Development of Biological Criteria for Wyoming Streams and their Use in the TMDL Process[a]

Benjamin K. Jessup[1], James B. Stribling[2], Jeroen Gerritsen[2], Kurt King[3], and Jack Smith[4]

1.) Tetra Tech, Inc., 98 Ferguson Avenue, Burlington, VT 05401, 2.) Tetra Tech, Inc., 10045 Red Run Blvd., Owings Mills, MD 21117, 3.) Wyoming Department of Environmental Quality, Land Quality Division, Suite D, 1043 Coffeen Ave., Sheridan, WY 82801, 4.) Wyoming Department of Environmental Quality, Water Quality Division, 250 Lincoln St., Lander, WY 82580.

ABSTRACT

The Wyoming Department of Environmental Quality (WYDEQ) initiated development of regionally-calibrated biological criteria for assessing the ecological condition of streams across the state. They intend to use the biological criteria to evaluate the overall ecological condition of their streams and watersheds, and to justify listing and delisting streams for their Total Maximum Daily Load (TMDL) requirements. Approximately 300 stream sites were classified by ecological similarity (membership in groups of relative physically- and chemically-homogeneity) and stream reference status (degree of anthropogenic impact, measured by physical and chemical characteristics). Biological attributes of the "least impacted" sites were assessed to distinguish groups with naturally similar macroinvertebrate taxonomic composition and community structure. Four regions of relative biological homogeneity were identified from among eight ecoregions. Variability of 70 metrics was tested within and among each bioregion and an index was formed from the metrics that responded consistently to stress. Stressed biological communities, identified by lower index scores, were found in 87% of the physically or chemically impacted sites. In the mountains, eight metrics were included in the index: insect taxa, non-insect taxa, Ephemeroptera taxa, percent Ephemeroptera, percent Oligochaeta, percent five dominant taxa, Hilsenhoff Biotic Index, and percent scrapers. In non-mountainous regions, ten metrics were included: total taxa, Ephemeroptera taxa, Plecoptera taxa, Trichoptera taxa, percent Plecoptera, percent non-insects, percent ten dominant taxa, biotic condition index (BCI CTQa), percent collector-gatherers, and scraper taxa. Strategies for applying the index as a biological criterion in natural resource assessment, use support designations, and the TMDL process are suggested.

a.) This paper is condensed from the full report prepared for U. S. EPA Region 8 (Stribling et al. 2000).

Presenter's Biosketch: Benjamin K. Jessup. Mr. Jessup has developed multimetric biotic indices for states throughout the mid-Atlantic, New England, and northern Rocky Mountains with Tetra Tech for 3 years. He has applied multimetric methods not only for state monitoring efforts, but also in risk assessments, impact assessments, and stressor diagnosis. He has compared multimetric and multivariate biotic assessment methods using Wyoming and Idaho data. Mr. Jessup earned his masters degree from the University of Maryland studying trout population dynamics in an urbanizing stream system. Before joining Tetra Tech, his efforts at biotic indicator development focused on aquatic macrophytes in the Chesapeake Bay watershed.

I. INTRODUCTION

Through the 303(d) and Total Maximum Daily Load (TMDL) framework outlined in the Clean Water Act (CWA) of 1972 (PL 92-500), those waters considered to be impaired and threatened must be improved to meet their designated uses. The definition of impairment by natural resource management or regulatory agencies is typically based on attainment or non-attainment of numeric chemical water quality standards. Similar frameworks and definitions of impairment are part of the National Water Quality Inventory (§305[b]), the Nonpoint Source Pollution Prevention program (§319), and are required by the Clean Water Action Plan. If those standards are not met (or attained), then the waterbody is considered to be impaired.

Establishment of ecologically meaningful impairment criteria for biological and physical habitat conditions (as opposed to chemical criteria) is recommended to evaluate use attainment in regard to aquatic life uses. One of the primary goals of the CWA is the maintenance and restoration of biological integrity, which covers biological, physical, and chemical quality. The widely-accepted definition of biological integrity is recognized by both the scientific AND regulatory communities as key to enhancing the protection of natural resources (Schneider 1992). It is defined as the ability of an ecological 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*" (Karr et al. 1986, Gibson et al. 1996).

