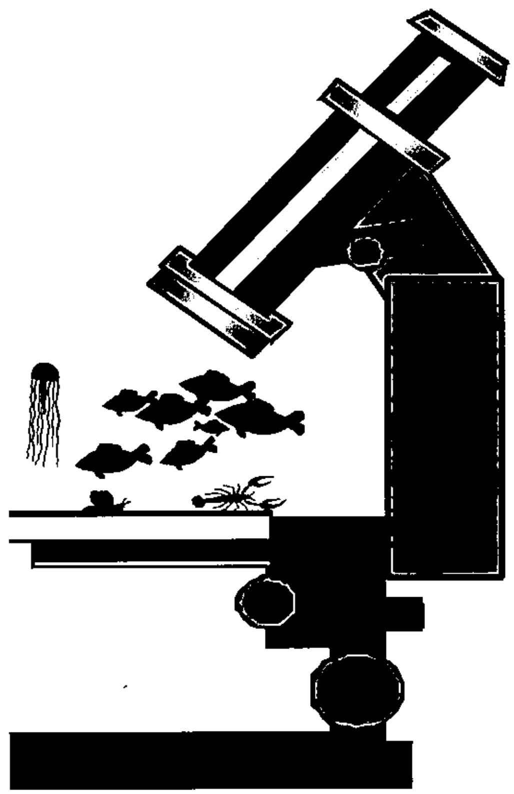
Biological Habitat Quality Indicators for Essential Fish Habitat Workshop Proceedings

14–15 July 1997
Charleston, South Carolina

Edited by

S. Ian Hartwell



U.S. Department of Commerce
National Oceanic and Atmospheric Administration
National Marine Fisheries Service

NOAA Technical Memorandum NMFS-F/SPO-32
September 1998

Biological Habitat Quality Indicators for Essential Fish Habitat Workshop Proceedings

14–15 July 1997
Charleston, South Carolina

Edited by
S. Ian Hartwell
NOAA/NMFS
Habitat Conservation Office
1315 East-West Highway
Silver Spring, Maryland

NOAA Technical MemorandumNMFS-F/SPO-32
September 1998

U.S. Department of Commerce
William M. Daley, Secretary

National Oceanic and Atmospheric Administration
D. James Baker, Under Secretary for Oceans and Atmosphere

National Marine Fisheries Service
Rolland E. Schmitten, Assistant Administrator for Fisheries

Table of Contents

Abstract

A national workshop to address development of biological indicators for habitat quality in Essential Fish Habitat (EFH) was held July 14-15, 1997 in Charleston, S.C. The workshop was attended by biologists from the National Marine Fisheries Service (NMFS) Office of Habitat Conservation, NMFS Science Centers, the National Ocean Service (NOS) Strategic Environmental Assessment Division and the Oceanic and Atmospheric Research (OAR) Sea Grant Program. The meeting included presentations by researchers from universities and Federal and state agencies who are performing bioindicator research and development in aquatic environments. These included projects in several benthic and pelagic estuarine habitats on the Atlantic and Gulf of Mexico coasts, coastal embayments, benthic habitats on the continental shelf in the Atlantic and Pacific Oceans, and rivers and open waters of the Great Lakes. In addition to ecological considerations, application of bioindicators to management needs, monitoring issues, and delineation of habitats into ecosystem units were addressed. Conceptual approaches for development of bioindicators of habitat quality for EFH, identification of current areas of research needs, and settings for potential pilot program initiation were developed. It was concluded that the Index of Biotic Integrity (IBI) approach will be useful by generating multimetric information to describe habitat quality in quantitative terms and for technical ecological assessment and research. Parameters for assessment metrics were developed for each of three general habitat types, vegetated, benthic, and pelagic. Areas requiring additional research include basic natural history information on species selected in the metric development process, quantification of their response to anthropogenic stress, and methods for delineating reference areas.

Executive Summary

Background

A major activity within the National Marine Fisheries Service (NMFS) is the implementation of the Essential Fish Habitat (EFH) requirements of the Magnuson-Stevens Fishery Conservation and Management Act of 1996. This legislation mandates that the regional Fishery Management Councils, in coordination with NOAA, amend each the 39 fishery management plans (FMPs) to include the best available information on habitat delineation for each of the approximately 600 managed species. The amended FMPs will include options and recommendations to minimize adverse effects on EFH and identify conservation and enhancement measures. These will include recommendations on activities or regulations that may impact water quality, so that NOAA can protect, conserve, restore and enhance essential habitats for each life stage of all managed species. The ultimate goal is to maintain the natural productivity of fish habitats at levels which will sustain populations at harvestable levels into the future.

A key requirement of the habitat assessment activities is an assessment of habitat quality. Habitat is defined as the combination of chemical, physical and biological components of the water and substrate in the local or regional ecosystem. The ultimate indicator of habitat quality is the response of the biological community to the interaction of stresses and resources available at a particular location and time frame. The biological community acts as the integrator of habitat quality. Coupled with habitat delineation, chemical analyses and physical characterization, biological indicators allow assessment of alteration of the environment including eutrophication, nonpoint source pollution, contamination, SAV loss, etc. Therefore, assessment of the condition of the biological community is an indicator of habitat quality, and can also be utilized to track preservation and/or restoration efforts. The value of biological criteria and biological assessment techniques has been demonstrated by their broad applicability not only to existing efforts to protect, restore, and manage aquatic resources, but in determining where management and restoration resources should be invested. Biological habitat quality indicators need to be developed for several types of marine environments to measure habitat quality in a variety of habitat types. The term biological integrity originates from the Federal Water Pollution Control Act amendments of 1972 and has remained a part of the subsequent reauthorizations. Efforts to construct a workable, practical definition of biological integrity have provided the supporting theory necessary for development of standardized measurement frameworks, techniques, and criteria for determining compliance with that goal. In 1981, Karr and Dudley defined biological integrity as "the ability of an aquatic ecosystem to support and maintain a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of the natural habitats of a region". This definition alludes directly to measurable characteristics of biological communities which are found in the least impacted habitats of a region. It was this definition and the underlying ecological theory which provided the fundamental basis for the development of numeric biological criteria in fresh water. The U.S. Environmental Protection Agency (EPA) essentially adopted this definition in the national program guidance on biological criteria. Biological criteria and attendant monitoring and assessment designs provide a means to incorporate broader concepts of water resource integrity while preserving the appropriate roles of the traditional chemical, physical and toxicological approaches developed over the past three decades.

Biological evaluation of aquatic habitat integrity is made possible by monitoring aquatic communities directly. Community bioassessments differ from approaches which rely principally on target species or indicator organisms by utilizing the aggregated information across multi-species assemblages. The aggregation of key community attributes functions as an indication of the more complex ecosystem elements and processes which can not be measured directly or completely. At the same time information about individual species is not lost in the process and can be accessed at any time. Furthermore properly designed bioindicator systems extract ecologically relevant information and provide a synthesized, numerical result that can be understood by non-biologists.

Conceptually, an Index of Biotic Integrity (IBI) utilizes a set of categories which reflect important ecological parameters, for example, diversity, abundance and trophic function. Within each category a variety of metrics are generated. Diversity can be quantified by number of species, species richness or one of several diversity indices. Abundance can be assessed in terms of numbers of organisms or biomass. Functional metrics reflect interactions between community segments, for example predator prey ratios, density of predator species, etc. Other categories can be included, such as condition indices, prevalence of lesions, proportion of pollution tolerant species, etc. Each metric is a site-specific measurement which can be general or very specific. Individual metrics are given a rating score on a numerical scale which reflects its value relative to a reference value. The individual scores are then summed on a site-specific basis so sites can be compared to each other, or a reference site. An alternative approach to selecting and scoring the metrics has been to designate reference sites a-priori based on other factors such as demonstrated lack of chemical contaminants, eutrophication, physical disturbance etc. An array of candidate metrics are then calculated and a final subset are selected based on multivariate statistical evaluation of the data. This approach allows application of IBI assessment to habitats in which functional relationships of resident organisms are not fully understood, due to high complexity or lack of knowledge.

There are a variety of important technical questions that have to be addressed before this approach can be widely employed in marine environments. Many of these problems have been solved for fresh water environments. Some states utilize IBI monitoring for habitat quality assessment while others have integrated IBIs into water quality monitoring as a regulatory tool for enforcing water quality permits. Marine environments are more complex than streams, requiring the development of different approaches within the IBI framework. A variety of pilot projects in marine environments have been initiated which have experimented with different methods and procedures.

Technical Presentations

The NMFS Bioindicator Workshop was organized to itemize and begin to address these parameters, establish a consensus of scientific approach to development of biological habitat quality indicators in EFH, and identify research and monitoring needs for future projects. The focus was directed toward practical applications of bioindicators in marine systems and research needs to support this development. Subsequent discussion groups addressed which biological parameters would be practical and meaningful to measure in each of three broad categories of habitat (benthic, pelagic and vegetated).

Technical presentations included a conceptual overview of the IBI approach in a variety of marine systems, as well as an extensive evaluation of the practical application of bioindicators to statewide water quality monitoring and regulation in the Great Lakes region. In addition to water quality regulatory permit and enforcement activity, state agencies in Ohio have recognized the utility of bioindicators in the implementation of non-Clean Water Act management activities such as endangered species protection, targeted fisheries management, hydro-modification and wetland dredge and fill permit evaluations.

Shallow tidal creeks serve as conduits through which many pollutants enter estuaries. Creek sediments act as a repository for toxic chemicals and other contaminants. It is largely unknown how effective wetland management policies and programs are at protecting tidal creek habitats, or how to restore degraded creek habitats. A South Carolina Marine Resources Research Institute study has initiated development of a data base to develop the information needed to characterize the ecological values, identify major pollution threats associated with watershed development, assess the cumulative impacts and develop environmental quality criteria for sustaining nursery functions of tidal creeks and associated marsh habitats Results indicate that monitoring efforts for tidal creeks should focus on the upper reaches of primary tidal creeks and should include measures of the health of resident organisms, water and sediment quality, and selected population and community parameters of resident species.

Habitat delineation methods and data base development for IBI derivation are compatible with current data base and GIS activities currently under way between NMFS and the National Ocean Service (NOS) in response to implementation of the Magnuson-Stevens Act. The primary data layers currently in place are estuarine salinity zones and USGS Hydrologic Cataloging Units. Additional data for coastal and offshore spatial units and EPA river reach files are being included. A complete database exists only for the contiguous states at the present time.

Estuarine and marine habitats are more complicated than freshwater streams due to their larger scale, and diverse biological and physical components, including a wide phylogenetic diversity of biota, highly transient species, strong physical and chemical gradients in water and sediment quality, and a strong interaction between the pelagic and benthic communities. Nevertheless, the basic principles of IBI development can be applied to these systems. Estuarine fish bioindicators have been, or are being developed, in Connecticut, Massachusetts, Chesapeake Bay, North Carolina, Florida and Texas. Investigations on the transferability of fish community bioindicator metrics for submerged aquatic vegetation (SAV) habitats developed for Cape Cod estuaries and tested in Chesapeake Bay, and from Chesapeake Bay pelagic habitats to coastal embayments have been instructive. The degree of modification to the metrics that was necessary to adapt the systems to different regions was relatively straight forward.

Fish and benthic invertebrate IBIs have been developed in freshwater environments to assess transitional zones going from lentic to lotic habitats (termed lacustuaries) and for near shore open-water habitats of the Great Lakes, analogous to estuaries and coastal zone habitats. The bioindicator systems have been demonstrated to be capable of quantitatively tracking habitat quality and are responsive to habitat quality changes resulting from watershed and riparian area management activities.

In some, but not all locations, benthic invertebrate bioassessment schemes have adapted somewhat different approaches than those utilized for fish community assessment. The benthic indicator development projects have employed complex mathematical schemes to develop metrics, due to the more complex and less well understood biological communities associated with benthic invertebrate communities. Current development projects in the New York/New Jersey harbors, the Virginian province Chesapeake Bay, SE Atlantic, and Gulf of Mexico rely heavily on the EPA Environmental Monitoring and Assessment Program (EMAP) data. Chemical contamination data has been used extensively to guide definition of reference sites.

Coastal benthic efforts on the Atlantic and Pacific continental shelves have taken divergent approaches from estuarine studies due to the more diffuse nature of impacts in off-shore habitats. However, gradients of habitat degradation can be identified and quantified. A great deal more development and research will be necessary to address the myriad of habitats present in off-shore areas.

Metric Development

Three discussion groups were formed, for the purpose of coming to consensus on an array of biological metrics which would be practical and meaningful to measure in each of three broad habitat types (benthic, pelagic, vegetated). Important attributes of metrics included consideration of ecological relevance, practicality, and demonstrated relationship to anthropogenic degradation of aquatic habitats. The vegetated habitat category included submerged aquatic vegetation (vascular plants and algae), emergent wetlands, mangrove and kelp habitats. The benthic habitat category included soft (unconsolidated sediment), hard (surfaces to which benthic organisms can attach) and live bottom substrates (physical structure of the habitat was composed of, or built by, oysters, coral or benthic assemblages with significant three-dimensional relief). The water column habitat included the open water column habitats of freshwater streams, estuaries, near shore and coastal waters. A total of 36 potential metrics in four categories (diversity, abundance, function and condition) were enumerated. There was considerable overlap between metrics in the three habitat types in the diversity, abundance and condition categories. It is instructive that there was very little overlap in functional metrics. Functional roles of a species in a habitat is much more site specific than other parameters. Overall, the metrics used in current programs do not cover as wide a range as the potential metrics considered in the break-out groups. The range and specificity of metrics utilized in fish IBI projects are greater than those used in benthic invertebrate projects.

Conclusions

Based on knowledge gained from preliminary studies, the IBI approach will be useful for assessing habitat quality in two primary ways: it brings together multimetric information to describe the quality of the biological community in simple, yet quantitative terms, and can be used for technical ecological assessment or to formulate research hypotheses for testing. The approach was specifically designed to assess environmental harm resulting from anthropogenic stressors. In addition to the regulatory need for site specific biological measurements, it is useful to be able to represent the condition of complex ecosystems concisely, by means of composite indices or simple graphics, so that managers and non-specialists can readily evaluate and compare information, establish goals, and set priorities for remediation or protection.

It is not necessary to sample all subunits of an ecosystem. This would not be possible in any case, as all gear is selective to some degree. Assuming the ecosystem is integrated at some level, assessment of specific habitat types and/or locations should be adequate if methods are carefully selected.

NMFS should move forward to identify appropriate attributes that would constitute biological indicators of habitat quality for the following habitat types: SAV, riparian, estuarine benthic/water column, coastal benthic, and coral reef habitats. Ongoing activities around the nation that are involved in developing and applying biological indicators, biodiversity indices, and IBIs should be inventoried. A list of habitat priorities should be developed for investigation and feasibility studies.

NMFS must develop partnerships with other Federal, state, university and private groups that are involved or interested in developing and applying indices of biological integrity. Maximum use of ongoing programs should be made.

One difficulty with the application of IBI techniques to complex marine systems has been the relative lack of intimate knowledge of the ecological roles and interactions of specific species and/or functional guilds, compared to fresh water systems. Therefore, a basic element of any future IBI development work is simple taxonomic and natural history documentation of the species selected for use as markers of stress. Data gaps in life histories of critical species, including the degree of natural variation, must be identified and resolved.

A related problem is the definition of what constitutes a reference condition. A-priori selection of 'reference' sites based upon one set of parameters (e.g. contaminants) have not been tested for efficacy in habitats which may have been impacted by other stressors (e.g.

eutrophication). Ideally, a credible index should be responsive to any form of habitat degradation. A comparative assessment of the mathematical methods for derivation of reference sites and results has not been performed.

While the cumulative index may contain qualitative elements, the quantitative behavior of properly developed metrics in relation to each other, and our ability to assess them in relation to anthropogenic impacts is instructive. The detailed information from individual metrics is not lost in the process. The IBI approach provides a framework for assessing habitat quality in a consistent, technically defensible method. It has a demonstrated utility in fresh water environments as a technical assessment method and as a management tool.

1.0 Introduction

A major activity within the National Marine Fisheries Service (NMFS) is the implementation of the Essential Fish Habitat (EFH) requirements of the Magnuson-Stevens Fishery Conservation and Management Act. This legislation mandates that the NMFS amend each of its 39 Fishery Management Plans (FMPs) to include the best available information on habitat delineation, habitat needs, human impacts, and mitigative measures so that NMFS can protect, conserve, restore and enhance key habitats for each life stage of each managed species.

Implicit in the exercise of identification and delineation of EFH is that monitoring habitat quality is part of the process. While the productivity of fisheries is one of the ultimate management objectives of NMFS, the strength of the Magnuson-Stevens Act is that it is directed at protecting fisheries "habitat". Habitat assessment is easier than site-specific productivity assessment of multiple species at various times and specific life stages. Habitat quality assessment addresses the fundamental question of "how much of the habitat is still unimpaired?" and "what alterations are being imposed on that which remains?". A means to evaluate and monitor the ecological integrity of habitat is essential if that habitat is to be managed for fishery production or restored to productive habitat (Figure 1.1). Therefore, the ultimate measure of EFH habitat quality is a measure of the condition of the biological community which inhabits it. This requires that the habitat be spatially and temporally delineated, and a method to continually assess biological condition be applied to it. Habitat quality monitoring programs should be incorporated into FMPs.

Sophisticated measures of habitat quality must be devised that reflect environmental conditions and which are sufficiently robust to be used in a wide variety of physiographic regions. A variety of habitat classification schemes have been devised for different regions and habitats (Brinson 1995, Davis and Harper 1996, Monaco et al. 1997, NOAA 1995, Osborne et al. 1991, Dethier 1990). These have incorporated a variety of indicators including basic water quality, physical and chemical parameters, and population metrics. These indicators are region specific in some cases. Indicators of biological integrity reflect parameters beyond those which only define the chemical and physical characteristics of the habitat, and should be used in concert with them to assess total habitat quality (Fausch et al. 1990). This will allow tracking the impacts of specific habitat stressors, such as contaminants, eutrophication, and wetland loss, and linkage of those stressors to ecological response. The ultimate indicator of habitat quality is the response of the biological community to the interaction of the stresses and resources available at a particular location and time frame. The biological community acts as the integrator of habitat quality.

Many routine programs that provide data for the current generation of 'indicators' are related to monitoring for human health or regulatory control programs, as opposed to actual measures of environmental quality (EPA 1996). Frequently, they only reflect the size of regulatory programs (e.g.,number of permits), rather than actual loadings to, or impacts on the environment. Without substantial data manipulation and the imposition of significant assumptions, monitoring data may not be amenable to translation into actual metrics of impact (Warner et al. 1991).

To be useful for NMFS management application, a habitat quality assessment approach must be linked to, or at least correlated with, fishery production in that habitat. When NMFS is called upon to engage in ecosystem management decisions that affect fisheries habitat, and where tradeoffs for other resource demands, such as water use, forestry, development, resource extraction, etc. are being considered, the role of NMFS is to estimate what is at risk in terms of fishery production. In addition to fisheries, the role of the National Oceanic and Atmospheric Administration (NOAA), including NMFS, is to act as stewards of the marine habitat and to protect the multiple benefits derived from the functioning of the intact ecosystem beyond direct production of food (e.g.,biodiversity, recreation, natural products, etc.). Habitat quality assessment paradigms must reflect the impact of various anthropogenic activities on fisheries productivity and the integrity of the marine ecosystem which supports that productivity.

This Proceedings document is organized into seven sections which provide a brief introduction to biological indicators, workshop objectives, technical presentations, workgroup products, conclusions and recommended follow-up activities. The technical presentations are presented as project summaries. The interested reader may contact the primary authors for more detailed documentation. The presentations were grouped into four categories, including practical applications and data base development, fish community studies, benthic community studies and continental shelf studies. The presentation on application of biological indicators in the state of Ohio is considerably longer than the other sections. While the other examples are of no less interest, the state of Ohio has successfully incorporated the basic scientific assessment of biological community condition into the very practical, real world regulatory framework for water quality monitoring and enforcement. It illustrates that biological indicators can be utilized for environmental management and regulatory needs, and how this has been accomplished in at least one state.

References

Brinson, M. 1995. The HGM approach explained. National Wetlands Newsletter. Nov-Dec:7-16.

Davis, T.J. and J.R. Harper. 1996. Estuarine mapping and classification system for British Columbia. Resource Inventory Committee, British Columbia Ministry of Environment, Victoria, BC, Canada.

Dethier, M.N. 1990. A marine and estuarine habitat classification system for Washington State. Wash. Nat. Heritage Prog., Dept. Nat. Res., Olympia, Wash.

Fausch, K.D., J Lyons, J.R. Karr and P.L. Angermeier. 1990. Fish communities as indicators of environmental degradation. In Biological Indicators of Stress in Fish S.M. Adams (ed.), Am. Fish. Soc. Symposium 8, Bethseda, Md.

U.S. Environmental Protection Agency. 1996. Environmental indicators of water quality in the United States. U.S. EPA, Office of Water, Wash., DC. EPA 841-R-96-002.

Monaco, M.E., S.B. Weisberg and T.A. Lowery. (in press). Summer habitat affinities of estuarine fish in USA mid-Atlantic coastal systems. Fish. Manag. and Ecol.

