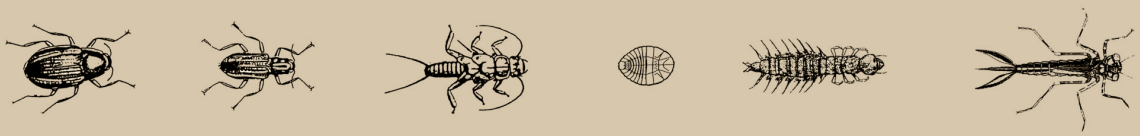
Development and Application of the Mississippi Benthic Index of Stream Quality (M-BISQ)



Mississippi Department of Environmental Quality
Office of Pollution Control
2380 Highway 80 West
Jackson, Mississippi 39204

June 30, 2003

MDEQ

Development and Application of the Mississippi Benthic Index of Stream Quality (M-BISQ)

(Section 104(b)(3) Water Quality Cooperative Agreement Grant, CP974381-01)

Prepared for:

Water Quality Assessment Branch
Mississippi Department of Environmental Quality
2380 Highway 80 West
Jackson, Mississippi 39204

Prepared by:

Tetra Tech, Inc.
10045 Red Run Blvd., Suite 110
Owings Mills, Maryland 21117-6102

Under Contractual Agreement MDEQ-01-ID-002 FTN/WLBE between Mississippi Department of Environmental Quality and FTN/WLBE Joint Venture, 3 Innwood Circle, Suite 220, Little Rock, AR 72211

June 30, 2003

ABSTRACT

In 2001, Mississippi Department of Environmental Quality initiated development of a geographically calibrated multimetric index for the purpose of re-evaluating §303(d) listed streams found throughout the state except the Mississippi Alluvial Plain. Biological, chemical, physical, and landscape data were collected for over 450 sites during the winter index period (Jan-Mar) of 2001. Sites used for calibrating and testing indices and associated metrics were selected using quantitative landscape, physical, and chemical criteria. They were selected to represent the least-disturbed and most-disturbed conditions. Five site classes (Northwest, Black Belt, Northeast, West and East) were developed for the state based on variability of physical, chemical, and biological characteristics of the streams sampled across the study region. Separate indices were developed for each of these site classes and used to evaluate the impairment status of streams found within each site class. A total of 455 streams were evaluated using the Mississippi Benthic Index of Stream Quality (M-BISQ). This report describes the steps involved in developing the M-BISQ and presents results from each step. Appropriate management uses of final indices as well as possible future analyses are recommended.

EXECUTIVE SUMMARY

The total maximum daily load (TMDL) process requires that water resource systems (such as, streams, rivers, lakes, reservoirs, and wetlands) be evaluated for overall ecological condition and, if assessed as degraded, improved to meet their designated use(s). As of 1999, approximately 700 waterbodies in Mississippi had been listed as degraded (i.e., §303[d] listed), however, little or no quantitative data were used in establishing approximately 550 of these listings. Therefore, the Mississippi Department of Environmental Quality (MDEQ) initiated a project to re-evaluate the state's §303(d) listed streams using biological data along with other physical and chemical information. These data were calibrated according to statistically-based reference points representative of desired least-disturbed conditions, and are summarized in the Mississippi Benthic Index of Stream Quality (M-BISQ). This IBI-type index can be used for assessing the overall ecological condition of sites, as well as contributing to evaluation of the effects of nutrient enrichment, sedimentation, habitat impairment, and land use conversions. The M-BISQ will be used in establishing restoration and remediation goals, tracking the effectiveness of restoration and remediation activities, and developing watershed management strategies.

Developing the M-BISQ involved the following steps: 1) develop database, 2) delineate preliminary site classes, 3) develop criteria for designation of least-disturbed sites (least-disturbed (a) [LDa] and least-disturbed (b) [LDb], where LDb criteria are slightly less stringent), and most-disturbed sites (MD), 4) calculate metrics, 5) delineate final site classes, 6) test metrics, and 7) develop index. In step 1, over 450 stream locations (§303(d) listed and potential LDa streams) were sampled over a 6-7 week period during a winter index period spanning January – March, 2000. Potential LDa sites were selected based on their location in areas of extensive forest cover, or agency knowledge of the stream or watershed. Data collected in the field included field chemistry (pH, water temperature, specific conductance, TDS, turbidity, and dissolved oxygen), water grab samples for laboratory analytical chemistry (COD, TOC, TP, TKN, NH_3, nitrate/nitrite, total alkalinity, and total chlorides), physical habitat (visual-based habitat quality assessment and modified 100-particle Wolman pebble count), and benthic macroinvertebrates (multiple-habitat approach). All data were entered into EDAS (Ecological Data Application System) for data management and analysis. In step 2, 10 preliminary classes were developed based on the variability of physical and chemical parameters among potential LDa sites. In step 3, LDa, LDb, and MD site criteria were developed for each of these preliminary classes. Spatial distribution of biological metric values (calculated in step 4) and multivariate analyses were used to describe the variability of benthic assemblages of the LDa and LDb sites to develop five final site classes, or bioregions. The Northwest bioregion was composed of the northern sections of Level 4 ecoregions 74b and 65e; Black Belt was ecoregion 65a; the Northeast site class was composed of ecoregions 65b, i and j; West bioregion was composed of ecoregions 74a, b, and c; and, the East bioregion was made up of ecoregions 65d, r, and f. The discriminatory ability of biological metrics was statistically evaluated (step 6) through comparisons of LDa and MD site metric values. The best performing metrics within each site class were standardized and incorporated into final indices (step 7) and resulted in five indices (one for each bioregion), each with 6 or 7 metrics, as follows:

BIOREGIONS				
Black Belt	**East**	**Northwest**	**Northeast**	**West**
Metrics				
No. Collector taxa	% Caenidae	No. Chironomidae taxa	% Clingers	Hydropsychidae/ Trichoptera
Beck's Biotic Index	No. Tanytarsini taxa	% Clingers	% Diptera	Beck's Biotic Index
No. Plecoptera taxa	% Filterers	% Ephemeroptera (no Caenidae)	% Filterers	No. Sprawler taxa
Total taxa	Beck's Biotic Index	No. Filterer taxa	% Tanytarsini	% EPT (no Caenidae)
No. Sprawler taxa	Hilsenhoff Biotic Index	Beck's Biotic Index	Hilsenhoff Biotic Index	No. Coleoptera taxa
No. Coleoptera taxa	% EPT (no Caenidae)	Hilsenhoff Biotic Index	No. Trichoptera taxa	No. Predator taxa
% Caenidae	% Clingers	% Tanytarsini		

Discrimination efficiencies (DEs) for the five indices, which describe the ability of an index to detect impairment (higher percents = better detection ability), ranged from 89-100%; the average DE was 92%. A comprehensive quality assurance/quality control (QA/QC) program was maintained that included training, field and laboratory audits, a series of QC checks with documented results. This report includes a data quality assessment that partitions variability and attempts to isolate error sources. Index scores will be used for assessing the status of §303(d)-listed streams (i.e., whether listing or de-listing should occur), and as an indicator to be used in long-term stream and watershed monitoring.

ACKNOWLEDGMENTS

This work was completed under contractual agreement number MDEQ-01-ID-001 FTN/WLBE between the Mississippi Department of Environmental Quality and the FTN & Associates/WLBE Joint Venture (Work order number 01-0001 FTN/WLBE-003). The work was partially funded under Section 104(b)(3) Water Quality Cooperative Agreement Grant Number CP974381-01. The primary authors of this report were David W. Bressler and James B. Stribling of Tetra Tech, Inc. (Owings Mills, Maryland). Matt Hicks was the MDEQ (Jackson, Mississippi) Project Manager and David Loch was the MDEQ QA Project Manager. Along with Matt and David, Randy Reed, Jeff Thomas, and Kevin Pigott provided solid technical and management review and input. The Program Coordinator was Ms Glenda Miller, USEPA, Region IV, and the Technical Coordinator was Mr. Ed Decker, USEPA Region IV.

Fieldwork was conducted by Chip Bray, Charles Cockrell, David Felder, Matt Hicks, Pete Howard, David Loch, Kevin Pigott, and Barbara Viskup of MDEQ; Lindsey Lewis and Robert Gregory of FTN & Associates (Little Rock, Arkansas); Deion Quinn of CivilTech, Inc. (Jackson, Mississippi); and Marcus Bowersox, David Bressler, Nathan Goodman, and Kristen Pavlik of Tetra Tech. Harold Bishop, Coleman Leach, Paul Grantham and Cliff Hornbeak of MDEQ's Office of Land and Water Resources performed tape-down and flow measurements. Laboratory sorting/subsampling was conducted by Mike Beiser, Allan Bennett, Chip Bray, Greg Clark, Charles Cockrell, Emily Cotton, Jenny Geraci-Ulmer, Sean Jackson, Barrott Lambdon, Eric Pederson, Richard Peets, Allison Sherman, Brandy Smith, Jermaine Travis, Chuck Thompson, Katherine Williams, Natalie Young, and Nicole Young of MDEQ; and by Chad Barbour, Carmela Biddle, Carolina Gallardo, Matt Geiman, Colin Hill, Corinne Marino, and Mandy Richardson of Tetra Tech. Mike Winnell of Freshwater Benthic Services, Inc. (Petoskey, Michigan) performed all primary taxonomic identifications. Alice Dossett and Richard Peets of MDEQ performed taxonomic quality control. Erik Leppo of Tetra Tech provided primary database management efforts, and customized the architecture of the Ecological Data Application System (EDAS) to MDEQ's program. John Brandon and Natalie Guedon of MDEQ and Barun Pani of Cyber Citizen, Inc., provided technical input for EDAS migration to MDEQ. Gary Hennington (Information Management Systems, Inc.) and Nona Yates (MDEQ) conducted extensive GIS analyses. Ed Decker, Jim Harrison, and David Melgaard (USEPA Region 4, Atlanta, GA), Pat Downey (FTN & Associates), Leslie Barkley, Mike Beiser, Chip Bray, David Felder, Henry Folmar, Matt Hicks, Kevin Pigott, Dave Loch, Randy Reed, Jeff Thomas, and Doug Upton (MDEQ), and Charlie Cooper and Sam Testa (USDA-ARS National Sedimentation Laboratory, Oxford, MS) provided comments on site classification and index development. Hoke Howard of USEPA Region 4 provided study design advice. Esther C. Peters and Susan Adair of Tetra Tech Inc. (Fairfax, Virginia), prepared the Quality Assurance Project Plan, and executed most QC activities, including all field and laboratory audits. Brenda Decker and Linda Shook of Tetra Tech, Inc. provided clerical and administrative support.

Table of Contents

Table of Contents (Continued)

List of Tables

List of Figures

List of Figures (Continued)

1. INTRODUCTION

1.1 Background on MDEQ biological assessment program and current status

Section 303(d) of the Clean Water Act (CWA) of 1972 requires that restoration and remediation strategies be developed for degraded waterbodies (NRC 2001). Those strategies typically include calculation of total maximum daily loads (TMDLs) for individual stressors (pollutants) affecting the condition of the waterbody. The initial step in the TMDL process is the determination of waterbody impairment and listing of the waterbodies. Because the initial listing process for Mississippi involved use of low quality data, or no data at all, the state undertook a program to develop a more reliable biological assessment methodology to confirm waterbody listing or the need for de-listing. Since the initial listing of Mississippi waterbodies more reliable field sampling, laboratory processing, and data analysis methods have evolved. This has, in large part, arisen from the need to address the CWA's goal of protecting, restoring, and enhancing the biological integrity of aquatic ecosystems, and the need to use consistent, reliable, and defensible data for water resources management and regulatory purposes.

The definition of impairment by natural resource management or regulatory agencies is typically based on attainment or non-attainment of numeric water quality standards associated with a waterbody's designated use. If those standards are not met (or attained), then the waterbody is considered to be degraded. Since all waters of the U. S. are designated for aquatic life use (ALU), indicators that reflect overall biological condition (such as a properly-calibrated multimetric index) are appropriate for evaluating impairment or non-impairment.

One of the questions that arises in any biological monitoring and assessment activity is related to the uncertainty that may be associated with the data and interpretive results. Sampling and analysis protocols used to develop assessments are in themselves a series of steps or methods, each providing samples, data, or statistical results to the next, eventually leading to an assessment of the site, stream, or watershed. There is always a certain amount of sampling or measurement error associated with each step of the process (Taylor 1988, Diamond et al. 1996). To allow Mississippi DEQ to begin to evaluate and report uncertainty associated with these biological assessments, they have designed and instituted a new strategy that includes partitioning of data collection and analysis procedures so that variance can be characterized and potential error sources identified and corrected.

The purpose of this project was two-fold: 1) develop an Index of Biological Integrity (IBI)-type indicator for use in describing the impairment status of wadeable streams and rivers in all ecoregions of Mississippi except the Mississippi Alluvial Plain (Delta), and 2) re-evaluate the impairment status of the state's 303(d)-listed streams and provide listing/de-listing recommendations. This report describes the development of a geographically calibrated biological index for Mississippi streams which will be used to develop impairment ratings for 465 sites (including ~300, §303[d] streams) throughout the state. Methods for developing the index, results of site classification and index calibration, bioregional summaries, and recommendations for stream management are presented.

1.2 Background on multimetric indices

Biological assemblages including benthic macroinvertebrates, periphyton, and fish have all been successfully used for monitoring stream conditions (Karr et al. 1986, Hill 1997, Southerland and Stribling 1995). In particular, benthic macroinvertebrate assemblages have been effective for bioassessment because they:

- are good indicators of localized conditions because they are relatively sedentary
- integrate the effects of many short-term environmental variations because most species have a life cycle of several months to several years
- are made up of species that constitute a broad range of trophic levels and pollution tolerances
- can be sampled easily, requiring few people and inexpensive gear, and resulting in minimal detrimental effect on the resident biota
- serve as a primary food source for fish
- are abundant and diverse in most streams

Biological integrity, defined as the ability of a system *"to support and maintain a balanced, integrated, adaptive community of organisms having a composition, diversity, and functional organization comparable with that of natural habitats of the region*" (Frey 1977, Karr and Dudley 1981, Karr et al. 1986, Gibson et al. 1996), has been acknowledged by scientific and regulatory agencies as an important component of natural resource protection (Schneider 1992).

A multimetric index of biological integrity (IBI, Karr et al. 1986), when calibrated according to the natural variation across a study region (Omernik 1987, Omernik and Griffith 1991), provides an objective approach for evaluating the ecological condition of waterbodies. Biological measures may exhibit variability (Karr and Chu 1999), however, assemblage-level indices more closely approximate actual biotic community composition (Buikema and Voshell 1993) than measures such as presence/absence of indicator species, single species toxicity tests, or estimates of population or abundance (Hughes et al. 1998).

Variously called rapid bioassessment protocols (RBP), the Invertebrate Community Index (ICI), the Benthic IBI (B-IBI), the Stream Condition Index (SCI) (among others), indices of biological integrity have been developed for many regions of North America (Barbour et al. 1999, Ohio EPA 1989, Kerans and Karr 1994, Barbour et al. 1996), and have been commonly used for assessing water resource quality (e.g., Karr 1991, Southerland and Stribling 1995, Gibson et al. 1996).). Geographically-calibrated, biological, multimetric indices for assessment of ecological conditions have been endorsed by the U.S. EPA (Gibson et al. 1996), the National Water Quality Monitoring Council (formerly, the Intergovernmental Task Force on Monitoring Water Quality) (ITFM 1995), and are currently used by over 42 states (Davis et al. 1996). The goal of the State of Mississippi is to use biological condition, physical habitat quality, and chemical conditions as

indicators of ecological health, and for ecological health to be the basis of evaluating water resource quality. Other states have found multimetric indices to be robust in detecting where there is a problem, and where more detailed, diagnostic testing is warranted (McCarron and Frydenborg 1997), such as water column and sediment toxicity and analytical chemistry.

For a multimetric index to function properly, least-disturbed conditions must be established as a baseline to which study stream conditions are compared. Least disturbed (a) (LDa) sites are considered those that are least degraded in a study region as defined by landscape, physical, and chemical characteristics (Hughes et al. 1986). The composite biological conditions found at a suite of LDa sites are the reference conditions to which study data are compared (Gibson et al. 1996, Barbour et al. 1996). The database of LDa sites and the analyses performed in developing and calibrating LDa conditions provide a systematic framework for assessing ecological impairment of streams.

There are essentially seven steps in developing a multimetric index, however, the steps are often iterative. Developing the database (step 1) involves selecting sites, field sampling, laboratory processing, structuring the data management system, entering data, quality assurance procedures and any other activities necessary for assembling the data so that they can be analyzed. Determining preliminary site classes (Step 2), is the process of delineating naturally variable regions according to abiotic data (e.g., physical and chemical data). Step 3 develops LDa and LDb site criteria that are stratified according to the geographic framework of the site classes. Selecting LDa and LDb sites in this way ensures that non-degraded waterbodies with naturally high or low levels of particular physical or chemical parameters are not excluded from the reference pool. Step 4 is calculation of metrics that describe components of benthic assemblages including richness, composition, trophic, habit, and tolerance. The fifth step involves geographic calibration of metrics and indices through development of bioregions. Multivariate and visually based statistics are used to evaluate the variability of biological assemblages found at LDa and LDb sites. For each naturally variable region, metrics are tested for stressor discrimination efficiency (step 6), scored (i.e., standardized), and assembled into bioregion-specific indices (step 7).

2. METHODS

The analytical framework used in site classification, final metric selection, biological index development, and development of scoring criteria follows that used in other states and regions (Barbour et al. 1996, Maxted et al. 1998, Stribling et al. 1998), while being calibrated to Mississippi's ecological potential and database.

The approach used in constructing an IBI follows seven basic steps:

- Develop database
- Determine preliminary regional site classes
- Establish numeric criteria for LDa and MD sites
- Compile and calculate candidate metrics
- Determine naturally occurring bioregional delineations
- Test metrics
- Combine metrics into index

2.1 Develop database

2.1.1 Develop QAPP

A comprehensive Quality Assurance Project Plan (QAPP) was developed and approved by USEPA Region 4 prior to sampling and analysis to ensure that data of sufficient quantity and quality were collected and assessed to allow the MDEQ to meet its needs (MDEQ 2001). The QAPP includes a general framework for the entire project and detailed standard operating procedures (SOPs) for all of the field sampling, laboratory data analysis, data entry, data management, and quality control (QC) activities. It follows the framework outlined in USEPA (1999).

2.1.2 Site Selection

A total of 463 nontidal stream locations distributed throughout the state except the Mississippi Alluvial Plain were visited over an 8-week span during the winter index period (January – March) (Figure 2-1; Appendix F). Most of these sites were sampled for benthic macroinvertebrates, physical habitat, and chemistry; in some cases certain data were not collected due to adverse conditions. Approximately 300 sites were from Mississippi's §303(d) list (Table 2-1). Two types of sites were selected specifically for purposes of developing the index: (1) potential LDa sites with a low percentage of managed* land use; and (2) sites located in areas containing known and potentially severe stressor sources (MD sites). Other sites were located in areas of more moderate stressor inputs. Some of the potential LDa and MD sites were also §303(d)-listed waterbodies. Efforts were made to locate a sufficient number of sampling sites in as many ecoregions and watersheds as possible to aid in describing the spatial variability of the biological, physical, and chemical characteristics of Mississippi's streams and watersheds.

*Managed land use is defined as altered landscape (agriculture, silviculture, mining, urban, residential, commercial, or industrial)

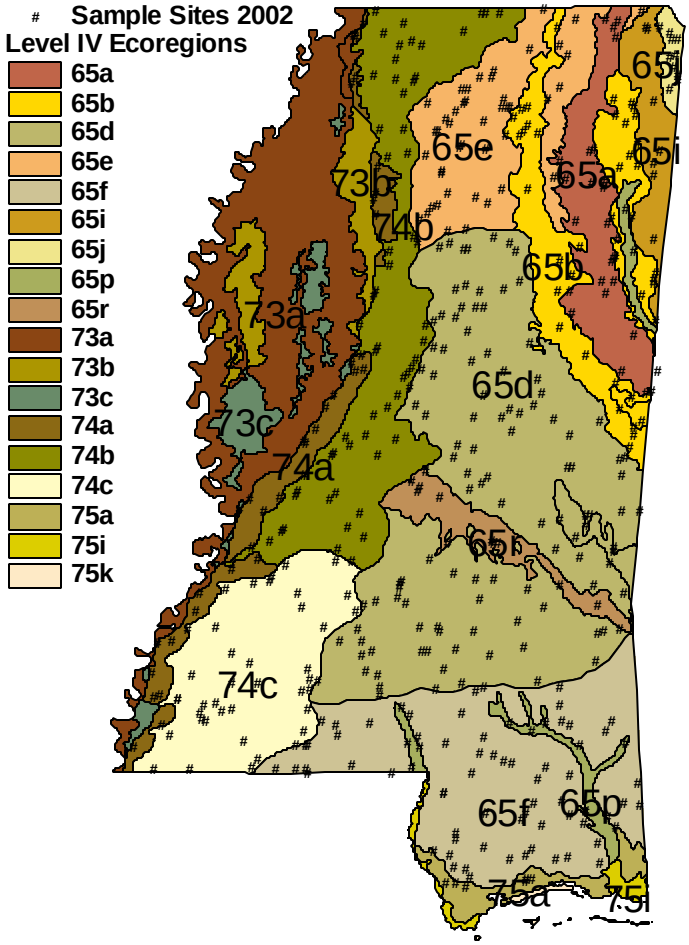


Figure 2-1. Level 4 ecoregions (Chapman et al. 2001 [draft]) overlain by sample sites.

2.1.3 Site Reconnaissance

Approximately 75 percent of the sites selected were visited by MDEQ staff prior to sampling to collect preliminary data on site locations, hazards, and potential sampling locations. The wadeability, representativeness, and accessibility of sites were noted. General physical habitat and surrounding environment was described; latitude/longitude coordinates were recorded using a Global Positioning System (GPS). Any extra equipment that would be needed by field teams was noted, as well as any other outstanding features at the site.

Table 2-1. Number of sites sampled and their use in different analyses.

	Number of Sites*	*Reason(s) for Sampling and Analysis*
All Sites	463	o listed on 1998 CWA 303(d) list as degraded; or o potential LDa sites; or o known MD sites
Potential LDa Sites	272	o land use/land cover of upstream drainage areas (GIS calculation of cover types using data from MS Land Cover Project)
Final LDa and LDb Sites	146	o sites meet specific, quantitative target levels of physical and chemical measurements (calculated from field-collected data)

*note that the number of potential LDa sites is a subset of all sites, and final LDa/LDb is a subset of the potential LDa sites

2.1.4 Sample Collection/Data Generation

2.1.4.1 *Benthic Macroinvertebrates*

For benthic macroinvertebrates, 525 samples were collected from 455 sites. To limit seasonal variability, all benthic macroinvertebrate sampling occurred within a restricted time frame (index period) of January–March. An additional, randomly chosen 70 samples were collected for QC purposes (MDEQ 2001).

Field sampling was completed in accordance with MDEQ-SOP-FLD-007, "Macroinvertebrate Collection in Low Gradient Glide/Pool Streams: Aquatic Dip Net - 20-Jab Method" (Appendix H [MDEQ 2001]). A list of equipment and expendable supplies used in the field is provided in Table 7 of MDEQ (2001). All samples were collected from multiple habitats using a D-frame net with 800×900 micron mesh net. Productive habitats including gravel/cobble, undercut banks and root material, snags/woody debris, and submerged aquatic vegetation, were sampled in the proportion in which they occurred (area-based) within the 100m reaches. Of the 20 total jabs used for the entire benthic collection process, 15 were proportionally-allocated to the above habitats. The other five jabs were allocated to sandy bottom substrate. If all of the 15 jabs allocated to productive habitats could not be used (i.e., these habitats were rare or absent) the remaining jabs were reallocated to the sandy bottom substrate habitats.

2.1.4.2 *Water Chemistry*

Instream chemical data (dissolved oxygen, pH, temperature, specific conductance, TDS, and turbidity) were collected from 453 sites using a multiprobe and turbidimeter in accordance with

MDEQ-SOP-FLD-004, "Operation of the Hydrolab DataSonde 4, YSI 6-Series Water Quality Multiprobe/Surveyor 4 and 610 Display Unit, and Hach Model 2100P Portable Turbidimeter" (Appendix F [MDEQ 2001]). Wet chemistry grab samples were collected from 459 sites using two 1-liter (squat quart) HDPE bottles according to MDEQ-SOP-FLD-005, "Water Quality Grab Sampling of Wadeable Streams and Shallow Surface Waters" (Appendix F [MDEQ 2001]) and analyzed in the lab for COD, TOC, TP, TKN, NH_3, nitrate/nitrite, total alkalinity, and total chlorides. Duplicate grab samples were collected at 48 randomly selected sites for a total of 507 grab samples.

