Workshop Abstracts

2000/2001 Aquatic Ecosystem Classification and Reference Conditions STAR Progress Review Workshop

September 24, 2002
Denver, CO

Table of Contents

Introduction

The mission of the U.S. Environmental Protection Agency (EPA) is to protect public health and safeguard and improve the natural environment—air, water, and land upon which life depends. Achievement of this mission requires the application of sound science to the assessment of environmental problems and to the evaluation of possible solutions. The National Center for Environmental Research (NCER) at EPA is committed to providing the best products in high-priority areas of scientific research through significant support for long-term research.

The Office of Research and Development's (ORD) Environmental Monitoring and Assessment Program's (EMAP) goal is to build the scientific basis, and the local, state, and tribal capacity, to monitor for status and trends in the condition of the Nation's aquatic ecosystems. Research into the development of new and better classification systems has been identified by EMAP as essential to improving the current EMAP approach. The research presented at this progress review represents some of the extramural component of EMAP. You may find extensive information about the EMAP program at http://www.epa.gov/emap.

In 2000 and 2001, NCER issued a Request for Applications (RFA) on the Development of Aquatic Ecosystem Classifications and Reference Conditions. The purpose of these solicitations was to support research that led to the development of functional, defensible classification schemes and associated reference conditions for use in the application of biocriteria to one or more of the following aquatic resources: wetlands, large rivers, ephemeral systems, reservoirs, lakes, streams, estuaries, near-shore coastal environments, and coral reef communities. A total of 6 grants have been funded under this program.

Annual progress reviews such as this one will allow investigators to interact with one another and to discuss progress and findings with EPA and other interested parties. Although the research described in this report has been funded wholly or in part by the EPA, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred. Any opinions, findings, conclusions, or recommendations expressed in this report are those of the investigators who participated in the research. If you have any questions regarding the program, please contact the program manager, Barbara Levinson, at 202-564-6911 or levinson.barbara@epa.gov.

To learn more about EPA's STAR Program, visit NCER's Web Site at: http://www.epa.gov/ncer.

Hierarchical Physical Classification of Western Streams: Predicting Biological Condition in Terms of Key Environmental Processes Bridging Local to Ecoregional Scales

Brian P. Bledsoe, N. LeRoy Poff, and Ellen E. Wohl
Colorado State University, Fort Collins, CO

A hierarchical classification framework is being developed for stream environments in the western United States that identifies the relative importance of key environmental characteristics at multiple spatial scales. Specific objectives of this research project are to: (1) develop an exhaustive, multiscaled physical habitat classification of western U.S. streams to derive predictive statistical models relating biotic condition to multiscaled environmental variables; (2) demonstrate the explanatory power and flexibility of the classification within and across diverse western U.S. ecoregions; (3) explicitly assess how well a classification that includes intermediate scale processes can predict biotic condition at localities without requiring extensive local (reach scale) habitat data; and (4) develop a systematic approach for objectively identifying and stratifying reference sites.

Regional Environmental Monitoring and Assessment Program (REMAP) data are being used on first through fourth order streams from six ecoregions in Colorado, Oregon, and Washington and new data as available from the EMAP Western Pilot Study to link a multiscale physical classification with expectations of taxa distribution and categorical abundance in western U.S. streams and rivers. Given the current emphasis on ecoregional and local scales in biomonitoring protocols, this approach will provide essential information on how environmental factors measured at intermediate scales can be used to further understand and predict biotic condition at individual sites distributed across the landscape.

In the Cascades Ecoregion in Oregon, up to 72 percent of the variation in macroinvertebrate genus richness at a site can be explained using no more than four environmental descriptors, at both the local and watershed scales. The model for macroinvertebrate richness for sites dispersed throughout the State of Oregon also performed well ($R^2 = 0.59$) using watershed scale factors such as land cover and disturbance indicators (e.g., road crossings). Over a geographic range comparable to a Level III Ecoregion (Omernik 1987), local scale metrics such as bed substrate descriptors best predict macroinvertebrate richness. Preliminary results also indicate that using combined watershed and local scale metrics results in better prediction of macroinvertebrate communities than using either scale independently (see Table 1).

By examining how much additional variation in community organization is explained by characterizing environmental factors at scales other than ecoregions and local habitat, this work provides essential information for weighing the costs and benefits associated with different levels of spatial resolution in monitoring. Thus, it directly supports the EMAP initiative.

The physical classification and the metrics describing the physiochemical environment at multiple scales across Oregon, Washington, and Colorado REMAP and EMAP sites will be further refined by developing intermediate-scale environmental descriptors prior to testing with biological data. Once a comprehensive matrix of multiscale hydrogeomorphic descriptors is developed, statistical analyses (e.g., multiple regression, CCA, CDA, Stochastic CART analysis, and other approaches) will be used to elucidate the key physical controls on biological communities. Last, the classification will be applied in collaboration with state-level water quality managers to further generalize the classification, improve user-friendliness, and demonstrate practical applications.