Operationally, application of an Index of Biological Integrity (IBI; Karr et al. 1986) calibrated on ecoregional and geographically finer patterns (Omernik 1987), provides an objective and repeatable process for judging ecological conditions (Gibson et al. 1996). Multimetric indices, such as the IBI, incorporate multiple biological community characteristics and measure the overall response of the community to environmental stressors (Karr et al. 1986, Barbour et al. 1995). This measure of the structure and function of the biological assemblages (a regionally-calibrated IBI) is an appropriate indicator of ecological quality, reflecting biological responses to changes in physical habitat quality, the integrity of soil and water chemistry, geologic processes, and land use changes (to the degree that they affect the sampled habitat).

Geographically-calibrated, biological, multimetric indices for assessment of ecological conditions have been endorsed by the U.S. EPA (Gibson et al. 1996) and 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 Wyoming is to use biological condition as an indicator of ecological integrity, and for ecological integrity to be the basis for determining stream impairment. Other states have found multimetric indices to be robust in detecting problems that warrant more detailed, diagnostic testing of water column and sediment toxicity or aquatic chemistry (McCarron and Frydenborg 1997).

Over a five-year period (1993-97), the Wyoming Department of Environmental Quality (WY DEQ) has collected data on biological (primarily benthic macroinvertebrates), instream physical habitat, stream channel form, surrounding land use characteristics, and field and analytical chemistry from 301 sites statewide as part of its Reference Stream Project. This paper identifies preliminary biological and physical habitat criteria for streams in each of the five primary Wyoming ecoregions, discusses use of these criteria for determining stream impairment/non-impairment, and makes recommendations on the use of specific biological attributes to guide development and evaluation of total maximum daily loads (TMDLs) for nonpoint source pollutants. This document also demonstrates the potential use of the criteria in determining the attainment of designated aquatic life uses.

II. METHODS

The approach used in constructing a regionally-calibrated multimetric biological index follows 7 basic steps:

1) Develop database
2) Determine appropriate strata
3) Establish numeric criteria for reference and degraded streams
4) Compile metrics
5) Determine naturally occurring bioregional delineations
6) Test metrics
7) Combine metrics into a regionally-calibrated index

1. Database development

Biological, habitat, and water quality data received from WY DEQ were entered into EDAS (Ecological Database Application System, version 2.0 [Tetra Tech 1999]) for continual data management and ongoing analysis. Data, metadata, and other ancillary information reside in a series of relational tables specific to stations, samples, benthic taxa, chemistry, and habitat. During database development, inconsistencies in collection methods were identified and associated samples were eliminated. Preliminary data structure was examined for bias resulting from data collection dates, site elevation, stream channel morphology, and membership within a seemingly distinct sub-ecoregion in the southern extension of the Middle Rockies West.

2. Determine appropriate strata

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 were examined before looking at biological heterogeneity. The geographic framework for delineating regions of *relatively* uniform water chemistry and instream habitat features are Level 3 ecoregions (Omernik 1987). Within this report, spatial divisions based on physical and chemical classification are called "ecoregions", and those divisions based on instream biological differences are called "bioregions". Bioregions are addressed in step 5 ("Determine naturally occurring bioregional delineations").

The similarities among parameter value distributions were investigated using box and whisker plots of those parameters that were consistently recorded and biologically significant, including: conductivity, sulfate, turbidity, and habitat parameter scores. The ecoregions were combined or segregated according to these analyses, for establishment of physicochemical reference criteria.

3. Establish criteria for reference and degraded streams

Data from the Level 3 ecoregions and the potential sub-ecoregion (MRWnew) were combined or segregated depending on results of Step 2. Reference status decision thresholds were determined from the physical and chemical parameter distributions. "Reference", "sub-reference", "other", and "degraded" categories were established in order of increasing anthropogenic stress (or decreasing naturalness). To rate in the "reference" category, a site must satisfy *all* of the "reference" criteria. Likewise, the "sub-reference" sites passed all of the "sub-reference" criteria (slightly relaxed from reference criteria). In contrast, sites were classified as "degraded" when *any* of the criteria for that class were met. "Other" sites were those that did not meet specific criteria for any other class. Data from multiple years at the same site were averaged before being compared to these criteria. Criteria (threshold levels) for each parameter were established using basic knowledge of acceptable environmental conditions, examination of the distribution of observed values, application of parameter limits at each site, and subsequent evaluation of sample size sufficiency in the resulting categories.