2.1.4.3 *Physical Habitat and Hydrology*

Water surface elevation was obtained by lowering a plumb bob from the nearest bridge and recording the distance from the water surface to a reference point on the bridge (MDEQ-SOP-FLD-003, "Stream Stage Measurements (Tape Down Procedure)" - Appendix D [MDEQ 2001]). Physical habitat was evaluated at 463 sites using MDEQ-SOP-FLD-006, "Habitat Assessment for Low-Gradient Glide/Pool Streams" (Appendix G [MDEQ 2001]). Ten habitat parameters describing instream habitat, bank, and riparian conditions were visually assessed and rated on a scale from 0 to 20 with 0 being the poorest habitat and 20 being optimal. Habitat assessments were performed on the same 100-meter reach from which macroinvertebrate samples were collected. Care was taken to avoid disturbing the sampling habitat prior to macroinvertebrate sampling. The locations of the sites were recorded by sketching a map, recording the GPS coordinates, and taking at least one photograph of the location. In addition, habitat assessments were performed at 70 randomly chosen sites (same as biological QC sites). Inorganic substrate particle size distribution was assessed by performing a modified Wolman pebble count according to MDEQ-SOP-FLD-008, "Modified Wolman Pebble Count" (Appendix I [MDEQ 2001]).

2.1.4.4 *Landscape*

Drainage areas to the 463 sample sites were delineated with ESRI's ArcGIS 8.1 GIS using digital elevation models (DEM) and the National Hydrography Dataset (NHD). Land use/land cover (LULC) percentages within the drainage areas were calculated from the 1997 Mississippi Land Cover Project data (MDEQ 1997). LULC percentages within a variety of different-sized riparian corridors (50, 100, and 200 m wide; areas 1km upstream and whole drainage) were calculated. Site elevation and stream gradient data were also developed from DEMs.

2.1.5 Sample Processing

One of the chemistry grab samples was preserved using 5 mL of 5N* H_2SO_4 and both samples were chilled on ice immediately after collection through delivery to the lab (Appendix F [MDEQ 2001]). The preserved sample was analyzed for COD, TOC, and nutrients (TP, TKN, NH_3 and Nitrite + Nitrate). The unpreserved sample was analyzed for total alkalinity and total chlorides.

The benthic macroinvertebrate samples were field-preserved in 95 percent denatured ethanol with internal and external labeling (Appendix H [MDEQ 2001]). Methods of laboratory processing were based on Barbour et al. (1999). Biological laboratory sample processing

* Normal

involved two steps. The initial or primary sample processing step included sorting, sub-sampling, and sorting rechecks. Standardized 200-organism sub-samples were completed in the laboratory using a Caton gridded screen (MDEQ-SOP-LAB-001, "Laboratory Sorting and Sub-Sampling" [Appendix J [MDEQ 2001]). Sub-samples were shipped to the taxonomist using the procedures in MDEQ-SOP-LAB-002, "Macroinvertebrate Shipping" (Appendix K [MDEQ 2001]). The secondary or final phase processing included taxonomic identification and verification procedures, tabulation, and enumeration and is detailed in MDEQ-SOP-LAB-003, "Macroinvertebrate Taxonomy" (Appendix L [MDEQ 2001]). Identifications were primarly to genus level with selected taxa to species, family, or higher.

2.1.6 Data Entry

Biological, habitat, and water quality data were entered or loaded into EDAS (Ecological Data Application System, version 3.0 [Tetra Tech 2000]), which is on a Microsoft Access 97 platform, and has been customized for the MDEQ Biological Monitoring Program. Data, metadata, and other ancillary information reside in a series of relational tables including: stations, samples, benthic taxa, chemistry, habitat, and others. Laboratory analytical chemistry results were received from the MDEQ chemistry lab in electronic format (Excel spreadsheets) and were imported into EDAS. Locational, physical habitat, and ancillary watershed characterization data were entered directly from field datasheets. Biological data (taxonomic and enumeration results) were entered directly from handwritten datasheets. All data entered were compared directly with hand-written datasheets by someone who did not do the primary data entry for QC purposes.

2.1.7 Tolerance Value Development

Stressor tolerance values (TV) are ratings assigned to taxa intended to reflect their capacity to withstand adverse environmental changes (TVs are further defined in Section 2.4). Tolerance values (TVs) were developed for benthic macroinvertebrate taxa found as part of this project (Appendix A). A suite of stressor gradients was developed using principal components analysis (PCA) and represented various combinations of data from 32 physical, chemical, and landscape variables. The stressor gradient was selected that was most highly correlated with NMDS axis scores and index scores (tolerance metrics excluded). To confirm that the appropriate PCA axis was chosen as the stressor gradient, NMDS scores were regressed against different PCA axes. PCA axis 1 was most highly correlated with the NMDS axes that explained the greatest amount of variation in the biological data. This PCA axis was scaled so that relative taxa abundance values could be directly related to the stressor gradient to determine taxon-specific tolerance values. Reciprocal averaging was used to select tolerance values based on the point along the PCA axis where the highest relative abundances occurred. If taxa occurred at <15 sites in this dataset, they were assigned TVs using previously-documented values from MDEQ.

2.2 Determine Preliminary Site Classes

Detection of changes in the benthic macroinvertebrate assemblage due to anthropogenic stressors must occur independently of inherent differences due to natural factors. Therefore, natural variability in the physical and chemical site characteristics of the data were investigated before evaluating biological heterogeneity. The geographic framework for delineating regions of

relatively uniform natural features was Level 4 ecoregions (Figure 2-1; Table 2-2). Ecoregions are delineations of areas with similar climate, geology, soils, vegetation, topography, and hydrology (Omernik 1987), and have been accepted as a geographic framework for delineating regions of relatively homogeneous natural conditions (e.g., Barbour et al. 1996). Using Level 4 ecoregions as a framework, physical and chemical data, collected during this project allowed for further refinement of groupings called site classes.

Table 2-2. Ecoregions and subecoregions of Mississippi (Omernik 1987, Chapman et al. 2001).

Name	Numeric Designation
Blackland Prairie	65a
Flatwoods/Blackland Prairie Margins	65b
Southern Hilly Coastal Gulf Plain	65d
Northern Hilly Coastal Gulf Plain	65e
Southern Pine Plains and Hills	65f
Fall Line Hills	65i
Transition Hills	65j
Southeastern Floodplains and Low Terraces	65p
Jackson Prairie	65r
Mississippi Alluvial Plain	73
Bluff Hills	74a
Loess Plains	74b
Rolling Plains	74c
Gulf Coastal Flatwoods	75a

The first step in developing the preliminary site classes was to select potential LDa sites throughout the state based on the percentage of natural land use found within site drainage areas and riparian corridors. Land use/land cover (LULC) criteria were geographically-stratified so physically and chemically distinct areas would have LDa sites representative of a range of conditions. To be considered a potential LDa site, only one of the LULC target levels had to be met. LULC target levels were derived from professional judgement about responses of stream conditions to human influence. In areas where extensive landscape modification was predominant, the 75th percentile of natural riparian land use (50m wide, 1km long corridors) was used to specify criteria, or target levels. The variability of chemical parameters including conductivity, alkalinity, pH, nutrients, COD, TOC, and turbidity and physical parameters including total habitat scores, individual habitat scoring components, and substrate size

(developed from pebble count data) among these least disturbed sites was investigated using box and whisker plots, GIS analysis, and Principal Components Analysis (PCA). The relationship of elevation and stream gradient to possible variation in physical and chemical parameters was also investigated. The ecoregions were combined or segregated to form the preliminary site classes according to the variability observed among chemical and physical characteristics.

2.3 Select Least-Disturbed and Most-Disturbed Sites

Criteria for selecting LDa and MD sites were established for each preliminary site class. Thresholds were determined for the physical and chemical parameters so that sites could be categorized by status, i.e., as "LDa", "LDb", "other", and "MD", in order of increasing anthropogenic stress. The purpose of designating quantitative thresholds is to enhance the defensibility of site classes. The suite of potential LDa sites (selected in step 2 using LULC criteria) was refined using quantitative physical and chemical criteria stratified according to the preliminary site classes. The physical and chemical parameters that were used for selecting LDa sites were those which appeared, based on the preliminary classification, to show substantial variation across the state (i.e., those that dictated the delineation of the preliminary classes). This process is intended to identify sufficient number of LDa sites to be representative of least disturbed conditions within each site class.

For a site to be considered "LDa" or "LDb", it must meet all of the criteria. To classify for LDa status, sites in areas with extensive landscape modification only had to satisfy land use *or* habitat target levels as opposed to other regions of the state which had to meet both land use *and* habitat target levels. None of the sites used for site class delineation were specifically known to be impaired, i.e., the state had no previous monitored data indicating non-support of aquatic life use, though not all of these sites may have been previously monitored. LDb sites were selected to increase the number of sites to use for developing the final site classes (bioregions). Physical habitat was the only parameter for which LDb criteria were relaxed. LDa and LDb criteria for water chemistry, physical habitat, and LULC were developed as follows:

- Water chemistry

  LDa and LDb sites: 5^{th} or 95^{th} percentile value of potential LDa distribution + 90% confidence interval (CI) (only available for grab sample data, not for *in situ*; developed from precision estimates calculated from duplicate and repeat sampling – see Appendix B)

- Physical habitat

  LDa sites: 25^{th} percentile value of potential LDa distribution + 90% CI
  LDb sites: 25^{th} percentile value of potential LDa distribution (no CI)

- Land Use/Land Cover

  LDa and LDb sites: by proportion of land cover as forested, most frequently ≥ 60% or ≥ 70% (this is the same criterion used for preliminary site classification).

The criterion for physical habitat was more stringent than the chemical parameters because physical degradation of streams and watersheds is known to be predominant throughout the region.

Sites were classified as MD by satisfying *any* of the several criteria for that class; however, if land use criteria were exceeded, to be classified as MD a site also had to exceed a habitat criterion equal to the 25th percentile of the entire site distribution (without CI included). LULC data were approximately nine years old so requiring low habitat scores when LULC indicated degradation ensured that the MD sites were selected.

MD site LULC, physical, and chemical criteria were developed as follows:

- Water chemistry

  Established through basic knowledge of acceptable environmental levels

- Physical habitat

  25th percentile of the entire site distribution minus the 90 percent CI.
  25th percentile of the entire site distribution (without CI included) for sites that exceeded LULC criteria.

- Land use

  Range of highest percentages of disturbed land use/cover within drainage areas and riparian corridors representing most disturbed.

2.4 Compile and Calculate Candidate Metrics

Candidate metrics for testing and potential inclusion in the final biotic index were selected from previous studies throughout the U.S. (Gibson et al. 1996, Stribling et al. 1998, Barbour et al, 1996). Metrics, defined as

> *"calculated terms or enumerated values representing some aspects of biological assemblage structure, function, or other measurable characteristic that change in predictable ways with increased human influence"* (Fausch et al. 1990, Barbour et al. 1995, 1999, U. S. EPA 1997),

fall into six categories in the MDEQ dataset: taxonomic richness, composition, habit, tolerance/intolerance, feeding group and diversity. A total of 84 metrics within the six categories were calculated and considered for inclusion in the index. The general ecological meanings associated with each category are discussed below.

Taxonomic Richness. Metrics in this category are counts of the distinct number of taxa within selected taxonomic groups. High taxa richness usually correlates with increasing health of the assemblage and suggests that niche space, habitat, and food sources are adequate to

support survival and propagation of many species. Metrics in this category may be focused on overall taxa richness (e.g., total taxa) or richness within particular groups (e.g., EPT taxa, Insect taxa, Chironomidae taxa).

Composition. These metrics are based on the proportion of individuals in a sample belonging to a specified taxonomic group. Expressed as percentages, these metrics reveal the relative abundance of different groups of benthic macroinvertebrates, each of which may respond differently to environmental conditions and community dynamics.

Tolerance/Intolerance. Tolerance of a taxon is based on its ability to survive short- and long-term exposure to physicochemical stressors that result from chemical pollution, hydrologic alteration, or habitat degradation. Tolerance metrics characterize the relative sensitivity of the assemblage to perturbation by measuring numbers of pollution tolerant and intolerant taxa or percent composition. Different taxa are assigned tolerance values that are incorporated into indices such as the Hilsenhoff Biotic Index (HBI) or the North Carolina Biotic Index (NCBI) or metrics such as % intolerant organisms. Tolerance values developed as part of this project were used for calculating tolerance metrics.

Feeding Group. The functional feeding group designation for an organism reflects the dominant mode of feeding, not the specific nutritional source or benefits (Cummins and Klug 1979, Anderson and Cargill 1987, Merritt and Cummins 1996, Wallace and Webster 1996). Designations for each taxon include scrapers, predators, collector-gatherers, collector-filterers, shredders, and others. Specialized feeders, such as scrapers, are more sensitive organisms and are thought to be well represented in healthy streams. Generalists, such as collectors, have a broader range of acceptable food materials than specialists (Cummins and Klug 1979), and thus are more tolerant to pollution which may alter food sources.

Habit. These metrics describe morphological adaptations for maintaining position and moving about in the aquatic environment (Merrit and Cummins 1996). Habit categories include movement and positioning mechanisms such as swimmers, clingers, sprawlers, climbers, and burrowers.

Diversity. These metrics measure the relative representation of each taxon (or evenness) as a percentage of the most common taxa. Low evenness or high percent dominance by few taxa is an indication that environmental conditions favor a limited type of organism, which suggests the presence of stressors.

2.5 Develop Bioregional Delineations

Before human-induced changes in biological assemblages can be detected, the natural variation among assemblages must be understood. Variability in the macroinvertebrate assemblage may result from natural variability in the physical and chemical site characteristics across a geographic range. Much of the natural variability can be accounted for by dividing the area into ecological regions such as the preliminary site classes developed in step 2 and level 3 and 4

ecoregions (Omernik 1987). To calibrate the final index, however, it is necessary to assess natural biological variability, which does not necessarily coincide with abiotic variations.

The goal of any classification scheme is to form groupings that minimize within-group variability and maximize among-group variability. Two primary techniques, ordination and comparison of metric distributions, were used to justify separating or combining data from preliminary site classes and ecoregions into regions of relative biological homogeneity (bioregions). To minimize human-influenced biological variability, only LDa and LDb sites selected in step 3 were used to develop bioregions.

Alternative classification schemes were examined with multivariate ordination of the LDa and LDb sites based on their species composition, following methods outlined in Jongman et al. (1987) and Ludwig and Reynolds (1988). Ordination is a category of methods for reducing the dimensionality of multivariate information (many species in many sites) by placing sites or species in an order. The ordination method that we used, non-metric multidimensional scaling (NMDS), arranges sites along axes so points close together correspond to sites with similar taxonomic composition and abundance and points farthest apart are most dissimilar (Jongman et al. 1987). This approach is more robust in producing separation of classes than other ordination methods (e.g., Kenkel and Orloci 1986, Reynoldson et al. 1997). The most widely used technique is based on an ordination algorithm that produces dimensions explaining variation in the data, with the first explaining the most, continuing with the second in descending amounts of explained variation (Kruskal 1964; Kenkel and Orloci 1986). Values are plotted as two- or three-dimensional graphs depending on the perspectives that best illustrate site classes or similarity groupings. For this analysis, the Bray-Curtis percent dissimilarity coefficient was used:

$$BC = \left({}^{2W}\!\!\diagup\!\!{}_{(A+B)} \right)$$

where W is the sum of common taxa abundances and A and B are the sums of taxa abundances in individual sample units. A pair of samples with identical taxa abundances would have a coefficient of 0 and a pair of samples with no taxa in common would have a coefficient of 1. This ordination method has been shown to be robust for ordination of species composition (e.g., Kenkel and Orloci 1986, Ludwig and Reynolds 1988), and has been used successfully for classification of stream communities (e.g., Barbour et al.1996; Reynoldson et al. 1995; Stribling et al. 1998).

The site-by-site matrix of Bray-Curtis dissimilarity coefficients was used in the NMDS ordinations (McCune and Mefford 1995, Kruskal 1964). An acceptable ordination should have a stress coefficient (measuring the goodness-of-fit of the ordination to the original data) of less than 20%. Stress is lowered as additional dimensions are allowed in the ordination and three axes are commonly required. The final NMDS configuration was plotted (as a scatterplot in two dimensions) to identify groupings of sites with similar taxa composition (low Bray-Curtis dissimilarity). When plotted points are labeled by site characteristics (e.g., preliminary site classes or ecoregions) the association between taxa composition and site characteristics can be

visualized. Preliminary site classes or ecoregion groupings that overlap in the ordination plots could be combined into bioregions for subsequent analysis.

The second technique used to evaluate potential bioregions was assessment of box and whisker diagrams of metric distributions from LDa and LDb sites (Figure 2-2). Similar distributions of metrics (medians, inter-quartile ranges and overall ranges) between ecoregions indicate similar biotic assemblages and justify aggregation of ecoregions into a single bioregion. Likewise, substantial differences in distributions suggest distinct bioregions.

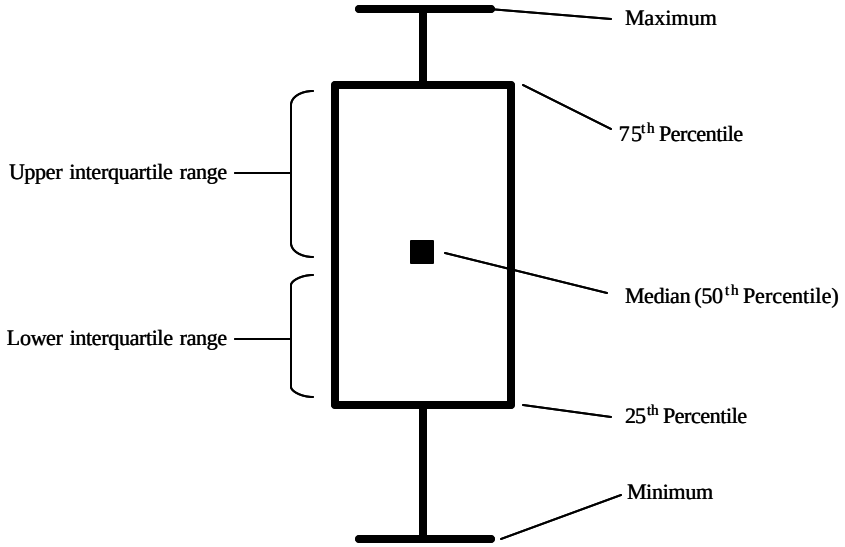


Figure 2-2. Example of box-and-whisker plot and its components

2.6 Test Metrics

The ability of metrics to detect impairment was assessed in two ways. Box-and-whisker plots were used to visually assess the ability of the metrics to distinguish between LDa and MD sites. This type of plot displays the median (central point), maximum and minimum values (whiskers), and 25^{th} and 75^{th} percentiles (box) of the LDa and MD site population. Decisions regarding the distinction of the populations of sites were made based on the degree of similarity between LDa and MD site distributions.

Discrimination efficiencies (DE) were used to quantitatively assess the ability of metrics to detect impairment. The DE is a numerical description of the degree of separation between metric value distributions of LDa and MD sites and is calculated as:

$$DE = 100 \times \frac{a}{b}$$

where a = the number of MD samples scoring below the 25th percentile of the LDa distribution and b = the total number of MD samples. A higher DE indicates better performance of a metric, or a better ability to distinguish between LDa and MD conditions. Most metrics decrease as stress increases; in these cases the 25th percentile of the LDa distribution (as described above) is used as the threshold. However, for metrics that increase with stress (e.g., HBI, % tolerant taxa), MD sites are classified correctly if the value is above the 75th percentile of the LDa distribution.

Within each bioregion, DEs were calculated for all metrics that show a clear response to stressors; those metrics that had unintelligible differences between distributions of LDa and MD sites in any of the bioregions were not considered as viable candidates for inclusion in the index and were therefore dropped from further analysis. Those metrics that responded to stress in opposite directions between bioregions were also dropped from the analysis. Therefore, metrics were not used in index formulation for several reasons:

1) obscure ecological meaning
2) lack of, or only weak, response to presence of stressors
3) irrelevance to ecosystems being studied
4) redundancy to other metrics being used (see step 7)

Exclusion of metrics occurred during different portions of the development process, particularly Steps 6 and 7.

2.7 Combine Metrics into Candidate Indices

A multimetric index is a simple additive approach for combining metric value information from different types of biological metrics into a single numeric assessment value. The process begins with metric scoring, then with averaging of the best performing (highest DEs) and most meaningful metrics.

2.7.1 Metric Scoring

To combine metrics into an index, metric values were standardized (i.e., scored) on a 100-point scale. The metric scoring strategy that was used in this analysis rated the metric values on a percentage scale from the least possible metric value to the highest observed metric value. For those metrics that decrease in value with stress (e.g., Tanytarsini taxa, EPT taxa, Beck's Biotic Index), the 95th percentile of the entire site distribution was considered the best value (i.e., the standard) to eliminate unusual outliers and avoid skewing the resultant scores. Metric values greater than or equal to this standard were given a score of 100, while those values less than the standard were scored as a percentage of the standard as follows:

$$score = \left(\frac{x}{x_{95} - x_{\min}} \right) \times 100 \quad for\ x \le x_{95}$$

where x is the metric value; x_{95} is the 95th percentile of the entire site distribution; and $x_{\min}$ is the minimum possible value (usually 0).

For those metrics that increase with stress (e.g., HBI, % Caenidae), the 5th percentile of the entire site distribution was used as the standard. All values less than or equal to this standard were given a score of 100. Values greater than the standard were scored as the percentage of the range from the maximum (worst) value to the 5th percentile (best) value:

$$score = \left(\frac{x_{\max} - x}{x_{\max} - x_5} \right) \times 100 \ \ for \ x \geq x_5$$

where x_5 is the 5th percentile value; and $x_{\max}$ is the maximum possible value (e.g., 100% for percentage metrics; 10 for HBI). For richness metrics the maximum observed value was used.

2.7.2 Index Selection

To avoid redundant information in the index, correlation analysis (Pearson Product Moment) was performed on all metrics. Those metrics with a correlation coefficient > 0.9 were considered redundant and were not used together in any index formulation. Metrics with correlation coefficients > 0.8 were used together only when absolutely necessary, for example, when no other metrics were available in a particular category.

Several test index formulations were made from suites of the best-performing metrics in each bioregion and from as many metric categories as practical. The index was calculated as an average of the proposed metric scores and a DE for the index was calculated as it was for each individual metric. Box and whisker plots of index scores for LDa and MD sites were also used to evaluate index performance. Configurations included metrics from six metric categories (taxonomic richness, composition, habit, feeding group, diversity and tolerance). Separate indices were developed for each of the five bioregions. Index configurations that had the highest DEs were chosen as final indices. When potential indices had the same DEs, separation of interquartile ranges, the presence of commonly used metrics, and the robustness of the configuration (i.e., the number of metrics) were used to decide on the final index configuration. Furthermore, metrics within index configurations were assessed with regard to whether the difference in LDa and MD metric values was ecologically meaningful (e.g., a difference of one taxon for a richness metric may not be important).

Precision of the five final indices was evaluated using the repeat and duplicate sample data. Precision estimates including root mean square error (RMSE), coefficient of variability (CV), and detectable differences (90% confidence intervals) were calculated for each index (Appendix B). Precision values of index scores for duplicate and repeat samples were similar to one another, therefore, these data were combined to derive an overall precision estimate for all replicated samples.

3. RESULTS

3.1 Database

All landscape, LULC, physical habitat, chemistry, and biological data assembled for this study are housed in the Mississippi EDAS (Microsoft Access 97) and are presented in Appendix F. Table 3-1 summarizes land use/land cover percentages for the five different bioregions by aggregated categories of land use. Out of 562 total taxa (Appendix C), new tolerance values were derived for 324 taxa. Another 149 taxa for which PCA-based tolerance values could not be developed (due to low numbers of organisms) were assigned tolerance values from previous lists (Appendix A).

3.2 Preliminary Site Classes

Using quantitative drainage area and riparian land use data calculated from GIS land use coverages (MDEQ 1997), 272 potential LDa sites were selected throughout the state (Table 2-1; Table 3-2). Based on box and whisker plots, PCA, and GIS analysis of a preliminary suite of potential LDa sites, the state was divided into six preliminary site classes. Upon selection of the potential LDa sites (i.e., the 272 described previously) the state was divided into nine preliminary site classes, excluding the Alluvial Plain. Chemical and physical parameters important to the class delineation included ammonia, chemical oxygen demand, chlorides, nitrate-nitrite, pH, specific conductance, TKN, TOC, TP, total habitat score, instream habitat, morphological habitat and average slope. The PCA loadings presented in Table 3-3 describe the variables that weighed most heavily on the PCA axis scores used for developing the preliminary site classes (Figure 3-1). A tenth preliminary site class was created from the northern part of site class 3 (Figure 3-2). Most of the class boundaries coincided with ecoregional boundaries; however, in several cases class boundaries cut through ecoregions or divided level 3 ecoregions along level 4 ecoregional lines (Figure 3-2, Table 3-4).