Table 1. Summary of key environmental descriptors for macroinvertebrate response variables ranked in order of importance, and R^2 explained by associated regression models for five geographic areas.

Colorado	CO (Screened for mine drainage)	Oregon	Cascades	Colorado (Screened) + Oregon
Landcover	Reach Descriptor	Local Basin Descriptor	Substrate	Landcover
Water Quality	Substrate	Watershed Disturbance	Reach Descriptor	Precipitation
Watershed Disturbance	Watershed Disturbance	Substrate	Local Basin Descriptor	Substrate
			Geology	Watershed Disturbance
0.370†	0.509*	0.496†	0.715†	0.432†

* = p < 0.01
† = p < 0.001

Assessing an HGM-Based Wetland Classification and Assessment Scheme Along a 1,000 km Gradient of the Appalachian Mountains: Hydrology, Soils, and Wetland Function

Charles A. Cole[1], Denice Heller Wardrop[1], Robert P. Brooks[1], and Christopher P. Cirmo[2]
[1]Pennsylvania State University, University Park, PA; [2]State University of New York, Cortland, NY

The objectives of this project are to expand a recently developed wetland classification and functional assessment protocol, created in Pennsylvania, to regions north and south within the Appalachian mountains and to determine if: (1) similar wetland types occur along a broad latitudinal gradient, and (2) wetland structure and function are similar (see Figure 1). The objective of the project is to test the ability of these tools to provide useful scientific and management information outside of their region of development.

A large reference wetland data set, developed in Pennsylvania over the past 10 years, will be expanded. Standard assessment protocols designed to measure a wide variety of structural variables for input into wetland functional assessment models will be used. Up to 20 extant reference wetland sites in Pennsylvania will be revisited to obtain updated information. Approximately 20 new reference sites will be developed in New York and 20 in Virginia to test the applicability of these models along a broad latitudinal gradient. All study sites will be instrumented with continuous water level recorders to develop comprehensive hydrologic data for site assessment. These data will be used in the assessment models and serve as a check on the classification of each wetland.

Site selection began in the summer of 2002 in New York and Virginia. To date, eight sites have been located in the Adirondacks and several more potential sites have been identified in the Catskills. Sites will be located in the George Washington and Jefferson national forests in Virginia, the southern assessment area. Training of personnel, including Principal Investigators and graduate students, has commenced to provide all personnel with a unified understanding of the study's protocol. Water level recorders have been installed in many of the Adirondack wetlands and will be placed in some Pennsylvania sites before winter. The Virginia sites are expected to be instrumented by early spring. Data collection for the functional assessment models has begun both in New York and Pennsylvania.

There are no preliminary findings to date. The selection of all sites will be finalized by the end of fall 2002, and as much information as possible will be collected before winter limits site accessibility. This will provide the first real test of the transferability of the classification scheme throughout the study region. The classification key will be adjusted if it proves to be inadequate. In addition, the bulk of the model data will be collected during the summers of 2003 and 2004.