NOAA. 1995. Environmental sensitivity index guidelines. NOAA Tech Memo., NOS, ORCA 92. Seattle: Hazardous Materials Response and Assessment Div., NOAA.

Osborne, L.L. et al. 1991. Stream habitat assessment programs in states of the AFS North Central Division. Fisheries, 16:28-35.

Warner, K.A., S.I. Hartwell, J.A. Mihursky, C.F. Zimmerman and A. Chaney. 1991. The lower Patapsco River/Baltimore Harbor contaminant data base assessment project. Prepared for Balt. Regional Planning Counc. Chesapeake Research Consortium, Solomons, Md. 128 pp.

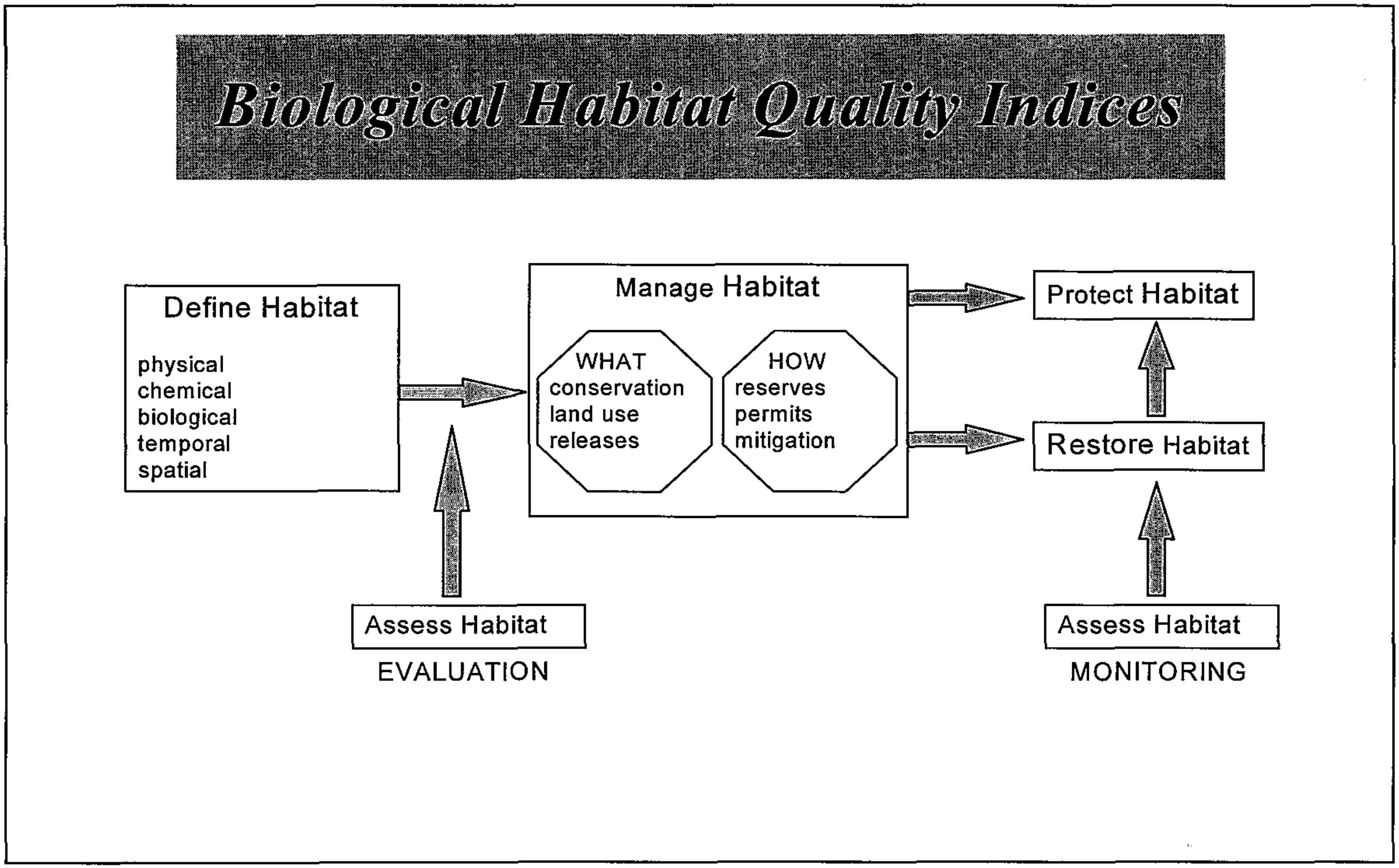


Figure 1.1 Diagrammatic representation of habitat management activities requiring habitat assessment.

2.0 Index of Biotic Integrity Overview

COUPLED WITH HABITAT DELINEATION, CHEMICAL ANALYSES AND PHYSICAL CHARACTERIZATION, BIOLOGICAL INDICATORS ALLOW ASSESSMENT OF ALTERATION OF THE ENVIRONMENT INCLUDING EUTROPHICATION, NONPOINT SOURCE POLLUTION, CONTAMINATION, SAV LOSS, ETC. BIOLOGICAL HABITAT QUALITY INDICATORS NEED TO BE DEVELOPED FOR SEVERAL TYPES OF MARINE ENVIRONMENTS TO MEASURE HABITAT QUALITY IN A VARIETY OF HABITAT TYPES.

The index of biological integrity (IBI) approach has been demonstrated to be an effective tool to reflect the cumulative response of the aquatic community to the total environment, with all the attendant interactions and compensatory limits of populations and communities (Karr and Chu 1997). Biological integrity can be represented by indices which integrate the interaction of the total environment with specific populations and communities. They may include multiple parameters which assess productivity, trophic interactions and species richness in the community (Figure 2.1). Bioindicators also have the potential to detect effects of trace level contamination and ephemeral events which may have long term effects on resident biota.

Assessing stream pollution was the driving force behind the original development of IBIs (Karr 1981). The IBI approach integrates a variety of other impact assessment methods which have been developed. These reflect a range of complexity, including indicator species or taxa, various species diversity indices, and multivariate methods (Deegan et al. 1997, Engle et al. 1994, Weisberg et al. 1997). Conceptually, an IBI utilizes a set of categories which reflect important ecological parameters, for example, diversity, abundance and trophic function. Within each category a variety of metrics are generated. Diversity can be quantified by number of species, species richness or one of several diversity indices. Abundance can be assessed in terms of numbers of organisms or biomass. Functional metrics reflect interactions between community segments, for example predator prey ratios, density of predator species, etc. Other categories can be included, such as condition indices, prevalence of lesions, proportion of pollution tolerant species, etc. Each metric is a site-specific measurement which can be general or very specific (e.g.,number of striped bass/km^2).

Each metric is then given a rating score on an ordinal scale (1, 2, 3 or 1, 5, 10 etc.). This step is very important as it normalizes the various metrics on a common scale (Figure 2.2). Thus, the measurements must be devised carefully, as they will be treated as being of equivalent ecological importance in the calculations, unless a weighting scheme is employed. In addition, they must reflect community response to stress. Assigning the score involves a good deal of ecological expertise (e.g., are 200 striped bass/km^2 twice as good as 100/km^2, or are they within the same range of habitat quality?). The individual scores are then summed on a site-specific basis so sites can be compared to each other based on percentile ranking of data relative to all stations, or relative to a reference site. Consistent sampling methods among sampling locations is crucial. An alternative approach to selecting and scoring the metrics has been to designate reference sites a-priori based on other factors such as demonstrated lack of chemical contaminants, eutrophication, physical disturbance etc. An array of candidate metrics is then calculated and a final subset is selected based on statistical evaluation of the data (Engle et al. 1994, Strobel et al. 1995) . This approach allows application of IBI assessment to habitats in which functional relationships of resident organisms are not fully understood, due to high complexity or lack of data.

There are a variety of important technical questions that have to be addressed before this approach can be employed in marine environments for gauging habitat quality in EFH. Many of these problems have been solved for fresh water environments (Karr and Chu 1997). Some states utilize IBI monitoring for habitat quality assessment (Ohio EPA 1988). Some have integrated IBIs into water quality monitoring as a regulatory tool for enforcing water quality permits. Marine environments are more complex than streams, requiring the development of different approaches within the IBI framework. A variety of pilot projects in marine environments have been initiated which have experimented with different methods and procedures (Deegan et al. 1997, Engle et al. 1994, Guillen 1997, Jordan et al. 1994, Lenat 1993, Linder et al. 1997, Nelson 1990, Weisberg et al. 1997). The NMFS Bioindicator Workshop was organized to enumerate and begin to address these parameters, establish a consensus approach to development of biological habitat quality indicators in EFH, and identify research and monitoring needs for future projects.

References

Deegan, L.A., J.T. Finn, S.G. Ayvazian and C.A. Ryder-Kieffer and J. Buonaccorsi. 1997. Development and validation of an estuarine biotic integrity index. Estuaries 20:601-617.

Engle, V.D., J.K. Summers and G.R. Gaston. 1994. A benthic index of environmental condition of Gulf of Mexico Estuaries. Estuaries 17(2).

Guillen, G.J. 1996. Development of a Rapid Bioassessment Method and Index of Biotic Integrity for Tidal Streams and Bayous located along the Northwest Gulf of Mexico. 1996. TNRCC Special Report. Houston, Texas.

Jordan, S. J., C. Stenger, M. McGinty, T. Arnold, S. Ives, D. Randall, B. Rodney, S. I. Hartwell. 1994. Estuarine habitat assessment and index of biotic integrity demonstration and testing. Draft Final Report to U. S. Environmental Agency, Office of Water. 17pp.

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

Karr, R.R. and E.W. Chu. 1997. Biological monitoring and assessment: Using multimetric indexes effectively. EPA 235-R97-001. Univ. Washington, Seattle, Wash.

Lenat, D.R. 1993. A biotic index for the southeastern United States: Derivation and list of tolerance values, with criteria for assigning water-quality ratings. J. N. Am. Benthol. Soc. 12:279-290.

Linder, C., M. McGinty, D. Goshorn and K. Price. 1997. Physical habitat and fish assemblages: an investigation of the near-shore areas of the Chesapeake Bay and Maryland's Coastal Bays. University of Delaware, Lewes, Delaware and Maryland Department of Natural Resources, Annapolis, Maryland.

Nelson, W.G. 1990. Prospects for development of an index of biotic integrity for evaluating habitat degradation in coastal systems. Chem. and Ecol. 4:197-210.

Ohio Environmental Protection Agency. 1988. Biological criteria for the protection of aquatic life: Vol I. The role of biological data in water quality assessment. Ohio EPA, Div. Water Qual. Monitoring and Assessment, Columbus, Ohio.

Strobel et.al. 1995. Statistical Summary: EMAP-Estuaries Virginian Province, 1990-1993. EPA/620/R-94/026.

Weisberg, S.B., J.A. Ranasinghe, D.M. Dauer, L.C. Schaffner, R.J. Diaz and J.B. Firthsen. 1997.An estuarine benthic index of biotic integrity (B_IBI) for Chesapeake Bay. Estuaries, 20: 149-158.

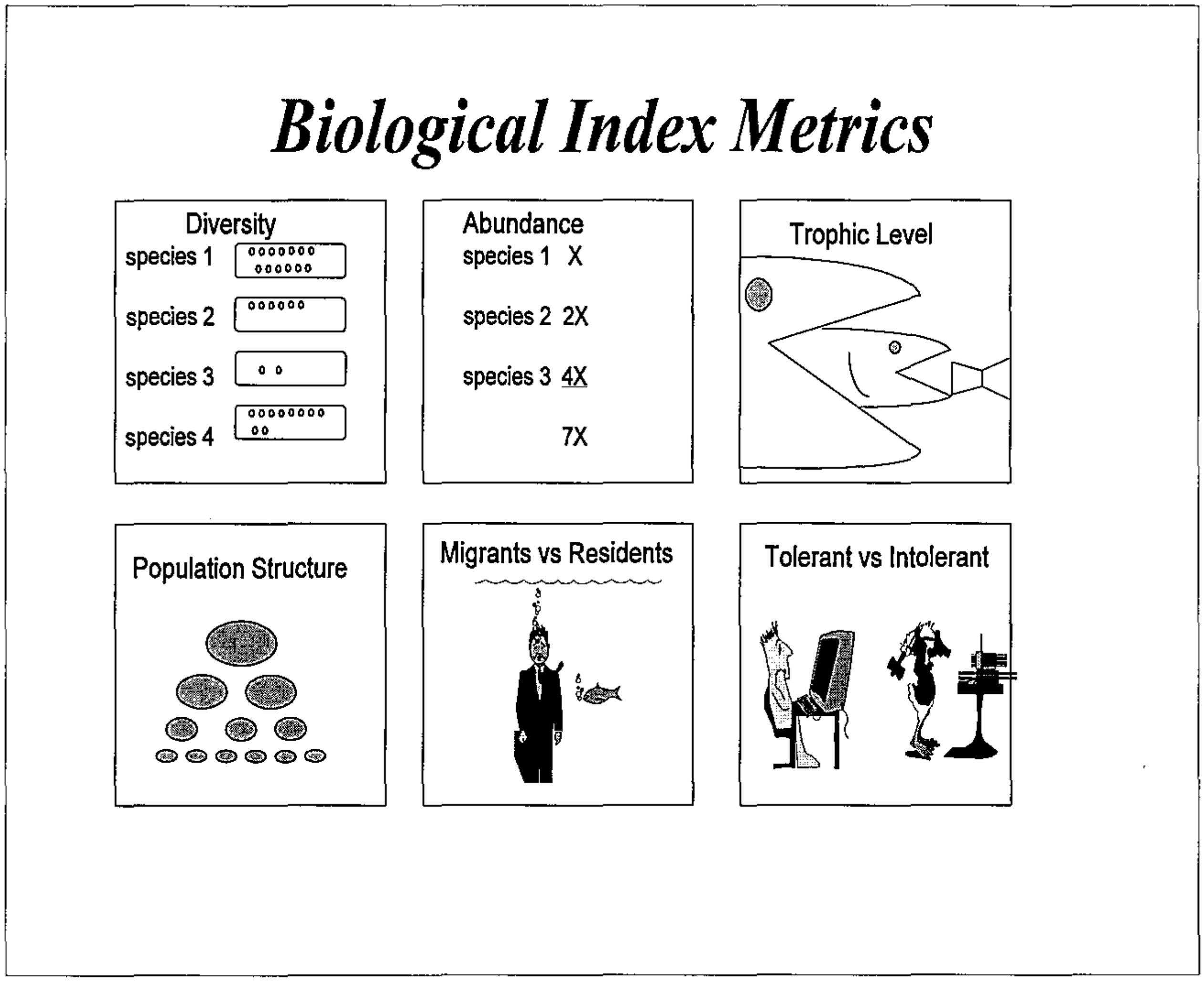


Figure 2.1 Typical metrics utilized in IBI assessments.

Index Calculation

		Criteria Values			Score
		low	med.	high	
	metric 1	x	y	z	a_1
Category I	metric 2	x	y	z	a_2
	metric 3	x	y	z	a_3
	metric 4	x	y	z	a_4
	metric 1	x	y	z	a_1
	metric 2	x	y	z	a_2
Category II	metric 3	x	y	z	a_3
	metric 4	x	y	z	a_4
	metric 5	x	y	z	a_5
	metric 6	x	y	z	a_6
Category III	metric 1	x	y	z	a_1
	metric 2	x	y	z	a_2

$$\texttt{Index} = \texttt{f}(a_n)$$

Figure 2.2 IBI calculation scheme. X, Y, and Z are cutoff values unique to each metric. Each score (e.g.,1, 5, 10) is assigned based on the value of measured environmental measures, relative to the criteria. The final index is the sum of the scores, which may be weighted in some fashion.

3.0 Workshop Objectives

The workshop was organized to develop a consensus within NMFS on the methods and utility of biological indicators for assessing habitat quality in EFH. The focus was directed toward practical application of bioindicators in marine systems and research needs. The workshop was arranged into one day of presentations by researchers in the field on conceptual approaches, experimental methods, pilot program results and, management applications. The following day, workshop participants broke into three discussion groups to deliberate which biological parameters would be practical and meaningful to measure in each of three broad habitat types (benthic, pelagic, vegetated) for the purpose of generating marine bioindicators. The three groups then compared and contrasted results. Finally, a general discussion of how to proceed with development of bioindicators for application to EFH was conducted. Several general questions were initially used as the basis for discussions of metric development and use.

1. What categories, metrics and calculation methods will work in which environments?
   Can the same metrics be used in an estuary in the Gulf of Mexico and an estuary in North Carolina or Maine? The specific measurements within each category or metric may have to be different in each case, because the biological communities will differ in species composition in different regions. Should metrics be expressed in terms of proportion or absolute numbers, e.g.,abundance.

2. Which environments are most feasible to assess this way?
   Within the limits of manpower and resources, how much effort will be required to obtain data with acceptable statistical power in differing habitat settings (e.g.,estuarine, open coastal, kelp bed).

3. How can individual metrics be assessed consistently over spatial and temporal regimes?
   Can the metrics be designed such that the need for 'professional judgement ' is eliminated? Can the metrics be designed so that a score of 28 in the Gulf of Mexico indicates the same ecological quality as a similar score from an estuary on the Atlantic coast or Alaska?

4. Can/should benthic-pelagic coupling be addressed?

5. What are the spatial scales over which bioindicators can be applied?
   How will habitat be delineated? How large an area in the delineated habitat should be assessed to confidently evaluate that habitat? Can results from a small area be extrapolated to surrounding habitat, or must the entire region be evaluated? What are the upper bounds of habitat area that can be assessed before localized impairment becomes undetectable?

6. Can estuarine results be coupled with coastal habitat units? Can watershed results be coupled with estuarine habitat units?

7. Can biological indicators be developed that are responsive to specific stressors?

8. Should redundancy be avoided or ignored in metric selection?

9. How should reference condition/sites be determined?
   How clean is clean?

WORKING DEFINITIONS

Assemblage — The association of interacting populations of organisms in a selected habitat.

Attribute — A measurable factor in the biological assemblage.

Biological Assessment — An evaluation of the condition of a habitat based on measurements of attributes of the biological assemblage.

Biological Integrity — The ability of an ecosystem to support and maintain a stable community of organisms having the structural and functional organization comparable to that of an undisturbed habitat within a region.

Category — A group of metrics which express a characteristic of habitat (e.g.,diversity, abundance, function, etc.)

Community — An assemblage of populations of organisms which either reside in, or utilize a specific habitat, within a particular time frame.

Ecological Integrity — The condition of an ecosystem as measured by the chemical, physical, and biological attributes.

Habitat — The combined ecological features of an area, including chemical, physical and biological components.

Index — An integrated expression of habitat condition incorporating multiple metrics.

Metric — A specific biological attribute, with a demonstrated empirical response to a gradient of anthropogenic disturbance.

4.0 Technical Presentations

4.1 ATTAINING ENVIRONMENTAL GOALS: BIOLOGICAL MONITORING AND ASSESSMENT IN THEORY AND PRACTICE

James R. Karr

University of Washington, Seattle, WA

Models guide much that we do in basic and applied ecology, including efforts to protect environmental quality. Models--whether conceptual, physical, or mathematical--can be wrong when they focus on the wrong endpoint or when they do not incorporate critical system components or processes. When models are not tested for their relevance in the real world, they can lead us astray as they squander both financial and environmental resources. It is especially regrettable when models lead us to ignore biological common sense or when scientists and managers focus on statistical significance rather than magnitude of effect and its biological consequence.

Because ambient biological monitoring focuses our attention on the most integrative endpoint (biological condition), we can use biological monitoring to test our models and assess the extent to which our policies protect ecological health. Biological monitoring has evolved rapidly during the twentieth century as knowledge has changed, and human-imposed stresses have become more complex and pervasive. Multimetric biological indices, like the index of biological integrity (IBI), integrate knowledge from earlier monitoring approaches while avoiding indicators that are flawed theoretically (Karr and Chu, 1997).

Developing effective multimetric biological indices involves five major activities:

1. Classifying environments to define homogeneous sets within or across regions (e.g., large or small streams, warmwater or coldwater streams).

2. Selecting measurable attributes that provide reliable and relevant signals about the biological effects of human activities.

3. Developing sampling protocols and designs that ensure that those biological attributes are measured accurately and precisely.

4. Defining analytical procedures to extract and understand relevant patterns in the data gathered.

5. Communicating the results to citizens and policy makers so that all concerned stakeholders can contribute to environmental policymaking.

References

Karr, J.R. and E.W. Chu. 1997. Biological Monitoring and Assessment: Using Multimetric Indexes Effectively. Univ. Washington, Seattle, WA. EPA 325-R97-001.