3.3 Criteria for Selecting Least -Disturbed and Most-Disturbed Sites

From the initial list of 463 sites, using the land use, physical and chemical target levels 83 LDa sites (Figure 3-3) and 63 LDb sites were selected for a total of 146 final LDa and LDb sites (Table 2-1; Table 3-2; Appendix F). A total of 72 MD sites were selected from the 10 preliminary site classes (Table 3-5; Figure 3-3; and Appendix F).

3.4 Candidate Metrics

A total of 84 metrics in six metric categories were calculated (Appendix F). Metrics were calculated using the lowest taxonomic level, usually genus. Metrics were also calculated using species level data (for those taxa identified to this level) but were not statistically different from genus-level metrics. Composition metrics were the largest category (N=31) and habit, trophic, and diversity metrics were the smallest groups (N=10, 10, and 1, respectively).

Table 3-1. Summary LULC for drainage areas and riparian corridors for five bioregions.

Bioregion	Land Use Category*	Complete Drainage Area	All Channels 100m Wide Corridors	1km Upstream Only 100m Wide Corridors	1km Upstream Only 50m Wide Corridors
East (n=205)	Forest	23.3	21.4	12.2	13.2
	Wetland	1.0	1.7	8.7	9.8
	Urban	3.0	1.8	10.1	9.4
	Agriculture	55.6	58.8	52.8	50.4
	Miscellaneous	16.9	19.4	16.1	17.2
Northeast (n=37)	Forest	52.7	59.0	51.8	53.6
	Wetland	8	7.3	20.8	21.5
	Urban	1	0.7	1.2	1.0
	Agriculture	35.9	15.9	22.2	9.7
	Miscellaneous	1.5	17.1	14.0	14.2
Northwest (n=91)	Forest	37.9	29.7	19.1	19.3
	Wetland	1.1	2.6	7.2	7.3
	Urban	1.0	0.5	0.4	0.4
	Agriculture	47.8	51.2	61.5	60.6
	Miscellaneous	12.2	16.0	11.9	12.4
	Forest	33.6	28.1	13.9	14.1
	Wetland	0.5	1.2	3.0	3.1
	Urban	1.0	0.5	0.5	0.5
	Agriculture	49.6	53.9	70.5	69.6
	Miscellaneous	15.3	16.3	12.0	12.7
West (n=96)	Forest	51.0	53.0	42.7	43.5
	Wetland	1.0	2.9	10.3	11.9
	Urban	1.5	1.0	0.9	0.8
	Agriculture	29.9	26.4	30.7	26.9
	Miscellaneous	16.7	16.7	15.5	17.0

*Forest and wetland are considered "natural" uses; urban and agricultural considered "managed". "Miscellaneous" constitutes small percentages of a variety of land uses.

3-2. Land use, physical, and chemical criteria used to select LDa and LDb sites for the 10 preliminary site classes.
= land use and land cover.

LULC Criteria					Physical and Chemical Criteria															Number of Sites			
al LULC		High Density																					
Riparian (100 m wide, 1 km long)	Riparian (50 m wide, 1 km long)	Drainage Area	Riparian (100 m wide, variable drainage long)	Riparian (100 m wide, 1 km long)	Total Habitat Score (>=) - Reference	Total Habitat Score (>=) - SubReference	Ammonia (mg/L) (<=)	Chemical Oxygen Demand (mg/L) (<=)	Chlorides (mg/L) (<=)	Dissolved Oxygen (mg/L) - LDa	Dissolved Oxygen (mg/L) - LDb	pH (<=)	pH (>=)	Specific Conductance (uS/cm) (<=)	Alkalinity (mg/L) (<=)	Total Kjeldahl Nitrogen (mg/L) (<=)	Total Organic Carbon (mg/L) (<=)	Total Phosphorus (mg/L) (<=)	NPDES proximity (km)	Reference sites	Subreference. Sites	Ref + Subref. Sites	LDa Sites / LDb Sites / LDa + LDb Sites
>70%		<3	<3	<3	157	133	0.31	22.45	5.08	4.00	0.45	7.30	5.49	88.00	23.52	0.91	6.33	0.10	>5	7	5	12	7.3
	>50%*	<3	<3	<3	101*	101*	0.26	39.45	17.08	4.00	0.79	7.73	7.47	386.00	141.82	1.29	8.33	0.12	>5	3	0	3	0.4
>70%		<3	<3	<3	146	131	0.37	24.45	7.18	4.00	0.28	5.30	7.00	179.00	23.12	1.22	8.33	0.13	>5	15	9	24	12.4
	>17%*	<3	<3	<3	106*	106*	0.40	24.45	9.08	4.00	1.96	6.90	6.20	102.00	31.22	1.76	6.33	0.34	>5	4	0	4	
	>67%*	<3	<3	<3	133	109*	1.53	60.45	13.78	4.00	0.97	8.12	5.59	372.00	160.82	4.10	18.33	0.55	>5	1	5	6	14.1
>70%		<3	<3	<3	121	97	0.30	20.45	481.48	4.00	0.47	8.28	6.64	1942.00	375.82	1.03	5.33	0.22	>5	4	7	11	3.1
>70%		<3	<3	<3	136	112	0.37	45.45	36.48	4.00	0.59	7.39	5.84	163.90	21.82	1.20	12.33	0.19	>5	21	8	29	0.5
>70%		<3	<3	<3	144	120	0.44	48.45	50.68	4.00	1.55	7.52	5.34	246.00	43.82	1.83	14.33	0.53	>5	12	13	25	69.6
>70%		<3	<3	<3	151	127	0.32	48.45	11.58	4.00	0.33	6.43	4.38	124.00	11.82	1.34	17.33	0.11	>5	11	12	23	12.7
>46	>53	<3	<3	<3	119	95	0.33	20.45	5.48	4.00	0.36	5.82	7.25	145.00	24.92	0.80	5.33	0.15	>5	5	4	9	
																			TOTALS	83	63	146	

in highly modified areas of the landscape, therefore, criteria had to be relaxed so that reference sites could be selected. For these areas the LULC
t level had to be met as opposed to the other classes where both the LULC *AND* the habitat levels had to be met to be considered a reference site.

n 65e was initially considered a modified area and LULC criteria were relaxed to include potential reference sites. However, upon investigation of
cal characteristics this ecoregion grouped with others that were not highly modified and, therefore, it was grouped in the same class as these.

Table 3-3. Principal Components Analysis loadings on first two axes used for developing preliminary site classes.

Parameter	Factor 1	Factor 2
Total Habitat Score	-0.81	0.13
Instream Habitat Score	-0.65	0.16
Morphological Habitat Score	-0.73	-0.05
Drainage Area ($km^{2)}$	-0.03	-0.17
Average Slope	0.11	0.68
Elevation (m)	-0.16	-0.17
pH	0.69	0.18
Log Ammonia (mg/L)	0.19	-0.49
Log Chlorides (mg/L)	0.60	0.02
Log Dissolved Oxygen (mg/L)	0.31	0.32
Log Nitrate/Nitrite (mg/L)	0.27	-0.29
Log Specific Conductance (uS/cm)	0.76	0.03
Log Total Kjieldahl Nitrogen (mg/L)	0.03	-0.82
Log Total Organic Carbon (mg/L)	-0.13	-0.73
Log Total Phosphorus (mg/L)	0.49	-0.45
ArcSin Square Root Silt/Clay	0.12	-0.51
ArcSin Square Root Sand	-0.18	0.32
ArcSin Square Root Gravel	0.10	0.43

3.5 Bioregional Delineations

Non-metric multidimensional scaling (NMDS) of taxonomic composition data from LDa and LDb sites suggested two bioregions roughly representing the western and eastern halves of the state (Figures 3-4 and 3-5). Metric value distribution across the state also suggested two bioregions (Appendix E). The exception to the east/west division was level 4 ecoregion 65a (Blackland Prairie) which was biologically more similar to the western section of the state than the east. The bioregional delineations followed preliminary class boundaries with the exception of class 7, which was divided between the two bioregions along the level 3 ecoregional boundary (Figure 3-4).

However, due to unique landscape characteristics in several areas of the state, the initial two bioregions were re-organized into five (Figure 3-6). The Northwest bioregion, is made up of preliminary site classes 4 and 10. Although classes 4 and 10 were initially in two different bioregions, when compared directly to one another they were not substantially different (Figure 3-7). Physiographic uniqueness suggested utility to maintaining the Northeast (preliminary class 1) and Black Belt (preliminary class 2) as distinct bioregions. The low number of LDa and LDb sites in these areas (n = 12 and 3, respectively) may have prevented being able to distinguish any biological differences from other areas. Field experience, as well as physical and chemical variability of the areas suggest that biological differences probably exist, therefore, these areas were delineated as distinct bioregions. For purposes of site assessment it was deemed better to compare study sites to LDa conditions in these particular regions rather than LDa conditions from the larger bioregions to which the Northeast and Black Belt bioregions initially belonged. The southern portions of the initial two bioregions were reorganized as distinct bioregions. As

expected, NMDS ordinations show overlap of the five bioregions (Figure 3-8). Table 3-4 shows the nesting of the Level 4 ecoregions within the preliminary site classes and final bioregions. Bioregional boundaries coincide with Level 3 and 4 ecoregional boundaries with a few exceptions.

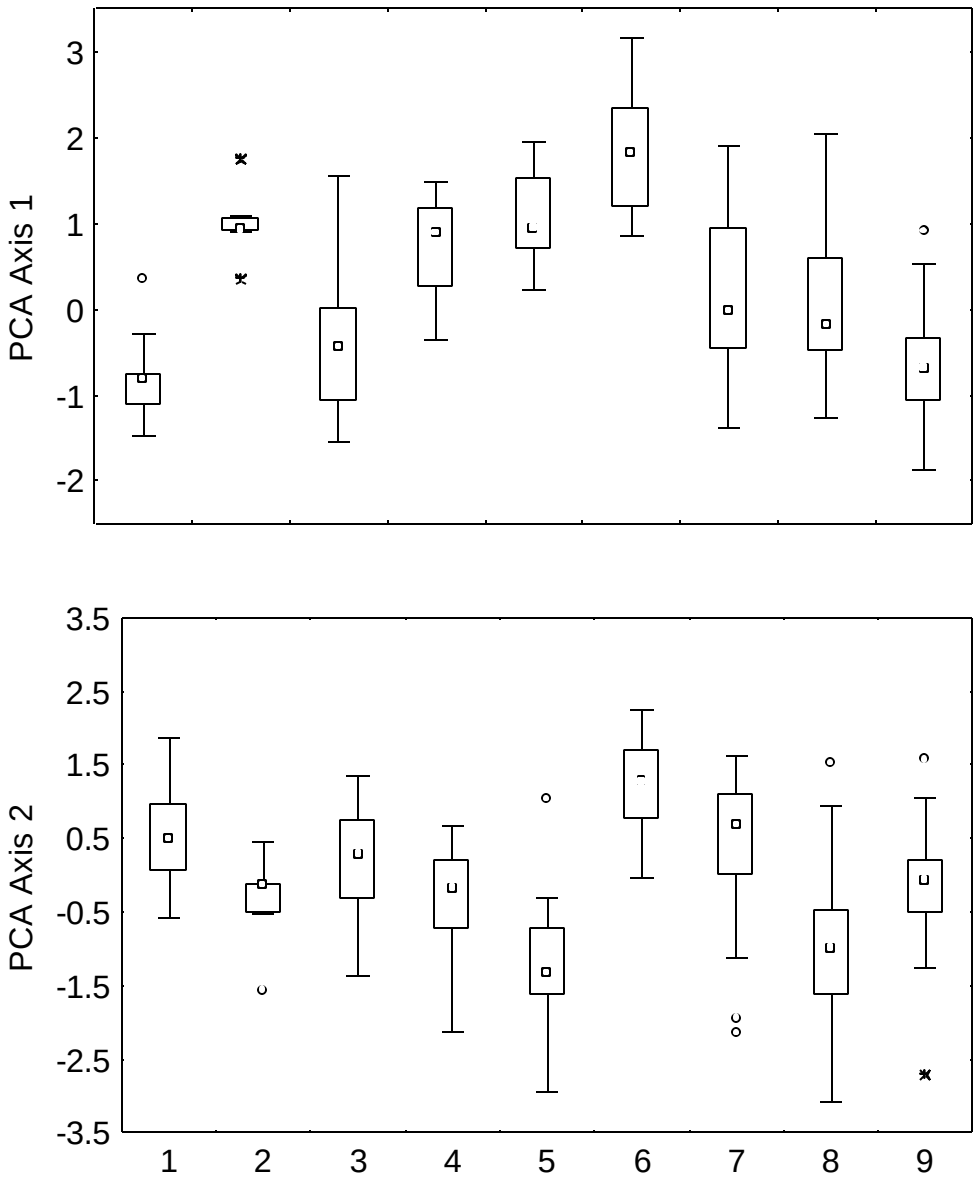


Figure 3-1. Distribution of PCA axes 1 and 2 among nine preliminary site classes. One additional class was added upon further analysis.

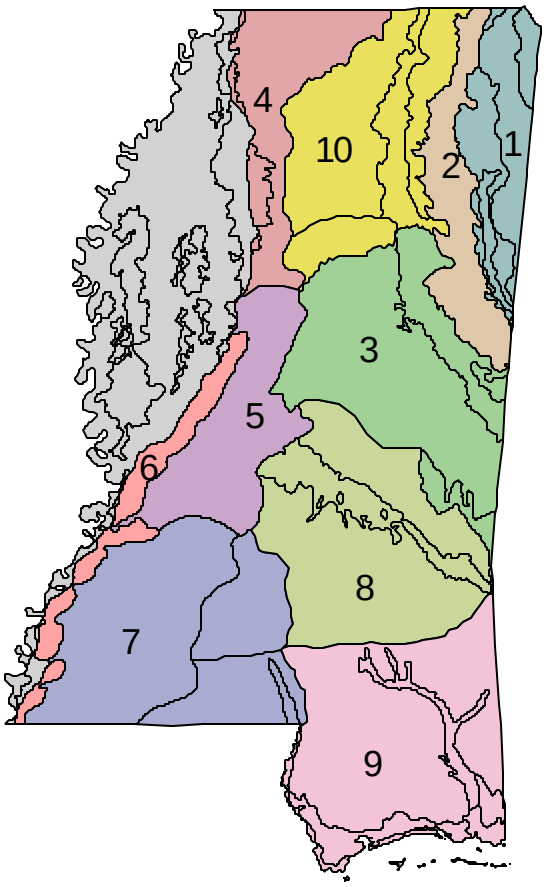


Figure 3-2. Map of 10 preliminary site classes developed based on patterns in physical and chemical data from potential LDa sites.

3.6 Metric Performance

Metric response to stressors in the five bioregions varied as represented by discrimination efficiencies (DE). Metrics were least efficient in the West, where the highest DE was 75 percent (Hydropsychidae/Trichoptera) and most efficient in the Black Belt where 14 metrics had DEs of 100 percent (Table 3-6). Overall, tolerance metrics performed the best (based on the number of metrics with high DEs) (Table 3-6). Other metrics that had consistently high DEs for most of the bioregions were the composition metrics: % Caenidae, % Ephemeroptera (no Caenidae) and % EPT (no Caenidae); the richness metrics: Tanytarsini taxa, Insect taxa, Chironomidae taxa and Total taxa; the habit metrics: % Clinger, Clinger taxa, and Sprawler taxa; and the trophic metrics: Filterer taxa, Collector taxa, and Predator taxa (Table 3-6; Appendix E).

Redundancy was tested among those metrics with the highest DEs and those with r-value > 0.80 were excluded from indices (Appendix F). Clinger metrics (i.e., Clinger taxa and % Clinger) and % Caenidae were often redundant with tolerance metrics (r >0.80). EPT metrics were often redundant (r >0.80) with individual E, P, or T metrics.

Table 3-4. Relationship of bioregions to preliminary site classes, ecoregions, and sample sites.

Bioregion	Preliminary Site Class	Level 4 Ecoregion	Number of Sites
Black Belt	2	65a	26
East	3	65b*	205
		65d*	
	7	65d*	
		65f	
		65p*	
	8	65d	
		65r	
	9	65f	
		75a	
Northwest	4	74a*	91
		74b*	
	10	65b*	
		65e	
Northeast	1	65b*	37
		65e	
		65i	
		65j	
		65p*	
West	5	74b*	96
	6	74a*	
	7	74c	

* Indicates that Level 4 ecoregion is split between either site classes or bioregions

Table 3-5. Land use/land cover (LULC), physical, and chemical criteria used to select MD sites for the 10 preliminary site classes.

Preliminary Site Class	LULC Criteria							Physical and Chemical Criteria							No. of sites
	Managed LULC				High Density Urban										
	Drainage Area	Riparian (100 m wide, whole drainage long)	Riparian (100 m wide, 1 km long)	Riparian (50 m wide, 1 km long)	Drainage Area	Riparian (100 m wide, whole drainage long)	Riparian (100 m wide, 1 km long)	Total Habitat Score (<=)	Total Habitat Score (<=)*	Ammonia (mg/L) (>=)	Nitrate/Nitrite (mg/L) (>=)	Total Kjieldahl Nitrogen (mg/L) (>=)	Total Phosphorus (mg/L) (>=)	Dissolved Oxygen (mg/L) (<)	
1	85	75	75		10	10	10	85	133	5	10	5	1	4	9
2	90	90	90		10	10	10	43	104	5	10	5	1	4	4
3	75	65	65		10	10	10	67	91	5	10	5	1	4	8
4	90	90		90	10	10	10	62	86	5	10	5	1	4	6
5	75	65		65	10	10	10	59	83	5	10	5	1	4	8
6	60	50	50		10	10	10	63	87	5	10	5	1	4	5
7	60	50	50		10	10	10	86	110	5	10	5	1	4	9
8	60	50	50		10	10	10	96	120	5	10	5	1	4	6
9	60	50	50		10	10	10	102	126	5	10	5	1	4	3
10	75	65	65		10	10	10	57	81	5	10	5	1	4	14
														TOTAL	72

*When a site exceeded one of the managed LULC criteria it had to also have a total habitat score lower than TOTHAB2 (25th percentile of the entire distribution [CI not used]) to be considered degraded

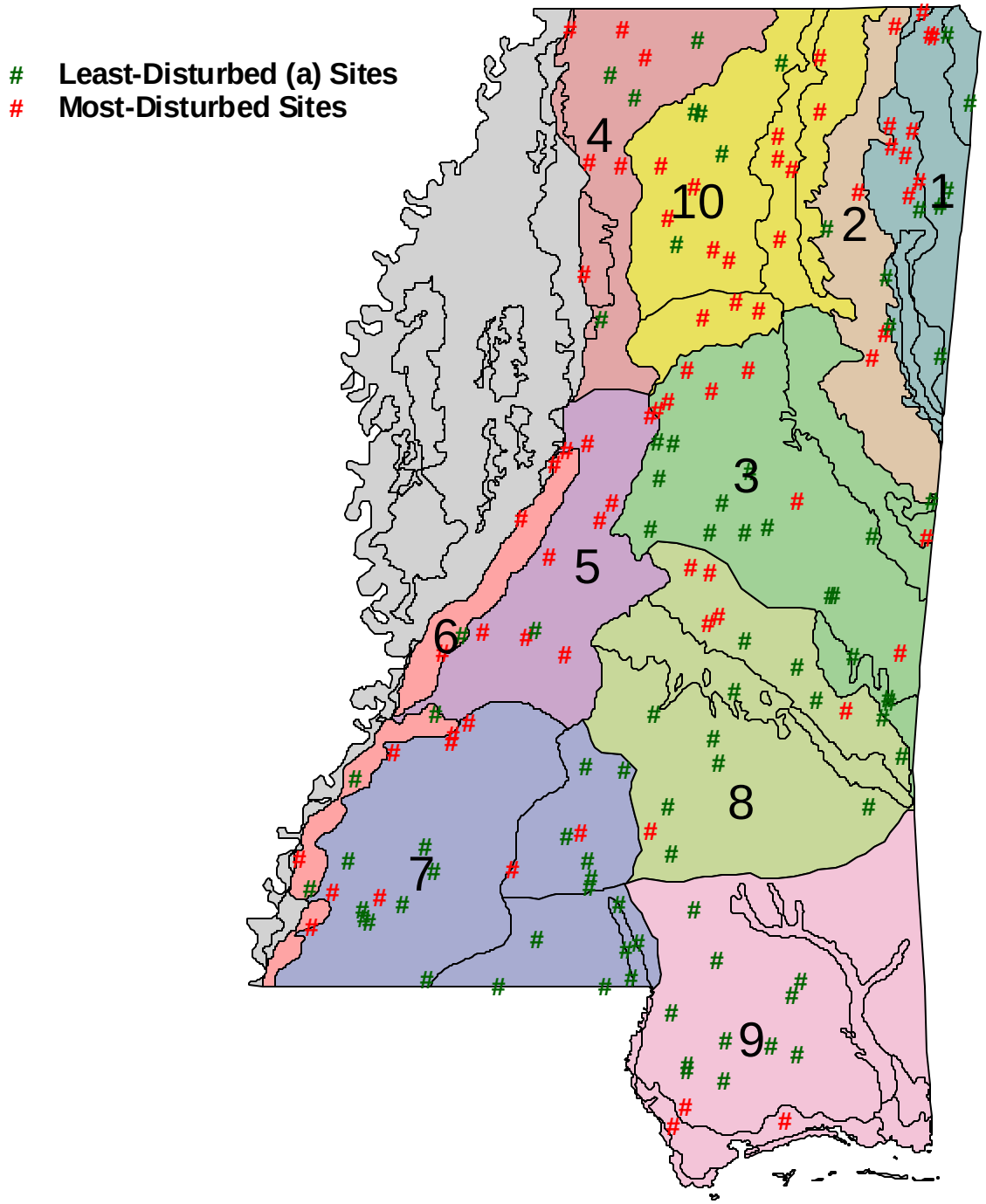


Figure 3-3. LDa and MD sites overlain on the 10 preliminary site classes.

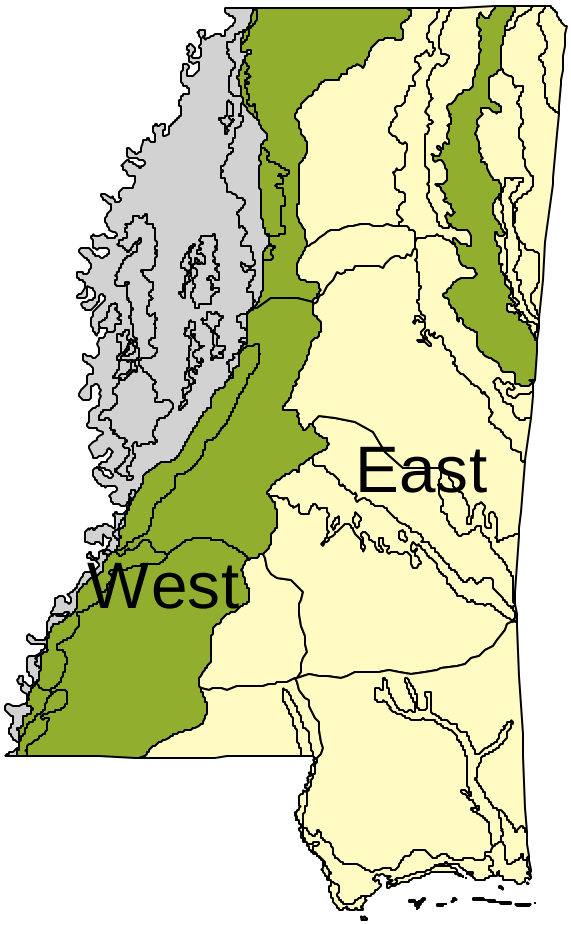


Figure 3-4. Map of initial division of state into two bioregions based on NMDS and box and whisker analyses.