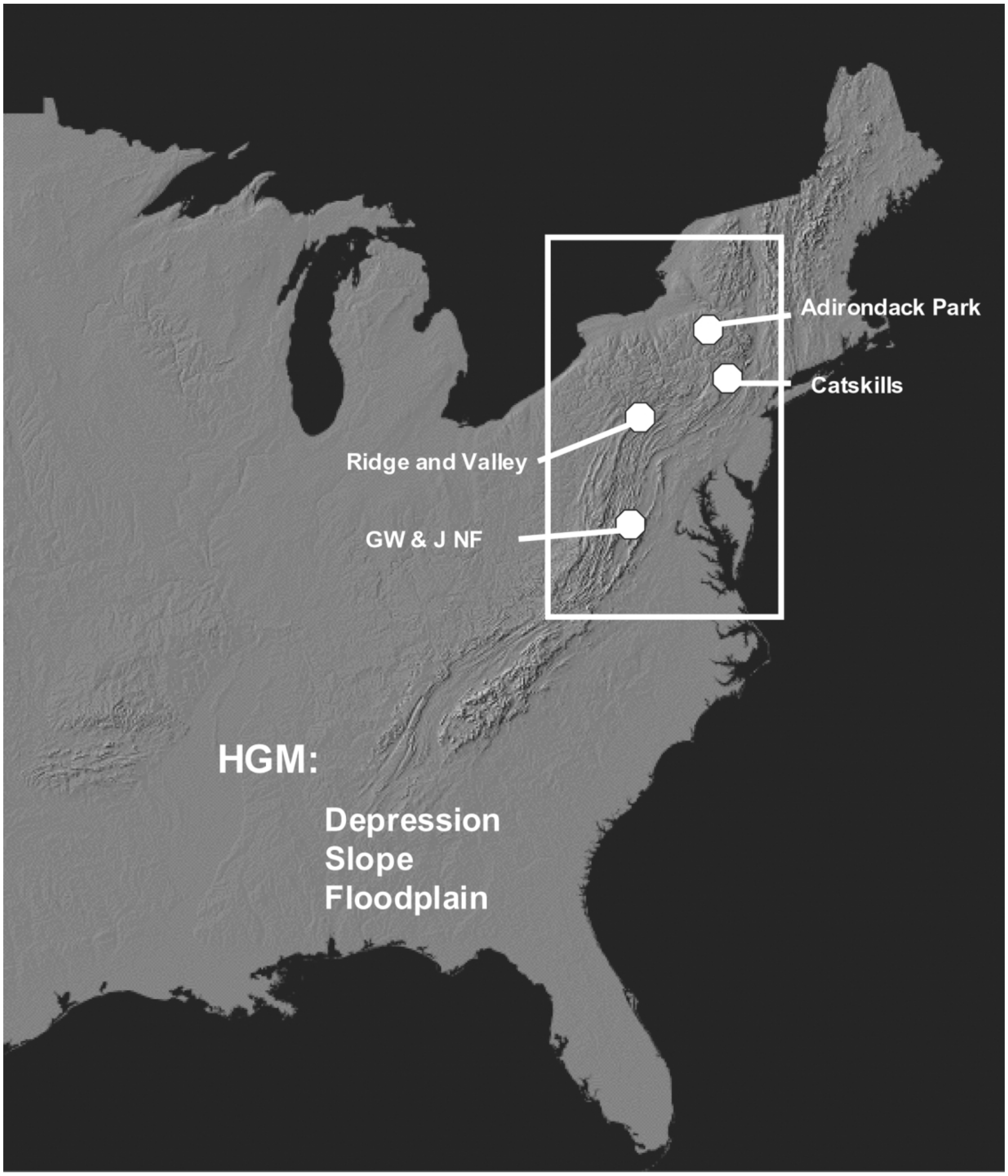

Figure 1. Wetland functional assessment across a 1,000 km gradient of the Appalachian Mountains.

An Empirical Evaluation of the Performance of Different Approaches To Classifying Reference Conditions in Streams

Charles P. Hawkins[1], R. Jan Stevenson[2], and David W. Roberts[1]
[1]Utah State University, Logan, UT; [2]Michigan State University, East Lansing, MI

The goal of this project is to identify the type and level of classification that is optimal for bioassessment purposes. Two primary questions are being explored: (1) How is the sensitivity of assessments affected by the approach used to classify sites (regionalization versus predictive discriminant model), the type of assemblage examined, and the spatial scale of classification? and (2) Are approaches to site classification transferable among regions? It is hypothesized that the classifications that most closely approximate local environmental conditions affecting individual organisms will be the best predictors of expected conditions.

All analyses are being conducted on two biological assemblages: algae and invertebrates. A combination of existing and new data are being used to evaluate performance of different classifications (e.g., ecoregions, catchment, type of stream reach, thermal strata) in terms of their accuracy and precision as well as their effect on the sensitivity of detecting impairment. The study area includes a 13-state region (CA, OR, WA, ID, NV, AZ, NM, UT, CO, WY, MT, SD, ND).

During the first year of the project, field sampling was completed at 324 sites located in 10 different states (OR, ID, MT, ND, WY, NV, UT, CO, AZ, NM). During this same time period, collaboration occurred with personnel from the U.S. Environmental Protection Agency (EPA) Region 8, the U.S. Department of Agriculture Forest Service, Oregon Department of Environmental Quality, and Washington State Department of Ecology to obtain data from approximately 200 additional sites in SD, ID, WY, OR, WA, and MT. Together with data collected previously from CA, OR, and WA, data from more than 1,000 potential reference sites have been collected for use in testing the hypotheses (see Figure 1). All of the invertebrate samples (excluding chironomid midges) have been processed and identified, and initial analyses have been started for the invertebrate data. The chironomid midge data will be available in about 6 months for the 324 sites that were sampled, once all samples have been mounted and identified. Approximately 50 percent of the diatom samples have been permanently mounted onto slides, and rapid counts (100 or 200 valves) have been completed for 115 slides. All field data have been entered into a computer database and initial quality assurance/quality control has been conducted on the database. In addition, compilation of landscape-level data has begun from which additional information will be extracted to characterize the environmental setting of each site. The following coverages are being produced within a Geographic Information System: digital elevation models (DEMs), bedrock geology, stream network, climate (temperature and precipitation), and land use. Approximately 340 water chemistry samples have been analyzed to date for NO_3, NH_3, TN, SRP, CL, and Si. Analysis for cations is underway. Much of the effort during the first year has been focused on several quality assurance issues. These issues include development of consistent criteria for defining reference sites (with EPA ORD Corvallis and EPA Region 8), development of a consistent taxonomy to be applied to periphyton, and analyses to quantify sources of error in the estimation of different biological indicators. Work is being done with both Forest Service personnel and EPA ORD scientists involved in similar large-scale monitoring and assessment projects to document and devise ways of controlling sources of error.