4. Compile metrics

Candidate metrics for testing and potential inclusion in the final biotic index were selected from those provided by WY DEQ and supplemented with others selected from previous studies (Gibson et al. 1996, Stribling et al. 1998). Metrics are the various measurements of the biological assemblage that respond predictably to anthropogenic stressors. They fall within seven categories in the WYDEQ dataset: taxonomic richness, composition, diversity, density, pollution tolerance, feeding group, and voltinism. A total of 70 metrics within the seven categories were considered for inclusion in the index.

5. Determine naturally occurring bioregional delineations

Two primary techniques were employed to justify segregating or combining data from ecoregions into bioregions. First, metric value distributions from "reference" and "subreference" sites were compared between ecoregions using box and whisker diagrams. Similar distributions (medians, inter-quartile ranges, and overall ranges) of metrics were considered indications of similar biotic assemblages and would suggest combination of ecoregions into a single bioregion. To confirm indications from the metrics, non-metric multidimensional scaling (NMDS) ordination of relative taxa abundances was applied. With this approach, sites that areplotted close together are more similar. A dissimilarity matrix was created using the Bray-Curtis index, that integrates the differences in taxa abundances among samples.

6. Test metrics

The discrimination efficiency (DE) is a numerical description of the degree of separation between metric value distributions of "reference" and "degraded" sites. It is calculated as the percentage of degraded samples that have metric values worse than the worst quartile of the reference metric value distribution. A higher DE indicates better performance of a metric, or a better ability to distinguish between unstressed and stressed conditions. Commonly, metric values decrease in response to increasing levels of stress (e.g., most richness metrics). The DE in this case is the percentage of "degraded" sites with values less than the 25th percentile of the "reference" site values. In metrics that increase with increasing stress (e.g., Hilsenhoff Biotic Index), the DE is the percentage of "degraded" sites with values above the 75th percentile of the "reference" values.

Within each bioregion, DEs were calculated for all metrics that show a clear response to stressors. Metrics that had unintelligible differences between distributions of "reference" and "degraded" sites in the bioregions were not considered as viable candidates for inclusion in the index and were dropped from further analysis. Such comparisons either showed no clear trend with increasing stress or had an insufficient range of metric values to be calculated.

7. Combine metrics into a regionally calibrated index

A multimetric index is an additive approach for combining varied metric value information into a single numeric assessment value. The process begins with metric scoring to standardize the values, then with averaging of the best performing and most meaningful metrics. Steps were taken to include all pertinent metrics in the index, while minimizing redundancy.

The metric scoring strategy rated the metric values on a percentage scale from the worst value to the optimal within each bioregion. To minimize the influence of outliers, the 95th percentile of all the metric values in the bioregion was considered the optimal. For reverse metrics, the 5th percentile was considered optimal and the worst value was the maximum metric value recorded in the entire data set. On the standardized 100-point scale, each metric contributes to the combined index with equal weighting.

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 index formulations (or suites of metrics) were tested using the best-performing metrics from as many 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 the individual metrics. The final index was selected from those formulations which included all ecologically-meaningful metrics from several metric categories and that had a high DE relative to other formulations.

Narrative site assessments were defined by a division of the range of “reference” site index scores based on an equal bisection of the range above the 25th percentile and an equal trisection below, within each bioregion. All scores above the 25th percentile of the reference data rated as “very good” or “good” and those below the 25th percentile were rated “fair”, “poor”, or “very poor”. “Degraded” sites that scored in the “fair”, “poor”, or “very poor” range for their bioregion were interpreted as being correctly classified.

III. RESULTS

1. Database Development

Data from 301 valid sites and 384 valid samples were incorporated into the EDAS database. As a result of analyses on data consistency, samples using coarser mesh (1000 µm vs. 500 µm) were considered invalid. The possible biases associated with elevation and sampling dates were inextricably linked, because sampling protocol required early sampling at higher elevations in order to avoid unsafe weather conditions later in the season. Because the differences in metric values could not be confidently attributed to either cause (elevation or sampling date), preliminary stratification by elevation (a geographic variable) or sampling date (a temporal variable) was not pursued, though the issue was re-examined during subsequent stratification steps. Significant differences in habitat features were detected between stream types within ecoregions. Stream channel morphology was also considered as a source of bias during the preliminary database development steps and subsequently in the stratification steps.