4.2 THE OHIO ENVIRONMENTAL PROTECTION AGENCY'S BIOLOGICAL MONITORING PROGRAM, IBI MEASURES AND THEIR POSSIBLE APPLICATION TO ESTUARINE ENVIRONMENTS

Roger F. Thoma and Chris O. Yoder

Ohio EPA, Columbus, OH

Introduction

The value of biological criteria and biological assessment techniques has been demonstrated by their broad applicability not only to existing efforts to protect, restore, and manage aquatic resources, but in determining where management and restoration resources should be invested. The majority of the attention given to biological criteria thus far has dealt with how they fit into existing Water Quality Criteria (WQC) and National Pollutant Discharge Elimination System (NPDES) permit frameworks. While this is certainly an important set of issues, it would be a mistake to emphasize only this one program area, as it has the demonstrated ability to be useful in virtually any issue involving the management of water resources where a goal is to protect, enhance, or restore aquatic communities and ecosystems. We define the management of aquatic resources here as being broader than the traditional purview of water quality management. Efforts to attain the goals espoused by the Clean Water Act (CWA) and other initiatives (e.g., maintenance and recovery of aquatic biodiversity) should recognize the potentially broad role that biological criteria and assessment have in each area. We believe that biological criteria and the attendant concepts of regionalization and reference sites have a broad applicability beyond the CWA.

The Ohio EPA water programs have relied extensively on ambient bioassessments since the late 1970s. The program areas within which biological criteria have found the most widespread uses are the biennial water resource inventory (305b report), water quality standards (aquatic life use classifications), NPDES permits (includes enforcement and litigation support), the construction grants program (now the State Revolving Loan Fund), the Ohio Nonpoint Source Assessment (CWA section 319), evaluation of wet weather flow impacts (stormwater, CSOs), the state certification of CWA section 404 permits (401 program) and petitioned ditches, ranking of Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) sites, and comparative risk. In addition the biological data has proved useful to other state agencies including the Ohio Department of Natural Resources (rare, threatened, and endangered species, scenic rivers, nonpoint source management, fisheries management) and the Ohio Department of Transportation (environmental impact statements).

There is little question that aquatic resources have been and continue to be degraded by a myriad of land use and resource use activities. Benke (1990) summarized the status of the nation's high quality rivers and streams concluding that fewer than 2% remain in this category. Judy et al. (1984) indicate that the declining status of surface waters across the U.S. is largely the result of nonpoint source impacts. A continued reliance on technology based and even water quality-based solutions to these problems will simply be insufficient. Water resources in Ohio and elsewhere have historically been and will continue to be impacted by human activities beyond those targeted by the NPDES permit process. These remaining problems are comparatively more complex and subtle, but are no less important or real. In fact, it is these more subtle and diffuse impacts which imperil aquatic resources to the point where additional species are declining in distribution and abundance, this in addition to those already declared as rare, threatened, or endangered (Ohio EPA 1992).

A monitoring approach, integrating biosurvey data that reflects the integrity of the water resource directly, with water chemistry, physical habitat, bioassay, and other monitoring and source information, must be central to accurately defining these varied and complex problems. Such information must also be used in tracking the progress of efforts to protect and rehabilitate water resources. The arbiter of the success of water resource management programs must shift from a heavy reliance on administrative activity accounting (numbers of permits issued, dollars spent, or management practices installed) and a pre-occupation with chemical water quality <u>alone</u> to more integrated and holistic measurements with overall water resource integrity as a goal. Biocriteria seems an essential component in making this shift.

Emphasizing aquatic life use attainment is important because: 1) aquatic life criteria oftentimes result in the most stringent requirements compared to those for the other use categories, (i.e., protection for the aquatic life use criteria should assure the protection other uses); 2) aquatic life uses apply to virtually all waterbody types and the diverse criteria (i.e., includes conventional pollutants, nutrients, toxins, habitat, physical, and biological factors, etc.) apply to all water resource management issues; and, 3) aquatic life uses and the accompanying chemical, physical, and biological criteria provide a comprehensive and accurate ecosystem perspective towards water resource management which promotes the protection of ecological integrity.

Finally, biocriteria can aid greatly the visualization of aquatic resource values and attributes. This is a critical need if we are to change the prevailing view of watersheds and streams as merely catchments and conveyances for municipal and industrial wastes, excess surface and subsurface drainage, or as obstacles to further land developments. In an effort to stem the virtually unabated loss of riparian habitat and watershed integrity, Ohio EPA has adopted a stream protection policy which sets forth guidelines under which various activities will need to be conducted in order to conserve biological integrity. Without biocriteria and the case examples developed over the past 14 years this would not have been possible and any opportunity to affect these degrading influences would have been lost.

While we have demonstrated how biological criteria can be developed and used within a state water resource management framework, some important challenges remain. The cumulative costs associated with environmental mandates, many of which consist of prescription-based regulations, have recently come into question. Both the regulated community and the public desire evidence of "real world" results in return for the expenditures made necessary by federal and state mandated requirements. Biological criteria seem particularly well suited to meet some of these needs in that the underlying science and theory is robust (Karr 1991) and biocriteria certainly qualifies as "real world".

While no single environmental indicator can "do it all", particularly in the more complex situations (*i.e.,* multiple discharges, habitat alterations, presence of toxic compounds, etc.), it is obvious that biological criteria have a major role to play. A lack of information from, or an over-reliance on any single class of indicators can result in environmental regulation that is less accurate and either under- or overprotective of the water resource. Accounting for cost is not only a matter of dollars spent, but is also a question of environmental accuracy and technical validity. In short, a credible and genuinely cost-effective approach to water quality management should include an appropriate mix of chemical, physical, and biological measures, each in their respective roles as stressor, exposure, and response indicators. Comprehensive monitoring designs using such cost-effective indicators must become a part of the "cost of doing business" and perhaps at the expense of programs where new evidence suggests that the resources devoted are disproportionate to the magnitude of the present problems (e.g., point sources vs. nonpoint sources).

Based on our experience over the past 17 years it is evident that including a biological criteria approach in a state's monitoring and assessment effort can foster a more complete integration of important ecological concepts, better focus water resource policy and management, and enhance strategic planning. Some specific examples include:

1. *Watershed Approaches to Monitoring, Assessment, and Management:* The monitoring and assessment design inherent to biological criteria is fundamentally watershed oriented and will yield information on a watershed basis.

2. *Integrated Point, Nonpoint, and Habitat Assessment and Management:* Biological criteria integrate the effects of all stressors over time and space, and the attendant use of chemical, toxicological, and physical tools enables the association of probable causes of observed impairments. This should provide a firm setting for the collaborative use of the same information for the management and regulation of both point and nonpoint sources (including habitat), two disciplines which have thus far been operated as independent programs.

3. *Cumulative Effects:* Biological communities inhabit the receiving waters all of the time and reflect the integrative, cumulative effect of various stressors. Such information provides a basis for management programs to evaluate different problems in relative terms.

4. *Biodiversity Issues:* The basic biological data provides information about species, populations, and communities of concern and also provides the opportunity to focus

beyond ecosystem elements, but include an assessment of processes as well.

5. *Interdisciplinary Focus:* Because of the inherently integrative character of the biosurvey monitoring and assessment design, a biological criteria approach provides the opportunity to bring ecological, toxicological, engineering, and other sciences together in planning and conducting assessments, interpreting the results, and using the information in strategic planning and management actions.

Examples of Non-Clean Water Act Uses

Biocriteria, because they measure the overall condition of aquatic communities and hence reflect the condition of the entire aquatic resource, are potentially useful outside the traditional purview of CWA programs. One of these areas is with nongame species, particularly the rare, endangered, threatened, and special status species listed by government agencies. Presently, in Ohio, 25 species are listed as endangered, 8 species as threatened, 13 species as special interest, 5 as extirpated, and 2 as extinct; this represents more than 30% of the Ohio fauna. Of the 41 species listed by Ohio EPA as extremely intolerant, intolerant, and sensitive (Ohio EPA 1987), 25 are listed as endangered, threatened, or special status. Sixteen additional species are in the process of significant declines, some of which are declining more rapidly than others (Rankin et al. 1992). This increases to more than 40% the fraction of the Ohio fish fauna which is potentially imperiled. If introduced species and those species that are on the fringe of their natural range are excluded, these percentages become even higher. These trends are potentially symptomatic of other environmental problems that could eventually emerge to affect attributes of surface waters which are of more direct human interest. Fish species that depend on relatively clean, silt free substrates, the continuous presence of good quality water, good instream cover, and headwater stream habitats seem to be most seriously affected. This information was provided by the biosurveys conducted by Ohio EPA over the past 14 years, thus the multiple use of the same data is exemplified. It also demonstrates the opportunity to utilize the dimensions of the data in ways which would otherwise become collapsed in the IBI evaluations. Nongame aquatic communities are not only indicators of acceptable environmental conditions for themselves, but also indicate that the water resource is of an acceptable quality for wildlife and human uses since they have the ability to integrate and reflect the sum total of disturbances in watersheds. While individual, site-specific watershed and water body disturbances themselves may seem trivial to some, the aggregate result of these individual impacts emerges in the form of a degraded and declining fauna on a regional or watershed scale. We will have a very difficult time demonstrating this problem if we do not employ monitoring and assessment efforts which generate this type of information in a scientifically credible manner which the public will accept.

Another potential use for biocriteria is in the management and assessment of lotic fisheries. The smallmouth bass (*Micropterus dolomieui*) is one of the most important game species in Midwestern rivers and streams. Furthermore, it is a species which requires little or no external support in the form of supplemental stocking. However, like any other valued fish species it does have specific habitat and water quality requirements. We examined the relationship between the occurrence and abundance of smallmouth bass with the Index of Biotic Integrity (IBI) throughout the state. The overall pattern is that this species reaches is highest abundance and occurs most frequently at sites with IBI scores at least in the fair range and preferably in the good and exceptional range. As expected, the species declines sharply as the IBI indicates increasingly degraded conditions (i.e., poor or very poor). This analysis demonstrates the relevance of the IBI to and correlation with tangible resource benefits of direct importance to resource users specifically and the public in general.

Activities requiring a permit under Section 404 of the CWA must be certified as meeting provisions of the WQS by the Ohio EPA. These are referred to as 401 certifications which largely pertain to wetlands and stream habitat impacting activities. These are the third leading cause of nonpoint source impact which has undoubtedly resulted in some of the most irretrievable impairments to aquatic life uses in Ohio. Biological evaluations of 401 certification issues has greatly increased since the adoption of numerical biocriteria and the attendant field evaluation techniques. This is presently the best legal means by which Ohio EPA can protect habitat quality. Biological criteria are especially useful in this process since habitat is a predominant factor in determining the ability of an ecosystem to support a structurally and functionally healthy assemblage of aquatic life. Furthermore, by using the result of the work that supported the development of the Qualitative Habitat Evaluation Index (QHEI; Rankin 1989), the ecological consequences of projects involving the degradation of lotic habitat can be predicted. This allows Ohio EPA to prevent unnecessary degradation of aquatic habitat and communities.

Major Factors That Determine Water Resource Integrity

Multiple factors in addition to chemical water quality are responsible for the continuing decline of surface water resources in Ohio (Ohio EPA 1997) and the U.S. (Judy et al. 1984; Benke 1990). These include the modification and destruction of riparian habitat, sedimentation of bottom substrates, and alteration of natural flow regimes on a watershed scale. Because biological integrity is affected by multiple factors, controlling chemicals alone does not assure the restoration of biological integrity (Karr et al. 1986). Biological criteria and attendant monitoring and assessment designs provide a means to incorporate broader concepts of water resource integrity while preserving the appropriate roles of the traditional chemical, physical and toxicological approaches developed over the past three decades.

The health and well-being of aquatic biota is an important barometer of whether we are achieving the Clean Water Act goal of maintaining and restoring the biological integrity of the nations's surface waters. This concept underlies the basic intent of state water quality standards. Yet this tangible end-product of Clean Water Act regulatory and water quality planning and management efforts is frequently not linked to source control and other performance measures. Simply stated, biological integrity is the combined result of chemical, physical, and biological processes in the aquatic environment. The interaction of these processes is readily apparent in the functioning of ecosystems. Thus management efforts which rely solely on comparatively simple, surrogate approaches to assessment and management carry a significant risk of failure in attempting to achieve the restoration of ecological integrity (Karr 1991). Therefore, ecological concepts, biological criteria, and attendant monitoring and assessment tools must be further incorporated into the management of surface water resources.

Understanding Biological Integrity: A Prerequisite to Biological Criteria

The term biological integrity originates from the Federal Water Pollution Control Act (FWPCA) amendments of 1972 and has remained a part of the subsequent reauthorizations. Early attempts to define biological integrity in practical terms were inconclusive. Although one of these efforts failed to produce a consensus definition or framework, several contributors urged that an ecological approach be employed (Ballentine and Guarria 1975). Biological integrity has since been considered relative to: 1) conditions that existed prior to European settlement; 2) the protection and propagation of balanced, indigenous populations; and, 3) ecosystems that are unperturbed by human activities. These criteria (especially 1 and 3) easily could be construed as referring to a pristine condition that exists in few, if any, ecosystems in the conterminous United States. Subsequent to this initial effort, a U.S. EPA sponsored work group concluded that biological integrity, when defined as some pristine condition, was difficult and impractical to define and measure (Gakstatter et al. 1981). The pristine vision of biological integrity was considered as a conceptual goal towards which pollution abatement efforts should strive, but the group also realized that past, present, and future impacts to surface waters may prevent the full realization in many parts of the U.S. More recently, efforts to construct a workable, practical definition of biological integrity have provided the supporting theory that necessarily precedes the development of standardized measurement frameworks, techniques, and criteria for determining compliance with that goal. Our concept of biological integrity follows the definition of Karr and Dudley (1981) "....the ability of an aquatic ecosystem to support and maintain a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of the natural habitats of a region". This definition alludes directly to measurable characteristics of biological communities which are found in the least impacted habitats of a region. It was this definition and the underlying ecological theory which provided the fundamental basis for the development of numerical biological criteria using a regional reference site approach in Ohio. U.S. EPA adopted a facsimile of this definition in their national program guidance on biological criteria (U.S. EPA 1991).

Biological Criteria

Biological evaluation of aquatic life integrity is made possible by monitoring aquatic communities directly. Community bioassessments differ from approaches which rely principally on target species or indicator organisms by relying on the aggregated information across multi-species assemblages. The aggregation of key community attributes functions as an indication of the more complex ecosystem elements and processes which are not measured directly and completely. At the same time information about individual species is not lost in the process and can be accessed at any time.

Six criteria that biological monitoring programs should be judged against have been defined (Herricks and Schaeffer 1985). These requirements are:

1. The measure(s) used must be biological.

2. The measure(s) must be interpretable at several trophic levels or provide a connection to other organisms not directly involved in the monitoring.

3. The measure(s) must be sensitive to the environmental conditions being monitored.

4. The response range (i.e., sensitivity) of the measure(s) must be suitable for the intended application.

5. The measure(s) must be reproducible and precise within defined and acceptable limits for data collected over space and time.

6. The variability of the measure(s) must be low.

Karr et al. (1986) evaluated the applicability of the IBI (based on stream fishes) to these criteria and found that it satisfied the six requirements. These evaluation mechanisms which are based on the recent improvements in ecological theory (re: Karr and Dudley 1981) provide a more comprehensive analysis of community information than do single dimension measures such as diversity indices, species richness, indicator species, numbers, biomass, etc. Furthermore the IBI type measures extract ecologically relevant information and provide a synthesized, numerical result that can be understood by non-biologists.

<u>Reference Condition and Reference Sites</u>

Although there is agreement that biological criteria should be based on data collected from reference sites, there exists technically different approaches. Two of these, the U.S. EPA Rapid Bioassessment Protocols (RBP; Plafkin et al. 1989) and the regional reference site approach (Hughes et al. 1986) are the most commonly used. The RBP specifies the selection of a <u>single</u> upstream or nearby reference site from which the results at unknown test sites are evaluated. One problem with selecting a <u>single</u> reference site is that the reference site could differ in more than the imposition of an impact. In regulatory applications one potential liability is that debates will center on whether the single reference site is sufficiently similar to the impacted site rather than focusing on whether the test site departs from the reference condition. If the single reference site is not representative of the impacted test sites then the resulting biological criteria will be either under or over-protective.

We have encountered situations in Ohio where insufficient knowledge about regional expectations resulted in misinterpretations about the severity of impacts in streams. A regional reference framework offers a substantial advantage for the interpretation of community responses beyond the derivation of biocriteria. By offering a more robust framework based on multiple and regionally attuned reference sites, the chance for deriving an inappropriate biocriterion is greatly reduced. Benefits will also be realized by having an approach within which the same framework and information can apply to the different programs in which the protection of aquatic life is a goal.

The selection of reference sites from which attainable biological performance can be defined is a key component in deriving biological criteria. Hughes et al. (1986) described at least seven different approaches that have been used to estimate attainable biological conditions in surface waters. Regional reference sites can have a dual role as the arbiter of regionally attainable biological performance (which is the basis for numeric biological criteria) and as an upstream reference (more commonly referred to as a control) for determining the significance of any longitudinal changes. It is important to realize this duality and the differences between each role.

Control sites are applied in the longitudinal upstream/downstream design characteristic of most water quality studies in lotic systems. While it is possible for reference sites to double as upstream control sites, the reverse is not always true. The following is a synopsis of the important and distinctive characteristics of each:

<u>Reference Sites</u>:

- "least impacted" sites are located throughout a homogeneous region (i.e., ecoregions);
- biological performance across multiple sites defines expectations and variability;
- benchmark levels (e.g., 25th percentile) of performance are used to establish numerical biocriteria within an established system of tiered aquatic life uses codified in the WQS;
- data from the reference sites are used to calibrate the IBI and ICI on a statewide basis - biocriteria are established on both an ecoregional and statewide basis;
- re-monitoring of reference sites occurs on a periodic basis (i.e., once every 10 yrs.) which provides the opportunity to make periodic adjustments to the indices, biocriteria, or both.

<u>Control Sites</u>:

- used primarily in an upstream/downstream format to evaluate longitudinal changes;
- does not serve as an arbiter of use potential or use attainment - however, the level of attainable performance for site-specific and antidegradation applications is defined;
- are important in point source monitoring and evaluation.

Ideally, reference sites for estimating attainable biological performance should be as undisturbed as possible and be representative of the watersheds for which they serve as models. Such sites can serve as references for a large number of habitat types if the range of physical characteristics within a particular geographical region are included (Hughes et al. 1986). It is for this reason, among others, that the selection of only the most pristine sites as references is inadvisable. To do so would artificially restrict reference conditions to only rarely occurring benchmarks for evaluating progress or deterioration (Hughes et al., 1986). While it is recognized that individual water bodies differ to varying degrees, the basis for having regional reference sites is the similarity of watersheds within defined geographical regions. Generally less variability is expected among surface waters within the same region than between different regions. This is because surface waters, particularly streams, derive their basic characteristics from their parent watersheds. Thus streams draining comparable watersheds within the same region are more likely to have similar biological, chemical, and physical attributes than from those located in different regions.

Framework for Deriving Numerical Biological Criteria

The derivation of biological criteria for Ohio surface waters is based on the biological community performance which can be attained at regional reference sites. The numerical biological criteria that result from the application of this framework represent the ecological structure and function that can reasonably be attained given present-day background conditions. Although these criteria are not an attempt to define pristine, pre-Columbian conditions, the framework design includes a provision for future changes to the criteria which would take place if changes in background conditions occur.

The framework within which biological criteria are established and used to evaluate Ohio rivers and streams includes the following major steps:

1. selection of indicator organism groups;

2. establish standardized field sampling, laboratory, and analytical methods;

3. selection and sampling of least impacted reference sites;

4. calibration of multi-metric indices (e.g., IBI, ICI);

5. set numeric biocriteria based on attributes of tiered aquatic life use designations;

6. reference site re-sampling (10% of sites sampled each year);

7. make periodic adjustments to the indices, biocriteria, or both as determined by reference site re-sampling results.

The key steps in this process are illustrated in Figure 4.2.1 I-VI, and presume that narrative statements of biological community condition (i.e., designated aquatic life uses) already exist in the WQS and that a regionalization scheme (e.g., ecoregions) is also included.

1. *Indicator Assemblages*

Our experience has shown that at least two assemblages should be monitored. Fish and macroinvertebrates were chosen as the routine organism groups for Ohio to monitor because both groups met the above criteria, have been widely used in environmental assessment, and there is an abundance of information about their life history, distribution, and environmental tolerances. The need to use two assemblages is apparent in the ecological differences between them, differences that tend to be complimentary in an environmental evaluation. The recovery rates differ between these two groups which can provide insights about whether or not a pollution problem has been completely abated. The value of having two assemblages independently showing the same result cannot be overstated and lends considerable strength to an assessment. However, differential responses can lead to the definition of problems that might otherwise have gone undetected in the absence of information from one or the other organism group. The differing sensitivities of the two groups is not the same to all substances or in every situation. Thus the resultant information can influence decisions to control certain substances or processes that might have been overlooked or underrated in an evaluation based on only one group. The use of these two groups is somewhat analogous to the use of a fish species and an invertebrate species as standard bioassay test organisms.

2. *Field and Laboratory Methods and Logistics*

The choice of field sampling methods is a cornerstone aspect of using and implementing bioassessments and biocriteria. Although a variety of methods and techniques are available, the choice of which ones to use should be dictated by the conditions that exist in a particular state or region. There are a number of equally valid techniques, some of which will undoubtedly work better in some habitats and/or regions of the U.S. than in others.

In selecting the appropriate field and laboratory methods there were several considerations. These include:

1. the need to produce assessments which are capable of discriminating the various impacts that occur in Ohio surface waters;

2. scientific validity; and,

3. cost-effectiveness.