It is still too early in the analyses to comment on the significance of the findings, but the level of cooperation and collaboration developed with different states, many of which view this project as directly relevant to their bioassessment needs, has been beneficial.

The majority of data analyses will be conducted during the upcoming project period. Only a modest amount of field activity has been planned for this year to fill in gaps in the developing database. The main activity will be associated with building predictive models for both invertebrate and diatom assemblages and creating the *a priori* classifications based on landscape attributes that will be compared with the predictive modeling results.

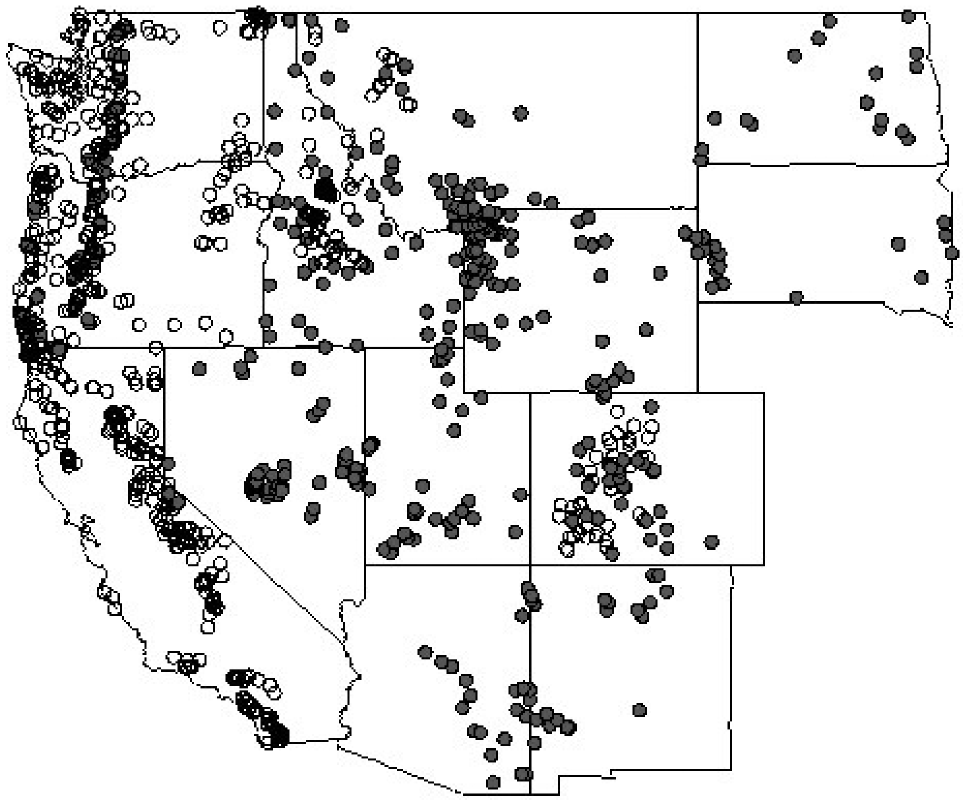

Figure 1. Distribution of 1,064 reference sites in the western United States that are being used to test hypotheses regarding useful approaches to classifying streams for bioassessment purposes. Filled symbols are locations of sites that were sampled during the summers of 2002 and 2003. Open symbols are locations of sites that either we or our collaborators sampled previously.

A Biologically Driven National Classification Scheme for U.S. Streams and Rivers

Alan Herlihy, Yangdong Pan, and Bob Hughes
Oregon State University, Corvallis, OR

Analyzing stream biological assemblage data at a national scale is extremely difficult and rarely attempted due to the problems of compiling the necessary database. The goal of this research project is to assemble a national database for the conterminous 48 U.S. states of stream/river fish, macroinvertebrate, and periphyton assemblages derived from regional scale synoptic surveys. The objectives of this project are to: (1) use our national database to develop 10-30 biologically driven national "classes" of stream systems; (2) within each class, separate natural from anthropogenic effects on stream ecological condition; and (3) establish quantitative relationships between catchment and riparian condition and water body condition (structure and function).