2. Determine appropriate strata

Distributions of conductivity, sulfate concentration, turbidity, and habitat scores (the overall score and individual parameters of embeddedness and riparian width) suggested possible groupings of ecoregions and a potentially new sub-ecoregion into relatively homogenous zones for site classification. The Southern Rockies, Middle Rockies West, and Middle Rockies Central have fairly similar distributions of these variables and were combined. The newly defined southern portion of the Middle Rockies West (MRWnew) was similar to the Wyoming Basin and was combined with it for purposes of site classification. The Northwestern Great Plains, Western High Plains, and Middle Rockies East are distinct from the others in at least one of the chemical and physical measures and were considered separately.

Within the ecoregions, elevation may be a critical factor as it affects conductivity and sulfate. In the Southern Rockies, these two parameters correlate with elevation significantly, but the sites that drive the relationship are affected by stressors (unnatural land uses and dams). In the Middle Rockies, where elevation also correlates significantly with conductivity and sulfate, the correlation is not as strong. Ultimately, the ecoregions were not subdivided into upper and lower sub-regions.

3. Establish criteria for reference and degraded streams

Site reference criteria were established based on conductivity, sulfates, turbidity, total habitat score, oil and grease, total petroleum hydrocarbons, oil sheen, and water odor. In the Southern Rockies, Middle Rockies West and Middle Rockies Central regions, five specific habitat parameters (embeddedness, channel alteration, bank vegetation, bank stability, and riparian width) and pH were used in addition to the others. Application of these criteria yielded 73 "reference" sites out of 301 total sites (Figure 1). The greatest percentage of reference sites (35%) was found in the Middle Rockies East and the lowest was in the Western High Plains (1 out of 10). "Degraded" sites make up the largest category overall.

4. Compile metrics

A total of 70 metrics in seven metric categories were calculated primarily by WYDEQ from benthic taxa lists. All of the voltinism and most of the taxonomic richness and composition metrics enumerate aspects of the aquatic insect assemblage, or present the proportion of individuals in the sample that are of a particular taxon, feeding, or life history strategy. Feeding group metrics are the largest category (n=24), and density metrics are the smallest (n=2).

5. Determine naturally occurring bioregional delineations

Using metric value distributions (box and whisker plots) and NMDS ordination of "reference" plus "subreference" site data, delineation of relative biological homogeneity (bioregions) was possible. The distributions of metric values showed a clear distinction between mountainous regions and other regions. Further examination in the mountain regions revealed that the Middle Rockies East (or Black Hills) are distinct from the other Rockies. The potentially new sub-ecoregion of the Middle Rockies West (MRWnew) differed from the other Rockies in the dominance metrics, but was otherwise similar. The Western High Plains have only a single reference site and are most logically grouped with the Northwestern Great Plains. The unimpaired sites from the Basin rate slightly better than the Plains sites in almost all metrics.

The NMDS ordination of 159 samples ("reference" plus "subreference") resulted in a three-dimensional configuration that is best viewed on the first and second axes (Figure 2). This plot, based on taxonomic composition similarities (the Bray-Curtis index), showed a clear distinction between mountainous and non-mountainous regions, similar to that seen with the box and whisker plots. The Middle Rockies East and Southern Rockies have samples that fall within the bulk of Basin and Plains samples. However, the center of mass of the Southern Rockies samples lies closer to the Central and Western Middle Rockies samples. The Middle Rockies East samples were mostly distant from the other samples. Separation between the Basin and Plains is subtle, because a few Northwestern Great Plains sites are dispersed among Basin samples. The centers of mass are offset enough to justify separation of the two bioregions. The single Western High Plains sample is far from all other sites, but was grouped with the Northwestern Great Plains by default. A site class (= bioregion here) cannot be technically justified with a singlesample, or very low number of samples. The resulting four bioregions include:

Bioregion 1 (the Middle West, new Middle West, Central and Southern Rockies),
Bioregion 2 (the Middle Rockies East [or Black Hills]),
Bioregion 3 (the Wyoming Basin), and
Bioregion 4 (the Western High and Northwestern Great Plains).

