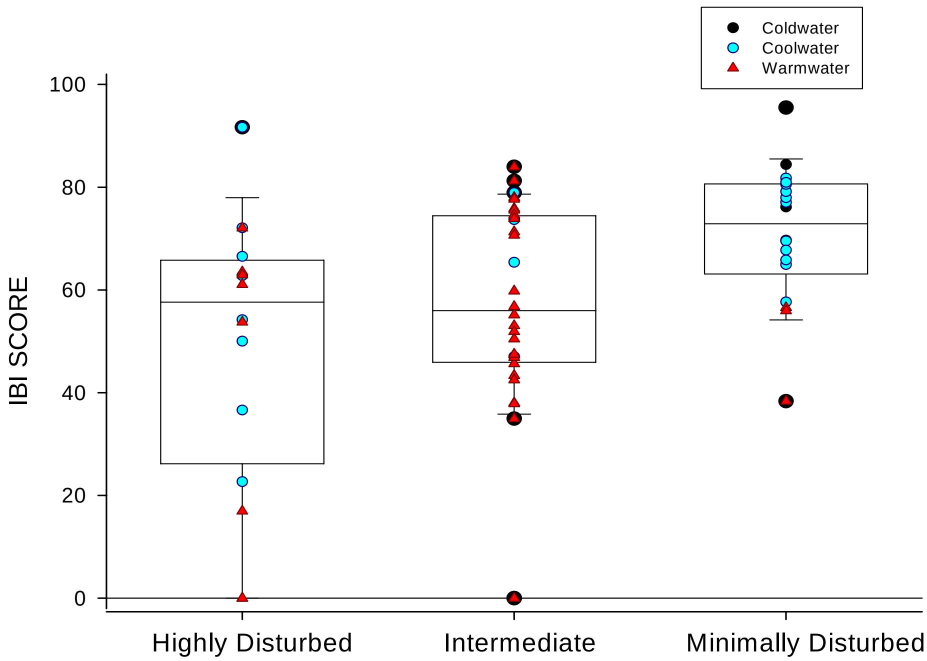


Figure 8. Response of index of biotic integrity to WV disturbance condition classifications with sites plotted according to their temperature class. Data are from EMAP surveys (n = 88 sites).

Relationships between IBI and stream temperature — The EMAP sites were categorized into cold, cool, and warm water streams based on the published temperature preferences of resident fishes. Based on these classifications, fewer reference sites were classified as warm water compared with cold and cool water (Figure 8). None of the disturbed sites were cold water. Most of the disturbed sites were warm, although cool sites were present in the intermediate and highly disturbed streams. These trends were supported by the CDFs of IBI scores for these sites (Figure 9A). Cold water sites consistently scored "fair" or better (i.e., IBI > 56), whereas cool and warm streams reflected a broader range of IBI scores and had numerous sites in the "poor" category (i.e., IBI < 55). The CDF of IBI scores for RARE sites differed from that of MAIA sites (Figure 9B). In general, RARE sites scored higher and had a narrower range than MAIA sites. Cool sites showed a broad range of scores similar to those at MAIA sites; however, warm sites all scored high with one exception.

Metric and IBI performance among temperature classes for RARE sites were directly compared (Figure 10). The small sample of cold streams (n = 4) limited identification of trends at these sites. One cold stream was not included in this analysis because no fish were collected at this site, presumably due to acid mine drainage (stream pH = 3.9). The only two metrics that increased from cool to warm streams were NSBEN_HAB_ NONTOL (number of nontolerant benthic habitat specialists) and NSCYP_NONTOL (number of nontolerant cyprinids). None of the other metrics showed a significant trend and the overall IBI score did not differ among temperature categories. This result was corroborated by a regression analysis of IBI scores and field temperature (recorded at the time of fish sampling) (Figure 11). The plot illustrates no relationship between temperature and IBI score ($r^2 = 0.01$).