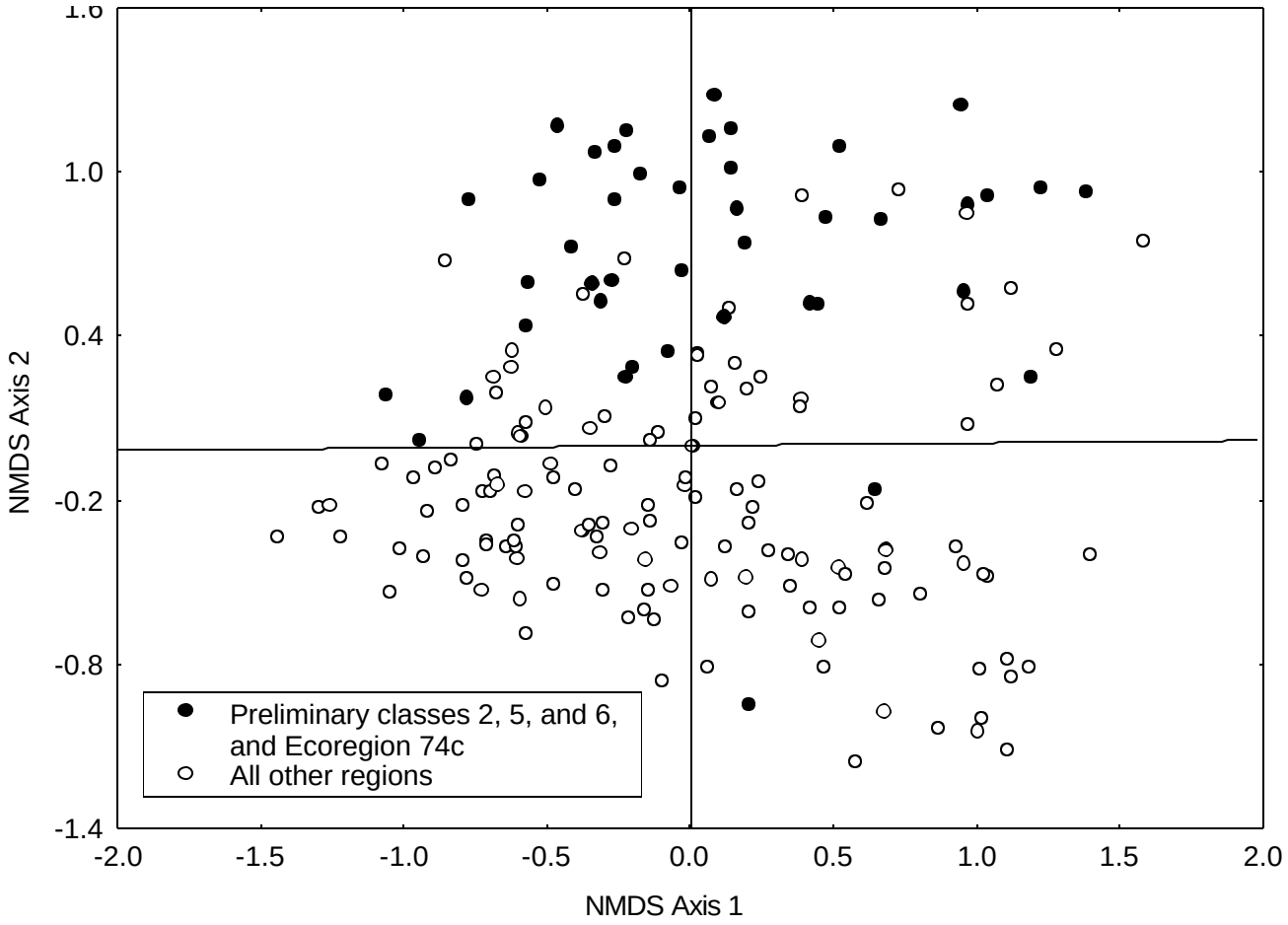


Figure 3-5 NMDS plot of axes 1 and 2 showing first grouping of sites into two bioregions across the state

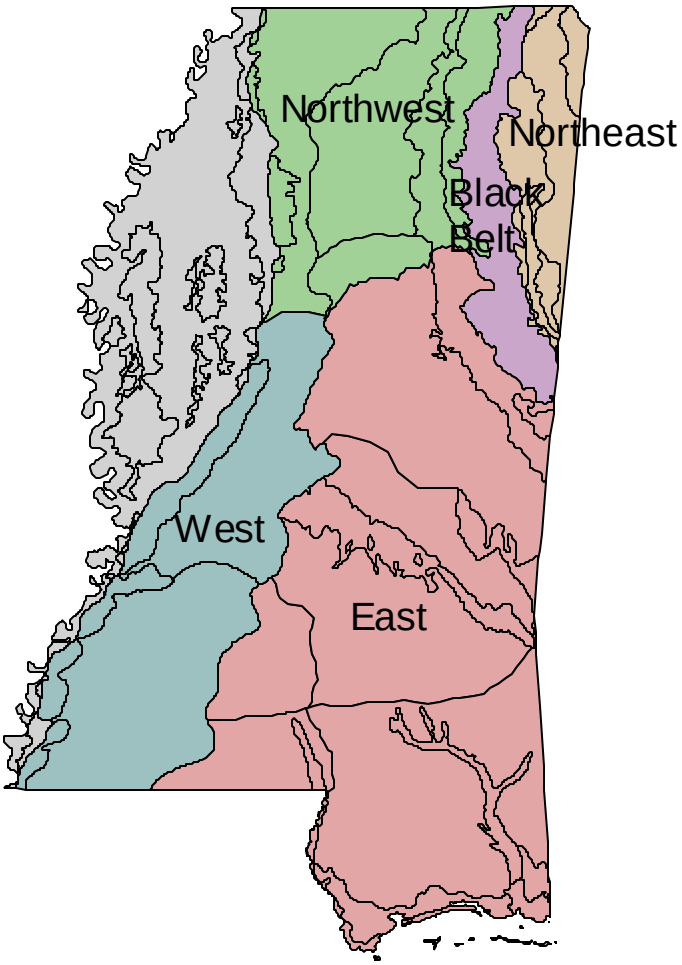


Figure 3-6. Final bioregional delineation developed based on NMDS ordination and box and whisker analyses.

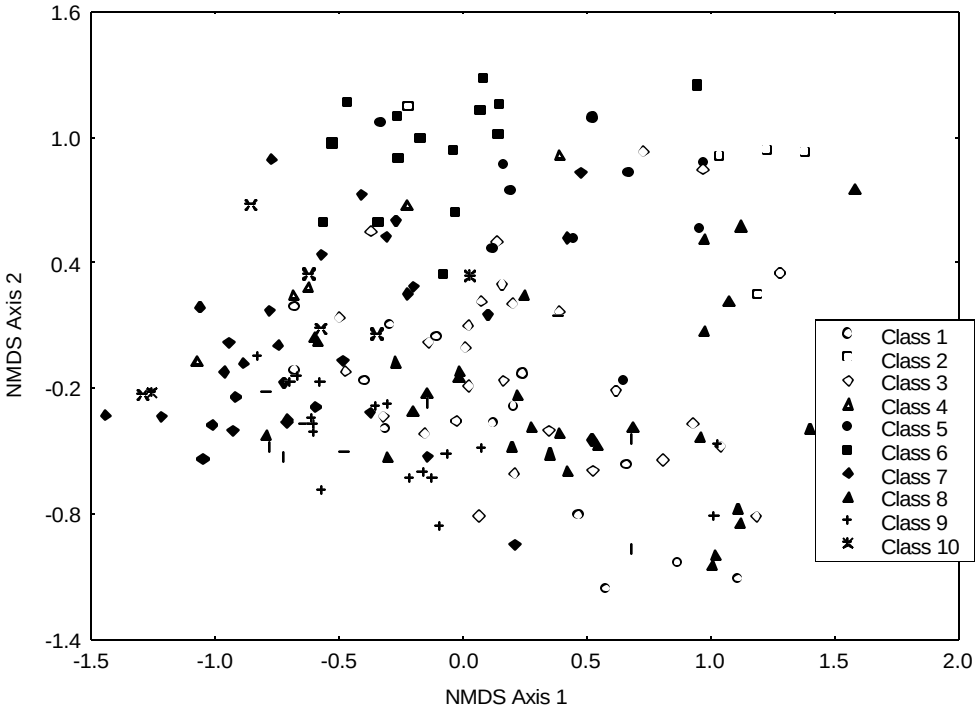


Figure 3-7. NMDS axes 1 and 2 scores grouped according to the 10 preliminary site classes.

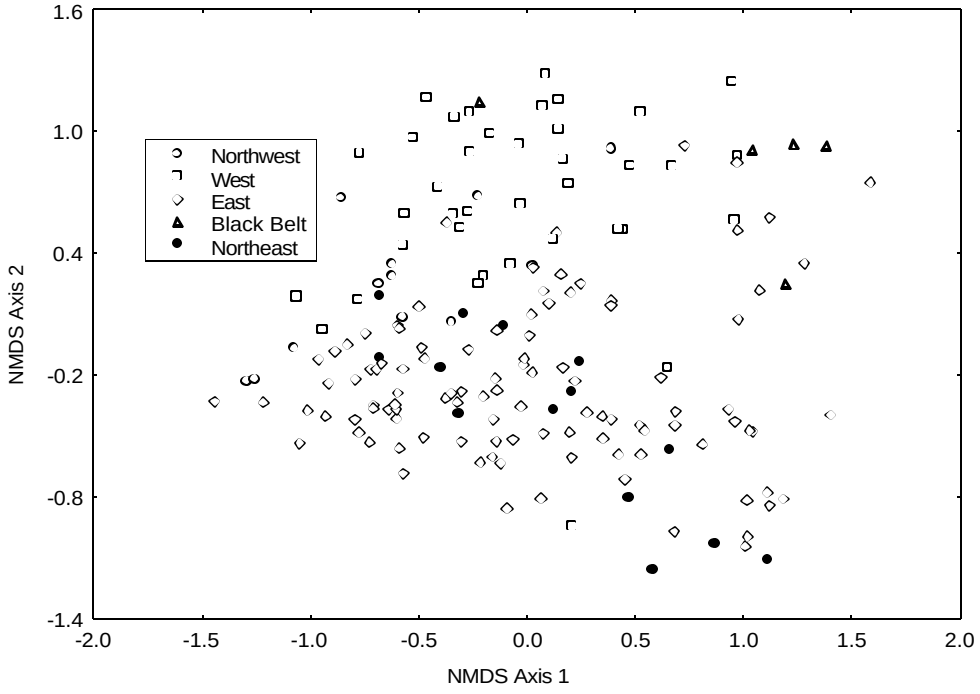


Figure 3-8. NMDS plot of axes 1 and 2 showing final grouping of sites in the five bioregions.

3.7 Biological Index Composition and Performance

Multiple index configurations for each of the five bioregions were tested to find the metric combination that resulted in the highest DEs. The suite of indices presented in Table 3-7 is a subset of the approximately 20-30 indices per bioregion that were actually tested. All of the indices tested in the Black Belt had DEs of 100 percent so the separation between interquartile ranges was used to distinguish between the indices (Appendix E). In the East, two indices had DEs of 89 percent; configuration 4 was chosen because it was composed of two commonly used tolerance metrics (HBI and Beck's) and was more robust being composed of 7 as opposed to the 5 metrics in configuration 1. In the Northwest region, index configuration 4 had the highest DE and also was composed of the most metrics. Three configurations in the Northeast had the same DEs so the separation of interquartile ranges was used to select the final index for this bioregion (Appendix E). Additionally, the tolerance metric (HBI) in configuration 4 was used in some of the other bioregion indices so it was chosen instead of the % Tolerant and Intolerant taxa metrics in configurations 2 and 3. Two index configurations in the West bioregion had 90 percent DEs and interquartile ranges were similar; however, configuration 4 seemed to show a slightly better separation between LDa and MD sites (Appendix E). Box and whisker plots for index configurations for other bioregions are presented in Appendix E. Box and whisker plots for final indices selected for each bioregion are presented in Figure 3-9.

Final M-BISQ scores for each bioregion area presented in Appendix G and descriptive statistics of the indices and metrics used in the indices are presented in Table 3-8. The confidence interval for repeat and duplicate samples combined was ±10.0 units, or points, on a 100-point scale (Table 3-9) (see Appendix B for descriptions of precision calculations). Site specific relative percent difference (RPD) calculations are presented in Appendix F.

3.8 Quality Assurance/Quality Control (QA/QC) and the Assessment of Data Quality

Overall variability (= total uncertainty, or error) of data from any measurement system results from accumulation of error from multiple sources (Taylor 1988, Clark and Whitfield 1994, Taylor and Kuyatt 1994, and Diamond et al. 1996). Error can generally be divided into two types: systematic and random. Systematic error is the type of variability that results from a method and its application or mis-application; it is composed of bias that can, in part, be mediated by using an appropriate quality assurance program. Random error results from the sample itself or the population from which it is derived, and can only partly be controlled through a careful sampling design. It is often not possible to separate the effects of the two types of error, and they can directly influence each other (Taylor 1988). The overall magnitude of error associated with a dataset is known as data quality; how statements of data quality are made and communicated, are critical for data users and decision makers to properly evaluate the extent to which they should rely on technical, scientific, information (Peters 1988, Costanza et al. 1992).

Table 3-6. Discrimination efficiencies (DEs) of all metrics tested within each bioregion. See Appendix F for metric definitions.

Northwest			West			East			Black Belt			Northeast		
Northwest			West			East			Black Belt			Northeast		
Category	Metric	DE	Category	Metric	DE	Category	Metric	DE	Category	Metric	DE	Category	Metric	DE
Habit	BRRWRTAX	65.0	Diversity	SHAN_2	35.0	Habit	SPRWLPCT	52.6	Composition	DOM2PCT	50.0	Composition	NONINPCT	44.4
Composition	CAENIPCT	95.0	Composition	HYD2TRI	75.0	Tolerance	NEWMHBI	94.7	Tolerance	BECKSBI	100.0	Composition	DIPPCT	88.9
Trophic	CLLCTPCT	65.0	Composition	TRICHPCT	35.0	Habit	SPRWLTAX	52.6	Richness	EPHEMTAX	50.0	Richness	ORTHOTAX	44.4
Tolerance	NEWPTOL	95.0	Tolerance	INTOLPCT	75.0	Tolerance	NEWBECK	89.5	Trophic	CLLCTTAX	100.0	Trophic	FILTRPCT	88.9
Trophic	CLLCTTAX	65.0	Richness	TRICHTAX	35.0	Composition	TNYT2CHI	52.6	Composition	EPTPCT	50.0	Habit	SWMMRTAX	44.4
Tolerance	NEWPINTO	91.7	Composition	EPTPCTNC	60.0	Tolerance	TOLERPCT	89.5	Habit	CLNGRTAX	100.0	Composition	NC_TANY%	88.9
Composition	NC_TANY%	65.0	Tolerance	BECKSBI	30.0	Tolerance	TOLERTAX	52.6	Composition	HYD2EPT	50.0	Composition	TRICHPCT	44.4
Composition	ENOCAEN%	90.0	Tolerance	NEWBECK	60.0	Tolerance	NEWINTTX	88.9	Richness	CRMOLTAX	100.0	Habit	CLNGRPCT	77.8
Composition	PREDPCT	65.0	Composition	HYD2EPT	30.0	Richness	TOTALTAX	52.6	Composition	HYD2TRI	50.0	Habit	BRRWRTAX	33.3
Habit	SPRWLPCT	90.0	Trophic	SHREDTAX	60.0	Tolerance	NEWPINTO	88.9	Richness	DIPTAXNC	100.0	Tolerance	HBI	77.8
Richness	PREDTAXR	65.0	Tolerance	NEWPTOL	30.0	Composition	DIPPCTNC	47.4	Composition	NC_TANY%	50.0	Richness	DIPTAXNC	33.3
Tolerance	NEWMHBI	85.0	Habit	SPRWLTAX	60.0	Habit	CLNGRTAX	84.2	Richness	DIPTAXR2	100.0	Tolerance	NEWMHBI	77.8
Trophic	SHREDPCT	65.0	Composition	NONINPCT	30.0	Composition	HYD2TRI	47.4	Tolerance	NEWPTOL	50.0	Composition	ENOCAEN%	33.3
Tolerance	NEWTOLTA	85.0	Richness	COLEOTAX	55.0	Tolerance	HBI	84.2	Composition	EPTPCTNC	100.0	Composition	TANYTPCT	77.8
Trophic	SHREDTAX	65.0	Composition	PREDPCT	30.0	Trophic	SHREDPCT	47.4	Trophic	SCRAPTAX	50.0	Composition	EPTPCTNC	33.3
Composition	TANYTPCT	85.0	Richness	PREDTAXR	55.0	Tolerance	BECKSBI	78.9	Richness	INSCTTAX	100.0	Richness	TANYTTAX	77.8
Richness	TANYTTAX	65.0	Trophic	SHREDPCT	30.0	Composition	TRICHPCT	47.4	Diversity	SHAN_2	50.0	Tolerance	INTOLTAX	33.3
Tolerance	NEWINTTX	83.3	Composition	CAENIPCT	50.0	Composition	CAENIPCT	78.9	Tolerance	INTOLPCT	100.0	Tolerance	TOLERTAX	77.8
Tolerance	INTOLPCT	60.0	Composition	TNYT2CHI	30.0	Richness	TRICHTAX	47.4	Trophic	SHREDPCT	50.0	Richness	PLECOTAX	33.3
Composition	CHIROPCT	80.0	Trophic	SCRAPPCT	50.0	Richness	EPTTAXR2	78.9	Tolerance	INTOLTAX	100.0	Tolerance	NEWPTOL	66.7
Richness	ORTHOTAX	60.0	Composition	DIPPCTNC	25.0	Richness	CHIROTAX	42.1	Habit	SPRWLPCT	50.0	Tolerance	BECKSBI	22.2
Habit	CLNGRPCT	80.0	Composition	BAET2EPH	45.0	Tolerance	INTOLTAX	78.9	Tolerance	NEWBECK	100.0	Trophic	SHREDTAX	66.7
Richness	TOTALTAX	60.0	Richness	DIPTAXR2	25.0	Richness	DIPTAXR2	42.1	Composition	CHIROPCT	25.0	Composition	DIPPCTNC	22.2
Composition	EPTPCTNC	80.0	Richness	CHIROTAX	45.0	Tolerance	NEWPTOL	78.9	Composition	PLECOPCT	100.0	Habit	SPRWLPCT	66.7
Composition	TRICHPCT	60.0	Composition	DOM1PCT	25.0	Richness	EPHEMTAX	42.1	Trophic	CLLCTPCT	25.0	Composition	DOM1PCT	22.2
Tolerance	HBI	80.0	Tolerance	HBI	45.0	Tolerance	NEWTOLTA	78.9	Richness	PLECOTAX	100.0	Tolerance	TOLERPCT	66.7
Composition	AMPHPCT	55.0	Composition	DOM2PCT	25.0	Richness	PREDTAXR	42.1	Habit	CLNGRPCT	25.0	Composition	DOM2PCT	22.2
Tolerance	NEWBECK	80.0	Richness	ORTHOTAX	45.0	Tolerance	INTOLPCT	73.7	Habit	SPRWLTAX	100.0	Composition	CAENIPCT	55.6
Richness	EPTTAXR2	55.0	Composition	EPHEMPCT	25.0	Diversity	SHAN_2	42.1	Composition	DIPPCT	25.0	Richness	EPHEMTAX	22.2
Tolerance	TOLERPCT	80.0	Tolerance	TOLERPCT	45.0	Habit	CLNGRPCT	68.4	Habit	SWMMRPCT	100.0	Richness	CHIROTAX	55.6
Composition	PLECOPCT	55.0	Richness	EPHEMTAX	25.0	Trophic	SHREDTAX	42.1	Composition	DOM1PCT	25.0	Composition	PLECOPCT	22.2
Richness	CHIROTAX	75.0	Richness	TOTALTAX	45.0	Composition	EPTPCTNC	68.4	Tolerance	TOLERPCT	100.0	Composition	HYD2TRI	55.6
Richness	PLECOTAX	55.0	Composition	ISOPCT	25.0	Composition	EPTPCT	36.8	Composition	EPHEMPCT	25.0	Diversity	SHAN_2	22.2
Habit	CLNGRTAX	75.0	Tolerance	NEWINTTX	42.1	Trophic	FILTRTAX	68.4	Richness	TOTALTAX	100.0	Trophic	SHREDPCT	55.6
Tolerance	TOLERTAX	55.0	Composition	NC_TANY%	25.0	Composition	ISOPCT	36.8	Trophic	FILTRPCT	25.0	Habit	SWMMRPCT	22.2
Richness	DIPTAXR2	75.0	Tolerance	NEWPINTO	42.1	Composition	PLECOPCT	68.4	Composition	AMPHPCT	75.0	Composition	TNYT2CHI	55.6
Tolerance	BECKSBI	50.0	Trophic	SCRAPTAX	25.0	Trophic	SCRAPTAX	36.8	Trophic	FILTRTAX	25.0	Richness	TOTALTAX	22.2
Trophic	FILTRTAX	75.0	Habit	BRRWRTAX	40.0	Richness	PLECOTAX	68.4	Composition	CAENIPCT	75.0	Richness	TRICHTAX	55.6
Composition	CRMOLPCT	50.0	Habit	SWMMRPCT	25.0	Habit	SWMMRPCT	36.8	Tolerance	HBI	25.0	Habit	CLMBRPCT	11.1
Tolerance	INTOLTAX	75.0	Composition	CHIROPCT	40.0	Trophic	FILTRPCT	63.2	Richness	CHIROTAX	75.0	Tolerance	NEWINTTX	50.0
Richness	DIPTAXNC	50.0	Composition	TANYTPCT	25.0	Composition	DOM1PCT	31.6	Tolerance	NEWMHBI	25.0	Composition	CRMOLPCT	11.1
Composition	TNYT2CHI	73.7	Composition	DIPPCT	40.0	Composition	NC_TANY%	63.2	Composition	COLEOPCT	75.0	Tolerance	NEWPINTO	50.0
Richness	TRICHTAX	50.0	Habit	BRRWRPCT	20.0	Composition	DOM2PCT	31.6	Tolerance	NEWTOLTA	25.0	Richness	CRMOLTAX	11.1
Richness	CRMOLTAX	70.0	Composition	EPTPCT	40.0	Composition	TANYTPCT	63.2	Richness	COLEOTAX	75.0	Composition	CHIROPCT	44.4
Composition	CRCH2CHI	47.4	Trophic	CLLCTPCT	20.0	Composition	HYD2EPT	31.6	Composition	NONINPCT	25.0	Composition	EPHEMPCT	11.1
Composition	DIPPCT	70.0	Trophic	FILTRTAX	40.0	Richness	TANYTTAX	63.2	Composition	CRMOLPCT	75.0	Habit	CLNGRTAX	44.4
Habit	BRRWRPCT	45.0	Trophic	CLLCTTAX	20.0	Composition	PREDPCT	31.6	Composition	OLIGOPCT	25.0	Composition	EPTPCT	11.1
Composition	DOM1PCT	70.0	Richness	INSCTTAX	40.0	Composition	DIPPCT	57.9	Richness	EPTTAXR2	75.0	Richness	DIPTAXR2	44.4
Richness	EPHEMTAX	45.0	Habit	CLNGRPCT	20.0	Trophic	SCRAPPCT	31.6	Trophic	SCRAPPCT	25.0	Composition	HYD2EPT	11.1
Composition	DOM2PCT	70.0	Tolerance	INTOLTAX	40.0	Composition	CHIROPCT	52.6	Composition	GASTRPCT	75.0	Richness	EPTTAXR2	44.4
Habit	SPRWLTAX	45.0	Habit	CLNGRTAX	20.0	Habit	BRRWRTAX	26.3	Trophic	SHREDTAX	25.0	Tolerance	INTOLPCT	11.1
Trophic	FILTRPCT	70.0	Tolerance	NEWMHBI	40.0	Trophic	CLLCTPCT	52.6	Richness	ORTHOTAX	75.0	Trophic	FILTRTAX	44.4
Habit	CLMBRPCT	40.0	Richness	DIPTAXNC	20.0	Composition	COLEOPCT	26.3	Habit	SWMMRTAX	25.0	Composition	OLIGOPCT	11.1
Richness	INSCTTAX	70.0	Habit	CLMBRTAX	35.0	Composition	ENOCAEN%	52.6	Composition	PREDPCT	75.0	Richness	INSCTTAX	44.4
Habit	CLMBRTAX	40.0	Richness	EPTTAXR2	20.0	Composition	ODONPCT	26.3	Composition	BAET2EPH	0.0	Composition	AMPHPCT	0.0
Composition	NONINPCT	70.0	Composition	COLEOPCT	35.0	Richness	INSCTTAX	52.6	Richness	PREDTAXR	75.0	Tolerance	NEWBECK	44.4
Composition	DIPPCTNC	40.0	Trophic	FILTRPCT	20.0	Composition	BAET2EPH	21.1	Composition	BIVALPCT	0.0	Composition	BAET2EPH	0.0
Diversity	SHAN_2	70.0	Tolerance	NEWTOLTA	35.0	Richness	ORTHOTAX	52.6	Composition	DIPPCTNC	50.0	Tolerance	NEWTOLTA	44.4
Habit	SWMMRTAX	35.0	Composition	ODONPCT	20.0	Richness	COLEOTAX	21.1	Habit	BRRWRPCT	0.0	Composition	BIVALPCT	0.0

Table 3-6 (cont'd). Discrimination efficiencies (DEs) of all metrics tested within each bioregion. See Appendix F for metric definitions.