As a first cut, all publically available stream ecological data from the U.S. Environmental Protection Agency's (EPA) Environmental Monitoring and Assessment Program (EMAP) and Regional Environmental Monitoring and Assessment Program (REMAP) projects will be compiled, formatted, and validated into an approximately 3,600-site national database. These data were all collected with the same field protocol, and sites were picked in a randomized, systematic fashion so they are representative of the study area. The national database will be augmented with high-quality state databases that were collected with similar methods. A set of least disturbed reference sites will be screened from the combined national database using available water quality and habitat data. A multivariate cluster analysis of the reference sites will identify 10-30 clusters, and a discriminant function analysis will be conducted to identify the environmental factors that best predict class membership. Within each new biologically derived class in the entire database, ordination analyses will be used to identify the major gradients in the assemblage data and relate them to the measured physical and chemical habitat. This will help separate natural from anthropogenic effects and establish the major functional relationships between stream biota and stream condition for each class. For 2-4 of the classes, catchment land cover data from remote sensing will be gathered to determine the scale and types of land use most related to the biological ordination in that class.

Most of the existing EMAP and REMAP data have been collected, and construction of the national database has begun. The status of the data collection efforts on a state-by-state basis are shown in Figure 1. Complete EMAP/REMAP statewide probability survey data have been obtained for 20 states, representing about one-half the area of the 48 states. REMAP data exist for portions of another 9 states. To fill in spatial gaps in the databases, a number of different state agencies have been queried for any existing data that will meet the study's needs. If these queries are successful, data from most of the country will be obtained (see Figure 1).

Analyzing a combined national EMAP database provides a unique opportunity to study classification, ecological gradients, relationships, and reference conditions at a national scale. At the end of the study, it will be possible to demonstrate patterns in stream biota and biological condition across the United States. Such a demonstration will aid in justifying a national stream assessment if the U.S. EPA and states choose to embark on one.

Available databases will be collected during the rest of 2002. All data will be formatted into a common format, and inconsistencies and differences in taxonomic names across datasets will be resolved. Also, differences in taxonomic resolution (i.e., family, genus, species levels) need to be resolved and coded in the data for future data analyses. The goal is to have a final national database ready by January 2003. The first half of 2003 will be spent on conducting the multivariate analyses of the national database and identifying preliminary national clusters.