6. Test metrics

Of the 70 metrics tested, 26 showed consistent response to stressors in all four bioregions. The discrimination efficiencies (DEs) ranged between 41 and 73%, with the tolerance indices performing best overall. None of the metrics in the density or voltinism categories performed well.

7. Combine metrics into a regionally calibrated index

Several index configurations were tested to find the metric combination that resulted in the greatest DEs in the bioregions and overall. Configurations included metrics from five categories (taxonomic richness, composition, diversity, tolerance, and feeding group). Two separate metric suites were selected for the Wyoming Stream Integrity Index (WSII), one to be used in the higher gradient bioregions (Rockies and Black Hills, WSII-H), and the other in the lower gradient bioregions (basin and plains, WSII-L). The greatest internal redundancy is with the metrics "Total taxa" and "% 10 dominant taxa" (r = -0.81) in the lower gradient index.

Median index scores are similar for the set of reference sites in the Rockies, Plains, and Basin bioregions, ranging from 70-75 (Figure 3). The lowest minimum for reference sites and the lowest values for degraded sites were seen in the Plains distributions. The probability of false negatives (indication of no impairment when, in fact, it does exist) for the index is 12.7% in all bioregions. In the "Rockies" bioregion, the probability of false negative, 17.1%; in the Plains, it is 4.5%; and in the Black Hills and Basins bioregions, it is 0%.

V. USE OF BIOLOGICAL INDICATORS TO ADDRESS MANAGEMENT AND REGULATORY CONCERNS

The biological index developed for the four bioregions of Wyoming can be used in several different resource management and regulatory activities. The primary activity includes stream bioassessment for internal agency goals such as prioritizing efforts. In such a context, watersheds could be targeted for restoration or preservation depending on the WSII scores. As stream restorations proceed, the WSII could be used to monitor resulting changes in biological condition, and thus, the effectiveness of the restoration. Likewise, the siting, design, and effectiveness of best management practices (BMPs) in urban areas (stormwater detention ponds, flow splitters, streambank stabilization, riparian vegetation or grassy swales) could be judged using the WSII.

Among other uses of the WSII are the application to Clean Water Act (CWA) legislation that requires periodic reporting of the condition of state waterbodies and the identification of waterbodies in need of restoration. Sections 305(b) and 303(d) of the CWA can be addressed using the biological information inherent to the WSII, as explained in the following sections.

1. Stream Bioassessments

Narrative assessments describe five stream condition categories (Figure 3, 4) based on the distribution of index values in reference sites of each bioregion. Greater than the 25^{th} percentile of the distributions is rated as "good" or "very good". WSII index values falling below the 25^{th} percentile are rated as "fair", "poor", or "very poor". This use of the biological criterion provides an estimate of resource status, an ordinal ranking of the condition of one site relative to the reference condition, and an indication of the necessity for more intensive, diagnostic sampling. Confidence intervals have been developed for the indices so that degrees of certainty can be associated with the WSII scores and ratings. Scores with confidence intervals that include the critical criterion should not be rated until further sampling is completed.

2. CWA Section 305(b) Assessments

Aquatic life use (ALU) assessments for the 305(b) program are based on the attainment of that designated use as defined in state water quality standards. There are three categories of attainment status: fully supporting, partially supporting, and non-attainment. For most states, a site can be rated as “fully supporting” when it is determined that aquatic life shall is “...as naturally occurs”. In large part, this leaves the final decision to the individual assessor, reducing objectivity in ALU decisions. Using the same reference condition index range and designating all index values falling above the 25th percentile as “fully supporting”; the upper 1/3 of the range below the 25th as “partially supporting”; and the lower 2/3 as “non-attainment” provides numeric thresholds to aid in those decisions (Figure 4). “Partially supporting” is an indication of impairment, but less severe than “non-attainment”. Sites falling within the confidence band would be treated as described above for bioassessment, that is, requiring additional intensive short-term information, or subsequent annual sampling.