These are inherently competing objectives. Elaborate and highly detailed assessments are not very cost-effective, yet the need for scientific validity prescribes an inherent level of rigor and complexity in the assessment and hence a higher cost. In contrast, assessments which lack sufficient detail and rigor may cost less, but lose in cost-effectiveness by lacking the scientific validity necessary to discriminate impacts which actually exist. Given the economic, social, and environmental consequences of the decisions being made with the data and results, it seems wiser to opt for a more complex and rigorous assessment.

3. *Criteria for Selecting Reference Sites*

The selection of reference sites is another cornerstone issue in biocriteria derivation. Should reference sites be selected primarily on a cultural basis without prior detailed knowledge of the reference site sampling results? Or, should the sampling results be used to assist in the selection of reference sites? We believe the latter approach may induce some unintentional bias into the biological criteria calibration and derivation process because of the inherent tendency to select the best sites instead of a more representative, balanced cross-section of sites that reflect both typical and exceptional communities. In extensively disturbed regions and uniquely undisturbed regions, the method of reference site selection will likely be less of an issue because of the relatively homogenous conditions. However, in regions that have a gradient of disturbances, the method of selection becomes more critical.

A notched box-and-whisker plot method was used to portray the reference site results for each biological metric by ecoregion. These plots contain sample size, medians, ranges with outliers, and 25th and 75th percentiles. Box plots have an important advantage over the use of means and standard deviations (or standard errors) because a particular distribution of the data is not assumed. Furthermore, outliers (i.e., data points that are two interquartile ranges beyond the 25th or 75th percentiles) do not exert an undue influence as they do on means and standard errors. In establishing biological criteria for a particular area or ecoregion we attempted to represent the typical biological community performance, not the outliers. The latter can be dealt with on a case-by-case or site-specific basis if necessary.

The Role of Reference Results in Biocriteria Derivation. The data obtained from sampling regional reference sites provides the basis for deriving numerical biological criteria. Reference sites serve a fundamental purpose by providing the database for calibrating the multi-metric indices and deriving the ecoregional numeric biocriteria. The reference database was used to establish the actual IBI, Invertebrate Community Index (ICI), and the Modified Index of well being (MIwb) biological criteria for each of three applicable aquatic life uses, three site types (IBI and MIwb), and across the five ecoregions of Ohio. This was done out of necessity on a statewide basis, but it could be organized differently. This is where broader calibration regions that extend beyond state boundaries could be useful.

It is imperative that reference sites meet the aforementioned criteria and thus be representative of the attainable biological community performance respective of habitat type within each ecoregion. The initial selection of reference sites occurred during the Stream Regionalization Project (SRP) of 1983-84. The results of this effort are reported in Larsen et al. (1986) and Whittier et al. (1987). While the 1983-84 SRP focused on watersheds with drainage areas of 10-300 square miles the reference site network was consequently supplemented with data from additional locations with drainage areas of 1-200 square miles sampled during 1981-89 (Ohio EPA 1987, 1989). These included reference sites on larger streams, mainstem rivers, and headwaters streams throughout the state. The transitional sections of Lake Erie tributaries, the Ohio River, and inland lakes and reservoirs were not included in this analysis. However, work is underway to address these areas within the next three to five years.

How Many Reference Sites Are Enough? We have frequently been asked this question as most are interested in deriving technically valid biological criteria at the lowest cost. Logically, enough reference sites must be selected to account for the range of natural variability among the least impacted reference sites within a region. Increased variability among reference sites, if it originates from natural sources and not sampling error, indicates the need to employ a stratification scheme among the reference sites for the purpose of biocriteria derivation. Stratification of natural variability is an essential component of biological criteria development if the resultant criteria are to become managerially useful. Our approach accomplished this through the use of tiered use designations, site types, and ecoregional stratification. Additional stratification variables could include mean annual temperature (e.g., warmwater versus coolwater streams; Lyons 1992) and gradient (e.g., low gradient versus high gradient streams; Leonard and Orth 1986).

High variability among reference sites without obvious natural causes could be a result of sampling problems which an increased number of reference sites would not correct. However, assuming proper stratification and a valid sampling approach we can then determine the minimum number of reference samples needed to arrive at a biocriterion (e.g., 25th percentile for a use designation) which adequately represents the potential biological performance of a region. The range of natural variability will not be encompassed with an insufficient reference database on which stratified expectations are to be based. This could result in biocriteria that are either under or over-protective of the biological performance defined by the designated aquatic life uses.

To illustrate the effect of reference site sample size on the Ohio EPA IBI biocriteria, we randomly selected sites from our reference database for each ecoregion and site type combination and, without replacement, recalculated the 25th percentile warmwater habitat (WWH) biocriterion after samples were added in increments of five. The procedure was performed for 50 trials over 15 different sets of reference sites (5 ecoregions X 3 site types per each ecoregion). The results were plotted on a three dimensional bar chart with the frequency at which a 25th percentile biocriteria value was randomly selected versus sample size. The analog of an asymptotic relationship of a 25th percentile IBI value with increasing sample size defined the minimum number of reference sites which are needed to achieve a biocriterion value which encompasses the inherent background variability.

Our criterion to determine when the analog to an asymptotic relationship was reached is where the variation in the 25th percentile value narrowed to one predominant value in terms of the number of observations per aggregation category. Of the 15 sets of reference samples tested (5 ecoregions X 3 site types per each ecoregion) this point ranged from a low of 10-15 samples for headwater sites in the Interior Plateau ecoregion to 75-80 samples for boat sites in the Eastern Corn Belt Plain (ECBP) ecoregion. The Huron/Erie Lake Plain (HELP) ecoregion appeared to require the fewest reference samples to reach the point of diminishing return and the ECBP ecoregion appeared to require the most reference samples. The other ecoregions tended to be intermediate between the HELP and ECBP.

Ecoregions with widespread and uniform land disturbance, such as the HELP ecoregion, require fewer samples to characterize the present reference condition while those with a greater degree of natural heterogeneity (i.e., ECBP) require the most samples. Most of the reference sites were sampled twice which makes the safe minimum number of sites for the Ohio ecoregions from as few as 5-8 sites per ecoregion per site type/stream size strata for the more homogeneous ecoregions to as many as 38-40 sites per ecoregion per site type/stream size strata for the more heterogeneous ecoregions. This may illustrate the need for further landscape stratification via sub-ecoregions. We believe that if uncertainty exists about the variability within an ecoregion more sites should be used than too few. In our experience this would be approximately 35-40 sites per ecoregion per site type. This may vary across the nation as these figures are most representative of the Midwestern U.S.

A failure to stratify variability where the clear need for a stratification scheme exists risks inaccurate biocriteria that may be under-protective of sites with greater biotic potential and over-protective of sites with lower biotic potential that otherwise would have been adequately protected by lower criteria. In contrast, attempts to stratify regions where little difference exists may lead to unnecessary regulatory complexity and an unsound and arbitrary scientific basis for biocriteria development.

The minimum number of reference sites also depends on the statistics upon which the criteria will be based. Extreme percentiles (e.g., 5th, 95th), because they represent the tails of distribution functions, are characterized by wider confidence bounds around the threshold statistic and will require a larger number of sites before a stable asymptote is reached, whereas the median of the same distribution will reach an asymptote with fewer samples (Berthouex and Hau 1991).

4. *Calibration of Multi-metric Indices for Drainage Area*

In order to establish biological criteria that are reflective of the legislative goal of attaining and restoring biological integrity in surface waters, a calibration of multi-metric indices is needed. The practical definition of biological integrity as the biological performance exhibited by the natural or least impacted habitats of a region provides the underlying basis for designing a reference site sampling network. This is not an attempt to characterize pristine or totally undisturbed, pre-Columbian environmental conditions as such exists in only a very few places, if any, in the conterminous U.S. (Hughes et al. 1982). The landscape and aquatic ecosystems of Ohio have been significantly altered during the past 150-200 years. This includes massive deforestation and conversion to agricultural and urban land use, extensive use of rivers and streams for wastewater discharges, extensive drainage and elimination of more than 90% of the wetlands, and extensive modification of stream and river habitats through channelization, impoundment, and encroachment on the riparian zone. Together these activities have radically altered the lotic ecosystems of Ohio, much of which is essentially irreversible. Thus expectations of how a biological community should perform are determined by the demonstrated attainability of natural communities at least impacted reference sites within a particular biogeographical region.

The reference site results were pooled on a statewide basis prior to constructing the drainage area scatter plots. Calibrating on a statewide (or other large area basis) as opposed to an ecoregion by ecoregion basis gives the resultant index important resolution between ecoregions. For example, it is useful to know that an index value of 30 means something different in the HELP ecoregion as compared to the WAP ecoregion while retaining comparability on a statewide basis. Having to deal with multiple, ecoregion-specific indices and resultant biocriteria values on a statewide and inter-regional basis would make communication and comparison much more difficult. Ideally, index calibration should occur on a broad spatial basis other than that defined by political boundaries. This is an area for further research and an opportunity for interstate cooperation.

5. *Set Numeric Criteria*

Once the task of calibrating the biological indices is completed the task of deriving the numerical biological criteria can proceed. However, on what basis were the decisions to select a baseline numerical criterion value for each index made? As was previously mentioned, Ohio EPA has employed a system of tiered aquatic life uses in the state WQS since 1978. These use designations are essentially narrative goal statements about the type of aquatic community attributes which are envisioned to represent each use. For the purposes of establishing numerical biocriteria, the two most important uses are WWH and Exceptional Warmwater Habitat (EWH). These use designations contain a narrative goal statement and specifies the numeric index thresholds which serve as the numeric biocriteria for each use. Numerical biological criteria for the WWH use designation, which is the most commonly applied aquatic life use in Ohio, were established as the 25th percentile value of the reference site scores by index, site type (fish), and ecoregion. The resultant numeric biocriteria for the WWH use vary by ecoregion in accordance with the narrative definition and the reference site results for each site type. It was felt that most of the least impacted reference results should be encompassed by the baseline WWH use designation for Ohio's inland rivers and streams. The selection of the 25th percentile value is analogous to the use of safety factors, which is commonplace in chemical water quality criteria applications, and has previous precedents such as the 75th percentile pH, temperature, and hardness used to derive unionized ammonia-nitrogen and heavy metals design criteria for wasteload allocations, using >20% mortality for determining significance in bioassay results, or even the 10^{-6} risk factor for human exposure to carcinogens. In this sense the 25th percentile acts as a safety factor in the derivation process. Choosing the 25th percentile as the minimum WWH criterion is conservative and reduces the influence of any unintentional bias induced by including potentially marginal sites.

Ohio EPA employs three indices as part of the numerical biological criteria: the IBI, the ICI, and the MIwb. The MIwb does not require a spatial calibration prior to use. However, both the IBI and ICI require calibration in order to establish individual metric scoring criteria tailored to the reference conditions. The sample value of each of the 12 IBI metrics is compared to the range of values from the least impacted reference sites located within the same ecoregion. Each IBI metric receives a score of 5, 3, or 1, based on whether the sample value approximates (5), deviates somewhat from (3), or strongly deviates (1) from the range of reference site values. The maximum IBI score possible is 60 (i.e., all 12 metric scores = 5) and the minimum is 12 (i.e.,, all metric scores = 1).

To determine the 5, 3, and 1 values for each IBI metric the reference site data base was first plotted against a log transformation of drainage area, the latter serving as an indicator of stream size. Other measures that have been used as an indicator of stream size include stream order (Fausch et al. 1984) and stream width (Lyons 1992). The decision to use drainage area was based on the availability and ease of calculation and relevance to stream size. Stream order was viewed as being too coarse (Hughes and Omernik 1981) and stream width is simply not representative of stream size given the widespread historical modification of streams throughout Ohio. In other regions of the U.S. these and other parameters may be appropriate for use in the calibration process. Additional dimensions could include temperature, gradient, elevation, and lake acres or shoreline distance. The one concept which continues to surface throughout this process is that these are decisions which can only be made reliably by regional experts.

The plots for each metric were visually examined to determine if any relationship with drainage area existed. If a relationship was observed a 95% line of best fit was determined and the area beneath trisected into three equal portions following the method recommended by Fausch et al. (1984). Wading and headwaters data was combined for in-common metrics to determine the slope of the 95% line even though scoring for these metrics was performed separately; all boat site IBI metrics were calibrated separately. The IBI metric scores (i.e., 5, 3, or 1) for a sample are determined by comparing the site value to the trisected scatter plots constructed from the reference site data base for each applicable metric. Certain metrics that showed no positive relationship with drainage area required the use of an alternate trisection method. Horizontal 5% and 95% lines were determined and the area between trisected. A bisection method was used only for the number of individuals metric. For two others (top carnivores, anomalies) the reference site data base was examined and scoring criteria established following Karr et al. (1986) and Ohio EPA (1987). The resultant 5, 3, and 1 values for these metrics are the same across drainage areas. A similar method of trisection was used by Hughes and Gammon (1987) for a modified IBI used in the lower 280 km of the Willamette River, Oregon.

The principal measure of macroinvertebrate community performance used by the Ohio EPA is the Invertebrate Community Index (ICI) which was originally developed by Ohio EPA (Ohio EPA 1987, DeShon, 1995). The ICI is an adaptation of the IBI concept to macroinvertebrate communities. The ICI consists of 10 structural and functional community metrics, each with four scoring categories of 6, 4, 2, and 0 points in order to result in scores which were comparable to the fish IBI scores. The point system is structured to operate the same as the IBI. The summation of the individual metric scores (determined by the relevant attributes of an invertebrate sample with consideration given to drainage area) results in the ICI value. To determine the 6, 4, 2, and 0 values for each ICI metric, the reference site database was plotted against drainage area. Each metric was visually examined to determine if any relationship existed with drainage area. When it was decided if a direct, inverse, or no relationship existed, the appropriate 95% line was estimated and the area beneath quadrisected. Certain percent abundance and taxa richness categories were not quadrisected since the data points showed a tendency to clump at or near zero. In these situations, a quadripartite method was used where one of the four scoring categories included zero values only, and, in two cases, the remaining scoring categories were delineated by an equal division of the reference data points.

A modified approach was necessary for determining the HELP ecoregion biocriteria. The HELP ecoregion is affected by significant and widespread historical land use and stream channel modifications dating to the 19th century. Setting the WWH criteria for the IBI and MIwb in this ecoregion involved detailed consideration of the extensive and essentially irretrievable physical stream habitat and watershed modifications. Based on the Qualitative Habitat Evaluation Index scores (Rankin 1989), the field observations of Ohio EPA biologists, and the descriptions of land use patterns (Whittier et al. 1987), none of the wading and headwaters reference sites in the HELP ecoregion reflected least impacted conditions relative to that observed in the other Ohio ecoregions. This distinction is made necessary by the widespread degree to which macrohabitats have been altered among the headwater and wadeable streams in the HELP ecoregion. Intensive row crop agriculture and attendant subsurface drainage practices (i.e., channel maintenance and tiling) have left few if any streams that match the intended definition of least impacted. As a result IBI and MIwb values from the wading and headwaters reference sites of this ecoregion reflect these environmentally degrading influences. Deriving the WWH wading and headwaters sites biocriteria involved an examination of IBI and MIwb results from <u>all</u> sites sampled during 1981-89 (Ohio EPA 1987, 1989). IBI and MIwb values that marked the upper 10% (90th percentile) of <u>all</u> sites sampled were selected as an alternative to the 25th percentile of the HELP reference sites which yielded lower values. The information contained in selected historical descriptions of streams in this ecoregion (Meek 1889, Trautman 1981, Kirsch 1895, Trautman 1939, 1981, Smith 1968, Trautman and Gartman 1974) was influential in making judgements about attainable WWH expectations in this ecoregion. Even with this adjustment the resulting IBI and MIwb criteria are the lowest in the state. Although the ICI values are likewise low for the HELP ecoregion the primary sampling technique is not nearly as affected by the macrohabitat modifications. Thus the 25th percentile from the reference sites was chosen as the WWH criterion for the ICI. Establishing biocriteria for the HELP ecoregion is an example of the dilemma posed by extensively disturbed areas - maintaining a balance between setting a goal for watershed restoration efforts and the pragmatic implications of maintaining present-day socioeconomic activities.

6-7. *Maintenance of the Reference Site Network and Periodic Adjustments*

The adoption of numerical biological criteria includes the task of maintaining the reference data base which includes a planned re-sampling of all sites within a prescribed time frame. A concern which is frequently expressed is that by basing aquatic community performance expectations on contemporary conditions defined by present day reference sites, aquatic life goals are somehow being frozen in time. This is why the concept of continual maintenance monitoring must be included as a part of the overall regional reference site approach. In Ohio, we have chosen to sample approximately 10% of the reference sites each year within the organization of the Five-Year Basin Approach. This will provide an opportunity to examine regional background aquatic community performance at periodic intervals (e.g., once/ten years) and make appropriate adjustments to the calibration of the multi-metric indices, the numerical biological criteria, or both.

Future Considerations and Potential Improvements

Calibration

The determination of the 95% line is one of the most important parts of the calibration process. While the line-of-best-fit method is presently accepted (Fausch et al. 1984), it is not a strict statistical derivation. As an experimental approach to possibly improve the objectivity of the 95% line determination we applied the technique described by Blackburn et al. (1992) in which a series of regression lines are determined across the upper surface of the wedge of points that result from the scatter plots of drainage area-dependent IBI metrics. Thus far we have determined this for the fish species richness metric. The results indicate a line that is not substantially different from the line-of-best-fit method. While this seems to initially confirm the line-of-best-fit method it appears to offer important advantages, the most obvious of which is a statistically objective method for determining the 95% line. One important drawback, however, is the inability of the statistic to determine when and where the slope of the line should change. This was done by visual interpretation for several of the IBI and most of the ICI metrics.

Calibration issues which need further examination include determining the degree of convergence between the 5, 3, and 1 lines at the lower drainage areas, the non-linear distribution of the scatter plots for the "percent of" metrics, and how to determine scoring for metrics which have no apparent relationship with stream size. Other considerations include the consistent designation of trophic guilds, tolerance rankings, refined metrics, refined metric scoring, and regional calibration. For example, differences exist in the designation of feeding and tolerance guilds between states which share similar faunas. In addition, criticism has been leveled at intolerant species designations as reflecting rare, threatened, and endangered status more so than true environmental tolerance. While we have dealt with most of these issues in Ohio, these and other issues will arise elsewhere, thus regional consistency in achieving a resolution of these issues will be needed.

The Ohio case example represents an effort to derive numerical biocriteria on a state-specific basis. This particular effort was constrained to data available or obtainable within the state boundaries. Political boundaries, however, seldom coincide with geographical or faunal region boundaries. Thus, consideration should be given to an alternative method to establishing what we term here as calibration regions. A calibration region is an area with a logical commonality with regards to faunal associations, species richness, waterbody type, and major drainage networks. Ideally, defining these areas would be done on the basis of regional attributes such as faunal similarity, aggregations of ecoregions and sub-ecoregions, and other relevant factors. For example a calibration region for the Midwestern U.S. might include the northern portion, or subsets therein, of the Ohio River drainage basin (all sub-drainages north of the Ohio River mainstem) which would include portions of five states. In order to begin coping with the regionally unique aspects of faunal composition, stream and river characteristics, and watershed characteristics, this type of framework seems essential if we are to maximize the utility and validity of biocriteria as water resource management decision criteria. Such a regional framework, while fostering interstate cooperation, would also provide a scientific forum for indicator selection and development, methods standardization, reference site selection, and calibration of multi-metric evaluation mechanisms. Stratification beyond this geographic level could be accomplished through the use of ecoregions, sub-ecoregions, and tiered aquatic life use designations (i.e., designated uses). This framework would also be adaptable to emerging national monitoring frameworks such as the U.S. EPA Environmental Monitoring and Assessment Program (EMAP), the U.S. Geological Survey National Water Quality Assessment (NAWQA), and the National Biological Survey (NBS). In fact, this seems to be a logical prerequisite to the analysis of the data from these efforts. Finally, regional calibration areas would provide a means of jurisdiction over the logistical and technical issues which inevitably arise within national monitoring and assessment programs.

Biological Index Variability

A frequent criticism of ambient biological data is that it is too variable to function as a reliable component of surface water resource assessment. Natural biological systems are indeed variable and seemingly noisy, but no more so than the chemical and physical components that also exist within aquatic ecosystems. Certain dimensions of ambient biological data are quite variable, particularly population or sub-population level parameters. Single dimension community measures can also be quite variable. The new generation community evaluation mechanisms such as the IBI and ICI are sufficiently redundant so as to compress and dampen some of the aforementioned variability. Rankin and Yoder (1990) examined replicate variability of the IBI from nearly 1000 sites throughout Ohio and found it to be quite low at least impacted sites. Coefficient of variation (CV) values were less than 10% at IBI ranges indicative of exceptional biological performance and less than 15% for the good performance range. This is lower than the variability reported for chemical laboratory analyses and inter-laboratory bioassays (Mount 1987). Variability as portrayed by CV values increased at IBI ranges indicative of increasingly impaired biological performance. Low variability was also found for the ICI with a CV of 10.8% for 19 replicate samples at a relatively unimpacted test site (DeShon, 1995). The variability of the MIwb was determined to be on the order of $\pm$0.5 MIwb units (Ohio EPA 1987). Other investigators have reported similarly low variability with other biological indices (Davis and Lubin 1989, Stevens and Szczytko 1990). Fore et al. (1993) used different statistical techniques and determined a variability of ± 3 IBI units using the Ohio database. Cairns (1986) suggested that differences in variability rather than differences in averages or means might be the best measure of stress in natural systems. Variability must begin to be recognized as a part of the signal rather than noise alone (Karr 1991). Not only is the variability of the measures used as biological criteria low, the degree of variability encountered can also be a useful assessment and interpretation tool.