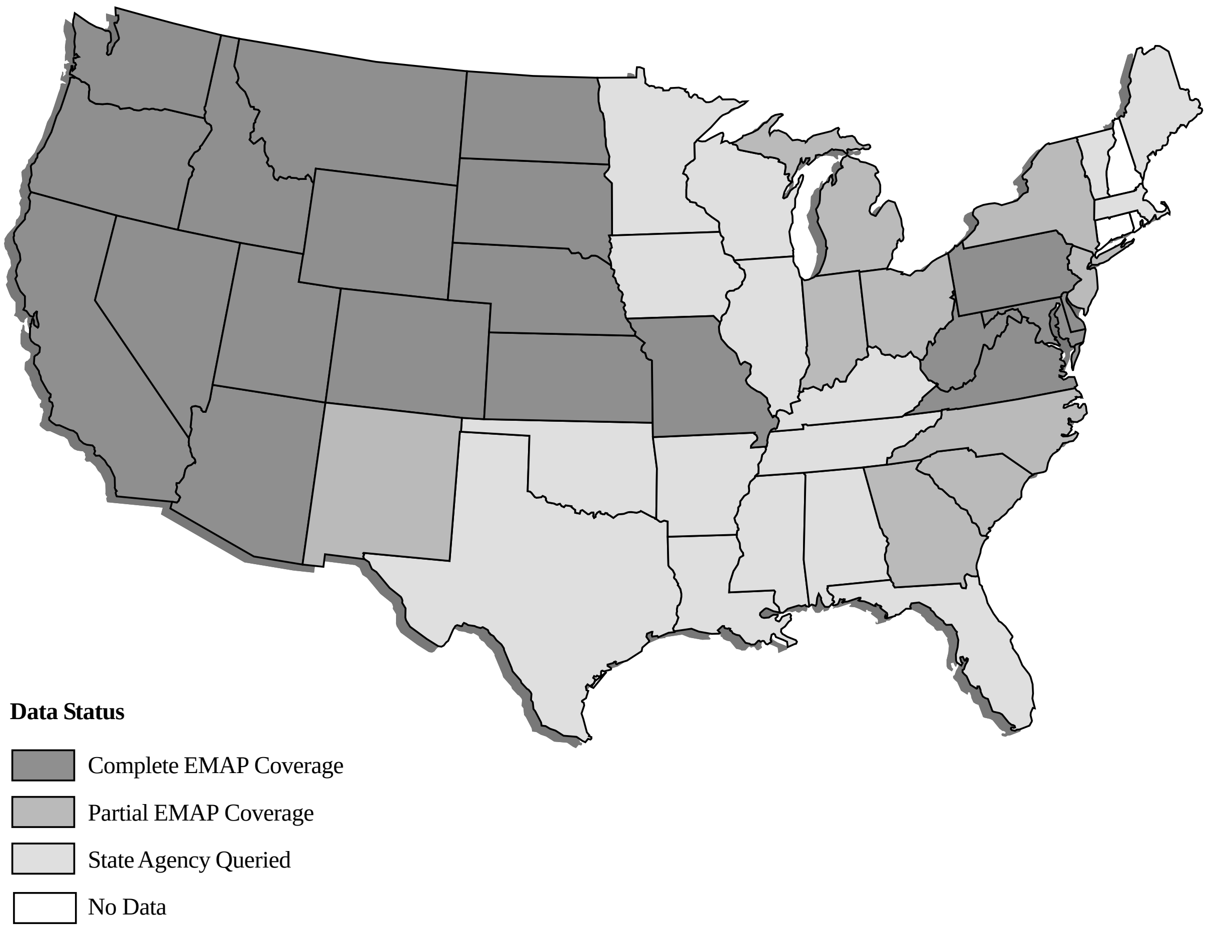

Figure 1. Status of the national stream biological assemblage database by U.S. state.

Development and Implementation of a Comprehensive Lake and Reservoir Strategy for Nebraska as a Model for Agricultural Dominated Ecosystems

John C. Holz, James W. Merchant, Anatoly A. Gitelson, Sherilyn C. Fritz, Kyle D. Hoagland, Istvan Bogardi, and Donald C. Rundquist
University of Nebraska, Lincoln, NE

In agriculturally dominated regions, land use practices have an unusually large impact on water bodies and, therefore, land use may reduce the utility of current ecoregion-based approaches to lake classification by dampening the signals that underlie the ecoregion framework. A team of water quality researchers has been assembled to develop a comprehensive classification scheme for agriculturally dominated ecosystems, using Nebraska as a highly representative model. Three objectives critical to achieving this goal are to establish: (1) a protocol for aggregating water bodies in agricultural ecosystems into classification strata and identifying reference conditions for these classes; (2) the role of remote sensing and the Geographic Information System in a classification strategy; and (3) a technology transfer link between the proposed classification system and end-users.

A water quality database for nearly 325 Nebraska lakes and reservoirs was established by sampling an additional 181 water bodies during the first 2 years of this project. All water bodies were sampled monthly from May through September for common limnological parameters (e.g., nutrients, clarity, chlorophyll). From this database, lakes and reservoirs are being classified hierarchically using a combination of rule-based and data-based approaches. For example, in the Sand Hills region of Nebraska, lake classes were defined by performing a factor analysis on the limnological data and plotting the significant factors to identify five groups of lakes with similar water quality characteristics. The factor analysis revealed that three significant factors explained more than 73 percent of the variability of the data, with alkalinity, conductivity, chlorophyll, nitrate+nitrite, Secchi depth, total suspended solids, orthophosphate, and total phosphorus loading significantly into the three factors. Interestingly, the Level IV Ecoregions based on soil type, native vegetation cover, topography, and geology do not accurately represent water quality in this region. Limited surface water inputs, combined with local hydrology, reduce the utility of landscape classification approaches in the Sand Hills region.

Improved methods also are being developed for integrating field data, data collected via airborne and close range remote sensing, data collected via operational and near future satellite remote sensing systems, and ancillary geospatial data in a multistage approach to lake classification. Biological indicators that integrate the lake conditions of each stratum will be developed from summer phytoplankton and zooplankton collections, and special emphasis has been placed on developing methods to remotely sense biological indicators of water quality based on the optical phytoplankton pigment structures of lakes. Algorithms for remote estimation of chlorophyll were developed and tested in close range remote sensing measurements in water bodies with different trophic status. The developed technique was very sensitive to even slight variation in chlorophyll concentration as well as in turbidity and Secchi depth. Standard error of chlorophyll concentration prediction was less than 5 mg/m^3 in the range of chlorophyll concentration from 10 to more than 194 mg/m^3 (see Figure 1). The technique also allows indication of presence of blue-green algae in water, although at this stage quantitative accuracy of the technique in phycocyanin estimation was not assessed. Measurements from aircraft (non-imaging) also showed high sensitivity of algorithms to chlorophyll and other optically active constituents.