Northwest			West			East			Black Belt			Northeast		
Composition	HYD2TRI	30.0	Composition	OLIGOPCT	20.0	Composition	CRMOLPCT	21.1	Habit	BRRWRTAX	0.0	Habit	BRRWRPCT	0.0
Composition	OLIGOPCT	30.0	Habit	CLMBRPCT	15.0	Richness	OLIGOTAX	21.1	Composition	CCO2CHIR	0.0	Composition	CCO2CHIR	0.0
Trophic	SCRAPPCT	25.0	Richness	CRMOLTAX	15.0	Habit	BRRWRPCT	15.8	Habit	CLMBRPCT	0.0	Trophic	CLLCTPCT	0.0
Richness	COLEOTAX	20.0	Habit	SPRWLPCT	15.0	Trophic	CLLCTTAX	15.8	Habit	CLMBRTAX	0.0	Trophic	CLLCTTAX	0.0
Richness	OLIGOTAX	20.0	Habit	SWMMRTAX	15.0	Richness	CRMOLTAX	15.8	Composition	CORBPCT	0.0	Habit	CLMBRTAX	0.0
Habit	SWMMRPCT	20.0	Richness	TANYTTAX	15.0	Richness	DIPTAXNC	15.8	Composition	CRCH2CHI	0.0	Composition	COLEOPCT	0.0
Composition	COLEOPCT	15.0	Tolerance	TOLERTAX	15.0	Composition	EPHEMPCT	15.8	Composition	ENOCAEN%	0.0	Richness	COLEOTAX	0.0
Composition	EPHEMPCT	15.0	Composition	CRMOLPCT	10.0	Habit	SWMMRTAX	15.8	Composition	ISOPCT	0.0	Composition	CORBPCT	0.0
Composition	EPTPCT	10.0	Richness	OLIGOTAX	5.0	Composition	NONINPCT	10.5	Composition	ODONPCT	0.0	Composition	CRCH2CHI	0.0
Composition	HYD2EPT	10.0	Composition	AMPHPCT	0.0	Composition	OLIGOPCT	10.5	Richness	OLIGOTAX	0.0	Composition	GASTRPCT	0.0
Trophic	SCRAPTAX	10.0	Composition	BIVALPCT	0.0	Composition	AMPHPCT	0.0	Composition	ORTH2CHI	0.0	Composition	ISOPCT	0.0
Composition	CCO2CHIR	5.6	Composition	CCO2CHIR	0.0	Composition	BIVALPCT	0.0	Composition	TANYTPCT	0.0	Composition	ODONPCT	0.0
Composition	ISOPCT	5.0	Composition	CORBPCT	0.0	Composition	CCO2CHIR	0.0	Richness	TANYTTAX	0.0	Richness	OLIGOTAX	0.0
Composition	BAET2EPH	0.0	Composition	CRCH2CHI	0.0	Habit	CLMBRPCT	0.0	Composition	TNYT2CHI	0.0	Composition	ORTH2CHI	0.0
Composition	BIVALPCT	0.0	Composition	ENOCAEN%	0.0	Habit	CLMBRTAX	0.0	Tolerance	TOLERTAX	0.0	Composition	PREDPCT	0.0
Composition	CORBPCT	0.0	Composition	GASTRPCT	0.0	Composition	CORBPCT	0.0	Composition	TRICHPCT	0.0	Richness	PREDTAXR	0.0
Composition	GASTRPCT	0.0	Composition	ORTH2CHI	0.0	Composition	CRCH2CHI	0.0	Richness	TRICHTAX	0.0	Trophic	SCRAPPCT	0.0
Composition	ODONPCT	0.0	Composition	PLECOPCT	0.0	Composition	GASTRPCT	0.0	Tolerance	NEWINTTX		Trophic	SCRAPTAX	0.0
Composition	ORTH2CHI	0.0	Richness	PLECOTAX	0.0	Composition	ORTH2CHI	0.0	Tolerance	NEWPINTO		Habit	SPRWLTAX	0.0

Table 3-7. Index configurations and DEs for five bioregions in Mississippi. Four potential index configurations (index #) are presented for each bioregion. The final index configuration chosen for each bioregion is in bold.

	Bioregion	Black Belt				East				Northwest				Northeast				West			
	Index #	1	2	3	**4**	1	2	3	**4**	1	2	3	**4**	1	2	3	**4**	1	2	3	**4**
Metric	DE	100	100	100	**100**	89	84	79	**89**	85	85	80	**90**	78	89	89	**89**	90	50	55	**90**
Burrower taxa																				x	
% Caenidae		x		x	**x**	x		x	**x**	x					x				x		
% Chironomidae											x										
Chironomidae taxa										x		x	**x**	x		x				x	
Collector taxa		x			**x**																
% Clinger							x	x	**x**	x	x		**x**	x		x	**x**				
Clinger taxa			x	x		x						x									
% Coleoptera			x																		
Coleoptera taxa					**x**													x			**x**
% Diptera																x	**x**			x	
Diptera taxa (no Chironomidae)		x		x																	
Diptera taxa			x								x										
Ephemeroptera (no Caenidae)										x	x		**x**								
% EPT (no Caenidae)			x						**x**									x	x		**x**
EPT taxa							x	x													
% Filterer						x			**x**							x	**x**				
Filterer taxa							x	x		x	x	x	**x**								
Hydropsychida/Trichoptera															x			x			**x**
Insect taxa		x		x															x		
Beck's Biotic Index		x			**x**				**x**			x	**x**					x		x	**x**
Hilsenhoff Biotic Index									**x**				**x**	x			**x**		x		
Intolerant taxa							x	x								x					
% Tolerant taxa								x			x	x			x						
% Intolerant taxa						x												x			
NC Tany%							x							x							
% Plecoptera			x																		
Plecoptera taxa		x		x	**x**																
Predator taxa				x														x			**x**
% Scraper																			x		
% Shredder			x											x	x						
Shredder taxa																			x		
% Sprawler															x						
Sprawler taxa					**x**													x	x		**x**
% Swimmer		x																			
% Tanytarsini										x		x	**x**			x	**x**				
Tanytarsini taxa						x	x		**x**						x					x	
Total taxa				x	**x**																
Trichoptera taxa																	**x**				

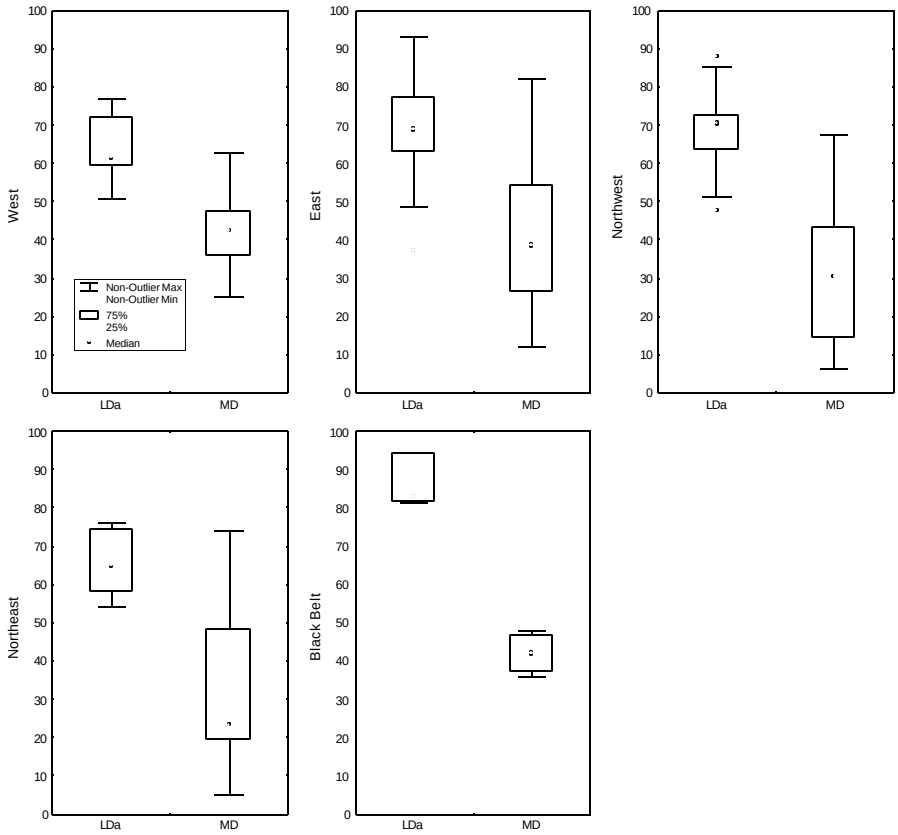


Figure 3-9. Comparison of LDa and MD index distributions within each bioregion. The wider the separation between box plots, the greater the discriminatory ability of the index.

One goal of the agency was to produce biological assessments using credible, technically defensible, and scientifically rigorous data (MDEQ 2001). Consequently, a comprehensive plan for ensuring the collection of such data was developed prior to project initiation (MDEQ 2001), and followed U. S. Environmental Protection Agency requirements for developing project plans (USEPA 1999). The Quality Assurance Project Plan (QAPP) describes, in detail, the procedures that are used for data collection, the technical rationale behind the procedures, and the series of activities and reporting procedures that will be used to document and communicate data quality. There are at least five data quality characteristics: precision, accuracy, representativeness, completeness, and comparability; assessments can be either quantitative or qualitative (Table 3-10). A stream assessment (in particular, a biological assessment) is a series of methods taken together as a protocol (Diamond et al. 1996, Barbour et al. 1999). The purpose of this section is to provide users of this report with an assessment of the data quality for each of the steps of the assessment process. Because detailed descriptions of methods are provided in the QAPP, and briefly in section 2.0 of this report, only specific critical methods information is presented below. If a particular data quality characteristic is not applicable (NA) to a method or protocol component, it is indicated as such.

Metrics and Index	Number of Sites	Minimum	5th %ile	25th %ile	Median	75th %ile	95th %ile	Maximum
Black Belt								
Beck's Biotic Index	26	0.0	0.0	1.0	2.5	7.0	10.0	11.0
Total taxa	26	13.0	14.7	22.6	28.2	31.0	39.5	44.1
Plecoptera taxa	26	0.0	0.0	0.0	0.0	1.0	3.0	3.0
Coleoptera taxa	26	1.0	1.0	2.0	3.6	4.9	6.0	6.7
% Caenidae	26	0.0	0.0	2.3	25.6	55.6	66.1	73.8
Collector taxa	26	6.0	6.8	8.8	10.6	13.8	17.0	17.9
Sprawler taxa	26	3.0	3.5	5.8	7.6	9.0	11.8	11.8
Index	26	30.2	35.0	44.9	50.0	69.0	83.5	94.5
East								
Beck's Biotic Index	204	0.0	3.0	11.0	18.0	26.0	38.0	43.0
Hilsenhoff Biotic Index	204	2.3	3.1	3.8	4.3	5.0	6.7	8.6
Tanytarsini taxa	204	0.0	0.0	1.0	2.0	3.6	4.9	6.0
% Caenidae	204	0.0	0.0	0.0	0.0	1.0	14.7	70.7
% EPT (no Caenidae)	204	0.0	0.0	6.8	13.2	21.5	38.7	90.3
% Filterer	204	0.0	2.0	14.1	26.9	41.0	58.0	81.2
% Clinger	204	0.0	5.3	36.2	52.9	66.2	80.9	87.8
M-BISQ	204	10.9	27.9	51.0	63.7	71.7	83.7	92.2
Northwest								
Beck's Biotic Index	91	0.0	1.0	5.0	8.0	14.0	25.0	31.0
Hilsenhoff Biotic Index	91	3.5	3.8	5.2	6.4	7.9	8.9	9.6
Chironomidae taxa	91	0.0	4.9	9.2	13.0	16.0	21.1	23.9
% Tanytarsini	91	0.0	0.0	0.5	3.0	10.1	30.9	46.3
Ephemeroptera (no Caenidae)	91	0.0	0.0	0.9	2.7	11.6	34.9	51.7
Filterer taxa	91	0.0	0.0	2.0	3.8	4.9	6.0	7.0
% Clinger	91	0.0	1.3	9.1	27.4	50.6	68.2	79.1
M-BISQ	91	6.0	9.5	26.2	40.1	59.3	81.3	87.8
Northeast								
Hilsenhoff Biotic Index	37	2.6	3.1	3.8	4.1	5.1	8.5	8.9
Trichoptera taxa	37	0.0	0.0	1.0	2.0	4.0	8.3	9.0
% Diptera	37	9.0	22.6	52.2	65.8	72.9	93.7	96.0
% Tanytarsini	37	0.0	0.0	2.4	6.9	22.3	41.0	45.1
% Filterer	37	0.0	0.5	12.9	32.3	45.7	89.1	92.0
% Clinger	37	3.3	4.3	37.2	55.7	66.2	93.3	95.5
M-BISQ	37	4.8	11.5	44.7	54.0	61.1	74.4	75.4
West								
Beck's Biotic Index	96	0.0	3.0	6.0	9.0	14.0	25.0	29.0
Coleoptera taxa	96	0.0	1.0	2.9	4.0	5.7	7.4	9.0
% EPT (no Caenidae)	96	0.0	0.0	2.1	7.2	16.0	39.5	80.5
Predator taxa	96	2.9	5.0	7.9	9.8	11.9	15.1	16.8
Sprawler taxa	96	4.0	4.9	7.0	8.9	11.2	14.1	18.0
Hydropsycidae/Trichoptera	96	0.0	0.0	0.0	55.1	94.5	100.0	100.0
M-BISQ	96	25.0	30.3	38.6	50.0	59.5	77.5	88.3

Table 3-8. Descriptive statistics for metric values and M-BISQ scores for all sites from the five bioregions.

Table 3-9. Precision statistics for metric values and index scores from biological repeat and duplicate sites. RPD values of 200 were excluded from the median RPD calculation to minimize the influence of low metric values which tend to skew the statistic.

	Repeat Samples (BR, n=34)					Duplicate Samples (BD, n=36)					Repeat + Duplicate Samples (n=70)				
rics	Mean	Estimated Standard Deviation (RMSE)	Coefficient of Variation (%)	Detectable Difference (90% confidence)	Median RPD	Mean	Estimated Standard Deviation (RMSE)	Coefficient of Variation (%)	Detectable Difference (90% confidence)	Median RPD	Mean	Estimated Standard Deviation (RMSE)	Coefficient of Variation (%)	Detectable Difference (90% confidence)	Median RPD
	59.6	6.2	10.4	10.2	12.6	55.3	6.0	10.8	9.8	7.8	57.4	6.1	10.6	10.0	10.5
ndex	17.4	4.2	24.0	6.8	22.9	17.4	3.3	19.3	5.5	16.9	17.4	3.8	21.7	6.2	22.0
ic Index	5.2	0.4	8.2	0.7	7.3	5.1	0.4	8.3	0.7	4.4	5.1	0.4	8.2	0.7	5.9
	40.0	7.0	17.4	11.4	17.6	40.3	5.3	13.3	8.8	12.8	40.1	6.2	15.4	10.2	14.3
a	1.7	1.0	59.3	1.7	48.6	2.0	1.0	52.9	1.7	22.3	1.8	1.0	55.9	1.7	39.8
a	2.7	1.3	48.3	2.1	37.8	2.5	1.5	60.0	2.5	40.0	2.6	1.4	54.3	2.3	38.1
taxa	14.5	3.6	24.6	5.8	24.9	14.4	2.9	20.2	4.8	19.3	14.5	3.2	22.5	5.3	22.7
a	2.6	0.9	35.3	1.5	27.2	2.6	0.8	30.9	1.3	30.8	2.6	0.9	33.1	1.4	28.6
a	3.7	1.4	36.7	2.2	42.9	3.5	1.3	37.6	2.2	34.7	3.6	1.3	37.1	2.2	40.0
	47.9	11.2	23.4	18.4	30.2	47.9	9.4	19.6	15.4	13.4	47.9	10.3	21.5	16.9	18.2
	12.0	6.9	57.8	11.4	32.9	11.8	4.8	41.0	7.9	44.4	11.9	6.0	50.1	9.8	41.9
tera (no Caenidae)	8.8	4.6	52.5	7.5	46.2	7.2	3.6	49.5	5.9	40.6	8.0	4.1	51.5	6.8	45.4
	9.1	4.1	45.1	6.7	36.3	8.8	6.0	67.6	9.8	47.0	9.0	5.1	57.3	8.4	41.9
enidae)	17.2	7.9	46.1	13.0	41.8	14.6	6.1	41.8	10.0	29.2	15.9	7.1	44.4	11.6	36.8
	22.4	9.1	40.7	14.9	39.5	20.6	8.2	40.1	13.5	28.7	21.5	8.7	40.4	14.2	35.7
	15.2	3.4	22.3	5.6	20.2	15.3	3.2	21.1	5.3	24.5	15.2	3.3	21.7	5.4	24.1
	4.8	1.2	25.3	2.0	22.9	4.6	1.2	25.8	1.9	29.0	4.7	1.2	25.5	1.9	28.4
	9.6	2.3	23.8	3.8	27.1	10.3	2.6	24.9	4.2	21.8	10.0	2.4	24.4	4.0	24.5
	43.3	13.3	30.7	21.8	28.3	41.1	8.6	20.9	14.1	17.7	42.2	11.1	26.4	18.3	20.0
	10.4	2.9	28.2	4.8	26.3	10.1	2.0	19.9	3.3	22.2	10.3	2.5	24.5	4.1	23.4
a/Trichoptera	38.7	26.3	67.9	43.1	15.4	39.5	29.7	75.1	48.7	10.8	39.1	28.1	71.8	46.0	13.3

Table 3-10. Error partitioning framework for biological assessment protocols. Performance characteristics may be quantitative (QN), qualitative (QL), or not applicable (na). Those characteristics in bold were addressed in this project.

Component Method or Activity	Performance Characteristics				
	Precision	**Accuracy**	**Bias**	**Representativeness**	**Completeness**
1. Field Sampling	**QN**	na	**QL**	**QL**	**QN**
2. Laboratory Sorting/ Subsampling	QN	QN	**QN**	QN/**QL**	na
3. Taxonomy	**QN**	**QN**	QL	na	na
4. Enumeration	**QN**	QN	QL	na	na
5. Data Entry	QN	**QN**	na	na	na
6. Metric calculation (e. g., Data Reduction)	na	**QN**	QL	na	na
7. Final Index and Site Assessment	**QN**/QL	**QN**	QL	**QL**	**QN**

Prior to initiation of fieldwork, all field and laboratory personnel reviewed standard operating procedures (SOPs) for activities they would be performing. Training workshops were held where all field and laboratory procedures were reviewed and demonstrated.

3.8.1 Field Sampling

3.8.1.1 Benthic Macroinvertebrates

Method overview. This sampling activity was performed with a long handled D-frame net (800 × 900 micron mesh) and a controlled level of effort (20 jabs) to sample multiple habitats over a 100m stream reach. Two types of duplicate samples were taken. After sampling the primary reach, a field team sampled a reach that was adjacent to it; this was termed a bioduplicate (BD). A field team would be assigned to resample a reach after another team had completed the primary sample; this was termed a biorepeat (BR). All sites for which duplicate and repeat sampling occurred were selected at random from the initial master site list. The designed rate of repeat sampling was approximately 15% (or 70 out of 475 sites); there were to be 35 BD samples, and 35 BR samples. The final totals were 36 and 34, respectively.
Precision was quantitatively evaluated in three ways: 1) the consistency of each of the field teams from one sample to the next in the same stream; 2) the consistency of the method when

applied by two *different* field teams at the same site; and 3) comparison of the two types of precision estimates

1) Intra-team consistency (reproducibility of a result)

Intra-team RPD (bioduplicates) for teams 1, 2, 4, 6, and 7 ranged from a median of 23-29% across all metrics (Figure 3-10), slightly higher than inter-team comparisons but with a much smaller spread. Twelve of 20 metrics had a median RPD spread of <20% among at least 3 of the 5 teams; they were: No. Chironomidae Taxa, No. Collector Taxa, No. Coleoptera Taxa, percent Diptera, percent Ephemeroptera without Caenidae, percent EPT without Caenidae, No. Filterer Taxa, HBI, Hydropsychidae/Trichoptera, No. Plecoptera Taxa, No. Predator Taxa, No. Tanytarsini Taxa, and Total Taxa. The highest intra-team RPDs were exhibited by percent Caenidae, percent Filterers, percent Tanytarsini, and No. Trichoptera Taxa.

2) Inter-team consistency (method precision)

Inter-team RPD (biorepeats) for teams 1, 2, 4, 6, and 7 ranged from 16-27%, slightly lower than for intra-team comparisons but with a larger spread. All metrics had at least three teams within a 20 percentage point spread (Figure 3-11). The largest overall spreads for inter-team comparisons (i. e., differences among teams overall) for metrics were percent Ephemeroptera without Caenidae, percent Filterers, No. Plecoptera Taxa, No. Tanytarsini Taxa, and No. Trichoptera Taxa.

3) Comparability of precision estimates developed using BR vs. BD

BR and BD sample pairs produced similar results. RPD across all teams showed substantial precision (=repeatability) for most of the metrics (Figure 3-10 and 3-11). CVs across all sample pairs for all teams exhibited good inter- and intra-team comparability (Table 3-9, Figure 3-12). The former seemed slightly worse than the latter, with the CV being slightly larger on 13 of the 20 metrics. The metrics with the highest CVs (>50%) (i. e., least precise) were No. Plecoptera Taxa (BR and BD), percent Tanytarsini (BR), percent Ephemeroptera (no Caenidae) (BR), No. Trichoptera Taxa (BD), percent Caenidae (BD), and Hydropsychidae/Trichoptera (BR and BD). The M-BISQ had an overall CV of 10.3 when inter-team and intra-team sample pairs were combined.

Overall, variability reflected seems to be low and acceptable; the majority of the metrics have RPD < 30%. We recommend that an MQO be established for each metric and the overall M-BISQ for use in future data quality assessments. It should be noted that the two teams, 3 and 5, had several changes in personnel over the sampling period, and had very few, if any BD or BR samples.

Accuracy is not directly applicable for field sampling in this project because it would require knowledge of all target organisms at a sampling location, which is not feasible with invertebrates. (NA)

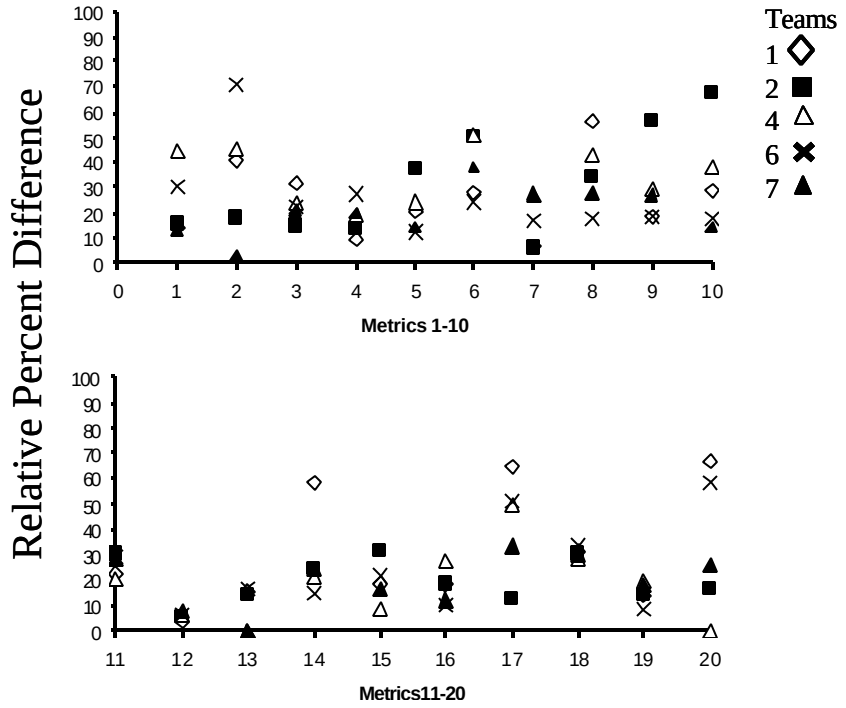


Figure 3-10. Intra-team relative percent difference (RPD) of individual metrics (bioduplicates).