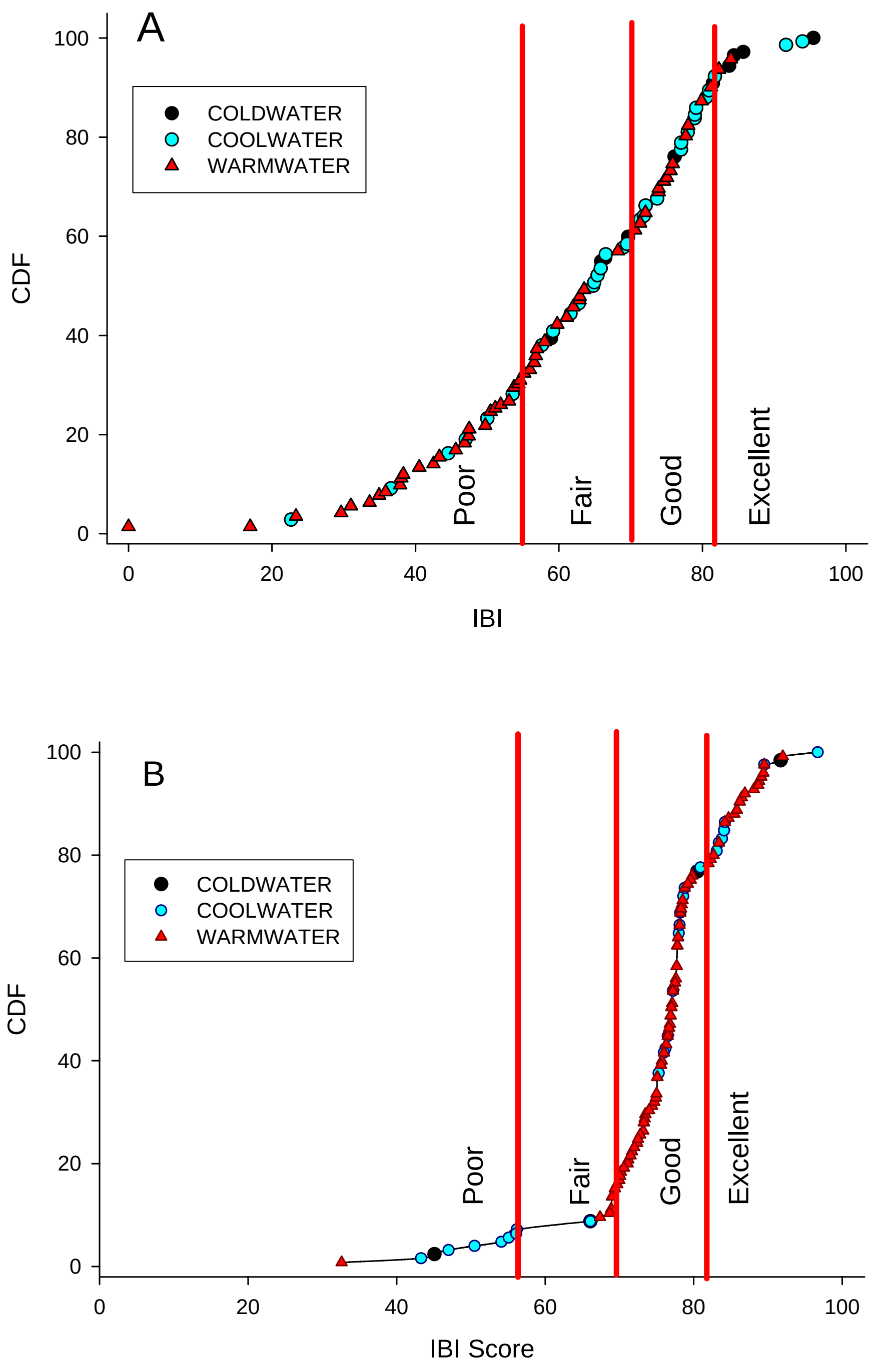


Figure 9. A. Cumulative distribution function (CDF) of index of biotic integrity scores for West Virginia from EMAP surveys (n = 69). Points are coded for stream temperature classification based on fish assemblage characteristics. B. CDF of index of biotic integrity scores for West Virginia RARE project (n = 84). Sites are coded for stream temperature based on data from temperature loggers. Red vertical lines show categories for narrative IBI criteria.

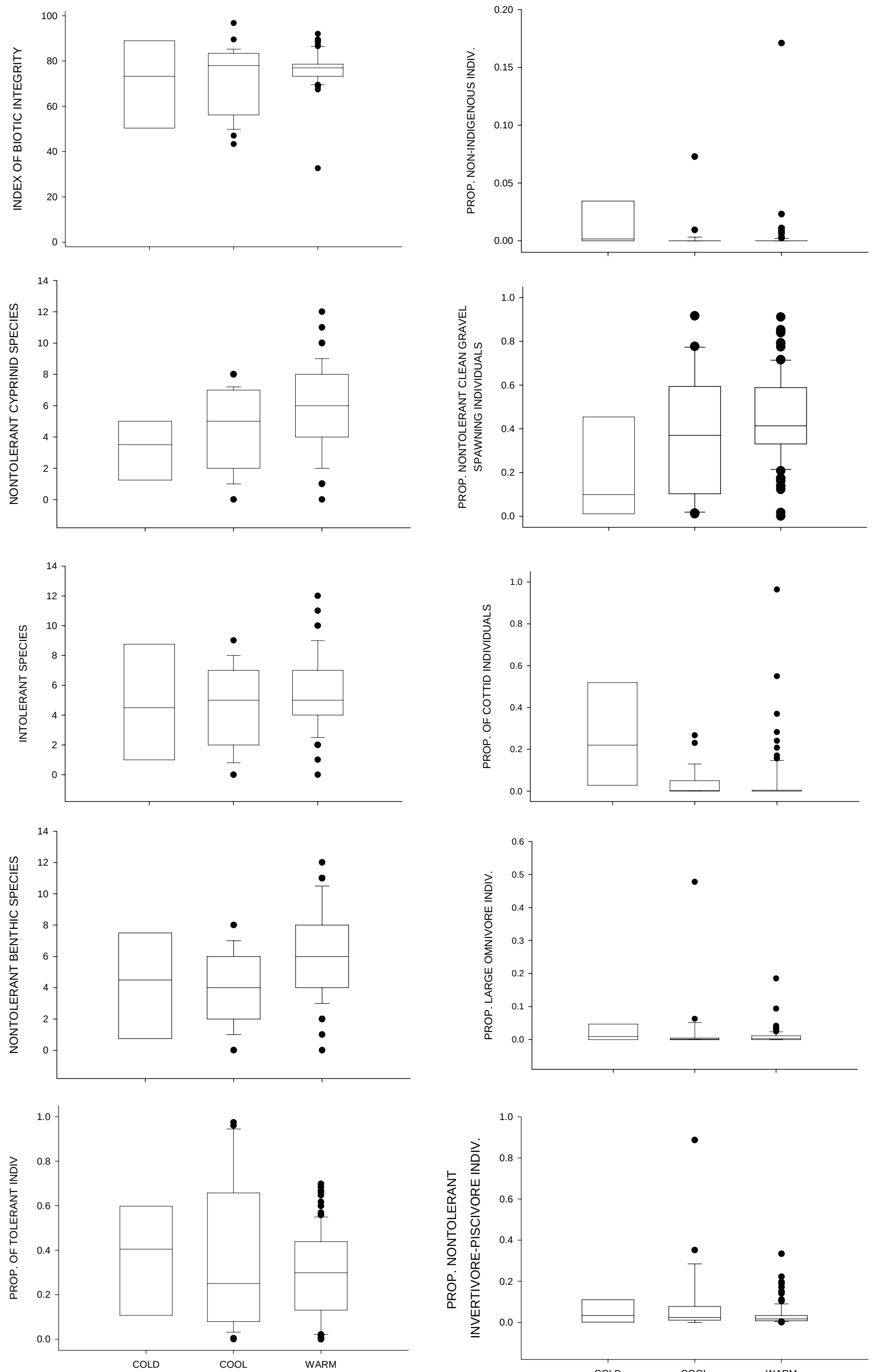


Figure 10. Box plots of IBI and component metrics for cold (n = 3), cool (n = 26) and warm (n = 54) RARE streams. Metric acronyms are defined in Appendix I.