Ohio EPA has addressed the variability inherent to biological measures in three general ways:

1. Variability is compressed through the use of multi-metric evaluation mechanisms such as the IBI and ICI.

2. Variability is stratified by the tiered use classification system, ecoregions, biological index calibration, and site type.

3. Variability is controlled through standardized sampling procedures which address seasonality, effort, replication, gear selectivity, and spatial concerns.

Initial Decisions and Other Considerations

There are a number of fundamental decisions which need to be made early in the development of biocriteria. This is a critical juncture in the process since these initial decisions will determine the overall effectiveness of the effort well into the future. Decisions about which sampling methods and gear to use, seasonal considerations, which organism groups to monitor, which parameters to measure, which level of taxonomy to use, etc. will need to be made. The axiom follows "....when in doubt choose to take more measurements than seem necessary at the time since information not collected is impossible to retrieve at a later date". This does not apply equally to all parameters. For example, seasonality is a well understood concept, therefore it is not necessary to sample in multiple seasons for the sake of data redundancy. However, parameters which require little extra effort to acquire should be included until enough evidence is amassed to evaluate its relative worth. One example in Ohio is with external anomalies on fish. A decision was made to record this information even though it was not immediately apparent what use this information would have. This one parameter has proven over time to be one of our most valuable assessment tools. For macroinvertebrates the issue of identifying midges to the genus/species level (as opposed to the family level) proved likewise to be a far sighted decision given the value of this group in diagnosing impairments. Samples could have been archived for later processing, but the logistical burdens that this would entail later on are even more undesirable.

Another important consideration is assuring that qualified and regionally experienced staff are available to implement the monitoring and assessment activities. Ecological assessment is no less in need of skill and experience than are other professions. However, biological field assessment is somewhat unique in that an equivalent level of expertise is needed in the field since many of the critical pieces of information are recorded and, more importantly, interpreted there. There is simply no substitute for direct experience in the field - this is not a job to be left to technicians. In addition, it is only prudent that the same professional staff who collect the field data also interpret and apply the information derived from the data in a "cradle to grave" fashion. Thus the same staff who perform the field work also plan that work, process the data into information, interpret the results, and apply the results via assessment and reporting. Such staff, particularly the more experienced ones, also contribute to policy development.

Logistics and Costs

The approach used by Ohio EPA to collect macroinvertebrate and fish community data is intended to secure an adequate sample, but not necessarily an exhaustive inventory. Fish relative abundance data is collected using standardized, pulsed D.C. electrofishing techniques. In an analysis of resources expended during FFY (Federal Fiscal Year; October 1 - September 30) 1987 and 1988 the following was revealed:

- 8.44 WYE (work year equivalents) were used to collect 1277 samples at 617 sites.
- An average of 0.014 WYE or 29.1 hours/site were expended to plan, collect, analyze, interpret data, and produce reports at an average cost of $740/site.
- This translates into 1-3 hours/sample with a field crew sampling 3-6 sites/day by working 10-14 hours/day.
- Post-field season laboratory effort ranges from 1-3 weeks.
- A field crew consists of one full-time biologist and two interns.

The approach used by Ohio EPA to collect macroinvertebrate community data is intended to secure an adequate sample, but not an exhaustive inventory of all taxa possible. Relative abundance data is collected using a standardized approach (artificial substrates; DeShon (1995). In an analysis of resources expended during FFY 1987 and 1988 the following was revealed:

- 5.02 WYE (work year equivalents) were used to collect 323 samples at 323 sampling sites, setting or retrieving 4-6 sites per day.
- An average of 0.015 WYE or 33.2 hours/site were expended to plan, collect, analyze, interpret data, and to produce reports at an average cost of $824/site.
- Laboratory effort is 12-20 hours/sample for artificial substrates and 2-6 hours for qualitative samples only.
- A field crew consists of one full-time biologist and one intern.

Concern is frequently expressed not only about the practical utility of biological field data, but the resources needed to implement such programs (Loftis et al. 1983, U.S. EPA 1985). Whole effluent toxicity evaluation has been advocated partly because it is viewed as more cost-effective than biological field evaluations (U.S. EPA 1985). Our experience with using a standardized and systematic application of biological field monitoring techniques integrated with the traditional chemical/physical and bioassay assessment techniques allows a detailed comparison of the costs involved with each component. Out of nearly 100 WYE (Work Year Equivalents) that were devoted to surface water monitoring and laboratory activities within the Division of Water Quality Planning and Assessment in FFY (Federal Fiscal Year) 1987 and 1988, 19.34 WYE or just over 19% of the total was devoted to ambient biological monitoring. When considered on the basis of agency-wide water programs this percentage is approximately 6%.

Table 4.2.1 gives the unit cost of the four monitoring components that are being compared and evaluated. Costs are broken down by sample collection, laboratory analysis, individual test, and evaluation as appropriate for each component. Included in the cost figures are all equipment, supplies, logistical, administrative, data analysis, and interpretation activities. Chemical/physical water quality costs were derived from grab samples taken from 3 to 8 times at each site during summer-fall low flow periods (mid-June through mid-October). Bioassay costs were on the tests routinely performed by Ohio EPA: 48 hour screening tests, 48 and 96 hour definitive tests, and seven-day acute/chronic tests. The seven-day tests were further subdivided between those analyzing daily composites (i.e., seven-day renewal test) and those designed to test one 24-hour "megagrab" sample.

An initial comparison of the cost of each component is evident from an examination of Table 4.2.1. Fish and macroinvertebrate evaluation costs/site are comparable. Obtaining chemical/physical water quality data was approximately twice that of either biological method alone, but only 5% more than both organism groups together. Comparison of bioassay costs was most appropriately done on a point source entity evaluation basis. For example, three to six biological sites are usually required to evaluate the impact from a single point source (a cost of $4692 to $9384 for both fish and macroinvertebrates), whereas a definitive and/or seven-day bioassay test costs $1848 and $3052, respectively. However, these bioassay costs represent those for a single test, not a complete, three test evaluation. U.S. EPA (1985) protocols specify three tests per evaluation per entity making the bioassay evaluation cost $5544 for a definitive evaluation and $9156 for a seven-day evaluation. Thus, sampling both fish and macroinvertebrates is comparable to a definitive or seven-day bioassay evaluation, and more so if a seven-day renewal test is employed.

Using an example situation, the cost to evaluate three entities discharging to a small river for acute and chronic toxicity using the seven-day static test would be $27,468 ($54,954 for a seven-day renewal test). Fish and macroinvertebrates sampled at 18 locations in the mainstem would cost $28,152. Furthermore, some of the 18 biological sampling locations would also be devoted to monitoring influences other than toxicity from point sources. For example the influence of factors that exert their effects by means other than toxicity (e.g., habitat, sediment, nutrient enrichment, flow alterations, low dissolved oxygen, etc.) will be apparent in the biological data and such results should play a key role in decision-making about water quality based effluent limits and other management needs. In this example, it was estimated that 12 of the biological sites would be necessary to determine the cumulative impact of the three point sources which results in a comparative biological component cost of $18,768. It is recognized that the chemical-specific and toxicity evaluations perform a uniquely essential function in attempting to separate relative contributions from interacting sources. This is an oft cited shortcoming of biosurvey results although response signatures are discernible in the data (Yoder 1991, Yoder and Rankin 1995a). The quality of the eventual decisions about water quality standards and discharge limitations (chemical or otherwise) would suffer significantly without the information provided by an integrated evaluation including chemical, biological, and toxicity measures (Yoder 1991).

States that do not operate extensive ambient bioassessment networks will need to be prepared for some rather sizeable start-up costs. While the cost analysis incorporated start-up costs for equipment and supplies, these were amortized over 5 or 10 years depending on the expected life of an item. Start-up equipment and supplies, for most states, could total from $200,000 to $500,000 depending on the number of field crews involved.

Data Management and Information Processing

Once field data is collected, processed, and finalized the next step is to reduce the data to scientifically and managerially useful information. The principal Ohio EPA data management system for fish, macroinvertebrate, and habitat data (Ohio ECOS) includes storage, processing, and analysis routines. Once data is tabulated in the field (fish and habitat) and laboratory (macroinvertebrates) and documented via chain of custody procedures, the data is entered directly into the electronic database. Basic information includes the field crew, waterbody name, date, and time. Site location is indicated by river mile (distance upstream from mouth) and latitude/longitude, both of which are determined from USGS 7.5 minute topographic maps. A basin-river code system is used to electronically identify individual streams, rivers, and lakes. Sampling information includes method or gear type and other information relevant to the use of each. Ohio ECOS generates data summaries and reports for a variety of community measures, community composition, or individual species/taxon analyses.

Conclusions

Biological criteria are an emerging and increasingly important issue for EPA, the states, and the regulated community. The use of biocriteria through bioassessments is growing nationwide as more states and local organizations shift their monitoring and assessment efforts in this direction. However, much remains to be done, particularly in the area of national and regional leadership. Technical guidance and expertise is needed to ensure a nationally consistent and credible approach and to resolve outstanding technical concerns listed by Yoder and Rankin (1995a). Resolving outstanding policy issues such as EPA's policy of independent applicability needs to be accomplished in such a manner as to encourage, not discourage, states to participate. In an era of declining government resources ways to accomplish the "increases" needed in biological monitoring to support the biocriteria approach must be developed. Based on our experience in Ohio the staffing of state programs should be a minimum of one work year equivalent for every 1200 miles of perennial streams and rivers. This estimate may vary by region and should additionally incorporate lake acres in states with a predominance of this water body type (Yoder and Rankin 1995b). The potential for bioassessments and biocriteria to modify the present capital and resource intensive system of tracking environmental compliance on a pollutant specific basis needs to be considered by EPA. This should prove to be a more cost and information effective approach to managing the nation's water quality programs.

Ohio EPA has been monitoring the condition of Ohio's surface waters intensively since the late 1970s. Biological assessment has always been emphasized and this was further formalized with the adoption of numerical biological criteria in 1990. Beginning in 1990 Ohio instituted a "5-Year Basin" approach to monitoring and NPDES permit reissuance. This schedule has been devised so that monitoring data is collected in advance of permit reissuance, implementation of best management practices for nonpoint sources, or other management actions which benefit from monitoring information. The 15 plus years of using an integrated biosurvey approach to monitor major sources of pollution has put Ohio EPA in a position to determine the effectiveness of water quality-based pollution controls. This effort has resulted in a shift away from a sole reliance on regulatory and administrative activities as the principal measures of success to the inclusion of environmental results oriented measures.

The eventual attainment of the CWA goal of biological integrity means more than achieving a higher level of species diversity, numbers, and/or biomass. In fact there are situations when increases in any one of these attributes may signify degradation. Managers also must strive for more than the protection target species, an effort which sometimes receives a disproportionate share of scarce resources. Merely conserving imperiled species, while nonetheless essential, is alone insufficient for maintaining and restoring biological integrity. Conservation policy needs to promote management practices which maintain and restore biological integrity, prevent endangerment, and enhance the recovery of species and ecosystems (Angermier and Williams 1993). The goals of water resource management must begin to focus additionally on the maintenance of self-sustaining and functionally healthy aquatic communities. Achieving this state of aquatic ecosystem integrity will "bring along" these other goals as well since functionally healthy communities include the elements of biodiversity and rare species that the more narrowly focused management efforts are striving to attain. Biological criteria can and should play an important role in meeting these challenges.

Ohio EPA has placed a high emphasis on monitoring as being much more than a data gathering activity by integrating the results into the entire water quality management process. Because of the investment made in monitoring over the past 15 years, we are now reaping benefits by being able to quantify improvements resulting from our regulatory efforts of the past 20 years, producing accurate estimates of goal attainment and non-attainment (Rankin et al. 1992), and in being able to approach new and emerging issues from a sound environmental basis. Invaluable insight into the potential uses of biological criteria, their advantages, and their limitations has been gained. A wide array of different types and degrees of environmental perturbation (both chemical and nonchemical) have been observed and evaluated. This experience provided the basis for many of the concepts and findings that are presented herein.

A growing body of information shows that other factors in addition to chemical water quality are responsible for the continuing decline of surface water resources in many cases (Judy et al. 1984, Rankin and Yoder, 1990). Because biological integrity is affected by multiple factors in addition to chemical water quality, controlling chemicals alone does not in itself assure the restoration of biological integrity (Karr et al. 1986). If we are to make progress in the restoration and protection of aquatic ecosystems our concerns must incorporate a broader focus on the water resource as a whole. The concepts inherent to biological integrity implicitly include such holism. Whole effluent toxicity testing offers an improvement over a strictly chemical approach, but alone lacks the ability to broadly assess ecosystem effects, particularly those caused by physical, episodic, and nontoxic chemical impacts. Biological criteria and the attendant biosurvey approach to monitoring and assessment provides a means to incorporate the broader concept of water resource integrity while preserving the traditional chemical/physical and toxicological approaches of the past three decades.

References

Angermier, P.L. and J.E. Williams. 1993. Conservation of imperiled species and reauthorization of the endangered species act of 1973. Fisheries 18(7):34-38.

Ballentine, R. K. and Guarrie L. J., Eds., 1975. The integrity of water: a symposium, U.S. Environmental Protection Agency, Washington, D.C., 230 pp.

Blackburn, T. M., Lawton, J. H., and Perry, J . N., 1992. A method for estimating the slope of upper bounds of plots of body size and abundance in natural animal assemblages, Oikos, 65: 107-112.

Benke, A.C. 1990. A perspective on America's vanishing streams. J. N. Am. Benth. Soc., 9 (1): 77-88.

Berthouex P.M & I. Hau. 1991. Difficulties related to using extreme percentiles for water quality regulations. Research journal of the water pollution control federation 63:873-879.

Cairns, J., Jr. 1986. Freshwater. In Proceedings of the workshop on cumulative environmental effects: a binational perspective. CEARC, Ottawa, Ontario and NRC, Washington, D.C. 175 pp.

Davis, W.S. and A. Lubin. 1989. Statistical validation of Ohio EPA's invertebrate community index, pp. 23-32. in Davis, W.S. and T.P. Simon (eds.). Proc. 1989 Midwest Poll. Biol. Mtg., Chicago, Ill. EPA 905/9-89/007.

DeShon, J.D. 1995. Development and application of the invertebrate community index (ICI), pp. 217-243. in W.S. Davis and T. Simon (eds.). Biological Assessment and Criteria: Tools for Risk-based Planning and Decision Making. Lewis Publishers, Boca Raton, FL.

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

Fore, L.S., J.R. Karr, and L.L. Conquest. 1993. Statistical properties of an index of biotic integrity used to evaluate water resources. Can. J. Fish. Aquatic Sci.

Gakstatter, J. and others. 1981. A recommended approach for determining biological integrity in flowing waters. U.S. EPA, Corvallis, Oregon, 26 pp.

Herricks. E.E. and D.J. Schaeffer. 1985. Can we optimize biomonitoring? Environmental Management 9: 487-492.

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

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

Hughes, R. M., Gakstatter, J. H., Shirazi, M. A., and Omernik, J. M., 1982. An approach for determining biological integrity in flowing waters, in *Inplace resource inventories: principles and practices, a National workshop*, Braun, T. B., Ed., Society of American Foresters, Bethesda, MD.

Hughes, R. M., and Omernik, J. M. 1981. Use and misuse of the terms watershed and stream order, in Warmwater Streams Symposium, Krumholz, L., Ed., American Fishery Society, Bethesda, MD.

Judy, R. D., Jr., P. N. Seely, T. M. Murray, S.C. Svirsky, M. R. Whitworth, and L.S. Ischinger. 1984. 1982 national fisheries survey, Vol. 1. Technical Report Initial Findings. U.S. Fish and Wildlife Service, FWS/OBS-84/06.

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

Karr, J. R., K. D. Fausch, P. L. Angermier, P. R. Yant, and I. J. Schlosser. 1986. Assessing biological integrity in running waters: a method and its rationale. Illinois Natural History Survey Special Publication 5: 28 pp.

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

Kirsch, P. H. 1895. A report upon investigations in the Maumee River during the summer of 1893, Bulletin of the U. S. Fishery Commission, 16, 315.

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

Leonard, P. M., and D.J. Orth 1986. Application and testing of an index of biotic integrity in small, coolwater streams. Trans. Am. Fish. Soc. 115: pp. 401.

Loftis, J. C., Ward, R. C., and Smillie, G. M. 1983. Statistical models for water quality regulation, Journal of the Water Pollution Control Federation, 55, 1098.

Lyons,J. 1992. The length of stream to sample with a towed electrofishing unit when fish species richness is estimated. N. Am. J. Fish. Manag. 12:198-203.

Lyons, J. 1992. Using the index of biotic integrity (IBI) to measure environmental quality in warmwater streams of Wisconsin. Gen. Tech. Rep. NC-149. St. Paul, MN: USDA, Forest Serv., N. Central Forest Exp. Sta. 51 pp.

Meek, S. E. 1889. Notes on a collection of fishes from the Maumee valley, Ohio, Proceedings of the U. S. National Museum, 2, 435.

Mount, D.I. 1987 (unpublished). Comparison of test precision of effluent toxicity tests with chemical analyses. U.S. EPA, Env. Res. Lab., Duluth, Minnesota.

Ohio Environmental Protection Agency. 1987. Biological criteria for the protection of aquatic life: volume II. Users manual for biological field assessment of Ohio surface waters. Division of Water Quality Monitoring and Assessment, Surface Water Section, Columbus, Ohio.

Ohio Environmental Protection Agency. 1989. Biological criteria for the protection of aquatic life. Volume III: Standardized biological field sampling and laboratory methods for assessing fish and macroinvertebrate communities. Ohio EPA division of water quality planning and assessment, Ecological assessment section, Columbus, Ohio.

Ohio Environmental Protection Agency. 1992a. Biological and habitat investigation of greater Cincinnati area streams: the impacts of interceptor sewer line construction and maintenance, Hamilton and Clermont Counties, OH, OEPA Tech. Rep. EAS/1992-5-1.

Ohio Environmental Protection Agency. 1992b. Ohio water resource inventory. Volume I, status and trends. E.T. Rankin, C.O. Yoder and D.A. Mishne (editors). Ohio EPA, Division of water quality planning and assessment, Ecological assessment section, Columbus, Ohio.

Ohio Environmental Protection Agency. 1997. Ohio water resource inventory, volume I, summary, status, and trends. Rankin, E.T., Yoder, C.O., and Mishne, D.A. (eds.), Ohio EPA Tech. Bull. MAS/1996-10-3-I, Division of Surface Water, Columbus, Ohio. 190 pp.

Plafkin, J. L. and others. 1989. Rapid Bioassessment Protocols for use in rivers and streams: benthic macroinvertebrates and fish. EPA/444/4-89-001. U.S. EPA. Washington, D.C.

Rankin, E. T. 1989. The qualitative habitat evaluation index (QHEI), rationale, methods, and application, Ohio EPA, Division of Water Quality Planning and Assessment, Ecological Assessment Section, Columbus, Ohio.

Rankin, E. T. and C. O. Yoder. 1990. A comparison of aquatic life use impairment detection and its causes between an integrated, biosurvey-based environmental assessment and its water column chemistry subcomponent. Appendix I, Ohio Water Resource Inventory (Volume 1), Ohio EPA, Div. Water Qual. Plan. Assess., Columbus, Ohio. 29 pp.

Rankin, E.T., C.O. Yoder and D.A. Mishne. 1992. Ohio water resource inventory, Volume I: Summary, Status and Trends, 1992. Division of Water Quality Planning and Assessment, Ecological Assessment Section. Columbus, Ohio.

Smith, P.W. 1968. An assessment of changes in the fish fauna of two Illinois rivers and its bearing on there future. Transactions Illinois State Academy of Science, 61(1): 31-45.

Stevens, J.C. and S.W. Szczytko. 1990. The use and variability of the biotic index to monitor changes in an effluent stream following wastewater treatment plant upgrades, pp. 33-46. in Davis, W.S. (ed.). Proc. 1990 Midwest Poll. Biol. Mtg., Chicago, Ill. EPA-905-9-90/005.

Trautman, M. B., 1939. The effects of man-made modifications on the fish fauna in Lost and Gordon Creeks, Ohio, between 1887-1938, *Ohio Journal of Science*, 39, 275.

Trautman, M.B. 1981. The Fishes of Ohio. Ohio State University Press. Columbus, Ohio. 782 pp.

Trautman, M.B. and R.K. Gartman. 1974. Re-evaluation of the effects of man-made modifications of Gordon Creek between 1887 and 1973 and especially as regards its fish fauna. Ohio J. Sci. 74(3): 162-173.

US Environmental Protection Agency. 1985. Technical support document for water quality-based toxics control. EPA-440-4-85-003. U.S. Environmental Protection Agency, Office of water, Washington, D.C.