3. CWA 303(d) Listing of Impaired and Threatened Waters

The use of highly quantitative measures for demonstrating reduced stressor loads, typically for individual pollutants, is inadequate for heralding improved biological condition. Instead, the WSII developed for Wyoming in this project is an appropriate indicator. For a stream site to be listed on the 303(d) list for a state, it must be determined as being either impaired or threatened. Metric and index values compared to decision thresholds can provide objective information for listing actions. When a stream or stream segment has been listed as impaired or threatened based on less than adequate information, site values falling above the 25th percentile of the reference distribution could justify its removal from the list (de-listing). If it falls in the bioassessment “fair” range (Figure 4), it could be listed but noted as only “partially impaired”. Placement on a “watch” list would enable special protection or expanded, more frequent monitoring. Sites rated “poor” or “very poor” would be considered impaired and would be listed. Once a TMDL-based stressor load reduction is implemented, a site can become “de-listed” only when the biological index values are found to be above the 25th percentile.

Using the TMDL Process in Conjunction with Watershed-Based Planning

Wyoming has selected an option for dealing with watershed pollution/degradation problems that allows increased public involvement. The state encourages local stakeholder groups to develop watershed plans that require an approval process similar to that for TMDLs. Development of the watershed plan- or TMDL-based management actions, when implemented, would ideally result in improvement of the aquatic life condition. Comparison of biological index values with the distribution of reference values will demonstrate the response of the stream biota to the implemented pollution reduction strategy. Use of bioassessment in the process is the same with each, except that with TMDLs “success” is when a site or watershed is de-listed; with the watershed approach, success is when a site upgraded to a “good” or “very good” rating from a “fair”, “poor”, or “very poor” rating.

VI. CONCLUSIONS AND RECOMMENDATIONS

As monitoring programs continue to gather information over time, databases used to develop and refine biological criteria expand. This means that, potentially, new reference sites are added, previously under-represented regions of the state become better known, and the definitions of stressor conditions become more refined. With the addition of data to fill geographic and temporal gaps, an increased understanding of the natural variability of Wyoming streams and watersheds is developed. Future sampling activity should target ecoregions and sub-ecoregions that have low numbers of reference and/or stressor sites. In particular, the Wyoming Basin and Western High Plains have few reference sites, whereas the Middle Rockies (MRC, MRW) are poorly represented by stressor sites. The site class findings presented here are in general agreement with those of Gerritsen et al. (2000, in press). The question regarding sample timing (or, index period) remains. To determine the effect of seasonal variability on the biological data, there needs to be repeated sampling at individual sites throughout the index period.

Principal uses of the multimetric indices and associated criteria are stream and watershed management and CWA program application. For the former, they will help detect impairment; prioritize streams in need of restoration or preservation; provide information on potential sources of stressors; and provide objective, ecologically-based targets for judging the effectiveness of restoration, BMPs, chemical controls, and overall stream and watershed management activities. Resulting assessments will also be useful in documenting and reporting the effects of nonpoint source degradation as part of the CWA Section 319 program, and for prioritizing and ranking watersheds in the Clean Water Action Plan. The Section 305(b) water quality inventory requires assessment of aquatic life designated use attainment, for which these biological criteria are the most appropriate measures.

Section 303(d) requires decisions on whether to list or de-list a site and whether a waterbody has reached a water resource quality target or goal. The biological criteria can help in that purpose. We recommend continued annual monitoring, as stated above, to fill gaps in information, to contribute to natural resource management decisionmaking, and to meet regulatory needs.

VII. LITERATURE CITED

Barbour, M.T., J.B. Stribling, and J.R. Karr. 1995. The Multimetric Approach for Establishing Biocriteria and Measuring Biological Condition. Pp. 63-76 In W. Davis and T. Simon, editors. *Biological Assessment and Criteria: Tools for Water Resource Planning and Decisionmaking*. Lewis Publishers, Ann Arbor, Michigan.

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.

Gerritsen, J., M.T. Barbour, and K. King. 2000. Apples, oranges, and ecoregions: On determining pattern in aquatic assemblages. Submitted May 1999 to the Journal of the North American Benthological Society.

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.

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, Virginia.

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.