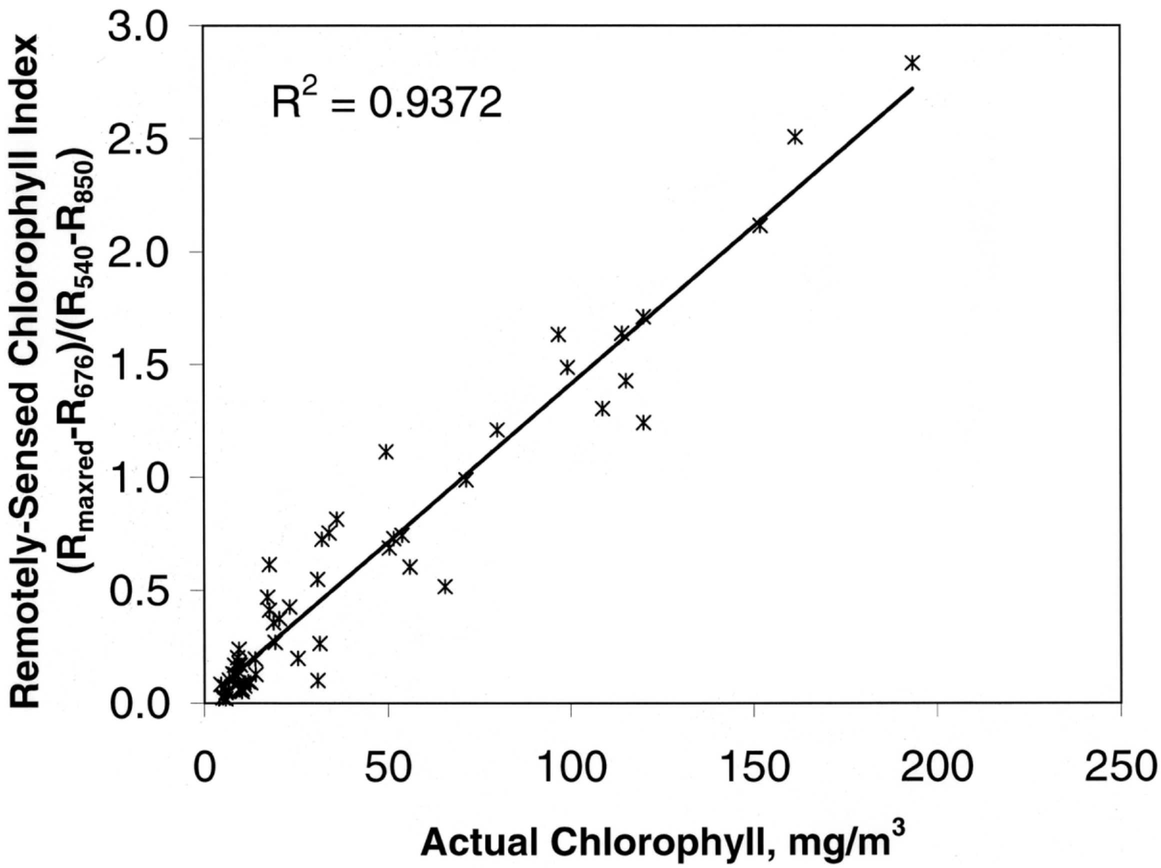

Figure 1. Actual chlorophyll concentrations versus an index of close-range, remotely-sensed chlorophyll concentrations in Nebraska lakes and reservoirs.

Protocols for Selection of Classification Systems and Reference Conditions: A Comparison of Methods

Lucinda Johnson[1], Jeffrey Schuldt[2], George Host[1], Jan Ciborowski[3], and Carl Richards[4]
[1]Natural Resources Research Institute, University of Minnesota, Duluth, MN; [2]University of Wisconsin, Superior, WI; [3]University of Windsor, Ontario, Canada; [4]Minnesota Sea Grant, University of Minnesota, Duluth, MN

The specific goals of this research project are to: (1) apply an *a priori* classification system to Great Lakes coastal ecosystems; (2) quantitatively identify reference sites within classification units using widely available spatial databases; (3) define reference biological conditions for classification units; (4) use biological data to test the efficacy of reference classifications; and (5) determine whether degraded biological conditions differ from reference conditions in a subset of classification units.

A hierarchically structured classification system will be implemented over a large region in the central United States to test a statistically based protocol for identifying reference conditions in a region. A combination of existing and newly collected data will be used to apply an *a priori* classification system to Great Lakes coastal ecosystems and identify regional reference conditions (see Figure 1). This classification will identify homogeneous units representing underlying abiotic conditions that control the structure and function of coastal ecosystems. A variety of existing national and statewide databases will be used to identify the number and type of pressure indicators within each unique classification unit. Pressure indicators will be used to assess the degree of anthropogenic influence on individual ecosystems and, consequently, identify candidate reference ecosystems. The biotic (macrobenthos and fish) communities and habitat structure will be characterized within these reference ecosystems. Several statistical methods will be used to test whether the biotic communities within each reference site are unique to a given classification unit. The same level of stratification will not be required to uniquely identify reference sites for all coastal ecosystem types; therefore, the classification system will be refined using agglomerative or divisive approaches.