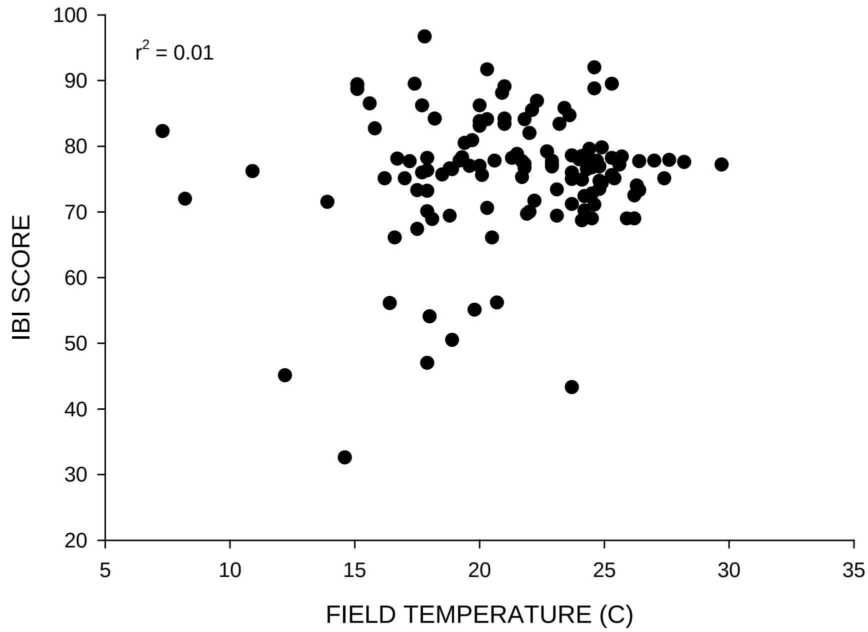


Figure 11. Plot of IBI and temperature for RARE sites. Temperature was recorded in the field when fishes were sampled.

Discussion

The recent MAHA, MAIA and RARE surveys have contributed a wealth of physical, biological and chemical data for West Virginia streams. These separate studies provide independent datasets that can be used to develop, test, and validate indices of biotic integrity or related hypotheses on anthropogenic alteration of stream communities. However, comprehensive analysis of the datasets was challenging due to differences in fish sampling methods. The analysis provided two lines of evidence supporting the use of fish data among datasets without normalizing for effects of temporal variability or sampling methodology. Seasonal, intra-annual, and interannual variation within sites was low. In addition, only marginal differences in selected metrics were observed between backpack electrofishing and the parallel wire technique. Because these differences were minor, the MAHA data were used to develop the IBI, the MAIA data was used to test IBI sensitivity to disturbance, and the MAIA and RARE data were used to assess IBI applicability across thermal regimes.

The IBI developed by McCormick et al. (2001), was used to demonstrate that IBI scores from EMAP studies in West Virginia mirrored those from the larger Mid-Atlantic region. Because the West Virginia sites showed a similar range of biotic condition to those in the Mid-Atlantic region, the metric selection methods of McCormick et al. (2001) were applied to the West Virginia MAHA data. Not surprisingly, the same metrics selected for the larger region were found to be sensitive to disturbance at the state-scale. Some metrics were modified slightly to improve their sensitivity to disturbance (e.g., tolerant taxa excluded from positive metrics). Tests of the IBI using the MAIA data showed that the IBI was sensitive to anthropogenic disturbance. IBI scores were negatively correlated with a number of disturbance variables measuring cumulative impacts to streams. These measures included a watershed disturbance class, disturbance score, condition class, and land use classification.

McCormick et al. (2001) did not determine if their IBI was equally applicable in cold, cool and warm water streams. A potential source of bias related to thermal regime is that cold water streams naturally have lower diversity than warm water streams. Thus, low scores in cold water streams would result from natural conditions rather than human disturbance. West Virginia IBI scores for cold, cool, and warm water streams from the MAIA dataset were compared and showed no bias for low scores in the cold and cool streams. Cold water streams were consistently classified as having "fair" or better condition. Likewise, cool streams showed a normal distribution of IBI scores from low to high.

The RARE dataset was also analyzed for relationships between IBI score and stream temperature. This analysis was hampered by two limitations of the RARE dataset. First, the number of cold water streams was too low (n = 4) to draw any conclusions for these streams. Second, IBI scores for RARE sites were skewed toward high scores, so the gradient of biotic condition was shorter than that observed for MAHA and MAIA sites. Given these shortcomings, warm sites did tend to score high, but that the scores of cool sites were distributed evenly from "poor" to "good". Direct comparisons of cool and warm streams showed no difference between most metrics or cumulative IBI scores. In addition, IBI scores from RARE sites were unrelated to stream temperature recorded at the time of fish sampling.

References

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Wadeable Streams and Rivers: Periphyton, Benthic Macroinvertebrates, and Fish, Second Edition. EPA 841-B-99-002, Office of Water, Washington, D.C.