U.S. Environmental Protection Agency. 1991. Environmental monitoring and assessment program. EMAP - surface waters monitoring and research strategy - fiscal year 1991, EPA/600/3-91/022. Office of Research and Development, Environmental Research Laboratory, Corvallis, OR. 184 pp.

Whittier, T. R., Larsen, D. P., Hughes, R. M., Rohm, C. M., Gallant, A. L. and J. R. Omernik. 1987. The Ohio stream regionalization project: A compendium of results. Report No. EPA-600/3-87/025. U. S. Environmental Protection Agency, Corvallis, Oregon.

Yoder, C.O. 1991. The integrated biosurvey as a tool for evaluation of aquatic life use attainment and impairment in Ohio surface waters. Pages 110-122 in Biological criteria: research and regulation, Proceedings of a symposium, 12-13 December 1990, Arlington, Virginia. EPA-440-5-91-005. U.S. Environmental Protection Agency, Office of water, Washington, D.C.

Yoder, C.O. and E.T. Rankin.1995a. Biological criteria program development and implementation in Ohio, pp. 109-144. *in* W. Davis and T. Simon (eds.). Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. Lewis Publishers, Boca Raton, FL.

Yoder, C.O. and E.T. Rankin. 1995b. Biological response signatures and the area of degradation value: new tools for interpreting multimetric data, pp. 263-286. *in* W. Davis and T. Simon (eds.). Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. Lewis Publishers, Boca Raton, FL.

Table 4.2.1 Cost comparison of macroinvertebrate community and fish community evaluations with chemical/physical grab sampling and acute and acute/chronic bioassay tests.

Sample Collection	Analytical Cost (Laboratory)	Cost per Test /Sample	Cost per Evaluation
	Macroinvertebrate Community		
	Artificial Substrates (includes qualitative sample)		
N/A	$397	$824	$824
	Qualitative Sample Only		
N/A	$150	$275	$275
	Fish Community		
	Cost per sample		
N/A	N/A	$340	$340
	Cost per site		
N/A	N/A	$340	$740
	Chemical/Physical Water Quality (4.6 samples/site)		
$1124[1]	$529[2]	$359	$1653
	Bioassay		
	Screening[3]		
$261	N/A	$1191d	$3573
	Definitive[4]		
$261	N/A	$1848	$5544
	Seven-day[5]		
$261	N/A	$3052	$9156
	Seven-day[6]		
$1973	N/A	$6106	$18318

[1]includes cost of sample collection and data analysis only; based on an average frequency of 4.6 samples/site in 1987 and 1988;
[2] analytical costs based on each sample being analyzed for 5 heavy metals ($7.00 ea.), 4 nutrients ($10.00 ea.), COD or BOD ($20.00 ea.), and 2 additional parameters ($20.00 for both); $115 per sample;
[3] 48 hour exposure to determine acute toxicity;
[4]48 and 96 hour exposure to determine LC50 and EC50;
[5]seven-day exposure to determine acute and chronic effects using a single 24-hour sample; cost based on analysis of one pipe only; costs for chemical analyses in sole support of the test are not included.
[6]seven-day exposure using a composite sample collected daily (renewal); other factors apply.

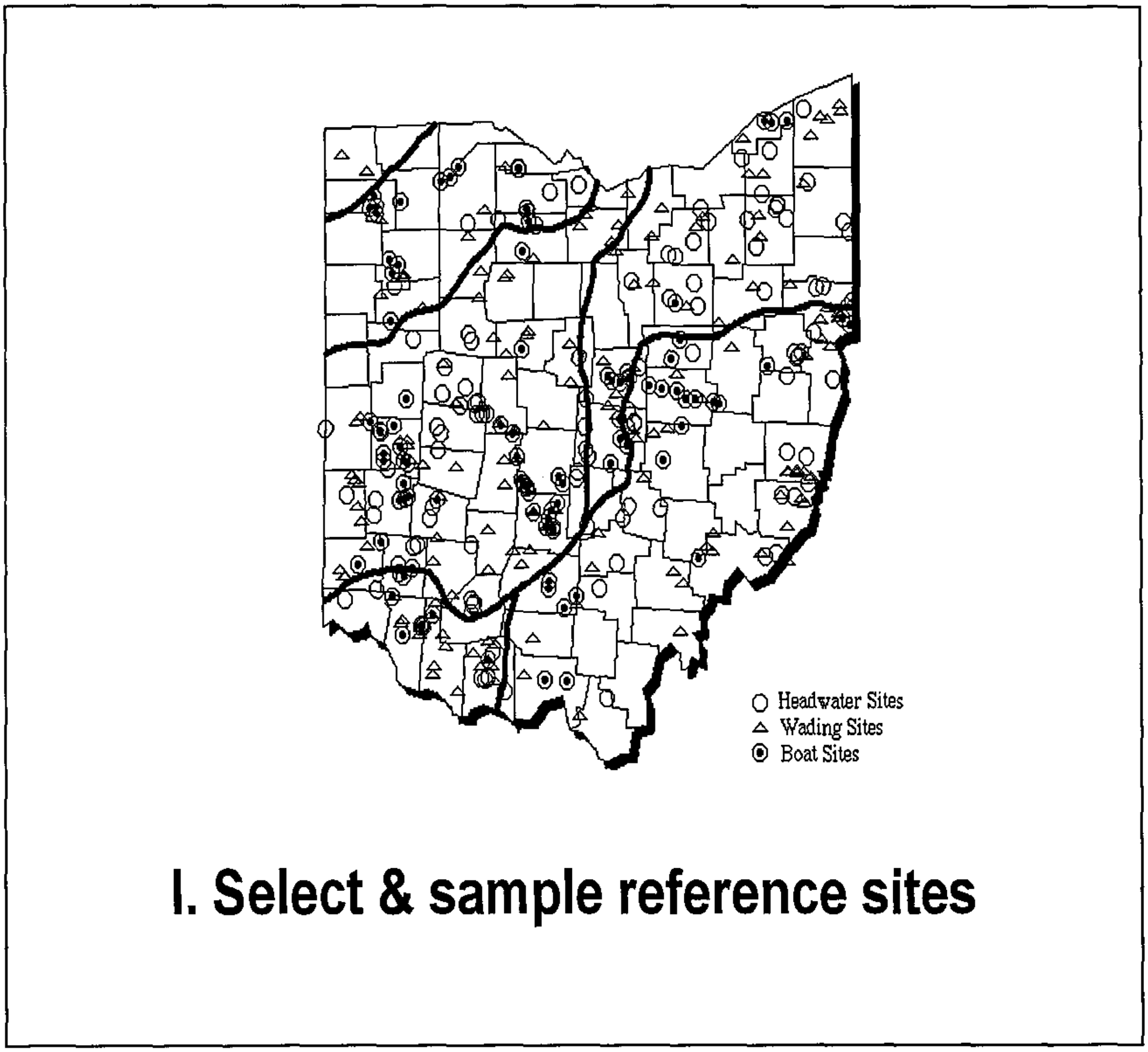
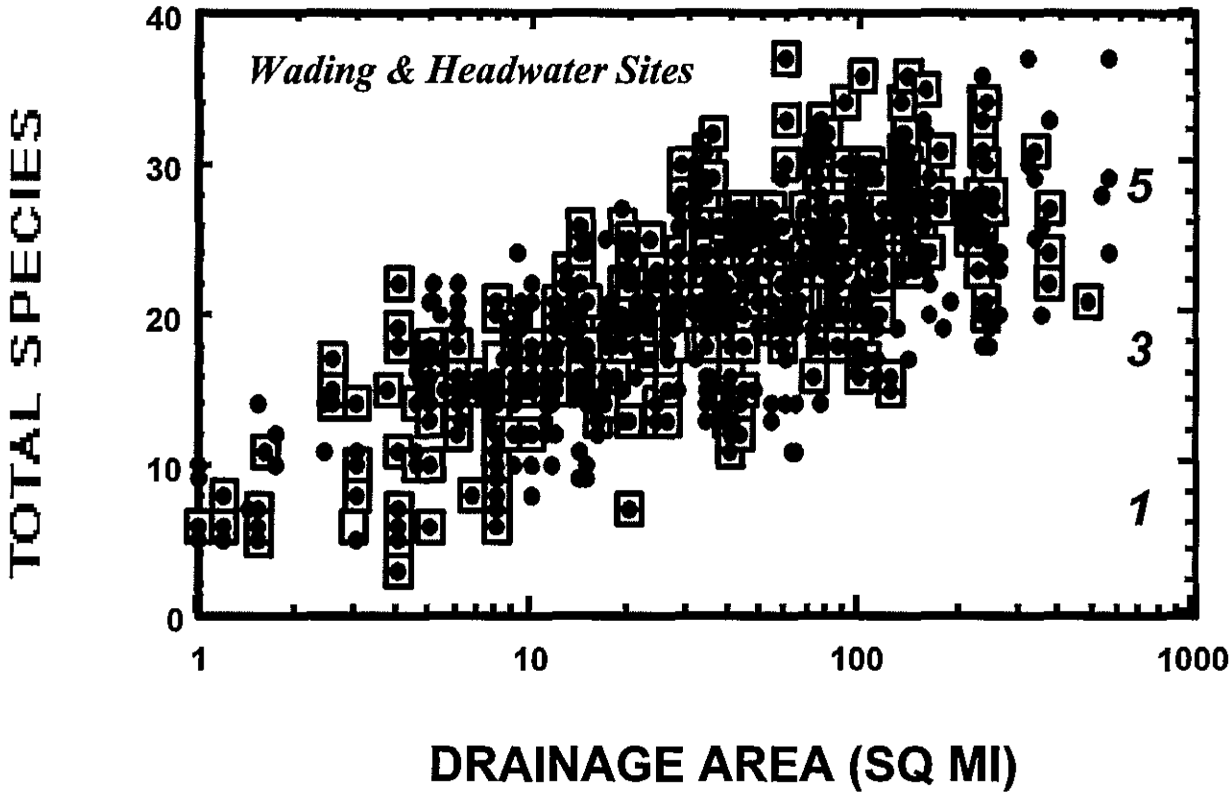


Figure 4.2.1 Key steps in the process of establishing biological criteria to evaluate Ohio rivers and streams.

Figure 4.2.1 (cont.)



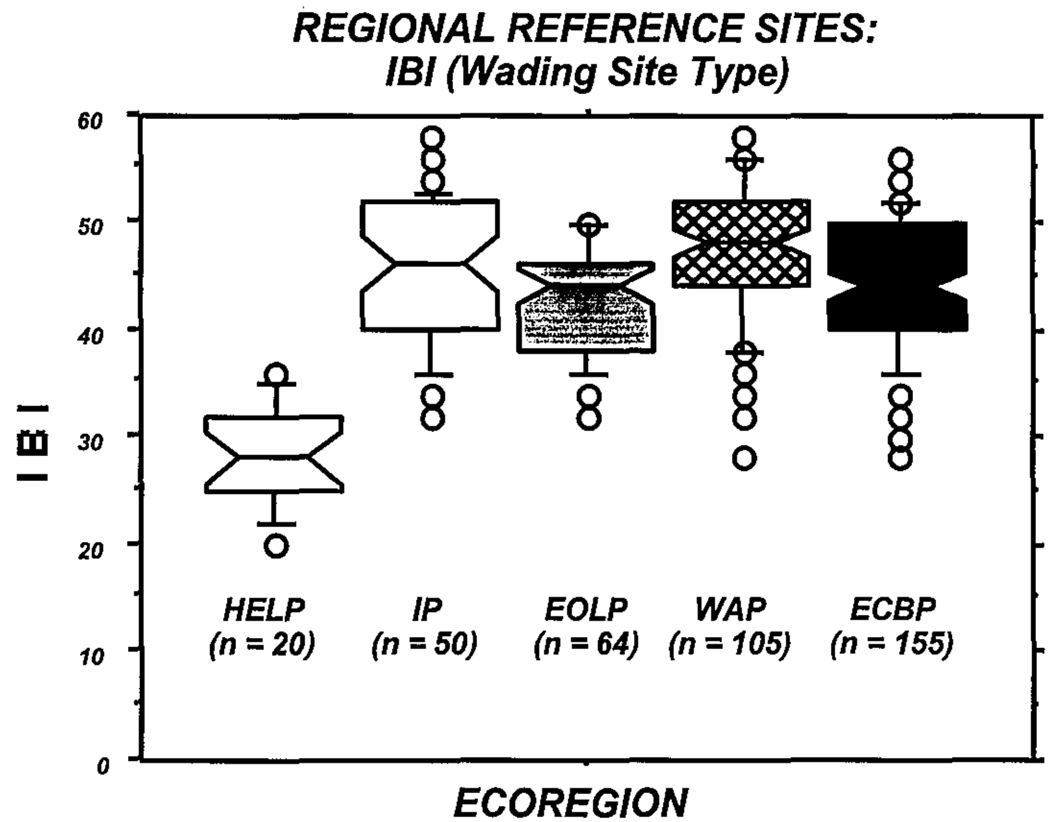
II. Calibration of IBI metrics

Figure 4.2.1 (cont.)

III. Calibrated IBI modified for Ohio waters

Metric	***5***	***3***	***1***
Number of Species	***Varies x Drainage Area***		
No. of Darter Spp.	***Varies x Drainage Area***		
No. of Sunfish Spp.	***>3***	***2-3***	***<2***
No. of Sucker Spp.	***Varies x Drainage Area***		
Intolerant Species			
>100 sq. mi.	***>5***	***3-5***	***<3***
<100 sq. mi.	***Varies x Drainage Area***		
%Tolerant Species	***Varies x Drainage Area***		
%Omnivores	***<19***	***19-34***	***>34***
%Insectivores			
<30 sq. mi.	***Varies x Drainage Area***		
>30 sq. mi.	***>55***	***26-55***	***<26***
%Top Carnivores	***>5***	***1-5***	***<1***
%Simple Lithophils	***Varies x Drainage Area***		
%DELT Anomalies	***>1.3***	***0.5-1.3***	***<0.5***
Relative Abundance	***>750***	***200-750***	***<200***

Figure 4.2.1 (cont.)



IV. Establish ecoregional patterns/expectations

Figure 4.2.1 (cont.)

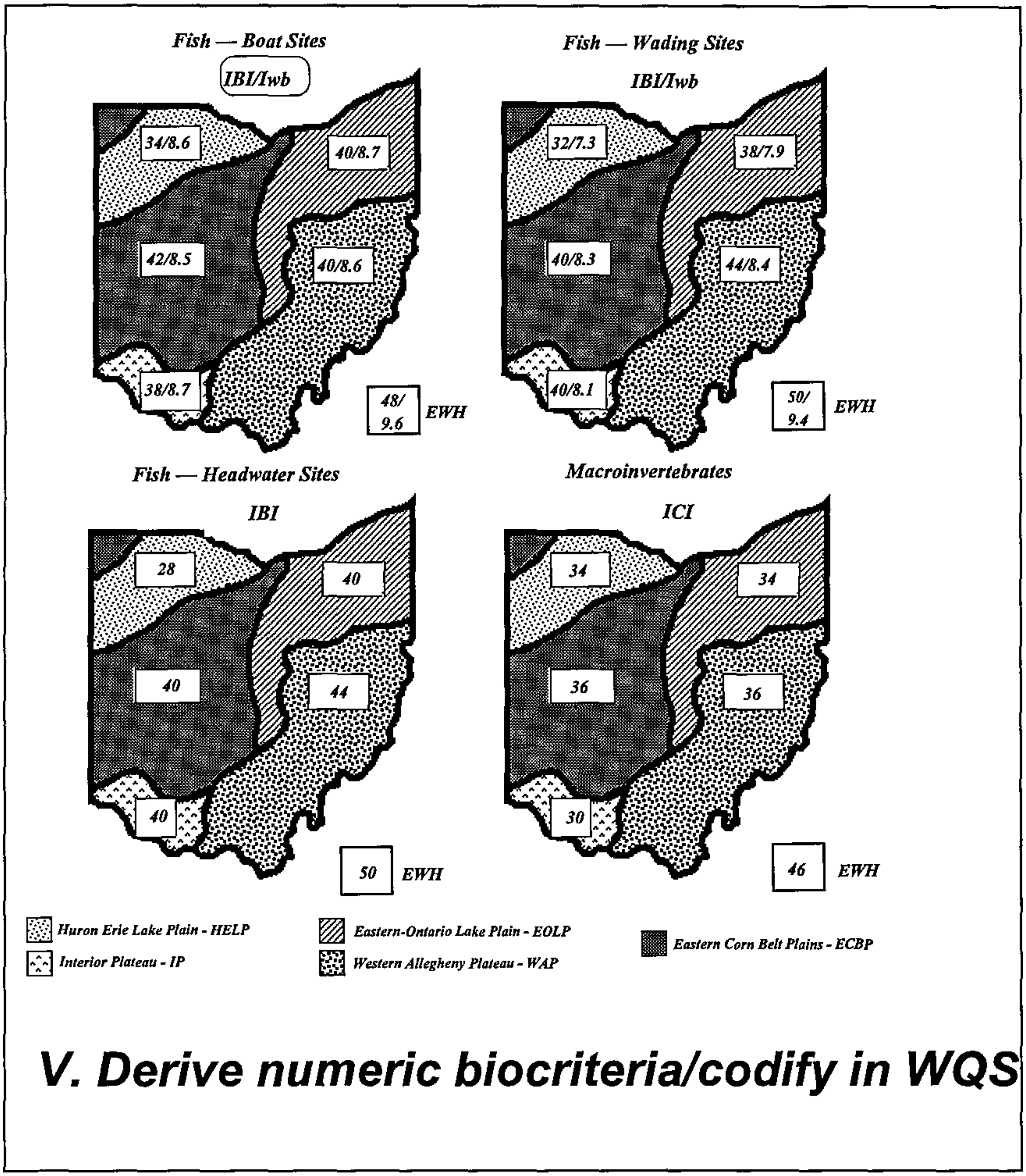
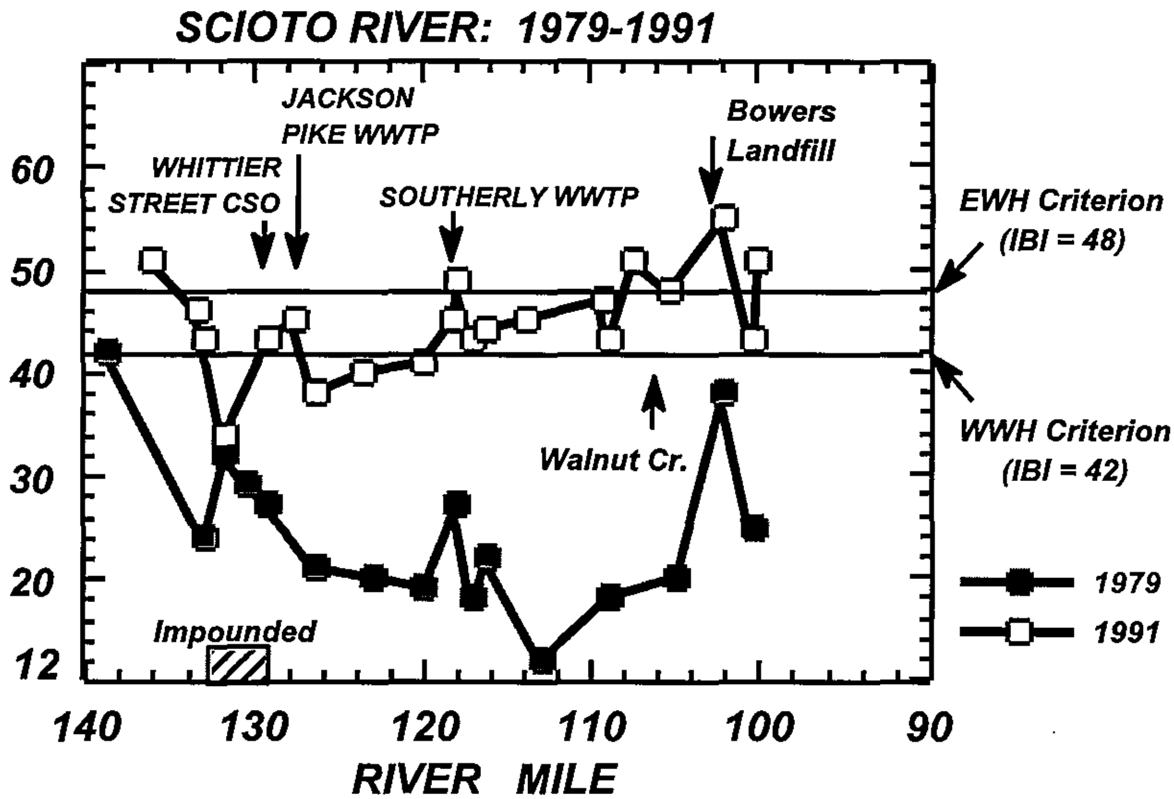


Figure 4.2.1 (cont.)



VI. Numeric biocriteria used in assessments

4.3 ASSESSMENT OF THE IMPACT OF WATERSHED DEVELOPMENT ON THE NURSERY FUNCTIONS OF TIDAL CREEK HABITATS.

A. F. Holland, G. H. M. Riekerk, S. B. Lerberg, L. E. Zimmerman, and D. M. Sanger,

S. C. Dept. of Natural Resources, Marine Resources Research Inst., Charleston, SC

Meandering shallow tidal creeks and the associated intertidal salt marshes are dominant features of Southeastern estuaries and provide critical nursery habitat for many species of fish, crabs, and shrimp. These shallow tidal creeks are also conduits through which many pollutants enter estuaries, with creek sediments serving as a repository for toxic chemicals and other contaminants discharged into estuarine environments.