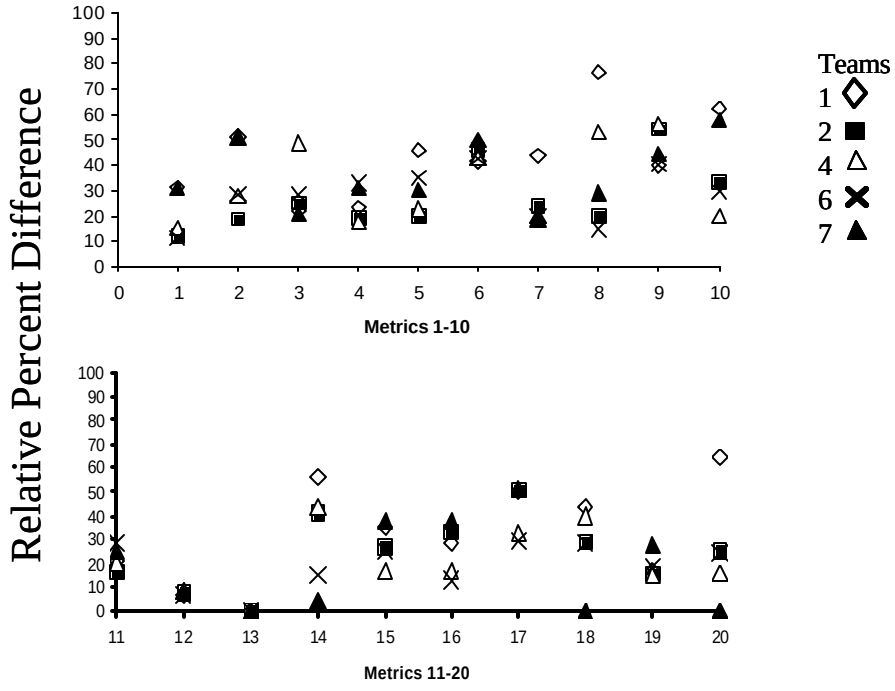


Figure 3-11. Inter-team relative percent difference (RPD) of individual metrics (biorepeats).

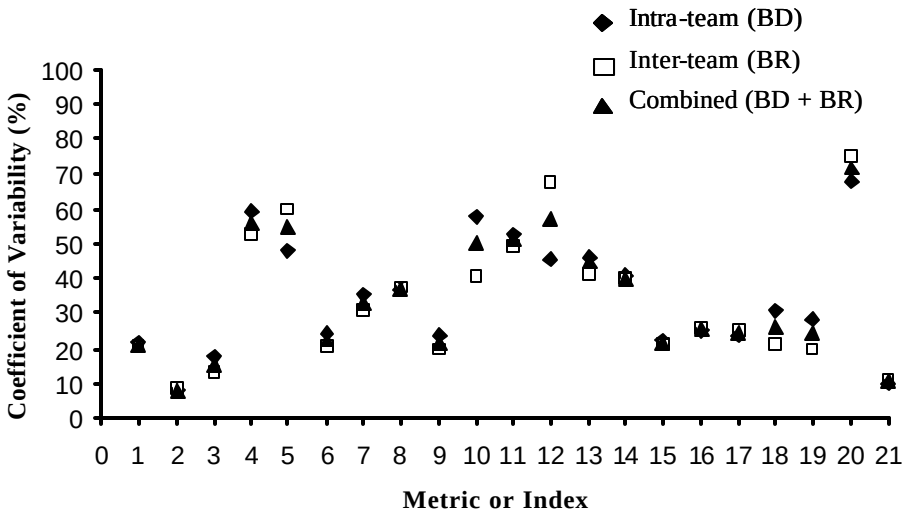


Figure 3-12. Comparison of CV across all metrics and the index using repeat sampling.

Bias control is attempted by allocating sampling effort among multiple habitats in proportion to their occurrence in the stream. The intent is to avoid over-sampling rare habitats and under-sampling abundant habitats.

Representativeness of the sampling approach is inherent in its design. The method targets multiple sub habitats (undercut banks, snags/woody debris, leaf litter, riffles, macrophyte beds, and sandy bottom), and, with the exception of sandy bottom, allocates sampling effort among the habitats in rough proportion to their occurrence through the 100m reach. This sampling approach is designed to produce a multi-taxon sample that reflects the benthic macroinvertebrate assemblage that the stream physical habitat has the capacity to support.

Completeness. There was a total of 475 sites for which sampling was planned for which the following sampling and data analyses were planned: benthic macroinvertebrates, field chemistry, laboratory analytical chemistry, physical habitat assessment, and pebble counts. Percent completeness for each is given in Table 3-11.

Table 3-11. Percent completeness of field sampling for different sample types.

Type of Sample	Number of Sites/Samples		Completeness (%)
	Planned	Sampled	
Benthic macroinvertebrates	475	455	95.8
Field Chemistry	475	453	95.4
Laboratory Chemistry	475	460	96.8
Physical Habitat	475	463	97.5
Pebble Count	475	463	97.5

3.8.1.2 Chemical

Method overview. Field duplicate grab samples were taken at 48 sites by six different field teams; the MDEQ Chemistry Laboratory performed all analytical procedures. All sample handling and laboratory analysis was performed as specified in the QA Project Plan (MDEQ 2001).

Precision. This characteristic was evaluated separately with reference to field collection and laboratory procedures. The precision of the laboratory analyses was evaluated by comparing value differences (range) between two duplicate values with an upper control limit (UCL) for that difference; the UCL was exceeded six times (Table 3-12) and is a rate considered acceptable. Field precision was characterized by calculating RPD for the field duplicates (Figure 3-13).

Table 3-12. Laboratory chemistry analytes and the number of control limit exeedences.

Analyte	UCL	No. Exceeding
Chloride (Cl)	0.3	1
Nitrate-Nitrate (N-N)	0.05, 0.12	0
Ammonia (NH_3)	0.1	2
Total Kjeldahl Nitrogen (TKN)	0.2	1
Total Phosphorus (TP)	0.06, 0.2	1
Alkalinity ($CaCO_3$)	3, 16	1
Chemical Oxygen Demand (COD)	8	0
Total Organic Carbon (TOC)	1	0

Analytes with greatest field consistency were alkalinity, chlorides, TOC, and N-N; the most variable (= lowest consistency) were NH_3, TKN, TP, and COD.

Accuracy. MDEQ uses percent recovery as assessment of the accuracy of chemical analysis, although it has been used as a measure of bias. Percent recovery for both reference standards and spiked duplicate samples never fell outside the range of 80-120%.

Bias. See accuracy above.

Representativeness. In part, this characteristic is demonstrated by comparison of duplicated grab samples. Non-representativeness of a sample would exhibit a larger number of exeedences than shown in Table 3-12.

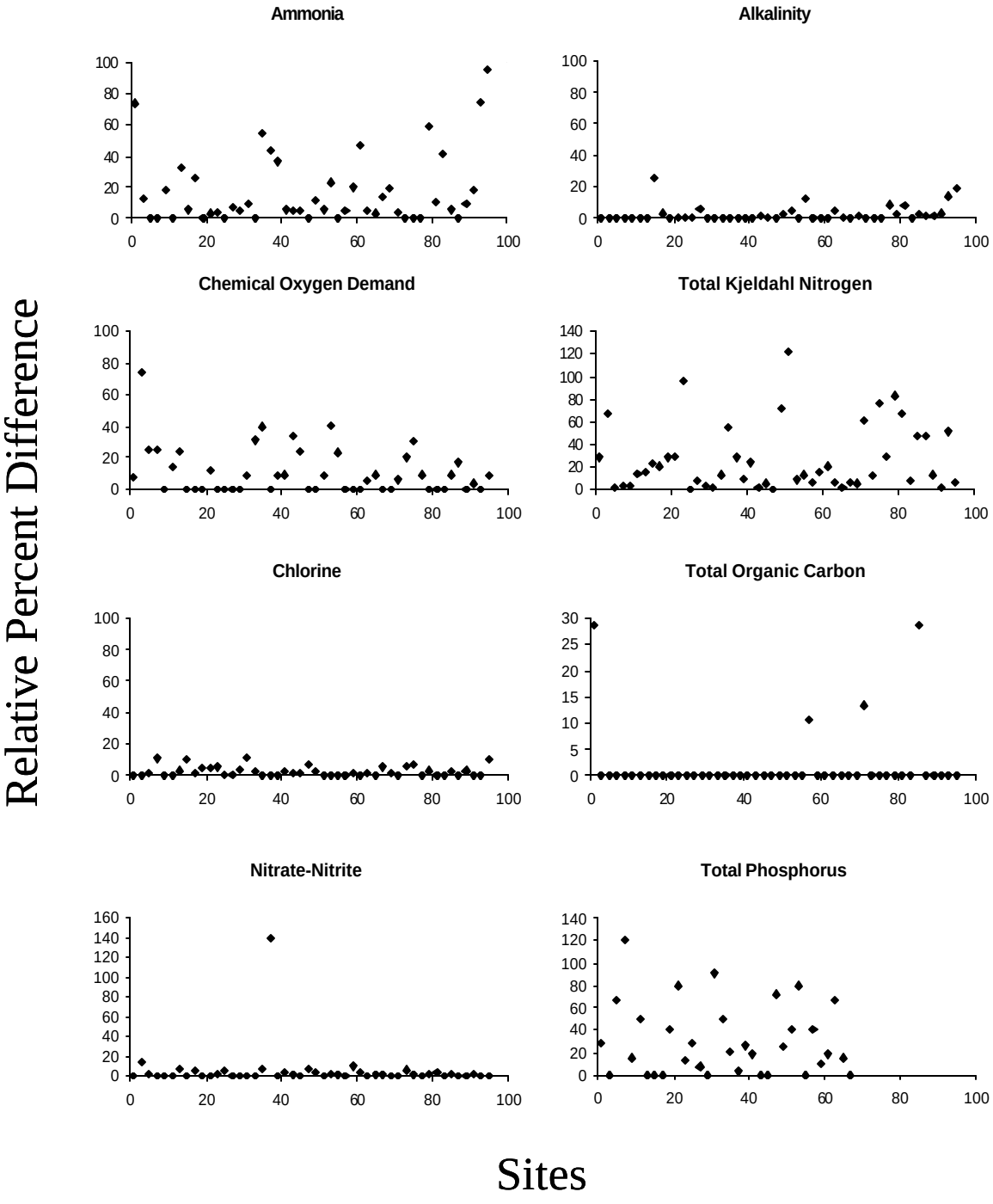


Figure 3-13. RPD of duplicate grab samples taken for laboratory analytical chemistry at 48 stream sites by six different field teams.

Completeness. Four hundred fifty three (453) chemistry grab samples out of 475 planned were taken, for a completeness of 96.8% (Table 3-11).

3.8.1.3 Physical

Method overview. The procedure for assessing physical habitat quality is based on that endorsed by the U. S. EPA (Barbour et al. 1999); it is visual-based and focuses on rating or scoring 10 different habitat components along a continuum of conditions. Each parameter is scored on a continuous scale of 0-20, with 0 being the worst condition, or most degraded; and 20 being the best condition, or most natural. This analysis evaluated inter-team variability by examining the difference in paired scores for 34 sites that were visited by a second (or repeat) team.

Precision. Overall inter-team RPD of the total habitat score at individual sites ranged from 0 (perfect agreement) to 41% (n=34) (Figure 3-14), with a median of 16%. Across all sample pairs, the CV was 11.2% and the 90% confidence interval was 23.6 (on a 200-point scale) (Figure 3-15). Five individual parameters had a CV>30% (bottom substrate/available cover, pool variability, sediment deposition, channel flow status, and bank stability). Only two had CV<20% (channel alteration, riparian vegetative zone width). Most of the total habitat RPD from field teams ranged from 0-25%, although some were occasionally as high as 35-40% (Figure 3-16).

Accuracy. Not applicable.

Bias. The level of bias with this method can be substantial if the operator is undertrained or has a minimum of experience. The level of training and experience among the field teams was not equivalent and likely directly influenced the variability of the final habitat scores.

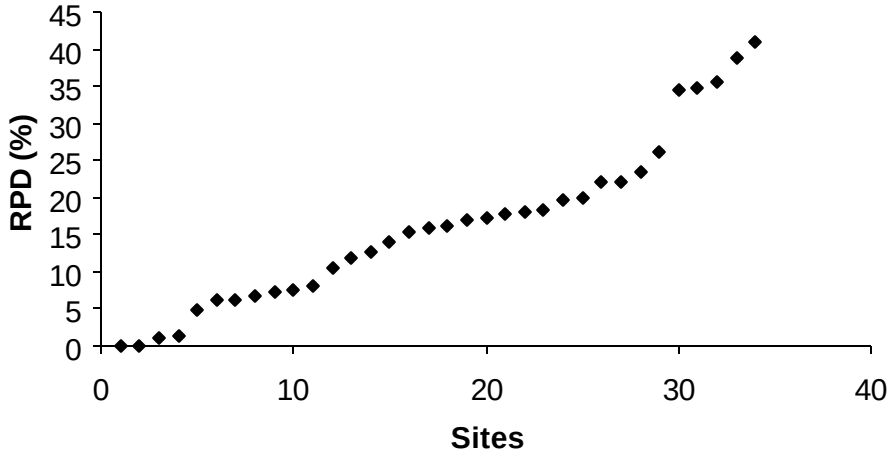


Figure 3-14. Inter-team relative percent difference (RPD) for physical habitat assessment at 34 sites.

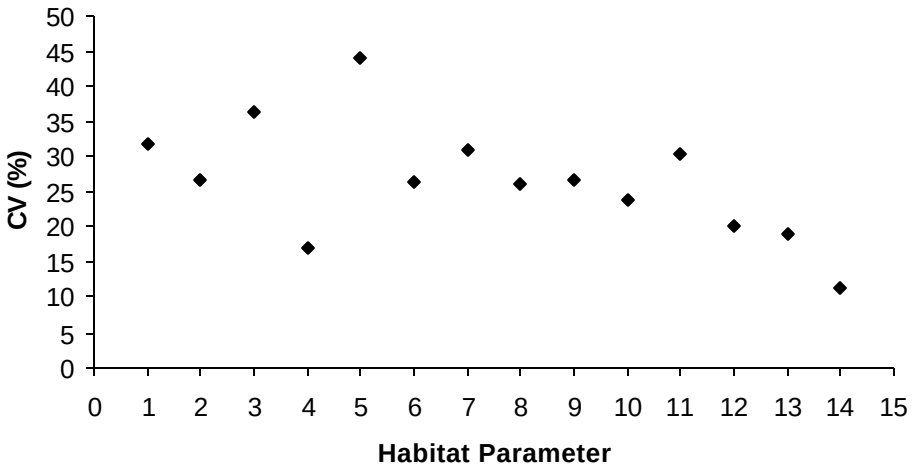


Figure 3-15. Coefficient of variability (CV) of repeated habitat assessment of 34 sites. Parameters 9-13 are split by right and left banks, and are thus, each scored on a 10-point scale individually. Parameter 14 is the aggregated total score.

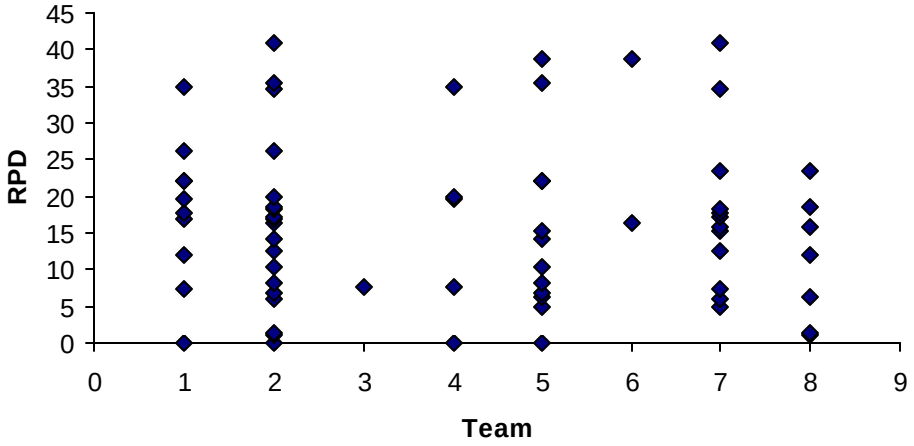


Figure 3-16. Range of intra-team relative percent differences (RPD) by field teams for total physical habitat scores.

Representativeness. This characteristic was not tested, but is intended to simultaneously represent the structural complexity of the stream channel morphology, its capacity to dissipate erosive flow energies, and its overall relative value as habitat for the stream biota.

Completeness. 463 habitat assessments were completed out of 475 planned for a completion rate of 97.5% (Table 3-11).

3.8.2 Laboratory Sorting and Subsampling

Method overview. The subsampling method involved using a 30-square Caton gridded screen, which allows separation of physically-defined amounts of sample material (leaf litter detritus, substrate particles) from the total sample, and then separation/removal of the organisms from that material. Enough gridded squares of material were removed and sorted, in turn, to reach the target number of organisms (200), by the rough count. Once the sort was complete, experienced laboratory personnel examined the remaining detritus to ensure that no organisms had been missed. If missed specimens were found, they were counted and recorded on the subsampling bench sheets. Each sample resulted in 3 "post-sorting" containers: 1) the 200-organism subsample, 2) the unsorted sample remains, and 3) the sample pickate (sort residue).

Precision of sorting and subsampling was not specifically evaluated; the performance characteristic is judged to be not applicable.

Accuracy of subsampling is directly (inversely) related to bias. Specifically, accuracy is not applicable to subsampling or sorting.

Bias of subsampling is evaluated using a performance characteristic similar to % recovery used in analytical chemistry laboratories, called % sorting efficiency, or PSE. An index is not calculated if the final count by the taxonomist is <160 and all 30 grids are sorted (i.e., the entire sample).

Inter-laboratory QC: A set of 54 samples randomly selected by MDEQ was shipped to a separate laboratory. These 54 pickate samples represented 10% of the 535 samples processed by the MDEQ laboratory. The pickate samples were received and examined for any specimens, according to MDEQ-SOP-LAB-001. They were initially assumed to be completely void of benthic macroinvertebrates. The QC laboratory performed sort re-checks under the same conditions as were used in the MDEQ Laboratory, no magnification (naked eye only), and additional artificial lighting, only if necessary. If organisms were found, they were removed and placed in a vial containing approximately 80% ethanol, and labeled with all of the originally required label data, and designated "pickate recoveries". When the pickate check was completed, the number of recoveries was noted on a data sheet. *Sorting efficiency* for a sample was calculated as:

$$\frac{A}{A+B} x100$$

where, A is the number of organisms found by the original sorter, and B is the number of missed organisms recovered by the QC laboratory sort checker. The laboratory sorting/subsampling measurement quality objective (MQO) for this project was to have a database *where ≤10% of the samples overall have a sort efficiency of <90%*.

Results. Thirteen (13) of the 54 samples failed; that is a 24% rate of failure of the 90% sorting efficiency threshold. This rate of failure exceeds the threshold by over 14 percentage points. Figure 3-17 is a control chart of the resulting sorting efficiencies from the 54 pickate samples.

The next step was to determine whether any pattern existed in the failures. Several potential factors were examined that may have effected the final sort efficiency: primary and secondary sorters, primary and secondary sort checkers, number of samples processed by individual sorters or checkers, number of grids sorted, date/day of subsampling start on a sample, and whether or not QC checks were performed on a sample. In some cases, sample sorting was begun by one sorter, but was completed by another, and those samples are shown as being completed by multiple laboratory staff. Likewise, the in-house QC check of the pickate occasionally had multiple staff checking the sort residue of a single sample.

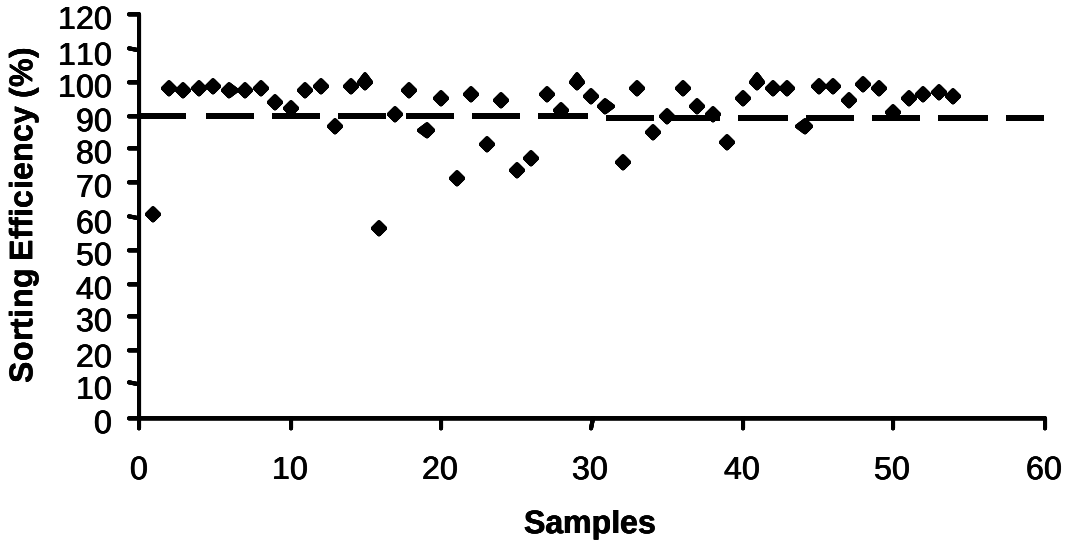


Figure 3-17. Control chart comparing per sample sorting efficiencies with the 90% threshold established for this project.

There did not appear to be a discernable pattern among the results of the checks. With the rate of sort efficiency failures being 24%, higher than the measurement quality objective of 10%, a corrective action was implemented. The corrective action required that the sort residue for all remaining samples be checked, and any specimen recoveries be added to the samples.

Corrective Action. The QC laboratory was tasked with picking all remaining organisms from each of the pickate samples, having them processed for taxonomic identification, and combining the results with the original sample data.

There were a total of 12,988 organisms found (termed "pickate recoveries") in the 515 pickate samples for an average of 25.2 missed organisms per sample. Examining the number of grids picked during the subsampling procedure, 62 samples had all 30 grids picked (in other terms, the entire sample), or 12% of the entire dataset. In the original re-check of the 10% randomly selected pickate samples, it was noted that there seemed to be a tendency for increasing sorting efficiency failure as larger numbers of grids were sorted. Examining the entire dataset exhibits a similar pattern: for samples requiring 10 or fewer grids to reach the 200 organism subsample target, there was a 24% rate of failure; for those requiring greater than 10 (up to 30), 59% of the samples failed. This could possibly be explained by efforts to expedite reaching the target by

rapid picking of only larger, more obvious organisms, and potentially overlooking more numerous, smaller ones. An additional possible explanation would be placing too much material (i. e., too many grids'-worth of detritus) into the sorting tray at one time, thus reducing the ability of the sorter to see the organisms; or that samples supposedly requiring more material to be sorted would be related to the smaller density of organisms in the sample, thus resulting in impatience developing in the sorter. Figure 3-18 shows the breakdown of the number of grids initially sorted to reach the target number. There was also no pattern of failure apparent relative to subsampling being performed early or late in the subsampling period.

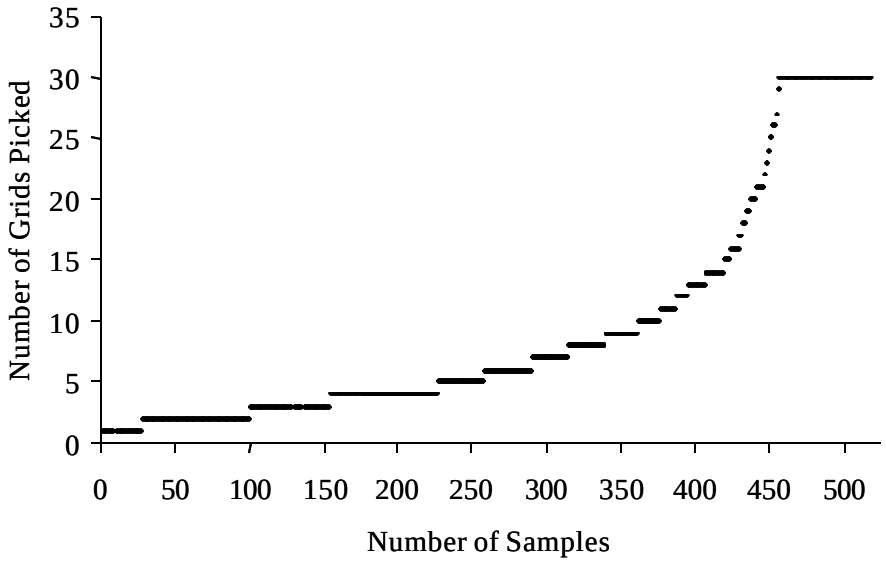


Figure 3-18. Number of grids required to attain the 200 organism subsample target level.

For 278 samples, addition of the pickate recoveries to the original subsamples resulted in a sample total in excess of 240 (200 organism target plus 20%), some even up to 1000. Rarefaction was used on "taxa richness"-based metrics to computer subsample to 240 organisms. These metrics included those that require counts of numbers of different taxa (either taxonomic, such as No. of Ephemeroptera taxa; functional-feeding-group-based, such as No. of filterer taxa; or habit-based, such as No. of clinger taxa).