Bryce, S. A., D. P. Larsen, R. M. Hughes, and P. R. Kaufmann. 1999. Assessing relative risks to aquatic ecosystems: a Mid-Appalachian case study. Journal of the American Water Resources Association 35:23-36.

Cincotta, D., F. Fulk, and N. E. Detenbeck. 2001. A Small Watershed Characterization, Classification, and Assessment for West Virginia Utilizing EMAP Design and Tools. Quality Assurance/Quality Control Plan. U.S. Environmental Protection Agency, National Exposure Research Laboratory, Cincinnati, OH.

Davis, W. S., and M. Scott. 2000. Mid-Atlantic Highlands Streams Assessment: technical support document. EPA/903/B-00/004. U. S. Environmental Protection Agency. Ft. Meade, Maryland.

Detenbeck, N. E., L. A. Jagger, S. L. Stark, M. A. Starry. WV REMAP Final Report: Watershed Classification Framework for the State of West Virginia. Washington, DC, 600/R-03/141.

Fausch, K. D., J. R. Karr, and P. R. Yant. 1984. Regional applications 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.

Herlihy, A. T., P. R. Kaufmann, and M. E. Mitch. 1990. Regional estimates of acidic mine drainage impact on streams in the Mid-Atlantic and Southeastern United States. Water, Air and Soil Pollution 50:91–107.

Herlihy, A. T., P. R. Kaufmann, M. R. Church, P. J. Wigington, Jr., J. R. Webb, and M. J. Sale. 1993. The effects of acidic deposition on streams in the Appalachian Mountain and Piedmont region of the mid-Atlantic United States. Water Resources Research 29:2687–2703.

Herlihy, A. T., J. C. Stoddard, and C. B. Johnson. 1998. The relationship between stream chemistry and watershed land use in the mid-Atlantic region, U.S. Water, Air and Soil Pollution 105:377-386.

Herlihy, A. T., D. P. Larsen, S. G. Paulsen, N. S. Urquhart, and B. J. Rosenbaum. 2000. Designing a spatially balanced, randomized site selection process for regional stream surveys: the EMAP Mid-Atlantic pilot study. Environmental Monitoring and Assessment 63:95–113.

Holton, G.D., and C. R. Sullivan. 1954. West Virginia's electrical fish-collecting methods. Progressive Fish Culturist 16(1): 10-18.

Hughes, R. M. 1995. Defining acceptable biological status by comparing with reference conditions. Pp. 31-47 in W. S. Davis and T. P. Simon (eds.) Biological assessment and criteria: tools for water resource planning and decision making. CRC Press. 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.

Jenkins, R. E. & N. M. Burkhead. 1994. Freshwater Fishes of Virginia. American Fisheries Society, Bethesda, Maryland.

Jones, K.B., K.H. Ritters, J.D. Wickham, R.D. Tankersly, R.V. O'Niell, 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, Washington, D.C. EPA/600/R-97/130.

Karr, J. R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6(6):21-27.

Karr, J. R., and D. R. Dudley. 1981. Ecological perspectives on water quality goals. Environmental Management 5:55-68.

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 Publications No. 5, Champaign, Ill.

Kaufmann, P. R. and E. G. Robison. 1998. Physical Habitat Characterization. Pp. 77-118 in J. M. Lazorchak, D. J. Klemm, and D. V. Peck, editors. Environmental Monitoring and Assessment Program - Surface Waters: Field Operations and Methods for Measuring the Ecological Condition of Wadeable Streams. EPA/620/R-94/004F.

Kaufmann, P.R., P. Levine, E.G. Robison, C. Seeliger, and D.V. Peck. 1999. Quantifying physical habitat in wadeable streams. U.S. Environmental Protection Agency, Research Triangle Park, NC. EPA 620/R-99/003.

Kazyak, P.F. 1994. Maryland Biological Stream Survey Sampling Manual. Prepared by Versar, Inc., Columbia, MD, for Maryland Department of Natural Resources, Chesapeake Bay Research and Monitoring Division.

Lazorchak, J. M., D. J. Klemm, and D. V. Peck, editors. 1998. Environmental Monitoring and Assessment Program - Surface Waters: Field Operations and Methods for Measuring the Ecological Condition of Wadeable Streams. EPA/620/R-94/004F. U.S. Environmental Protection Agency, Washington, D.C.

Leonard, P. M., and D. J. Orth. 1986. Application and testing of an index of biotic integrity in small, coolwater streams. Transactions of the American Fisheries Society 115:401-414.

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.

McCormick, F. H. and R. M. Hughes. 1998. Aquatic vertebrate indicator. 1998. Pp. 159-180 in J. M. Lazorchak, D. J. Klemm, and D. V. Peck, editors. Environmental Monitoring and Assessment Program - Surface Waters: Field Operations and Methods for Measuring the Ecological Condition of Wadeable Streams. EPA/620/R-94/004F.

McCormick, F. H., R. M. Hughes, P. R. Kaufmann, D. V. Peck, J. L. Stoddard, and A. T. Herlihy. 2001. Development of an index of biotic integrity for the Mid-Atlantic Highlands region. Transactions of the American Fisheries Society 130:857-877.

Miller, D. L., P.M. Leonard, R.M. Hughes, J.R. Karr, P.B. Moyle, L.H. Schrader, B.A. Thompson, R.A. Daniels, K.D. Fausch, G.A. Fitzhugh, J.R. Gammon, D.B. Halliwell, P.L. Angermeier & D.J. Orth. 1988.1988. Regional applications of an Index of Biotic Integrity for use in water resource management. Fisheries 13(5):12-20.

Omernik, J. M. 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77:118–125.

Overton, W. S., D. L. Stevens, and D. White. 1991. Design Report for EMAP, Environmental Monitoring and Assessment Program. EPA/600/3–91/053. U.S. Environmental Protection Agency, Corvallis, Oregon.