Resource management and regulatory agencies responsible for protecting estuarine environments do not know if the policies and programs they have implemented are adequately protecting tidal creek habitats. These agencies also lack the knowledge required to restore degraded creek habitats.

In 1994, the South Carolina Marine Resources Research Institute initiated a study, called the Tidal Creek Project (TCP), to develop the information needed to: (1) characterize and define the ecological values of tidal creeks and associated marsh habitats;(2) identify the major pollution threats to tidal creeks associated with watershed development; (3) assess the cumulative impacts of watershed development on tidal creek habitats including the living resources that use them as nurseries; and (4) develop environmental quality criteria for sustaining tidal creek nursery functions. This study was funded jointly by the Charleston Harbor Project (1994-1996) and the Marine Recreational Fisheries Advisory Board (1995-1996).

Approach

The general study approach used was to sample and contrast the physical, chemical, and ecological characteristics of tidal creeks draining relatively pristine, undeveloped watersheds (called reference creeks) and creeks draining highly developed watersheds (called developed creeks). Associations between physical, chemical, and ecological characteristics of creeks and the various types of human development and land cover that occurred were also evaluated. This sampling approach is generally referred to as the comparative watershed assessment approach.

Creeks in the developed watershed class were selected to represent the major types of development that occur in the South Carolina coastal zone including: (1) industrial development, (2) urban development, (3) suburban development, and (4) agriculture. Creeks in the reference class were either predominately forested and/or salt marsh. Watersheds of similar sizes and physical characteristics were evaluated from both the reference and developed classes. The tidal creeks sampled included representatives of the major salinity zones (brackish water to near full strength sea water) and sediment types (sand, mixed, and mud sediments) that occur in South Carolina.

The accuracy, precision, representativeness, completeness, and comparability of the information produced by the TCP were evaluated through a formal Quality Assurance (QA) Program. This QA program was designed to ensure the information produced by the TCP was adequate for addressing study objectives and developing environmental policy. A computerized relational data base system was also established to facilitate efficient storage, retrieval, and analysis of the data produced. This data base provides a means through which the data can be accessed by other researchers or regulatory and resource management agencies. A copy of the TCP data base will be provided to state and federal agencies upon request.

Findings

Salinity was identified as the major factor controlling the distribution and abundance of living resources in shallow tidal creeks. Salinity fluctuated over greater ranges and was generally more variable in creeks with developed watersheds than in reference creeks. The increased variability and extreme fluctuations in the salinity of developed creeks appeared to be related to the "flashier" runoff associated with the increased amount of impervious surface in developed watersheds (e.g., roofs, roads, parking lots). Creeks that were dominated by salt marshes and limited freshwater inputs had relatively stable salinity distributions.

Dissolved oxygen (DO) concentration is a fundamental requirement for maintaining balanced, indigenous populations of fish, shellfish, and other aquatic biota in shallow tidal creeks. Pollution related decreases in DO is generally considered to be the greatest threat to the environmental quality of estuaries. DO in tidal creeks fluctuated with phase of the moon, time of day, and stage of the tide. DO in both reference and developed creeks frequently did not meet state water quality standards (4 mg/l), with the lowest and most stressful DO to living aquatic resources occurring during early morning and night-time low tides. DO in developed creeks was less predictable and had larger amounts of unexplained variance than DO in reference creeks. About 68% of the variance in the DO of reference creeks was associated with natural cycles. Only about 20% of the variance in DO of developed creeks could be attributed to natural cycles. Living resources inhabiting developed creeks were exposed to stressful low DO more frequently than living resources inhabiting reference creeks. Tidal creek ecosystems in both reference and developed watersheds appeared to consume more DO than they produced. Point-in-time measurements of tidal creek DO does not adequately represent the exposure of living resources to stressful low DO events.

Sediment characteristics were also identified as an important environmental factor influencing the distribution of the living resources in shallow tidal creeks. Sediments in developed creeks were generally composed of more sand and had larger site-to-site variation in physical characteristics than reference creeks. The greater sand content and more variable sediment characteristics in creeks located on developed watersheds were probably associated with alterations in erosion and deposition processes associated with watershed development.

Tidal creek sediments are repositories for pollutants. Trace metal and organic contaminant concentrations in sediments of the upper reaches of developed creeks, particularly those with industrialized watersheds or long histories of high density urban and suburban development, were enriched with chemicals to levels known to adversely affect living resources. Enrichment levels ranged from 2-10,000 times the values observed in reference creeks or deeper areas of South Carolina estuaries. Contaminants of particular concern were copper, lead, chromium, polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), and older pesticides, including DDT and chlordane. Low density suburban development did not result in sediment contaminant levels that were of concern. The single agriculture watershed sampled did not provide an adequate representation of sediment contaminants in agricultural watersheds because pollution inputs, mainly pesticides, are episodic and do not persist in sediments.

The distribution of contaminants in tidal creeks varied with the type of development and kind of contaminant. For example, PAHs, which are mainly derived from street runoff and specific point sources, generally had the highest concentrations in sedimentary environments of upper reaches of creeks. Pesticides concentration in at least one suburban watershed was frequently highest in the salt marsh adjacent to houses.

Sediment bioassays indicated that the cumulative amounts of chemicals in sediments of the upper reaches of developed creeks, particularly industrialized creeks, were adversely affecting key biological processes. Sediment bioassays from reference creeks did not suggest exposure to these sediments resulted in acute or chronic impacts on living resources.

The kind of benthic prey available to fish, crabs, and shrimp using tidal creeks as nurseries varied with salinity and sediment characteristics. Human activities associated with watershed development did not adversely affect the biodiversity of benthic organisms in creeks. Long-term salinity distributions and estuary-wide water quality were more important in controlling biodiversity of benthos in tidal creeks than were the local processes occurring within creeks.

The abundance of benthic organisms in tidal creek habitats was mainly controlled by salinity, sediment characteristics, and location within tidal creeks. These three factors accounted for between 7 and 84% of the variance in the abundance of benthic populations. After accounting for the effects of salinity, sediment characteristics, and location within a creek on benthic distributions, both increases and decreases in the abundance of benthic populations were found in developed watersheds. The greatest differences occurred in the upper regions of developed creeks where benthic population abundances were generally reduced, particularly at sites with a long history of industrial or urban development.

Results of a benthic recruitment experiment demonstrated that benthic resources maintained high population levels in creeks by continually recruiting to bottom sediments over the summer. This continual recruitment over the summer provided a renewable source of food for fish, shrimp, and crabs using tidal creeks as nurseries. Salinity, sediment characteristics, location within creeks (upper or lower reaches), and predation by fish and shrimp all had large influences on benthic recruitment success and colonization processes. After accounting for the variation in recruitment due to these natural factors, human alterations of tidal creek watersheds were found to adversely affect the recruitment processes for the numerically dominant benthic organism reproducing during the summer. Recruitment of these organisms was greatly reduced in developed creeks.

Mummichogs and grass shrimp, the preferred prey of many species of recreationally important fish including juvenile red drum, spotted seatrout and flounder, were the dominant fish and crustaceans collected in seine samples from tidal creeks during the summer. Penaeid shrimp and spot were the dominant recreationally important living resources that were found in tidal creeks. Much of the variation in the abundance of fish and crustaceans that occurred from creek-to-creek was associated with variation in sediment characteristics and salinity distributions. After accounting for creek-to-creek variation due to salinity and sediment distributions, no differences in abundance of the numerically dominant species of fish and crustaceans and the kinds/diversity of the fish and crustaceans were found between developed and reference creeks. The abundance of selected key species were, however, reduced in specific creeks with long histories of industrial and urban development.

Although no differences in abundance of numerically abundant fish were observed between creeks located in developed watersheds and reference creeks, the numerically dominant resident fish (i.e., mummichogs) collected from creeks with developed watersheds generally were characterized by poorer physiological condition (i.e., skinnier) and had blood that was not as vigorous as fish from reference creeks. The differences in the blood vigor between developed and reference creeks was most pronounced in male fish and suggests that immune system of resident fish is compromised in developed watersheds.

Fish and crustaceans in size ranges sought by fishermen were rarely collected from tidal creeks. These biota are apparently not be able to tolerate the low DO and other environmental conditions that occur in tidal creeks during summer.

Conclusions and Recommendations

The cumulative impact of development has adversely affected the health of individual resident fish and altered distributions of the type of prey available to fish, shrimp, and crabs that use shallow tidal creeks as nurseries. These alterations, however, do not appear to be substantial enough to adversely affect the populations of recreationally and commercially important living resources that use creeks as nurseries. The number of creeks that are affected in South Carolina is small and the regions of creeks that are the most severely affected is confined to the headwaters which is not the preferred nursery habitat for living resources. Living resources from adjacent habitats continually repopulate impacted regions of creeks.

We believe the alterations to tidal creeks identified above are "early warnings" of more widespread degradation that will occur if the pollution inputs are not reduced. It is interesting that these are the same symptoms that were identified for Chesapeake Bay and other Northeastern estuaries in the early to mid 1970s before it became obvious that the living resource populations of the Bay were declining.

The data base that has been created for primary tidal creeks provide critical baseline information for a broad range of tidal creeks located in developed and undeveloped watersheds. This data base is a research platform for designing and conducting a broad array of future environmental research. Scientists from other institutions and geographical areas are being encouraged to use these data as part of their assessment and research programs.

Because tidal creek ecosystems are consumers of DO, they require adequate amounts of DO to sustain their functions. Water quality management agencies should ensure that DO allocation schemes provide sufficient DO to tidal creeks.

Factors that contribute to low DO in tidal creeks have not been identified or evaluated. Currently, we do not know if the observed alterations to DO dynamics in developed tidal creeks is associated with increased loadings of oxygen consuming pollutants, increased loadings of nutrients (nitrogen and phosphorous) that stimulate excessive growth of primary producers, modifications to the hydrodynamics of tidal creeks from development of the watershed, and/or some other unidentified cause. Until the low DO in tidal creeks can be linked to contributing factors, it is unlikely that policies which prevent DO problems can be developed. A DO budget for tidal creeks and the associated salt marshes to define their relative importance as consumers and identify the major factors controlling low DO conditions needs to be developed. Development of a DO budget is a critical step in the development of DO standards that will ensure that nursery functions provided by tidal creeks are sustained as South Carolina's coastal watersheds are developed.

Additional research on the chronic, sublethal effects of chemical contamination to the health of individual organisms in tidal creeks needs to be conducted. Priority research topics include evaluation of the effects of contamination on immune systems, genetic adaptations of resident living resources to chronic exposure of high levels of chemical contaminants, bioaccumulation/trophic transfer of contaminants as a means of export, and *in situ* effects of contaminant exposure on survivorship, growth, and production of valued living resources (e.g., juvenile red drum).

Based on the data collected to date, status and trends monitoring efforts for tidal creeks should focus on the upper reaches of primary tidal creeks and should include measures of the health of resident organisms, water and sediment quality, and selected population and community parameters of resident living resources. The objective of tidal creek monitoring programs should be to assess the proportion of creeks that have degraded characteristics.

4.4 A PROPOSED SPATIAL FRAMEWORK FOR ESSENTIAL FISH HABITAT DATA COLLECTION AND ANALYSIS

Mark Monaco and Paul Orlando

NOAA, Strategic Environmental Assessment Division, Silver Spring, MD

A spatial framework was proposed for organizing and analyzing data to describe and identify Essential Fish habitat (EFH) for the Nation's riverine, estuarine, and coastal waters. A prerequisite for implementing habitat management approaches is a comprehensive set of spatial units for mapping the areal extent of fish species, habitats, and watershed stressors in rivers, estuaries, and offshore areas.

To enable collection and organization of EFH data within Fishery Management Plans (FMPs) at national, regional, and local spatial scales, the proposed framework must extend from rivers to the continental shelf (Figure 4.4.1). Thus, the proposed spatial framework includes four geographic areas: rivers, estuaries, estuarine watersheds, and offshore zones. For estuaries and watersheds, spatial boundaries have already been defined by the National Ocean Service (NOS) through its *Coastal Assessment and Data Synthesis Framework (CA&DS)*. The CA&DS includes more than 130 estuaries and watersheds along the Atlantic, Gulf of Mexico, and Pacific coasts (Figure 4.4.2). The EPA River Reach file of stream segments as well as offshore segments (e.g., biogeographic zones, NMFS depth strata, 10 minute grids) can be added to the CA&DS to complete the required spatial coverages.

The existing framework reflects the evolution and maturity of a national program for estuaries initiated in 1985. The framework has been widely distributed and is currently used by numerous agencies including EPA, USGS, MMS, and several state agencies. Information available through Federal and state agencies can be readily "tagged" to CA&DS spatial units and integrated with NOAA data sets already in the CA&DS. Numerous data sets describing estuarine resources, habitat, and watershed uses have already been developed using the proposed spatial framework. Nearly all of these are national data sets. Because the thematic data are aggregated by common spatial units, they can be used to make comparisons, rankings, statistical correlations, and other analyses related to resource use and environmental quality.

The CA&DS is available as a digital (ArcInfo E00) product. Since 1985, NOS has used the CA&DS to assemble national data sets on estuarine resources, habitat, water quality, and watershed activities. Data regarding species distributions and their associated habitats could be organized by any of the existing spatial units or any that will be added to the CA&DS. An example product that integrates the CA&DS with biological information is NOS's Estuarine Living Marine Resources program (ELMR) program (e.g., Jury 1994, Monaco and Christensen 1997). More detailed descriptions of these spatial units and available data sets are given below.

Existing Spatial Units for Estuaries and Watersheds

Two fundamental "building blocks" in the CA&DS are estuarine salinity zones and USGS Hydrologic Cataloging Units (HUCs). At present, these are the smallest geographic units in the CA&DS and are readily aggregated into larger units that define estuaries and watersheds, respectively.

Estuarine Salinity Zones

Each estuary is subdivided into three zones between the head of tide and its ocean boundary based on average annual and depth-averaged salinity conditions. These zones correspond to the following salinity regimes: Tidal Fresh (0.0 to 0.5 ppt), Mixing (0.5 to 25 ppt), and Seawater (> 25 ppt). Two major NOAA data sets use the salinity framework to aggregate information. These include the ELMR data set for species distribution and abundance and the National Estuarine Eutrophication Survey of dissolved oxygen, nutrient concentrations, algal blooms, and ecological shifts.

While the existing 3-zone salinity structure provides a consistent and logical approach for synthesizing biotic information in estuaries, more refined spatial and temporal salinity units may be useful for some EFH applications. To that end, refined salinity distributions have already been completed for approximately 50 estuaries in the South Atlantic and Gulf of Mexico regions. For these systems, seasonal salinity contours have been constructed at 5 ppt increments from the head of tide to the ocean boundary for both surface and bottom layers of the water column. Seasonality was defined by the 3-month high salinity period, the 3-month low salinity period, and the two transitional periods. These refinements are required for approximately 80 North Atlantic, Mid-Atlantic, and West Coast estuaries.

Watersheds

Physical boundaries for estuarine watersheds were based on the USGS HUC system. Typically, catalog units occupy about 700 square miles and represent all or part of a surface drainage or a distinct hydrologic feature. For each estuary, the watershed includes all catalog units that drain to the estuary. In large watersheds, a distinction is made between the portion of the drainage area that is immediately adjacent to tidally-influenced waters (termed the "Estuarine Drainage Area", EDA) and the more distant regions adjacent to tidal-fresh streams (termed the "Fluvial Drainage Area", FDA). All 130+ estuaries in NOAA's National Estuarine Inventory have EDAs. Nearly one-half of the 130+ estuaries have FDAs. Several major NOAA data sets use the EDA/FDA framework to organize information. Among these are pollutant sources and loadings, population trends, land use, wetland distributions, and physical/hydrologic data.

Spatial Units Now Being Added to the CA&DS

Coastal and Offshore Spatial Units

This component of the proposed EFH spatial framework is relatively undeveloped, but can readily accommodate any proposed organizational units. We propose that the coastal and offshore EFH data be organized by depth strata and or grid cells as most of state and Federal monitoring programs organize or sample fisheries data by depth zones or grids. For example, the NMFS northeast coast bottom trawl surveys use approximately 57 depth strata for fishery independent monitoring, while the joint NOAA and Canadian Dept. of Fisheries and Ocean, East Coast of North America Strategic Assessment Project organizes environmental data by 10 X 10 minute grid cells. It is likely that a combination of approaches will be required to define spatial structures in coastal and marine areas due to the diversity of habitats, oceanographic currents, sampling programs, and data availability across the Nation. In addition, EFH in marine areas could be aggregated/organized by oceanographic features, such as large marine ecosystems (LMEs) (Sherman et al. 1990).

EPA River Reach File

To help accommodate riparian issues, anadromous fish habitats, and other freshwater-related concerns, EPAs River Reach file is being added to the CA&DS. This system, which divides rivers into reach segments, includes nearly all but the smallest streams within a watershed. The system is hierarchical and encodes river reaches as primary, secondary, tertiary, or quaternary depending on how far removed the stream is from the major tributary.

Comments and suggestions made at the IBI workshop will be incorporated into a complete draft to be reviewed by agencies and institutions involved in the EFH initiative (e.g., NMFS, American Fisheries Society, Atlantic States Marine Fisheries Commission). We suggest the way forward on this work is to incorporate comments from the community on the feasibility and usefulness of developing a consistent spatial framework to collect and organize data and information to support the EFH initiative. Ultimately, consensus should be obtained on the spatial structures necessary to meet the EFH objectives of: 1) describing, identifying, and mapping EFH; 2) inventorying habitat impacts; and 3) developing corrective actions to conserve and enhance habitats

---

References

Jury, S.H., J.D. Field, D.M. Nelson, and M.E. Monaco. 1994. Distribution and abundance of fishes and invertebrates in North Atlantic estuaries. ELMR report no. 13. Silver Spring, MD: National Oceanic and Atmospheric Administration, Strategic Environmental Assessments Division. 221 pp.

Monaco, M.E. and J.D. Christensen. 1997. The NOAA/NOS Biogeography Program: Coupling species distributions and habitat. In: Boelert, G.W. and J.D. Schumacher (eds.), Changing oceans and changing fisheries: Environmental data for fisheries research and management. NOAA Technical Memorandum, NOAA-TM-NMFS-SWFSC. pp. 133-138.

Sherman, K., Alexander, and B.D. Gold. 1990 Large marine ecosystems. AAAS Press. Pub. No. 90-30s. Washington, DC. 242 p.

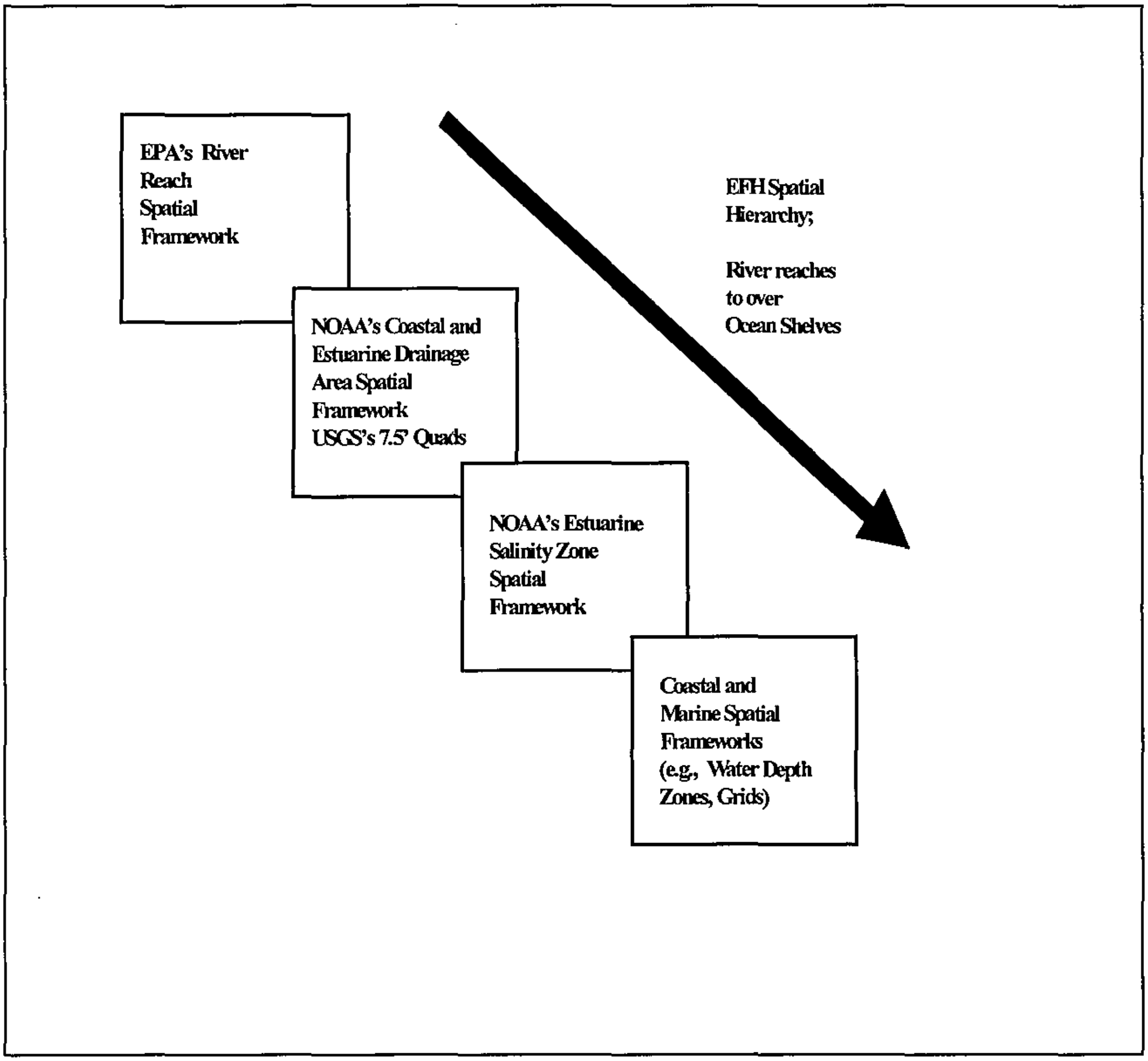


Figure 4.4.1 Proposed spatial framework to support EFH data and analysis.