Representativeness. Two aspects of the sample handling and laboratory processing method in part, ensure representativeness. First, the initial laboratory handling of the sample, specifically the effort to thoroughly mix the sample in a bucket by swirling in a water-filled bucket, and, second, the randomization process for original selection of grids for sorting. An important aspect of subsampling representativeness would be whether those samples where the 200 organisms level was attained in a low number of grids (e.g., 1 or 2). If the sample was well mixed prior to spreading, it is possible that the selected grid(s) are not characteristic of the sample overall.

There were 25 and 72 samples that attained the subsampling target in 1 and 2 grids, respectively. This was not evaluated.

Completeness. Not applicable.

3.8.3 Taxonomic Identification and Enumeration

Method overview. Identifications were performed by a taxonomic laboratory (Freshwater Benthic Services, Inc.) using the most up-to-date technical literature. Taxonomy was performed to hierarchical levels as specified in the MDEQ QAPP (MDEQ 2001), mostly to genus, some to species, and others to higher levels (i. e., tribe, subfamily, family, order, or class). Ten percent (10%) of the project samples (n=535) were randomly chosen by MDEQ for re-identification, resulting in 54 samples. Once the primary identifications were completed for all 54 samples, the vials and slides were shipped in return to the MDEQ lab. They were sent with site information only (i. e., without identifications), thus representing blind samples. The MDEQ lab performed re-identifications. Another aspect of sample processing that is related to and affected by taxonomy is enumeration, or the direct counts of individuals in a sample, both in total and separated by individual taxa.

Precision. The 54 randomly-selected samples are the properties that were measured using two different "methods", the taxonomists. Enumeration is performed simultaneously with identification.

Enumeration. Final specimen counts for samples are dependent on the taxonomic identifications (ID), not the rough counts obtained during the initial sorting activity. Comparison of counts uses "Percent Difference", where

$$(\left|Lab1 - Lab2\right| / Lab1 + Lab2\ ) \text{ x } 100$$

Although there were several samples where total counts are substantially different, most differences were low (Figure 3-19), with a mean of 4.7%. Different counts seemed to have mostly originated from differences in slide-mounted worms and midges, some apparently having cleared to the point of not being visible to the second lab. There were a number of instances where specimens were lost or misplaced during sample handling. Overall, the differences in counts, while initially problematic, does not appear to present a serious problem with the lab processing. Nonetheless, procedures should be investigated that would allow maintenance of sample integrity during both initial and follow-up processing.

Taxonomy. Side-by-side comparison between the taxonomic results delivered by the two labs was performed. The process entailed examination of the list of names for each sample and the number of organisms each lab found for each name. For each sample, the number of agreements was determined, divided by the number of comparisons, and subtracted from 1 to give *percent taxonomic disagreement*, or PTD. Precision of taxonomic identification was assessed by

comparing genus-level taxonomic results from two independent taxonomists, and was calculated as:

$$PTD = \left[1 - \left(\frac{comp_{pos}}{comp_{tot}}\right)\right] \times 100$$

where $\underline{comp}_{pos}$ is the number of agreements, and $\underline{comp}_{tot}$ is the total number of taxonomic comparisons. The lower the PTD value, the more similar are sample taxonomic results, and the greater is the overall taxonomic precision.

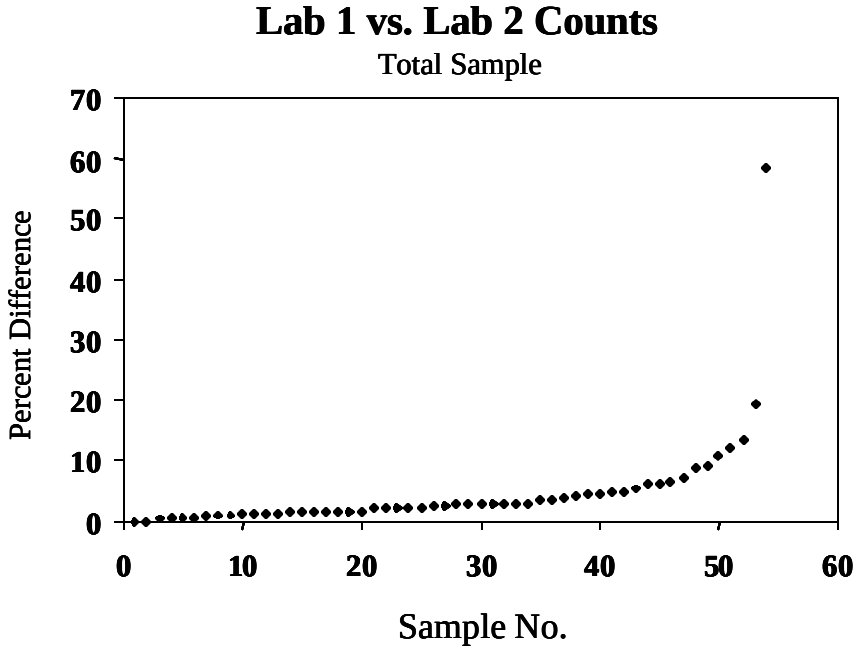


Figure 3-19. Comparison of sample enumeration for 54 samples by two laboratories. The mean difference is 4.7%.

This number quantifies the precision with which the taxonomic database is developed. The original comparison resulted in a mean PTD of 26%, well above the project goal (measurement quality objective) of <15% for the overall dataset. Further examination of the lists revealed several areas of consistent disagreement, which, if combined or aggregated to higher taxonomic levels, would substantially lower the rate of disagreement. Several of these combinations were performed and the PTD calculated for each (Figure 3-20). By aggregating selected chironomid, amphipod, and oligochaete taxa in Composite 5 mean PTD improved from 26% in the original to

11% in the fifth scenario. The groups and taxonomic levels where there seemed to have been the more frequent and major disagreements between the two labs are:

- Amphipoda genera
- Oligochaeta genera
- Chironomidae
    - Psilometriocnemus vs. Hydrobaenus vs. Parametriocnemus
    - Cricotopus vs. Orthocladius vs. Cricotopus/Orthocladius
    - Polypedilum species
    - Rheotanytarsus vs. Paratanytarsus

The original taxonomy was used in all analyses.

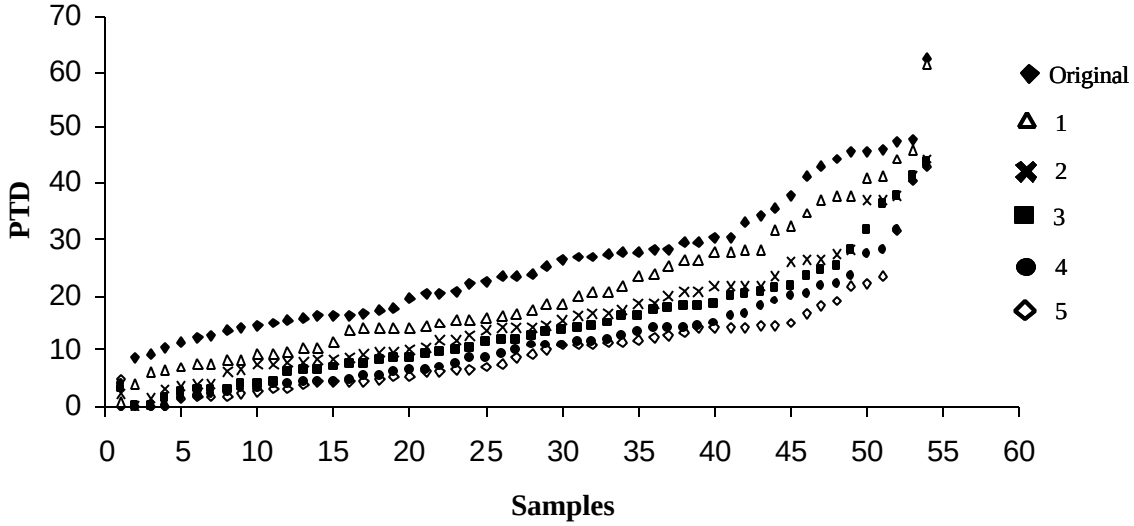


Figure 3-20. Percent taxonomic disagreement (PTD) between two laboratories for five different scenarios of aggregating identifications to higher taxonomic levels.

Accuracy. Definition of accuracy requires specification of an analytical truth (Taylor 1988, Clark and Whitfield 1994). For taxonomy that could be 1) the most up-to-date technical literature/keys, 2) an identified reference collection verified by specialists in different taxonomic groups, or 3) specimen by specimen comparison with museum-based type specimens. All taxonomy in this project was completed using technical literature specified in the QAPP (MDEQ 2001). The reference collection assembled by Freshwater Benthic Services, Inc. for this project contains specimens representing 562 total taxa, and is housed in the MDEQ Biology Laboratory, Pearl, Mississippi. Specialists in several groups will verify selected individuals of different taxa, as decided upon by the Biology Laboratory staff. Option 3 is not feasible, nor considered necessary, for this project.

Bias. This type of error in taxonomy would be problematic if there were consistent misinterpretation of technical keys, misunderstanding of morphological features, or poor processing of samples (including slide mounts of Chironomidae and Oligochaeta). Occasional problems with poor slide mounts were noted, but the extent to which these effected error in the taxonomic analysis was not evaluated.

Representativeness. Not applicable.

Completeness. Completeness of taxonomic analyses is dependent on how well the taxonomist is able to determine the identity of individual specimens, and the frequency of attainment of the targeted hierarchical level. For example, if the final resulting ID for a specimen was at the family level, where the QAPP called for genus level as the target, then that could be said to be a non-complete identification for that taxon. The reason it was left at a more coarse level might have been that it was an early instar with underdeveloped morphological features, or a damaged or poorly mounted specimen. This aspect of the taxonomy was not evaluated.

3.8.4 Data Entry

Method Overview. All data were entered into EDAS (Ecological Data Application System, version 3.0, MSAccess 97, customized for MDEQ). Data types entered included header information, comments, Section 1 riparian zone/instream features, sediment/substrate, water quality, habitat types, habitat assessment, pebble count, taxonomic data, and analytical and field chemistry. There were a total of 377 data entries per site/sample, and 201,695 total for the project (n=535 samples).

Precision. Not applicable.

Accuracy. The accuracy of the data entry was checked by direct comparison of original datasheets (handwritten in the field or laboratory) with printouts from the database. All data entries (100%) were checked by an individual *other than* the primary data entry technician. Notations on the initial printouts were kept when data entry errors were discovered, and marked when corrections were made in the database. To develop an estimate of the rate of data entry error, 50 sites were randomly selected and the (corrected) errors totaled. There were a total of 279 errors discovered and corrected during this QC check, a rate of 1.5%. The incidence of error was greatest for the pebble count data (15.3%); and the rate of error least for sediment/substrate; habitat types; and analytical and field chemistry (0%). All errors were corrected.

Bias. Not applicable.

Representativeness. Not applicable.

Completeness. Not applicable.

3.8.5 Metric Calculation

Method Overview. In structuring the biological portion of the database, it was necessary to relate several sources of non-primary, or secondary, data to each taxon. Three tables were developed that organized tolerance values, functional feeding groups, and habit, and are contained within EDAS. Tolerance values were developed as described in Section 2.1.7 and Appendix A. Functional feeding group and habit designations were taken primarily from Merritt and Cummins (1996) and Barbour et al. (1999). Eighty-two metrics were calculated for each of 524 samples using structured queries in EDAS.

Precision. Not Applicable.

Accuracy. A subset of metric values was hand-calculated using only the taxonomic and enumeration data, and then comparing them to those that resulted from the EDAS queries. The purpose of this QC activity is to ensure that the metric calculation queries are performing operations as intended. It resulted in 695 metric values being recalculated by hand out of 42,968 values. If differences were found, each value was checked for error in the calculation process (hand calculator vs. computer algorithm), and corrections made.

The framework for this QC procedure goes through three steps, and resulted in pattern that was a combination of systematic and random characteristics (Figure 3-21). Step 1 selected one metric for a multiple samples (systematic, every third sample, 154 calculations); Step 2 was a recalculation of 82 metrics for a single site/sample (82 calculations); and Step 3 was "diagonal" through the dataset, so that every site had at least one metric calculated, some had multiple values calculated (459 calculations). For Step 1 the HBI calculation was selected as it represents one of the more complicated queries with greater potential for error. Site 357 was randomly selected for Step 2. The pool of samples to check (n=454) excluded organism re-identifications, field duplicates, and field replicates.

Step 1. Of the 154 calculations 19 were labeled as incorrect (12.3%). Upon calculation by a second individual 11 calculations were found to be correct and the other 8 were correct to three significant digits. Corrective Action: None.

Step 2. Eight of 82 calculations had errors (9.7%) and it was determined that there were problems in the database calculations tied to tolerance values. The core metrics affected were Beck's Biotic Index, Hilsenhoff Biotic Index, and Percent Tolerant Individuals. Corrective Action: The queries were corrected in the database and these new values were then subjected to the same QC check. The re-check of the miscalculated metric values confirmed that the problems were corrected.

Step 3. Six values of 454 were labeled as being in error (1.3%). Upon calculation by a second individual only 4 calculations were found to be in error. The affected metrics and samples were % Non Insects (Site 20), % EPT no caenids (Sites 184 and 261), and % Ephemeroptera no caenids (Site 335). In each of these cases, one individual of the genus Haemonais (a worm) was incorrectly mapped to the genus level ID for Habrophlebiodes (a mayfly). The genus Haemonais occurs in the database as 3 different identifications (Haemonais, Haemonais waldvogeli, and

Haemonais variant) in a total of 38 instances in 37 different samples. The mapping error was fixed in the database. This caused the only minor discrepancies in the calculations. The initial sites/metrics

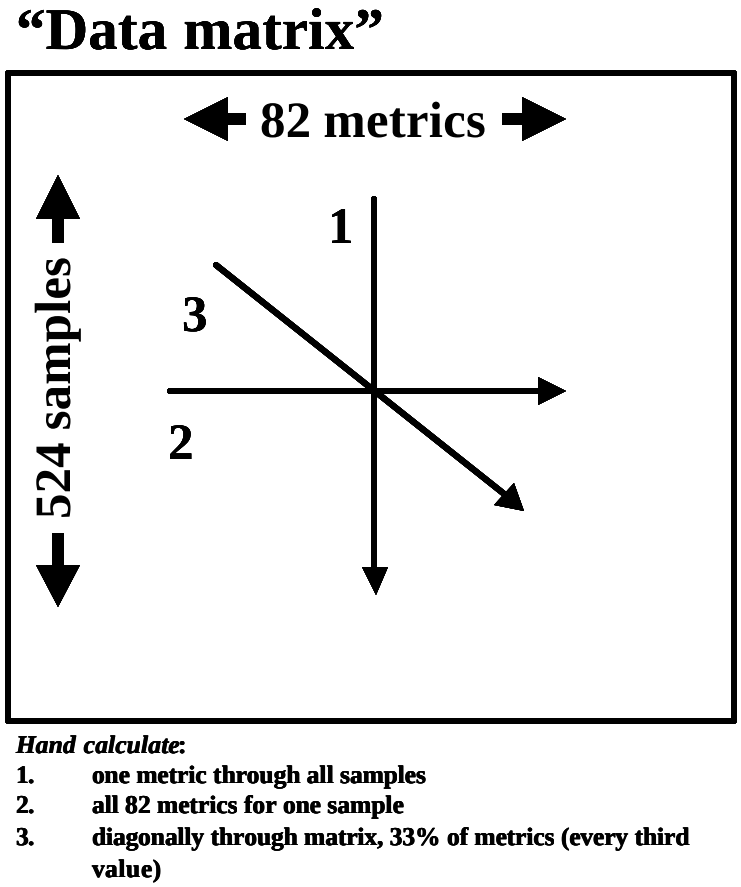


Figure 3-21. Pattern for selecting cells in the data matrix to recalculate by hand; it results in 414 values out of 31,030 being recalculated. This QC check procedure ensures that the interaction between metric calculation queries and raw data is performing as expected.

A total of 695 calculations were checked out of a possible 42,968 (1.6%). Of the 695 calculations checked 11 had errors (1.6%) that were subsequently corrected.

Bias. Not Applicable.

Representativeness. Not Applicable.

Completeness. Not Applicable.

3.8.6 Final Index (M-BISQ) and Site Assessment

Method Overview. The final index is an aggregation of metrics. Two kinds of repeat sampling (intra-team bioduplicates and inter-team biorepeats) provided data to calculate estimates of variance or precision (relative percent difference, coefficient of variability, and detectable difference) at both the metric and index levels. Objective definition of MD sites, and testing the capacity for metrics and indices to detect those sites as degraded (using discrimination efficiency) is characterization of index accuracy.

Precision. Table 3-8 and Figure 3-12 show the results of all repeat sampling on metric and overall index precision. Ten metrics demonstrated good precision (repeatability) with CV<30%; they are: Beck's Biotic Index, Hilsenhoff Biotic Index, Total Taxa, No. Chironomidae Taxa, Percent Diptera, No. Collector Taxa, No. Filter Taxa, No. Predator Taxa, Percent Clingers, and No. Sprawler Taxa. Six metrics had a CV>50% (No. Plecoptera Taxa, No. Trichoptera Taxa, Percent Tanytarsini, Percent Caenidae, and Hydropsychidae/Trichoptera). The overall index had a CV of 10.3% and a 90% confidence interval of "10.0 index units.

Accuracy. The analytical truth used for calculating accuracy of the M-BISQ was the number of sites designated as "MD" using physical and chemical data. The percentage of designations where MD sites were correctly identified as degraded by the M-BISQ is the discrimination efficiency (DE) (see sections 2.6 and 2.7.2 for discussion of DE). If an index correctly categorized all sites as biologically degraded, it can be said to have an accuracy of 100%; 15 out of 30 would be an accuracy of 50%; and so forth. Thus, accuracy calculations must be performed for each site class since the analytical truth is the set of MD sites designated for each class. The accuracy of the M-BISQ is 90% for the Northwest bioregion, 100% for the Black Belt, 89% for the Northeast, and 90% for the West and East bioregions, respectively (Table 3-6).

Bias. An artifact of calculating DE is that high values (e. g., between 95-100%) can be associated with low numbers of MD sites. That is, if a dataset has a high number of MD sites, and also a high DE, confidence can be placed in the result. Conversely, if a high DE is obtained with a low number of sites, the result should be accepted only with lower confidence. The Black Belt and the Northeast bioregions only had 26 and 37 sites, and DEs of 100 and 89.

Representativeness. These biological assessments must be discussed first in terms of scale: areal and site-specific. In this dataset, the percentage of sites within a watershed, bioregion, or across the state, should not be considered representative of all streams or watersheds within that group. A large proportion of the streams (if not all of them) were selected based on some existing knowledge, expected land cover conditions, or their status relative to Mississippi's §303(d) list of impaired waters. For these stream assessments to be considered representative of a broader area than the stream itself, and thus be able to be combined into a mean or median watershed (or other areal) condition, the site selection process would need to be random or stratified-random.

However, they can be considered representative of the individual stream because of the manner in which samples were taken, that is, the field collection procedure is designed to sample the benthic macroinvertebrate fauna the stream physical habitat has the capacity to support (see

section 3.8.1.1). Sampling effort is not intentionally skewed toward an individual habitat type; it is distributed across specific habitat types in proportion to their occurrence within a reach. Also, direct interpretation of the results is in the context of best attainable conditions within a regional stream type.

The index score was not calculated if the final count for a subsample was <160 organisms and all 30 grids had been sorted (i.e., the entire sample). This is intended to minimize the bias that may be associated with performing assessments with inadequate samples and data.

Completeness. Biological assessments were completed for 95% of the 455 streams sampled. Inadequate numbers of organisms (<160) prevented assessments from being completed at 22 sites.

4. BIOREGIONAL SUMMARIES

4.1 East

The East bioregion, composed of seven ecoregions (Table 3-4), is the largest of the five bioregions and had the most sample sites (205 sites) (Figure 2-1). Physical habitat and chemistry are variable within this bioregion as evidenced by the number of preliminary site classes that are contained with this bioregion (Table 3-4). In particular, the southern part of the bioregion is characterized by an abundance of low pH blackwater streams. Chemical parameters including, COD, NN, TKN, TOC, and TP were highest in the central part of this bioregion. Generally, however, the loam and clay soils tend to be leached and, thus, most areas have low fertility (ADEM/MDEQ, 1995). Stream substrate consisted of higher amounts of gravel in the southwestern part of the east bioregion, while silt/clay was prominent in the central region. Overall, though, sand was the most prevalent substrate type (Median = 66%) (Appendix F). Surrounding natural land uses were more abundant and physical habitat was of higher quality in the east bioregion than the other bioregions (Figure 4-1 and Figure 4-2). The highest index scores in this bioregion were found in the southern half with the exception of the streams in the coastal region in the far south which had low index scores. Biologically-impacted sites were more abundant in the northern part of the bioregion.

The most degraded site in the east bioregion was Lewis Creek (site 174), which had a M-BISQ score of 11 (Appendix G). This stream is located in the northwest part of the bioregion and has a highly modified riparian zone made up of 74 percent managed land uses (i.e., anthropogenic land uses) (Appendix F). The least degraded stream found in the east bioregion was Tilton Creek (site 464), which had an M-BISQ score of 92. Located in the southern half of this bioregion, this stream had high quality habitat (177) and abundant gravel (56%) (Appendix F). Several LDa sites in the east bioregion had relatively low index scores. These sites included the Strong River (site 319; M-BISQ = 49) and Pinishook Creek (site 272; M-BISQ = 49), both located in the southern, middle, and northern areas of this bioregion, respectively.

Dead Tiger Creek (site 521) and Scooba Creek (site 566) had low M-BISQ scores (33 and 28, respectively) but high habitat scores (129 and 146, respectively). Standing Pine Creek (site 262) and Pretty Branch (site 396), both of which were classed as MD sites, had high M-BISQ scores (70 and 80, respectively) but poor habitat quality. Nine samples from this bioregion had insufficient data and thus could not be assessed[1] (Appendix G).

4.2 West

The West bioregion (Figure 3-6) is represented by ecoregions 74a, 74c and the southern half of 74b and contained 96 sample sites (Table 3-4). The northern part of this bioregion is more heavily human-influenced, mostly in the form of agricultural lands, than the southern part. This northern section is the preliminary site class 5 which was designated as a separate site class because of differences in chemical and physical characteristics. Additional biological data for LDa sites from this region may suggest that it should be designated as a separate bioregion.

[1] Samples contained less than the 160-organism target level for site assessment. Index scores were calculated but results were not be used for evaluation of impairment.

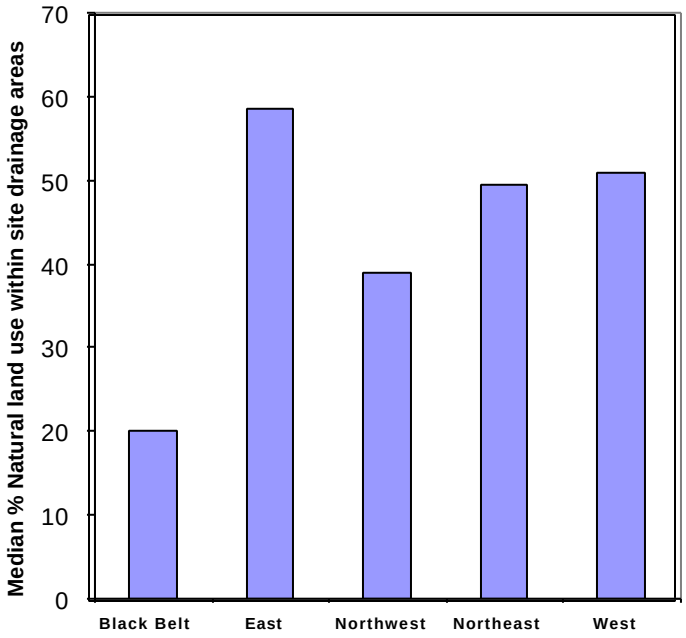


Figure 4-1. Median percent natural land use (i.e., forest and wetland) found in the five bioregions.