Pflieger, W. L. 1975. The Fishes of Missouri. Missouri Department of Conservation. Jefferson City, Missouri.

Plafkin, J. L., M. T. Barbour, K. D. Porter, S. K. Gross, and R. M. Hughes. 1989. Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic Macroinvertebrates and Fish. U.S. EPA Office of Water, EPA/440/4-89/001, Washington, D.C.

Raitz, K.B., Richard Ulack, and T. R. Leinbach. 1984. Appalachia: A Regional Geography Land, People, and Development. University of Kentucky Press, Lexington, KY.

Roth, N., M. Southerland, J. Chaillou, R. Klauda, P. Kazyak, S. Stranko, S. Weisberg, L. Hall Jr., and R. Morgan III. 1998. Maryland Biological Stream Survey: development of a fish index of biotic integrity. Environmental Monitoring and Assessment 51:89-106.

SAS Institute, Cary, NC. Software applications. 2001.

Scott, M. and L. Hall. 1997. Fish assemblages as indicators of environmental degradation in Maryland coastal plain streams. Transactions of the American Fisheries Society 126:349-360.

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. CRC Press, Boca Raton, Florida, USA.

Smogor, R. A., and P. L. Angermeier. 1999. 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, New York, New York.

Trautman, M. B. 1981. The Fishes of Ohio. Ohio State University Press. Columbus, Ohio.

Urquhart, N. S. 1982. Adjustment in covariance when one factor affects the covariate. Biometrics 38:651–660.

United States Department of Agriculture (USDA). 1996. America's Private Land, a Geography of Hope. USDA, National Resources Conservation Service.

USEPA. 1995. Stream fisheries impacted by acid mine drainage in Maryland, Ohio, Pennsylvania, Virginia and West Virginia. *http://www.epa.gov/reg3giss/datalib/metamd95.txt*.

USEPA. 2002. Consolidated Assessment and Listing Methodology: Toward a Compendium of Best Practices. Office of Wetlands, Oceans, and Watersheds. Washington, DC. http://www.epa.gov/owow/monitoring/calm/calm_contents.pdf.

Wehrly, K. E., M. J. Wiley, and P. W. Seelbach. 2003. Classifying regional variation in thermal regime based on stream fish community patterns. Transactions of the American Fisheries Society 132(1):18-38.

Whitney, G. C. 1994. From Coastal Wilderness to Fruited Plain. Cambridge University Press. Cambridge.

Woods, A.J., J.M. Omernik, D.D. Brown, and C.W. Kiilsgaard. 1996. Level III and IV Ecoregions of Pennsylvania and the Blue Ridge Mountains, the Ridge and Valley, and Central Appalachians of Virginia, West Virginia, and Maryland. EPA/600/R-96/077. U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Corvallis, OR. 50p.

Appendix I. Candidate metrics with descriptions, ranges, means and standard deviations. Metrics in boldface represent original McCormick et al. (2001) variables (some names were changed for analysis as indicated). Italics designate revised metrics adopted here.

Metric	Description	Range	MEAN	STD
NSP	No. of species including unknowns	33	10.70	7.68
NUMSPEC	No. of species	32	10.60	7.61
NUMNATSP	No. of native species	32	9.97	7.38
NFAM	No. of native families	7	3.54	1.68
NSCENT_NONTOL	No. of non-tol. Centrarchids	4	1.22	1.40
NSCENT	No. of Centrarchids	6	1.70	1.83
NSCATO	No. of sucker species	4	1.10	1.05
NSCATO2	No. of sucker species No white suckers	3	0.63	0.71
NSCATO_NONTOL	No. of non-tolerant sucker species	3	0.73	0.88
NSCYPR	**No. of cyprinid species**	**12**	**4.42**	**2.91**
NSCYPR_NONTOL	*No. of non-tolerant cyprinid species*	*9*	*2.75*	*2.36*
NSDART	No. of darter species	11	1.96	2.40
NSICTA	No. of ictalurid species	3	0.19	0.53
NSBENT_HAB	No. of benthic habitat specialist species	19	4.99	3.94
NSBENT2	**No. of benthic habitat species excluding blacknose dace**	**16**	**3.70**	**3.37**
NSBENT_HAB_NONTOL	*No. of non-tolerant benthic habitat specialists*	*16*	*3.09*	*3.21*
NSCOLU	No. of pelagic species	14	4.99	3.91
NSBENT_INV	No. of benthic invertivore species	20	4.82	4.11
NSBENT_INV_NONTOL	No. of non-tolerant benthic invertivore species	18	3.46	3.62
NSINTOL	*No. of sensitive species (larger list)*	*11*	*3.03*	*2.94*
NSSENS	**No. of sensitive species (original McCormick et al. list = NSINTOL)**	**6**	**1.18**	**1.58**
NSTOLE	No. of tolerant species	9	3.69	2.17
NSCLNSPWNR	No. of clean gravel spawning species	15	3.90	3.29
NSCLNSPWNR_NONTOL	No. of non-tolerant NSCLNSPWNR	15	3.24	3.22
NSEXOT	No. of non-indigenous species	4	0.52	0.80
NTROPH	No. of trophic guilds	6	3.79	1.61
NREPROS	No. of reproductive guilds	4	2.66	1.37
PCATO	Prop. of sucker indiv.	0.37	0.05	0.07
PCATO_NONTOL	Prop. of non-tolerant sucker indiv.	0.37	0.03	0.06
PCENT	Prop. of centrarchids	0.90	0.08	0.16
PCENT_NONTOL	Prop. of non-tolerant centrarchids	0.33	0.05	0.07
PCOTTID	***Prop. of sculpins***	***0.72***	***0.06***	***0.13***
PCYPR	Prop. of cyprinids	1.00	0.61	0.30
PCYPR_NONTOL	Prop. of non-tolerant cyprinids	0.67	0.22	0.19
PCYPTL	Prop. of tolerant cyprinids	1.00	0.43	0.34
PATNG	Prop. of attacher non-guarders	0.53	0.08	0.12
PBCLN	Prop. of clean gravel broadcast	0.60	0.09	0.12