A comprehensive inventory has been completed of the coastal wetlands for the entire Great Lakes coastal region within the United States, excluding connecting channels. The Great Lakes coastal ecosystems were classified on a regional basis using the scheme of Keys et al.[1] at the EcoProvince level, and Omernik et al.[2] at the Ecoregion level. Five geomorphic units along the shoreline were identified: high energy zones, bays, coastal marshes, riverine wetlands, and protected wetlands. Land use, road density, point sources, and population density data were summarized for unique topographic receiving areas for each shoreline unit. Within classification units, reference ecosystems have been defined as those sites with the least amount of anthropogenic influence. Sampling of biotic (macrobenthos and fish) conditions in 100-120 reference ecosystems is underway.

A novel method of defining the bounds of the reference condition based on the joint influence of multiple stressors has been developed, and a significant contribution has been made to the research and management community in the Great Lakes by assembling a comprehensive inventory of coastal wetlands. Results from the field data collections are not available yet.

Sampling of high energy shorelines, protected wetlands, and riverine wetlands is underway. Quantifying stressors for coastal marshes and embayments, to be sampled next year, will continue.

---

[1]Keys Jr. J., Carpenter C., Hooks S., Koenig F., McNab W.H., Russell W., Smith M.L. Ecological units of the eastern United States—first approximation. Atlanta, GA: U.S. Department of Agriculture, Forest Service, 1995.

[2]Omernik J.M. Ecoregions of the conterminous United States. *Annals Assoc Am Geographers* 1987;77:118-125.

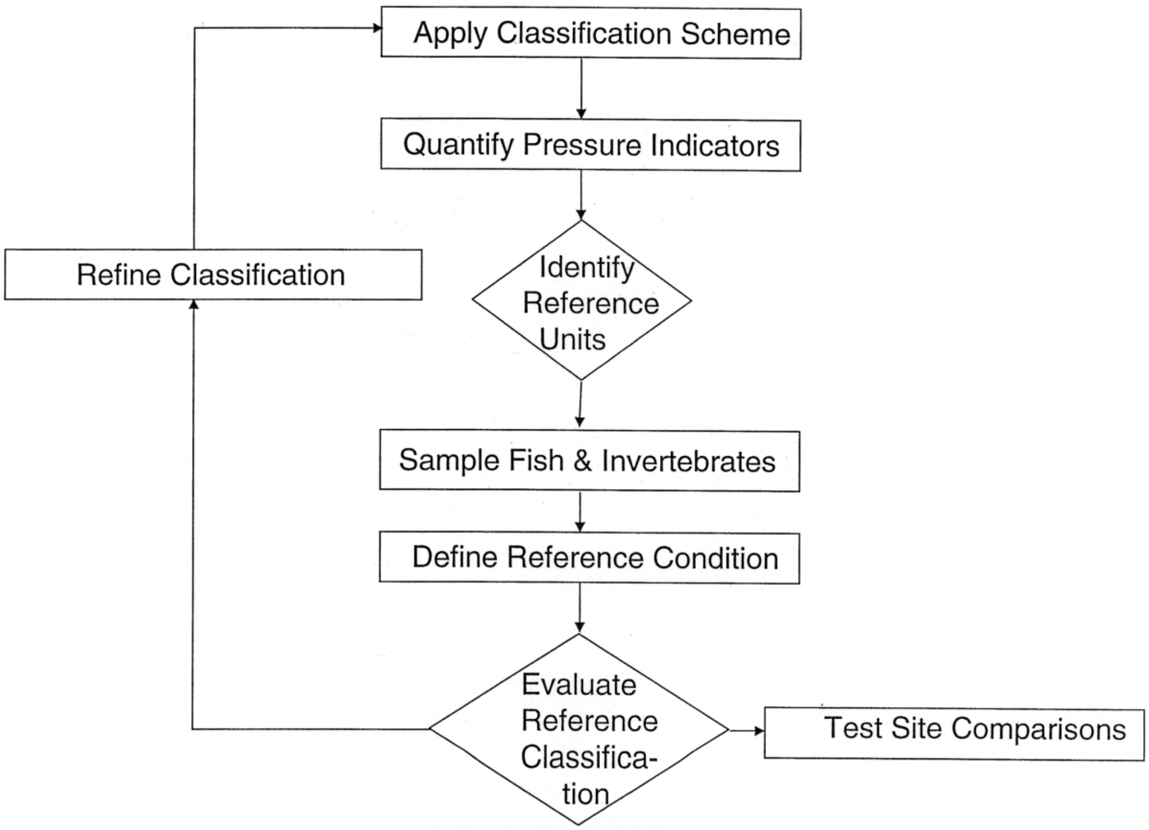

Figure 1. Conceptual model of reference area identification.

Index of Authors

United States
Environmental Protection Agency
Office of Research and Development
National Center for Environmental Research (8701R)
Washington, DC 20460

EPA/600/R-02/054
September 2002
www.epa.gov/ncerqa