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

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

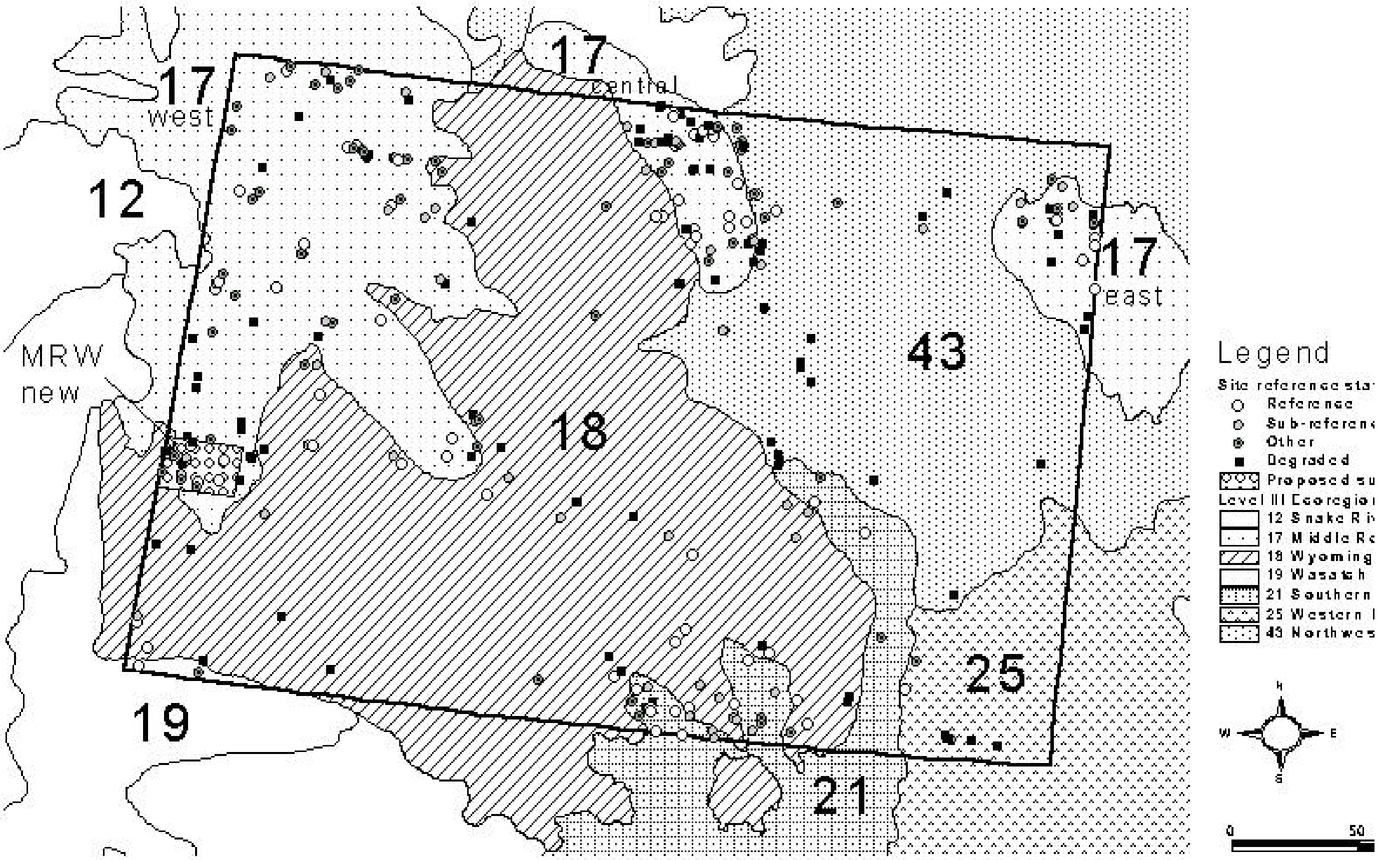
Rosgen, D. 1996. Applied River Morphology. Wildland Hydrology, Pagosa Springs, CO.

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.

Stribling, J. B., B. K. Jessup, J. Gerritsen. 2000. *Development of Biological and Physical Habitat Criteria for Wyoming Streams and Their Use in the TMDL Process*. Prepared for U. S. EPA Region 8, Denver, CO.

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.

Tetra Tech, Inc. 1999. Ecological Data Application System (EDAS). A User's Manual. Prepared by Tetra Tech, Inc., Owings Mills, Maryland.


17
west
17
central
12
17
east
MRW
new
43
18
25
19
21
Legend
Site reference sta
Reference
Sub-referen
Other
Degraded
Proposed su
Level III Ecoregio
12 Snake Ri
17 Middle Ro
18 Wyoming
19 Wasatch
21 Southern
25 Western
43 Northwes
N
W
E
S
0
50