Metric	Description	Range	MEAN	STD
	spawners			
PBCST	Prop. of broadcast spawners (tolerant)	0.13	0.01	0.03
PCGBU	Prop. of clean gravel buriers	1.00	0.35	0.27
PNEST	Prop. of nest building indiv.	0.34	0.07	0.09
PNTGU	Prop. of nest guarders	1.00	0.39	0.24
PGRAVEL	**Prop. of gravel spawners**	**1.00**	**0.59**	**0.23**
PCLNSPWNR	Prop. of clean gravel spawners	1.00	0.44	0.26
PCLNSPWNR_NONTOL	*Prop. of non-tol. clean gravel spawners*	*0.87*	*0.26*	*0.25*
PCOLD1	Prop. of coldwater indiv	0.87	0.05	0.18
PCOLD2	Prop. of coolwater indiv	0.87	0.05	0.18
PNIS	**Prop. of non-indigenous species (original McCormick et al list = PEXOT)**	**0.60**	**0.05**	**0.12**
PEXOT	*Prop. of non-indigenous species (revised list)*	*0.60*	*0.05*	*0.12*
PINTOL	Prop. of sensitive indiv.	1.00	0.22	0.26
PTOLE	***Prop. of tolerant indiv.***	***1.00***	***0.51***	***0.34***
PTREPRO	Prop. of tolerant reproductive indiv.	1.00	0.40	0.24
PBENT_INV	Prop. of benthic invertivores	1.00	0.45	0.25
PBENT_INV_NONTOL	Prop. of non-tolerant benthic invertivores	0.81	0.24	0.20
PCARN	Prop. of carnivores	1.00	0.27	0.26
PHERB	Prop. of herbivores	0.50	0.07	0.10
PINVERT	Prop. of all invertivores	0.95	0.40	0.28
PMACOMNI	**Prop. of macro-omnivores (original McCormick et al list = PMACRO)**	**0.35**	**0.03**	**0.06**
PMACRO	*Prop. of macro-omnivores (revised list)*	*0.35*	*0.03*	*0.07*
PMICRO	Prop. of micro-omnivores	1.00	0.22	0.23
PMICRO2	Prop. of micro-omnivores (revised list)	1.00	0.22	0.23
POMNI_H	Prop. of omniv. + herbiv.	1.00	0.31	0.24
PPISC	Prop. of piscivores	0.17	0.01	0.03
PFISHBUG	Prop. of invertivore-piscivores	1.00	0.26	0.26
PPISCINV	Prop. of invertivore-piscivores (revised list)	1.00	0.28	0.25
PPISCINV2	**Prop. of invertivore piscivores excluding creek chub**	**0.87**	**0.10**	**0.18**
PPISCINV_NONTOL	*Prop. of non-tol. Invertivore-piscivores*	*0.87*	*0.09*	*0.18*
NUMFISH	Number of indiv.	1718	236.25	271.85

Appendix II. Spearman's correlation coefficients ($r > 0.2$; $p<0.05$) showing response of WV IBI score to chemistry, physical habitat, and land cover variables sampled as part of EMAP (n = 69 sites). Stressor variable acronyms are defined in Appendix III.

Stressor Variable	Spearman's r
AG_TOT	-0.2
DISTOT	-0.2
FOR_TOT	0.2
MINE_TOT	-0.26
NONRES	-0.33
TOT_RD	0.29
HOUSINGDENS_KM	-0.48
_TOT	-0.25
ALTD	-0.22
DOC	-0.31
MN	-0.39
NH4	-0.21
NTL	-0.2
PCT_FN	-0.31
W1_HALL	0.26
QR1	-0.25
QRDIST1	-0.35
QRPHALT2	-0.25

Appendix III. List of stressor variables compiled for WV stream sites.

Acronym	Description
ACID_CLS	Acid Dep. Condition (ANC)
AG_TOT	% watershed - agricultural lands
ALKCALC	Calculated Alkalinity (µeq/L)
ALTD	Total Dissolved aluminum (µg/L)
AMD_CLS	Acid Mine Drainage Condition (SO4,ANC)
AMDCLS	AMD Classification
AMDCLS2	Lumped AMD Classification
ANC	Gran Acid Neutralizing Capacity (ueq/L)
AREA_WS	Watershed area digitized from maps
AREASUM	Residual Pool Vert Profile Area (m^2/reach)
AREASUMC	Residual Pool Vert Profile Area (m^2/chan.)
AREAWSHA	Watershed Area in Hectares
ASPCTDEG	Est. aspect of watershed longest dim.
bank_veg	bank protective vegetation score
BAR_TOT	% watershed - barren land
CA	Calcium (µeq/L)
chan_alt	lack of channel alteration score
chan_fls	channel flow status score
chan_sin	channel sinuosity score
CHL	Amount of Chlorophyll a (mg)
CHL_M2	Chlorophyll a (mg)/m^2 of Stream Bed
CHL_MASS	Ratio of Choro-a(mg):Periphyton AFDM(g)
CL	Chloride (µeq/L)
cnd_bank	condition of banks score
CO3	Calculated Carbonate (µeq/L)
COND	Specific Conductance (µS/cm)
CONDTION	Sandys Site Condition Class (1=good)
CROWS_D	Straight line valley length of reach (m)
DIC	Dissolved Inorganic Carbon (mg/L)
DIST_CLS	Overall Disturbance Class
DISTOT	Sum of land use(URB_TOT+AG_TOT+MINE_TOT)
DOC	Dissolved Organic Carbon (mg/L)
ECOREG	Omernik Rev. Ecoregion ID
ECOREGL4	Omernik Level 4 Ecoregion ID (1996 ver.)
ELEV	Est. elevation of stream index site (m)
elevmax	Highest watershed elevation (m)
ELEVMEAN	Mean Watershed Elevation (m)
ELEVMIN	Min Watershed Elevation (m)
embedded	gravel not buried by fines score