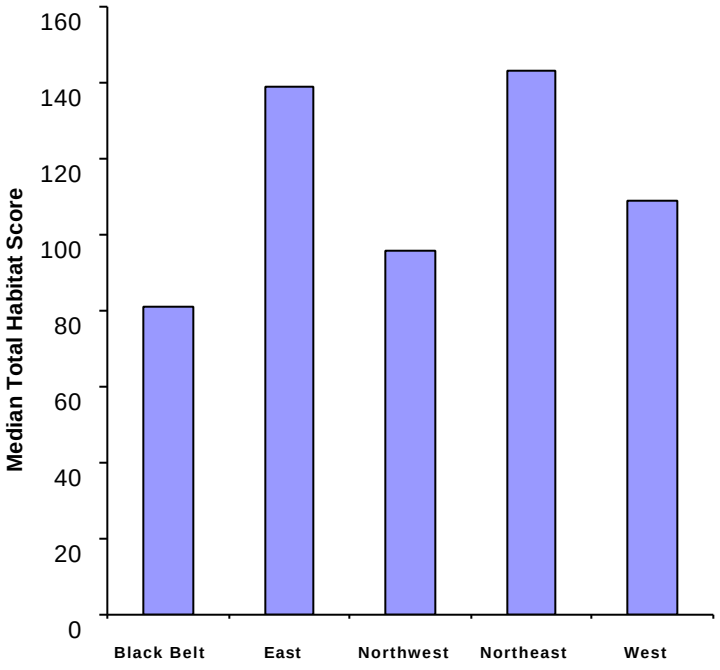


Figure 4-2. Median total habitat scores from five bioregions.

Qualitative field observations of physical and biological conditions at and around sites in this region suggest that this region may be distinct from the rest of the west bioregion, however, present data do not support that conclusion. The southern section of this bioregion has more forested areas, however, logging and associated runoff and erosion in some national forests poses a threat to stream integrity. The western portion of this bioregion has higher levels of chlorides, specific conductance, alkalinity, and pH than the rest of the bioregion. These elevated chemical measurements are likely due to inputs of brine historically used in oil drilling common to this region. Chemical parameters including TOC, TKN, COD, and N+N were all higher in the north likely due to the prominence of agricultural lands in this region. Habitat quality was lower in the north and silt/clay substrate was prominent. Gravel substrate was also more abundant in the southern and far western parts than in the northern half. Sand was the most prevalent substrate found in this bioregion (median = 52%) (Appendix F).

The most impacted site in the west bioregion was Hays Creek (site 163; M-BISQ = 25) located in the far northeastern portion of the bioregion (Table 3-4). The stream was surrounded by mostly agricultural lands and had poor habitat quality (Appendix F). The least degraded site in this bioregion was Brushy Creek (site 371; M-BISQ = 88), located in the southern half of the bioregion. This stream had a moderate habitat score (108); however, gravel was an abundant component of the substrate (41%). Bayou Pierre (site 357) and Porter Creek (site 300) both located in the central part of the bioregion, were classified as MD sites because of a large proportion of managed land within riparian corridors and low habitat scores; however, these sites had relatively high M-BISQ scores (57 and 62, respectively). It is possible that these surrounding land uses may have improved since the land use data layer was developed or that the habitat was lower than what the scores suggested (i.e., scores were at the lower range of the ± 24 point confidence interval). Ford's Creek (site 327) and Big Creek (site 305) both had high habitat scores but low M-BISQ scores (both 38). Three LDa sites including the East Fork Amite River (site 553) located in the south, and Dowd Creek (site 362) and Limekiln Creek (site 298) both located in the central part of the bioregion, had relatively low index scores (M-BISQ=57, 52, and 52, respectively). Four samples in this bioregion had insufficient biological data (<160 organism count), therefore, M-BISQ scores could not be used for assessment (Appendix G).

4.3 Northwest

This area of the state has experienced many years of intensive and widespread farming, deforestation, and direct alterations to stream channels (Thorne 1997, Watson et al. 1997, Van Wilson 1997, and Shields et al. 1998). Many streams in this region (Figure 3-6) are entrenched due to extensive and severe downcutting that resulted from historic channelization of major rivers. Ongoing channel adjustment is apparent throughout the region and is evidenced by severe incisions, widespread bank instability and mass wasting, channel widening, and alternate aggradation and degradation of stream bottoms (Shields et al. 1998, Thorne 1997). As part of these geomorphic processes, headcuts are migrating upstream in many watersheds, and extreme in-channel bank and bed erosion is leading to several hundred thousand tons of sediment being mobilized (Simon and Darby 1997, Grissenger and Murphy 1986). The scarcity of LDa quality sites in this region made it difficult to assess natural variability among the different sites;; however, as more data from sites in this bioregion are collected, it may be possible to detect natural variation and further refine the current bioregion. Chemical parameters, including TP,

COD, TKN, N+N, and specific conductance, were highest in the northwest bioregion. Specific conductance was highest in the far eastern portion of this bioregion. Habitat conditions were poorest in the east and gravel substrate was more abundant in the far western part of this bioregion. Sand was the most abundant substrate (median = 74%) (Appendix F). The biologically least- disturbed streams are found in the center of the bioregion, while the most degraded are found in the east, northwest, and south.

The most degraded site was McIvor Canal (site 89; M-BISQ = 6) located in the western half of this bioregion. This stream was surrounded by mostly managed land and had poor habitat quality (Appendix F). The least disturbed site was Little Spring Creek (site 34; M-BISQ = 88) located in the central part of the bioregion. This stream had high quality habitat (142) and had a high percentage of natural land use within riparian corridors. Several MD sites including Little Tallahatchie River (site 55), Yocona River (site 112), and Hudson Creek (site 87) had relatively high M-BISQ scores (67, 64, and 61, respectively). All of these sites had high percentages of anthropogenic land uses within riparian corridors and relatively low total habitat scores. Two LDa sites, Hickahala Creek (site 18) located in the north and Cane Creek (site 158) located in the south, had relatively low index scores (M-BISQ = 50 and 47, respectively).

Duncans Creek (site 110) had one of the lowest M-BISQ scores (14) in this bioregion, however, habitat was relatively high (116). White's Creek (site 3) also had a high habitat score (150), but low M-BISQ score (31). Pigeon Roost Creek (site 13) and Clear Creek (site 86) both had high M-BISQ scores (74 and 72, respectively), however, habitat scores were low (92 and 85, respectively). Five samples from this bioregion could not be assessed due to low organism numbers (Table 3-4).

4.4 Black Belt

The Blackland Prairie (ecoregion 65a), or Black Belt, is distinctly different from other areas in this part of the state (Figure 3-6), and is characterized by chalk bedrock with a thin soil overburden (Hicks and Haynes 2000). Flat agricultural lands, catfish ponds, and channelized, highly entrenched streams characterize this bioregion. The soils are composed of chalks and marls making them dark and nutrient rich. Historically, streams in this region have been recorded as having high turbidity and alkalinity, which was supported by field and analytical chemistry gathered in this study. Conductivity and alkalinity were higher in this bioregion than in surrounding areas and habitat quality was generally poor (Figure 4-2). Sand was the most prevalent substrate (median = 38%), however, silt/clay was also abundant (median = 35%).

The most impaired stream was Hang Kettle Creek (site 195; M-BISQ = 30) located in centrally (Appendix G). This stream, like many in this bioregion, was surrounded by agricultural lands and had poor physical habitat (Appendix F). Additionally, the substrate was composed of mostly silt/clay (85%). The three LDa sites in this bioregion were the three best sites as measured by index scores in this bioregion. Two of these, Tallabinella Creek (site 129; M-BISQ=84) and Spring Creek (site 196; M-BISQ=94), both located in the central area, had large sections of stream bed that were composed of hard pan clay. The other LDa stream, Ash Creek (site 285; M-BISQ=82), is located in the far south near the border with the east bioregion.

Catalpa Creek (site 207) had one of the highest habitat scores in the Black Belt bioregion; however, the M-BISQ score was one of the lowest (35). Trim Cane Creek (site 188), Chiwapa Creek (site 568), and Tuscumbia River Canal (site 548) all had low habitat scores but had some of the highest M-BISQ scores (73, 72, and 64, respectively). Three sites in the Black Belt had insufficient biological data (i.e., sample numbers less than 160 organisms) and, thus, could not be assessed (Appendix G).

4.5 Northeast

The Northeast bioregion (Figure 3-6) is composed of ecoregions 65b, i, and j, and is characterized by rolling hills and transitional areas to the Blackland Prairie. The far northeast portion of this bioregion has the most topographic relief and the streams contain more gravel and cobble than others in the state (median = 19%). The rest of the bioregion is flatter with more agricultural lands with streams exhibiting poorer habitat, less gravel and cobble, and more sand. Overall, sand was the most prevalent substrate (median = 60%) (Appendix F). Most of the sites with high index scores are located in the east, while most of the degraded sites are located in the west.

The most disturbed stream in the Northeast bioregion was Twentymile Creek (site 80; index = 5) located along the border with the Black Belt bioregion. This site was surrounded by mostly anthropogenic land and had poor habitat quality (Appendix F). The least degraded site was an unnamed tributary to the Tennessee-Tombigbee waterway (site 65; index=75) located in the northeastern section of this bioregion. This site had a high percentage of surrounding natural land and a high physical habitat score relative to other sites in the Northeast (Appendix F).

Indian Creek (site 66) had high quality habitat but a low M-BISQ score (29). Little Yellow Creek (site 64) had poor habitat quality and was classified as a MD site, however, the M-BISQ score was one of the highest in this bioregion (73). One site in this bioregion could not be assessed due to low organism numbers (Appendix G). One LDa site, Yellow Creek (site 205), located in the southern portion of this bioregion, had a fairly low M-BISQ score (54).

4.6 Importance of Error

For the sites where habitat quality and M-BISQ score did not appear to correspond (e.g., high index, low habitat) it is important to recognize that error in habitat assessments or biological sampling or processing may be a factor contributing to discrepancies. QA/QC procedures were used to reduce error, however, precision estimates such as the 90% confidence interval show variability between habitat and biological data. In cases where the maximum variability in index and/or habitat variability occurs, habitat and index scores may not correspond, simply due to this variability. For instance, in cases where habitat score was high but index score was low, it is possible the habitat score was at the high end of the ±24, 90% confidence interval (Table 3-9) and that the M-BISQ score was at the low end of ±10.0 confidence interval resulting in a discrepancy due at least in part to expected data variability, as opposed to an ecological effect such as chemical stress.

5. DISCUSSION

5.1 Shortcomings of indices

Index performance may also be related to the quality and quantity of LDa and MD sites found in each bioregion. The Black Belt and Northeast bioregions, in particular, have few LDa sites, which could inhibit selection of the most discriminating metrics because of the potential for random error among LDa or MD sites. The more sites available for investigating metric performance, the less potential there is for a few sites to influence the overall LDa and MD site metric value distributions. Quality of LDa sites may also influence the performance of metrics and indices in areas such as the Northwest bioregion where LDa sites represent "best attainable" conditions, as opposed to more natural conditions. The intensity of degradation found at MD sites can also affect our ability to select the most efficient metrics. In bioregions such as the East, where highly degraded sites are relatively rare, the difference between LDa and MD sites may not be as great as in other areas. This occurrence can make it more difficult to evaluate the discriminatory ability of metrics and, thus, more difficult to choose the best metrics.

Despite these types of shortcomings in metric and index performance, the data presented in this report indicate that the indices in all bioregions were able to detect impairment. All five indices exhibited distinct separations between LDa and MD sites indicating they were performing correctly; however, the distance of separation varied among bioregions (Figure 3-9). The strength of separation between LDa and MD boxplots is *directly* affected by how good the LDa sites are, and, how bad the MD sites are. Because there are ranges of variability in both, there will be differences in the magnitude of separation

5.2 Potential future analyses

To confirm that current indices were selected appropriately and that they are correctly identifying degraded streams an independent dataset should be evaluated. It is recommended that the data from the approximately 100 wadeable stream sites sampled in 2002 be evaluated using the same techniques used in this study as a confirmation of the metrics and indices used for evaluating streams sampled in 2001. This process would involve calculating the appropriate metrics and indices for each sample and comparing DEs to those from the original study.

Tolerance values, bioregions, and indices should all be evaluated for potential revisions as more data are collected. As more data from LDa sites become available, additional biological variations across the state may be evident and may indicate that current bioregions should be divided or re-combined to represent natural variation in biotic assemblages. If new bioregions are developed, additional indices may also then become necessary. Additionally, indices may need to be refined as metric performance characteristics are further analyzed using data from new LDa and MD sites. Tolerance values can also be refined as more physical and chemical data are collected and as stressor gradients are refined.

5.3 Management recommendations

In addition to its focus on use in evaluating streams for CWA §303(d) purposes, the M-BISQ can be used in various other resource management and regulatory activities including: helping to prioritize streams by severity of stressor loads; identifying stressor sources; and providing objective, ecologically-based methods for judging the effectiveness of restoration, TMDL, chemical controls, and other management activities. An important component to developing effective restoration practices is to identify the most critical stressors in degraded waters. This stressor identification process can be performed using the M-BISQ along with abiotic data and should be the next step following §303(d) listing/de-listing. Once stressors have been identified, management activities (e.g., TMDLs) can be geared to address particular stressors. The M-BISQ can then be used to evaluate the effectiveness of management practices.

As monitoring programs continue to gather information over time, databases used to develop and refine biological criteria expand. This means that, potentially, new LDa sites are added, previously under-represented regions of the state become better understood, and definition of MD conditions are refined. The process established here for updating the stream biological database should be repeated with future sampling data. As more data are collected an increased understanding of the natural variability of Mississippi streams and watersheds can be developed. This information should be used to refine LDa and MD criteria, bioregional boundaries, tolerance values, and M-BISQ organization.

6. Literature Cited

ADEM/MDEQ. 1995. Alabama/Mississippi Pilot Reference Site Project, 1990-1994. Alabama Department of Environmental Management, Field Operations Division, Ecological Studies Section; Mississippi Department of Environmental Quality, Office of Pollution Control, Biological Services Section.

Anderson, N. H. and A. S. Cargill. 1987. Nutritional Ecology of Aquatic Detritivorous Insects. Chapter 30, IN F. Slansky and J. Rodriguez (editors), *Nutritional Ecology of Insects, Mites, Spiders and Related Invertebrates*. John Wiley and Sons, New York, NY. 1016 pp.

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 decisionmaking. Lewis Publishers, Boca Raton, FL.

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 strams using benthic macroinvertebrates. Journal of the North American Benthological Society 15(2): 185-211.

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

Buikema, A. L. and J. R. Voshell. 1993. Toxicity studies of freshwater benthic macroinvertebrates. Pages 344-398, IN, D. M. Rosenberg and V. H. Resh (editors). *Freshwater Biomonitoring and Benthic Macroinvertebrates*, Chapter 10. Chapman and Hall, NY.

Chapman, S., J. Omernik, M. Beiser and D. Johnson. 2001. Draft: Ecoregions of Mississippi. U.S. EPA, National Health and Environmental Effects Research Laboratory, Corvallis, OR. [ftp://ftp.epa.gov/wed/ecoregions/ms/ms_eco.e00].

Clark, M. J. R. and P. H. Whitfield. 1994. Conflicting perspectives about detection limits and about the censoring of environmental data. Water Resources Bulletin 30(6): 1063-1079.

Costanza, R., S.O. Funtowicz, and J.R. Ravetz. 1992. Assessing and communicating data quality in policy-relevant research. Environmental Management 16(1): 121-131.

Cummins, K. W. and M. J. Klug. 1979. Feeding ecology of stream invertebrates. Annual Review of Ecology and Systematics 10: 147-172.

Davis, W. S., B. D. Snyder, J. B. Stribling, and C. Stoughton. 1996. Summary of State Biological Assessment Programs for Streams and Wadeable Rivers. EPA 230-R-96-007. U. S. Environmental Protection Agency; Office of Policy, Planning, and Evaluation. Washington, DC.

Diamond, J. M., M. T. Barbour, and J. B. Stribling. 1996. Characterizing and comparing bioassessment methods and their results: a perspective. Journal of the North American Benthological Society 15(4): 713-727.

Fausch, K. D., J. Lyons, J. R. Karr, and P. L. Angermeier. 1990. Fish communities as indicators of environmental degradation. Pp. 123-144, IN S. M. Adams (editor). *Biological Indicators of Stress in Fish*. American Fisheries Society Symposium 8, Bethesda, MD.

Frey, D. G. 1977. Biological integrity of water – an historical approach. Pages 127-140, IN R. K. Ballentine and L. J. Guarraia (editors). *The Integrity of Water*. Proceedings of a Symposium, March 10-12, 1975, U. S. Environmental Protection Agency, Washington, DC.

Gibson, G.A., M.T. Barbour, J.B. Stribling, J. Gerritsen, and J.R. Karr. 1996. *Biological criteria: Technical guidance for streams and rivers*. EPA/822-B-94-001. U.S. EPA, Office of Science and Technology, Washington, DC.

Grissinger, E. H. and J. B. Murphey. 1986. Bank and bed adjustments in a Yazoo bluffline tributary. pp. 1003-1012, *In* Wang, S. S. Y., H. W. Shen, and L. Z. Ding. Proceedings of the Third International Symposium on River Sedimentation. The University of Mississippi, University, MS.

Hicks, M.B. and C.G. Haynes. 2000. An annotated list of Trichoptera in the Black Belt prairie region of West centeral Alabama. Entomological News 111(3): 215-222.

Hill, B.H. 1997. The use of periphyton assemblage data in an index of biotic integrity. Bulletin of the North American Benthological Society 14: 158.

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

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.

ITFM (Intergovernmental Task Force on Monitoring Water Quality). 1995. The Strategy for Improving Water-Quality Monitoring in the United States. Technical Appendixes. Final Report of the Intergovernmental Task Force on Monitoring Water Quality. U. S. Geological Survey OFR 95-742. Reston, VA.

Jongman, R.H.G., C.J.F. ter Braak, and O.F.R. van Tongeren, editors. 1987. Data Analysis in Community and Landscape Ecology. Pudoc Wageningen, Netherlands.

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

Karr, J. R., and E. W. Chu. 1999. *Restoring Life in Running Waters. Better Biological Monitoring*. Island Press. Washington, DC. 206 pp.

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

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

Kenkel, N.C. and L. Orloci. 1986. Applying metric and nonmetric multidimensional scaling to ecological studies: some new results. Ecology 67: 919-928.

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

Kruskal, J.B. 1964. Nonmetric multidimensional scaling: a numerical method. Psychometrika 29: 115-129.

Ludwig, J.A. and J.F. Reynolds. 1988. Statistical Ecology: A Primer on Methods and Computing. John Wiley & Sons, NY.

Maxted, J., M.T. Barbour, J. Gerritsen, V. Poretti, N. Primrose, A. Silvia, D. Penrose, and R. Renfrow. 2000. Assessment framework for mid-Atlantic coastal plain streams using benthic macroinvertebrates. Journal of the North American Benthological Society 19(1):128-144.

McCarron, E. and R. Frydenborg. 1997. The Florida bioassessment program: an agent of change. Human and Ecological Risk Assessment 3(6): 967-977.

McCune, B., and M.J. Mefford. 1995. PC-ORD. Multivariate Analysis of Ecological Data, Version 2.0. MjM Software Design, Gleneden Beach, OR, USA.

MDEQ. 1997. Mississippi Land Cover Project. Prepared by the Space Remote Sensing Center. Stennis Space Center. Stoneville, MS. For Mississippi Department of Environmental Quality. Jackson, MS. (For further information, contact Mr. Gary Hennington [601-961-5183]).

MDEQ. 2001. Quality Assurance Project Plan for 303(d) List Assessment and Calibration of the Index of Biological Integrity for Wadeable Streams in Mississippi. February 15, 2001. Mississippi Department of Environmental Quality, Jackson, MS.

Merritt, R.W. and K.W. Cummins. 1996. *An Introduction to the Aquatic Insects of North America*. 3rd edition. Kendall/Hunt Publishing Co., Dubuque, IA.

National Research Council (NRC). 2001. *Assessing the TMDL Approach to Water Quality Management*. National Academy Press, Washington, C. 109 pp.

Ohio Environmental Protection Agency. 1989. *Addendum to biological criteria for the protection of aquatic life, volume II: users manual for biological field assessment of Ohio surface water.* Ohio EPA, Division of Water Quality Planning and Assessment, Ecological Assessment Section, Columbus, OH.

Omernik, J. M. 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77(1): 118-125.

Omernik, J. M. and G. E. Griffith. 1991. Ecological regions versus hydrologic units: frameworks for managing water quality. Journal of Soil and Water Conservation 46(5): 334-340.

Peters, J. A. 1988. Quality Control Infusion into Stationary Source Sampling. Chapter 22, pages 317-333, IN Lawrence H. Keith (editor), *Principles of Environmental Sampling*. ACS Professional Reference Book. American Chemical Society. Columbus, OH.

Reynoldson, T.B., R.C. Bailey, K.E. Day, and R.H. Norris. 1995. Biological guidelines for freshwater sediment based on **BE**nthic **A**ssessment of **S**edimen**T** (the BEAST) using multivariate approach for predicting biological state. Australian Journal of Ecology (1995) 20: 198-219.

Reynoldson, T.B., R.H. Norris, V.H. Resh, K.E. Day, and D.M. Rosenberg. 1997. The reference condition: A comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. Journal of the North American Benthological Society 16: 833-852.

Schneider, E. D. 1992. Monitoring for Ecological Integrity: the State of the Art. Pages 1403-1419, IN D. H. McKenzie, D. E. Hyatt, and V. J. McDonald (editors). *Ecological Indicators. Proceedings of the International Symposium on Ecological Indicators*. October 16-19, in Ft. Lauderdale, Florida. Elsevier Science Publishers, Ltd. Essex, England.

Shields, F. D., Jr., S. S. Knight, and C. M. Cooper. 1998. Rehabilitation of aquatic habitats in warmwater streams damaged by channel incision in Mississippi. Hydrobiologia 382: 63-86.

Simon, A. and S. E. Darby. 1997. Disturbance, channel evolution and erosion rates: Hotophia Creek, Mississippi. pp. 476-481, IN Wang, S. S. Y., E. J. Langendoen, and F. D. Sheilds, Jr. (editors). *Management of Landscapes Disturbed by Channel Incision: Stabilization, Rehabilitation, and Restoration.* University of Mississippi, University, MS. ISBN 0-937099-05-08.

Southerland, M. T. and J. B. Stribling. 1995. Status of Biological Criteria Development and Implementation. Chapter 7, 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, FL.

Stribling, J. B., B. K. Jessup, J. S. White, D. Boward, and M. Hurd. 1998. *Development of a Benthic Index of Biotic Integrity for Maryland Streams.* CBWD-MANTA - EA-98-3. Maryland Department of Natural Resources, Chesapeake Bay and Watershed Programs. Monitoring and Non-Tidal Assessment Division Annapolis, MD.

Taylor, John K. 1988. Defining the Accuracy, Precision, and Confidence Limits of Sample Data. Chapter 6, pages 102-107, IN Lawrence H. Keith (editor), *Principles of Environmental Sampling*. ACS Professional Reference Book. American Chemical Society. Columbus, OH.

Taylor, B. N. and C. E. Kuyatt. 1994. Guidelines for evaluating and expressing the uncertainty of NIST measurement results. NIST Technical Note 1297. National Institute of Standards and Technology, U.S. Department of Commerce, Washington, DC. 24 pp.

Tetra Tech, Inc. 2000. Ecological Data Application System (EDAS). A User's Manual. Prepared by Tetra Tech, Inc., Owings Mills, Maryland, Prepared for USEPA, Office of Water and Oceans and Wetlands, Washington, DC. *(For more information on this report, contact Robert King [USEPA] at 202-566-1177).*

Thorne, C. R. 1997. Fluvial processes and channel evolution in the incised rivers of north-central Mississippi. pp. 23-34, IN Wang, S. S. Y., E. J. Langendoen, and F. D. Sheilds, Jr. (editors). *Management of Landscapes Disturbed by Channel Incision: Stabilization, Rehabilitation, and Restoration.* University of Mississippi, University, MS. ISBN 0-937099-05-08.

U. S. EPA. 1997. Monitoring Guidance for Determining the Effectiveness of Nonpoint Source Controls. U. S. Environmental Protection Agency, Office of Water. EPA 841-B-96-004.

U.S. EPA. 1999. EPA Requirements for Quality Assurance Project Plans. EPA QA/R-5. U.S. Environmental Protection Agency. QA Division. Washington, DC. Interim Final. November.

Van Wilson, K. 1997. Geomorphic response to channel modifications of Skuna River in north Mississippi. pp. 482-487, IN Wang, S. S. Y., E. J. Langendoen, and F. D. Sheilds, Jr. (editors). *Management of Landscapes Disturbed by Channel Incision: Stabilization, Rehabilitation, and Restoration.* University of Mississippi, University, MS. ISBN 0-937099-05-08.

Wallace, J. B. and J. R. Webster. 1996. The role of macroinvertebrates in stream ecosystem function. Annual Review of Entomology 41: 115-139.

Watson, C. C., N. K. Raphelt, and D. S. Biedenharn. 1997. Historical background of erosion problems in the Yazoo Basin. pp. 115-119, IN Wang, S. S. Y., E. J. Langendoen, and F. D. Sheilds, Jr. (editors). *Management of Landscapes Disturbed by Channel Incision: Stabilization, Rehabilitation, and Restoration.* University of Mississippi, University, MS. ISBN 0-937099-05-08.