epif_sub	epifaunal substrate score
EXOT_CLS	Condition Class based on nonnative fish
FE	Total Iron (mg/L)
FEN_SECT	Fenneman physiographic section designation
FISH_D	Reach Length (m) -- as the fish swims
FOR_TOT	% watershed - forest
frq_riff	riffle frequency score
grazing	vegetative grazing disturbance score
HCO3	Calculated Bicarbonate (μeq/L)
HOUDENKM	Housing unit density (housing/km^2)
in_cover	instream cover score
K	Potassium (μeq/L)
LRBS_TST	Log_{10}[Relative Bed Stability] - Fast estimate
LSUB_DMM	Substrate-Mean Log_{10}(Diameter Class mm)
LTEST	Log_{10}[Erodible Substr Dia.(mm)]-Fast estimate
LTROFF_M	Approx. meters of annual runoff
LWD_CLS	LWD Condition (XFC_LWD)
LWSKM2	Log_{10} watershed area (km^2)
MG	Magnesium (μeq/L)
MINE_TOT	% watershed - mines/quarries/gravel pits
MN	Total Manganese (mg/L)
NA	Sodium (μeq/L)
NH4	Ammonium (μeq/L)
NO3	Nitrate (μeq/L)
NONRES	% watershed - non-residential urban lands
NRP	Number of residual pools in reach
NTL	Total Nitrogen (μg/L)
NTL_CLS	Nutrient Condition (NTL)
NTROPH	Number of trophic guilds
NUTRCLS	Nutrient Classification
PCAN_C	Riparian Canopy Coniferous (Fraction of reach)
PCAN_D	Riparian Canopy Deciduous (Fraction of reach)
PCAN_E	Rip Canopy Broadlf evrgrn (Fraction of reach)
PCAN_M	Rip Canopy Mix Conif-Decid (Fraction of reach)
PCAN_N	Rip Canopy Absent (Fraction of reach)
PCT_BIGR	Substrate >= Coarse Gravel (>16 mm) (%)
PCT_BL	Substrate Boulders -- 250-4000 mm (%)
PCT_CA	Cascade (% of reach)
PCT_CB	Substrate Cobbles -- 64-250 mm (%)
PCT_FA	Falls (% of reach)
PCT_FAST	Fast Water Habitat (% riffle & faster)
PCT_FN	Substrate Fines -- Silt/Clay/Muck (%)
PCT_GC	Substrate Coarse Gravel -- 16-64 mm (%)

PCT_GF	Substrate Fine Gravel -- 2-16 mm (%)
PCT_GL	Glide (% of reach)
PCT_HP	Substrate Hardpan -- (%)
PCT_OM	Substrate Organic Detritus -- (%)
PCT_ORG	Substrate Wood or Detritus -- (%)
PCT_POOL	Pools -- All Types (% of reach)
PCT_RI	Riffle (% of reach)
PCT_SA	Substrate Sand -- .06-2 mm (%)
PCT_SAFN	Substrate Sand & Fines -- <2 mm (%)
PCT_SFGF	Substrate <= Fine Gravel (<=16 mm) (%)
PCT_SLOW	Slow Water Habitat (% Glide & Pool)
PCTCHARP	% of channel length that forms residual pools
PCTCHASD	% of channel length with sediments present
PCTRCHRP	Residual pool length proportion (% reach)
PCTRSED	Thalweg Sediment (<16mm) Pres.(% length of Thalweg)
PFC_ALG	Filamentous Algae Presence (% reach)
PFC_ALL	Any Types Fish Cover Present (% reach)
PFC_AQM	Aq. Macrophytes Presence (% reach)
PFC_BIG	LWD,RCK,OHB or HUM Fish Cover Pres (% reach)
PFC_BRS	Brush & Small Debris Presence (% reach)
PFC_LWD	LWD Presence (% reach)
PFC_NAT	Any Natural Fish Cover Present (% reach)
PFC_OHV	Overhang. Veg. Presence (% reach)
PFC_RCK	Boulders Presence (% reach)
PFC_UCB	Undercut Bank Presence (% reach)
PHSTVL	Closed System pH
PMID_C	Rip MidLayer Coniferous (Fraction reach)
PMID_D	Rip MidLayer Deciduous (Fraction reach)
PMID_E	Rip MidLayer broadlf evrgrn (Fraction reach)
PMID_M	Rip MidLayer Mix Con-Decid (Fraction reach)
PMID_N	Rip MidLayer Absent (Fraction of reach)
pool_sub	pool substrate characterization score
pool_var	pool variability score
PRECIP_M	Approx. annual precipitation (m)
PROJECT	EMAP or REMAP
PTL	Total Phosphorous (µg/L)
PTL_CLS	Nutrient Condition (PTL)
QR1	Riparian Quality Index
QRDIST1	Riparian Disturbance Index
QRPHALT2	Riparian Habitat Condition
QRVeg1	Riparian Vegetation Index
QRVeg2	Riparian Vegetation Index 2 (Understory layer)
RBPMEAN	mean of all nonmissing scores

RBPSUM	sum of all nonmissing scores
RD_DEN	Road density (m/ha)
REACHLEN	Length of sample reach (m)
REF	RBP Habitat/Chem Reference Site (Y/N)
REFIAN	Ian NABS Paper Ref Site (N>100) (Y/N)
REFPHIL	Ref Site by Chem/RBP/Quant. Phab.
REFSANDY	Ref Site by Chem/RBP/Phab/Condtion
RIP_CLS	Riparian Condition (QRPHALT2)
ripa_veg	width of riparian vegetation zone score
ROUGHNES	Terrain Roughness (unitless)
RP100	Mean Residual Depth (m^2/100m)
RP100C	Mean residual area per 100 m of chan.
RPGT100	Residual Pools >100cm deep (number/reach)
RPGT50	Residual Pools >50cm deep (number/reach)
RPGT75	Residual Pools >75cm deep (number/reach)
RPMXAR	Max. RP profile area in reach (m^2/pool)
RPMXDEP	Maximum residual depth in reach (cm)
RPMXLEN	Max. residual pool length in reach (m/pool)
RPMXVOL	Max volume of any pool in reach (m^3)
RPMXWID	Max residual width of any pool in reach (m)
RPXDEP	Mean RP depth in reach (cm/pool)
RPXLEN	Mean length of residual pools (m/pool)
RPXVOL	Mean residual pool volume (m^3/pool)
RPXWID	Mean residual width of reach (m)
SECTNAME	Section name on Fenneman (1946) map
SED_CLS	Excess Sediment Condition (LRBS_BW5)
sedi_dep	lack of sediment deposition
SINU	Channel Sinuosity (m/m)
SIO2	Silica (mg/L)
SITE_ID	Site identification code
SLOPE	Approx. slope (HI_TO_LO / WSLTH)
SLOPMEAN	Mean Watershed Slope (%)
SO4	Sulfate (ueq/L)
SOBC	Sum of Base Cations (ueq/L)
STATE	Site State Location
STRAHLER	Stream Order (Strahler)
TEMPSTRM	Stream temperature (C)
TOLERNT9	Final IBI PTOLE Metric Score
TOT_RD	m road in watershed (1992 TIGER files)
TOTPLEN	Total residual pool length (m/reach)
TOTPLENC	Total residual pool length (m/chan.)
TOTPVOL	Total residual pool volume (m^3/reach)
TOTPVOLC	Total residual pool volume (m^3/chan.)

TOTSDLEN	Total RP length with sediment (m/reach)
TOTSDLNC	Total RP length with sediment (m/chan.)
TRASHED	Highly Disturbed Site by Chem/RBP/Sandy
TSS	Total Suspended Solids (mg/L)
TURB	Turbidity (NTU)
URB_TOT	% watershed - urban lands
velocity	presence of velocity/depth regimes score
VISIT_NO	Visit Number
W1_HAG	Rip Dist--Sum Agric Types (ProxWt Pres)
W1_HALL	Rip Dist--Sum All Types (ProxWt Pres)
W1_HNOAG	Rip Dist--Sum NonAg Types (ProxWt Pres)
W1H_CROP	Rip Dist--Row Crop (ProxWt Pres)
W1H_LOG	Rip Dist--Logging Activity (ProxWt Pres)
W1H_MINE	Rip Dist--Mining Activity (ProxWt Pres)
W1H_PSTR	Rip Dist--Pasture/Hayfield (ProxWt Pres)
W1H_PVMT	Rip Dist--Pavement (ProxWt Pres)
W1H_ROAD	Rip Dist--Road/Railroad (ProxWt Pres)
WETL_TOT	% watershed wetlands
WS_AREA	Watershed area (km^2)
WS_COND	Watershed Condition Class (Colleens)
WSDISTRB	Human Disturbance Level in Watershed
XC	Riparian Veg Canopy Cover
XCDENBK	Mean Bank Canopy Density (%)
XCDENMID	Mean Mid-channel Canopy Density (%)
XCEMBED	Mean Embeddedness--Channel only (%)
XCL	Riparian Canopy > 0.3m DBH (Cover)
XCM	Rip Veg Canopy+Mid Layer Cover
XCMG	Rip Veg Canopy+Mid+Ground Cover
XCMGW	Rip Veg Canopy+Mid+Ground Woody Cover
XCMW	Rip Veg Canopy+Mid Layer Woody Cover
XCS	Riparian Canopy <= 0.3m DBH (Cover)
XDEPTH	Thalweg Mean Depth (cm)
XEMBED	Mean Embeddedness--Channel+Margin (%)
XFC_ALG	Fish Cvr-Filamentous Algae (Areal Prop)
XFC_ALL	Fish Cvr-All Types (Sum Areal Prop)
XFC_AQM	Fish Cvr-Aq. Macrophytes (Areal Prop)
XFC_BIG	Fish Cvr-LWD,RCK,UCBorHUM(Sum Area Prop)
XFC_BRS	Fish Cvr-Brush&Small Debris (Areal Prop)
XFC_LWD	Fish Cvr-Large Woody Debris (Areal Prop)
XFC_NAT	Fish Cvr-Natural Types (Sum Areal Prop)
XFC_OHV	Fish Cvr-Overhang Veg (Areal Prop)
XFC_RCK	Fish Cvr-Boulders (Areal Prop)
XFC_UCB	Fish Cvr-Undercut Banks (Areal Prop)

XINC_H	Channel Incision Ht.-Mean (m)
XSLOPE	Channel Slope -- reach mean (%)
XUN	Undercut Distance--Mean (m)
XWD_RAT	Mean Width/Depth Ratio (m/m)
XWIDTH	Wetted Width -- Mean (m)
XWXD	Mean Width*Depth Product (m^2)
YEAR	Sample Year
ZN	Dissolved Zinc (mg/L)

United States
Environmental Protection
Agency

PRESORTED STANDARD
POSTAGE & FEES PAID
EPA PERMIT NO. G-35

Office of Research and Development
National Exposure Research Laboratory
Cincinnati, OH 45268

EPA/600/R-06/010
February 16, 2006