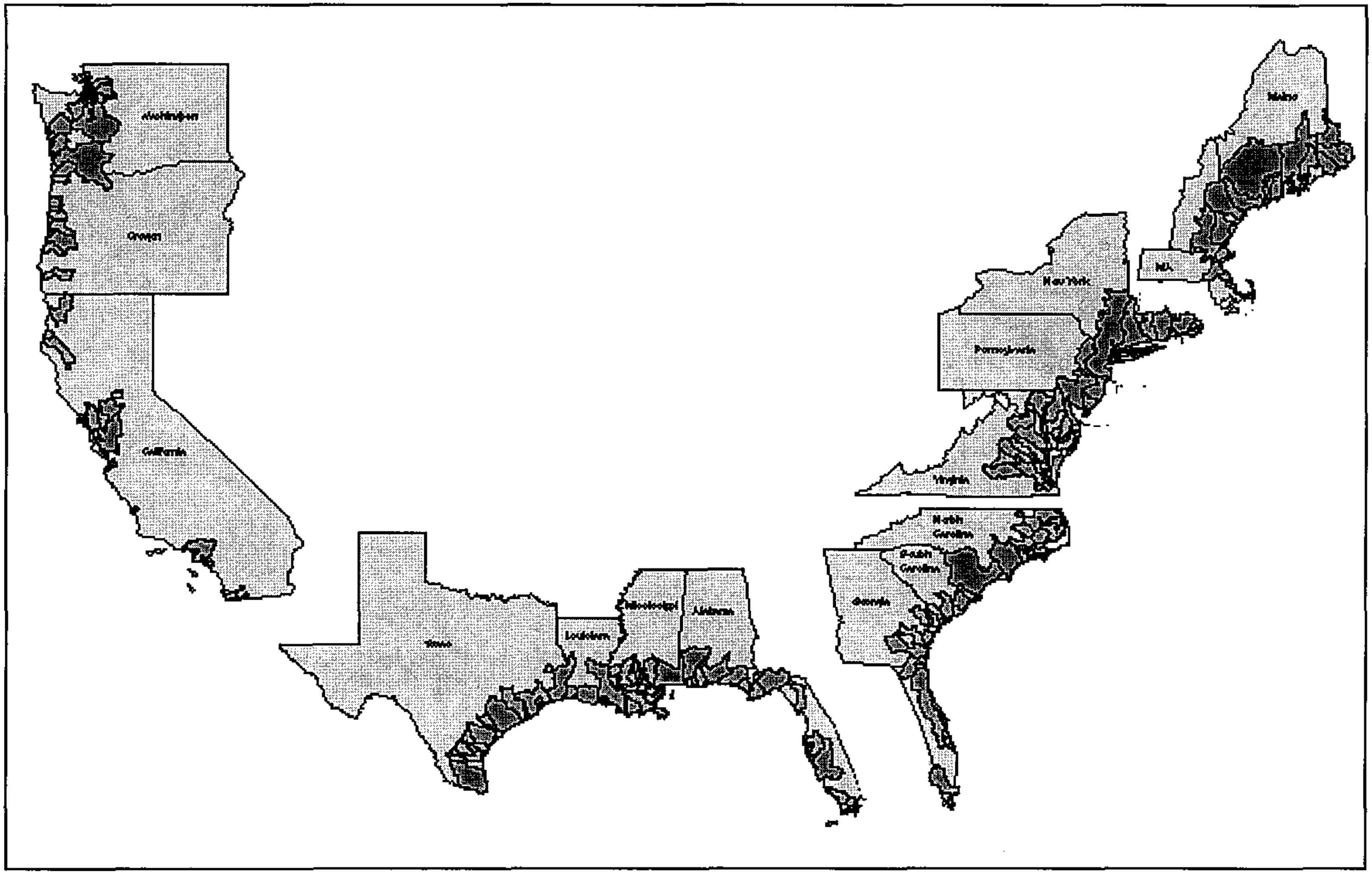


Figure 4.4.2 Estuarine drainage watersheds included in the Coastal Assessment and Data Synthesis Framework.

4.5 AN ESTUARINE INDEX OF BIOTIC INTEGRITY FOR CHESAPEAKE BAY TIDAL FISH COMMUNITIES

Margaret McGinty[1] and Cecelia Linder[2]

[1]Maryland Department of Natural Resources, 580 Taylor Avenue, Annapolis, MD
[2]University of Delaware, Lewes, DE

A fish Index of Biotic Integrity (IBI) was developed for tidal fish communities of several small tributaries to the Chesapeake Bay (Jordan et al., 1990, Vaas and Jordan, 1991, Carmichael et al., 1992a, b). It is based on the original IBI (Karr, 1981) and has a nine metric index that contains measures of *species richness* (number of species, number of species comprising 90% of the catch and the number of species in the bottom trawl), *trophic structure* (proportion of carnivores, planktivores, and benthivores) and *abundance* (number of estuarine fish, number of anadromous fish, and total fish with Atlantic menhaden removed). The IBI was tested using stepwise discriminant analysis to determine the weight of each metric relevant to the IBI score. This exercise showed that six of the nine metrics accounted for ~95% of variability. Of these six, the anadromous fish metric showed to be the most influential metric on the IBI score (partial $r^2=.59$). However, the anadromous fish metric also strongly correlated with spring flow ($r^2=.95$, $p=.0001$). Because IBI's are intended to identify biological impairment due to anthropogenic influences, it was undesirable that the most influential metric on the IBI was so strongly influenced by natural variation. With this realization, a reevaluation of the IBI was initiated to attempt to define metrics that were minimally influenced by natural variation. Following is a description of the procedures applied to redefining the estuarine fish IBI.

Data used for the reevaluation were from 12 tributaries sampled between 1989 and 1995. The data were divided so that two data sets were available for the effort, a development set, and a test set. Stations included in the IBI development set were those for which consistent monitoring has been done. Reference sites were established *a priori* based on reference criteria. We attempted to model the criteria established in the Virginian Province EMAP effort (Weisberg et al. 1992). Criteria for bottom dissolved oxygen concentrations, sediment toxicity, algal blooms, and land use features were proposed to identify reference and degraded sites within the test data sets. The criteria selected did not clearly discriminate between reference and degraded conditions. Cluster analysis was applied to the data to group data into two clusters, reference sites and degraded sites. The results of the clustering grouped the small, predominately urban tributaries, including a site near Sparrows Point into one group, and the other sites which included larger scale agriculture dominated into a second group.

The data from these sites were used to calculate and test approximately seventy possible metrics. Approximately twenty possible metrics were selected from box and whisker plots. A metric was selected if the mean of the reference group was different from the means of the degraded, and if the upper quartile of the degraded group did not extend past the mean of the reference group. We had limited success in designating meaningful metrics from this procedure. Possible reasons for this are that watershed scale is influencing the results, and that the reference criteria used to establish the reference and degraded sites do not significantly influence mobile fish communities.

Presently, data are being evaluated as originally done, where reference and degraded conditions are assigned based on dominant land use within the watershed. We are accounting for the influences of seasonal flow patterns and watershed scale. Thus far, eight metrics have been shown to be statistically meaningful in discerning the differences in land use. They are presently being examined for ecological significance. We are also exploring methods to account for flow influence to retain some or all of the original metrics.

References

Carmichael, J., B. Richardson, M. Roberts, and S. Jordan. 1992a. Fish Sampling in Eight Chesapeake Bay Tributaries. Maryland Department of Natural Resources, Tidewater Administration, Chesapeake Bay Research and Monitoring Division CBRM-HI-92-1. Annapolis, MD.

Carmichael, J., B. Richardson, S. Jordan. 1992b. Development and Testing of Measures of Ecological Integrity and Habitat Quality for Chesapeake Bay Tidal Tributaries. Final Report to Maryland Coastal Zone Management. Maryland Department of Natural Resources, Tidewater Administration, Chesapeake Bay Research and Monitoring Division. Annapolis, MD.

Jordan, S.J., P.A. Vaas, and J. Uphoff. 1990. Fish Assemblages as Indicators of Environmental Quality in Chesapeake Bay. IN Biological Criteria: Research and Regulation, 1990.

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

Vaas, P.A. and S.J. Jordan. 1991. Long Term Trends in Abundance Indices for 19 Species of Chesapeake Bay Fishes: Reflections in Trends in the Bay Ecosystem. In: J.A. Mihursky and A. Chaney (eds.). New Perspectives in the Chesapeake System: A Research and Management Partnership. Proceedings of a Conference. Chesapeake Research Consortium Publication No. 137. Solomons, Maryland, p. 539-546.

Weisberg, S. B., A. F. Holland, K. J. Scott, H. T. Wilson, D. G. Heimbuch, S. C. Schimmel, J. B. Frithsen, J. F. Paul, J. K. Summers, R. M. Valente, J. Gerritsen, R. W. Latimer. 1992. Virginian Province Demonstration Report, EMAP - Estuaries - 1990. EPA/620/r-93/006. U.S. Environmental Protection Agency, Washington, D.C. 20460.

4.6 ESTUARINE BIOTIC INTEGRITY INDEX

Melissa J. Weaver[1] and Linda A. Deegan[2]

[1]Ecology & Evolutionary Biology Dept., Univ. of Tennessee, Knoxville, TN
[2]The Ecosystems Center, Marine Biological Laboratory, Woods Hole, MA

An Estuarine Biotic Integrity Index (EBI) has been developed and validated for Southern New England (Deegan et al. 1993, Deegan et al. in press). The EBI is an assessment of the condition of estuarine ecosystems based on the abundance, diversity, and composition of the fish community. Fish integrate and reflect the condition of and linkages between ecosystems and serve as indicators of the biotic integrity of a whole region. The EBI focuses on submerged rooted aquatic vascular plant (SRV) habitats because they are critical habitats for fish and are sensitive indicators of anthropogenic stress. Both the EBI and its' metrics are well-correlated with habitat and water quality, but the EBI does better than its' individual metrics in predicting ecosystem health, as indicated by fish production (Deegan et al. in press). The EBI as developed in Southern New England consists of eight metrics which include both functional grouping and specific species as indicators of estuarine health: total number of species; number of estuarine spawners, estuarine residents, and nursery species; number of species which make up 90% of individuals; % benthic dependent (feeding, spawning, dwelling, etc.) based on the number or biomass of individuals; and, % with disease. Individual metrics and the overall index show a strong correlation with habitat degradation. Habitats that were classified as impacted on the basis of year-round measurements of chemical and physical characteristics (algal blooms, macroalgae, low dissolved oxygen, high nutrients, dredged channels) had highly modified fish communities. These changes in the biotic community were reflected in low EBI values. Differences between moderate and low quality habitats are most pronounced near the end of the summer reflecting the cumulative effects of habitat degradation. The EBI and its metrics are well-correlated with habitat quality in moderate quality embayments such as Waquoit Bay. Thus, the EBI can be used to evaluate the current status of Southern New England estuaries.

For the EBI to be useful, it must not only reflect the current status of fish communities, but it must track changes in habitat quality over time, be applicable over a wide range of estuaries and habitat quality within the same geographical region, and be transferable to other regions. To test whether the EBI reflects long-term changes in habitat we compared habitat quality and fish communities at sites in Waquoit and Buttermilk Bays from the late 1980's to the mid-1990's. The EBI provides corroborative evidence that habitat quality within Waquoit Bay has continued to degrade and that efforts to control nitrogen inputs into Buttermilk Bay have prevented further degradation of habitats and maintained stable fish populations (Chun et al. 1996). Evaluation of the applicability of the EBI across a wide range of habitat quality within a geographic range requires that the mechanisms that relate the fish community structure with stress, degradation, and loss of functions are similar throughout the quality range. In 1996 we sampled estuaries (23 sites from 13 embayments) in Buzzards Bay, Southern New

England for which there is extensive background information about nitrogen loading and other stressors. These sites include nearly pristine sites and severely degraded sites and extend the range of habitat degradation compared to the original study. These data will allow us to test the response of the fish community to more extreme conditions and in estuaries that differ not only in levels of anthropogenic stress but also in flushing rate, exposure to wave action, morphology, sediment, macroalgae and eelgrass abundance, amount of marsh edge, and fish species. Moreover, we will be able to test the general applicability of the EBI throughout the Southern New England region. To test the transferability of the EBI to the Mid-Atlantic Region, we sampled habitat quality and fish communities in the lower Western Shore of Chesapeake Bay (26 sites within 5 subestuaries) in summer 1995. Several of the metrics and the EBI itself are correlated with habitat quality and are successful at separating low quality sites from all other sites, but they were unable to discern differences between the fish communities in medium and high quality (nearly pristine) sites in the Chesapeake Bay. The EBI as originally developed was not directly transferable to the Mid-Atlantic Region but required modification in the selection of the EBI metrics and their classification levels. Further development with regards to the scale of sampling, aggregation of data, and analyses are required to standardize the EBI for use among regions.

Several of the metrics that comprise the EBI varied with anthropogenic stress in Chesapeake Bay in the same manner as for Southern New England estuaries. Fish abundance, biomass, and number of species declined with increased stress in both regions. Of the original set of eight metrics, the number of species, nursery species, resident species, and spawners, and the proportion benthic by number of individuals were correlated with habitat quality and were higher for medium than for poor quality sites in Chesapeake Bay. However, the total number of species and the number of species for these life history strategies were the same in moderately degraded sub-estuaries as in pristine sub-estuaries.

Because the Mid-Atlantic Region has a wider range of anthropogenic stress and an intrinsically more diverse and abundant fish community compared to Southern New England, we anticipated that other metrics may be more useful than the original metrics in discriminating between sites of differing habitat quality in the Mid-Atlantic. For example, we found very few specialized feeders in the Southern New England region so trophic metrics did not differ with habitat quality (most species were benthic invertivores). Chesapeake Bay fishes exhibit a broader array of food web position and feeding strategies and we would expect specialized feeders to decline with increased stress. In fact, the number of invertivores did better than most other potential metrics in discerning habitat quality and was included in the Chesapeake Bay index.

Further modifications were made to the EBI. In Chesapeake Bay as well as in New England, the count of individuals and total catch biomass for low quality sites were less than for medium quality sites. However, in both regions, the species dominance (number of species that comprise 90% of the catch) did not distinguish sites by quality, and there were few abnormalities among the individual fishes (less than 0.1%), and so these two metrics were not included in the calculation of the modified EBI for the Chesapeake Bay. In general, metrics based on the number of species in functional groups did better than those based on counts of individuals or biomass. Because the number of benthic species differed between medium- and high-quality sites, as well as between low- and medium-quality sites, it was added to the EBI. We elected to retain the proportion of benthic (count or number) because it was a good separator of low- and medium-quality sites.

Surprisingly, the absolute number of species in any habitat quality was not higher in the Mid-Atlantic Region compared to Southern New England, rather it was lower for several of the metrics, including the number of resident species. Although there were many more species in the total sampling catch in the Chesapeake Bay than in the Southern New England catch (about 59 versus about 35), the number of species that nurse, spawn, or reside permanently within the estuary was lower in trawl catches in Chesapeake than in New England. However, the number of species within a sampling site within the Chesapeake was sometimes higher than that in Southern New England. There was a much higher diversity among trawls within a site and within an embayment in Chesapeake Bay than in the Southern New England bays, but the EBI did not reflect this diversity. The EBI as originally developed apparently reflects the fish community diversity at a very local level perhaps due to the scale of sampling, and the manner in which the data were aggregated and analyzed. (The metrics were scored and the EBI calculated for each trawl and then averaged across trawls at each site.) Furthermore, the overall patterns in the metrics in relation to quality differences did not differ between riverine estuaries (York, James, and Rappahannock Rivers) and subembayments (Lynnhaven and Mobjack) within Chesapeake Bay. By calculating an index that integrates the habitat quality and quantity over each embayment and then throughout the subregion (lower Western Shore of Chesapeake Bay), we may have a more effective measure of habitat quality within Chesapeake Bay.

References

Chun, N. K., M. J. Weaver, and L. A. Deegan 1996. Assessment of fish communities in New England embayments: application of the Estuarine Biotic Integrity Index. Biological Bulletin 191 (October).

Deegan, L. A., J. T. Finn, S. G. Ayvazian,, C. A. Ryder 1993. Feasibility and application of the index of biotic integrity to Massachusetts estuaries (EBI). Final Project Report. Massachusetts Executive Office of Environmental Affairs, Department of Environmental Protection, North Grafton, Massachusetts.

Deegan, L. A., J. T. Finn, S. G. Ayvazian,, C. A. Ryder-Kieffer, and J. Buonaccorsi. Development and validation of an estuarine biotic integrity index. Estuaries: in press.

4.7 OHIO'S LAKE ERIE AND LACUSTUARY MONITORING PROGRAM

Roger F. Thoma and Chris O. Yoder

Ohio EPA, Columbus, OH

In 1993 the Ohio EPA began a project designed to develop numerical biological criteria for shoreline waters of Lake Erie and areas of tributary streams affected by lake levels, referred to as lacustuaries. The term lacustuary is a combination of lacustrine and estuary. Lacustuary is defined as a transition zone in a river that flows into a freshwater lake and is the portion of river affected by the water level of the lake. Lacustuaries begin where lotic conditions end in the river and end where the lake proper begins. They have hydrologic conditions similar to estuaries in that they are affected by tides (primarily wind driven, occasionally barometric) and are lentic habitats. Lacustuaries differ from estuaries since their chemical properties are less saline with salinity gradients going from higher upstream to lower at the lake interface (Brant and Herdendorf 1972). It is felt that the term lacustuary is needed to avoid confusion of terms and concepts that ensue when estuary is used for freshwater systems. Though there are some similarities, estuaries and lacustuaries differ in numerous important functions and should not be confused with each other.

This IBI project was conducted in the following steps: 1) sampling of the general habitat types found in the Lake Erie ecosystem using various sampling methodologies; 2) evaluation of sampler type efficiency and selection of the method to be used in each habitat type; 3) continued sampling using the selected methodology; 4) evaluation of potential metrics; 5) selection and calibration of IBI metrics; 6) continued sampling; 7) calculation of Lake Erie shoreline and lacustuary IBI scores; 8) evaluation of environmental conditions in Lake Erie and associated lacustuary areas. This study was built on data collected since 1982.

Ninety sites (324 individual collections) were sampled in Lake Erie from 1993 through 1996. Site selection reflected the habitat types found in the lake's nearshore areas and provided a thorough coverage (approximately one site for every 5 miles or 8 kilometers) of the area investigated. Sites were located along harbor breakwaters, sand/gravel beaches, the shores of the Lake Erie Islands, bedrock cliffs and modified shore lines with numerous types of structures designed primarily to prevent shoreline erosion. Wetland/bay-like habitats were sampled in Sandusky Bay, East Harbor State Park, and Presque Isle PA (11 sites). Lacustuaries were sampled at 125 sites (593 individual collections) from 1982 through 1996. Sites were located at the mouth, head, and midsection of each lacustuary.

All fish were collected using a 5.8-meter modified V-hull john boat. Lacustuary habitats were sampled during daylight and lake sites were sampled at night. Fish were identified to species, enumerated, examined for external anomalies, and either returned to the lake or preserved as voucher specimens. Weights were taken on a representative sub-sample if more than 15 individuals of a species were captured. All fish were weighed if 15 or fewer individuals of a species were captured. Each sampling site was 500 meters long and within 1 meter of the shore.

Metrics

A large number of metrics were examined to determine the metrics best suited for use in a Lake Erie IBI and lacustuary IBI. Examination of metrics for lake and lacustuary sites indicated the relative abundances and percent composition of fish in the two types of habitat should be evaluated separately. When metrics were selected, an effort was made to use groupings that maximized the range of values possible. Metrics with low breadth can result in a yes-no, present-absent evaluation instead of the intended strongly - moderately - little deviation assessment. Comments on selected metrics are listed below (LE- Lake Erie, L-lacustuary). A complete list is presented in Table 4.7.1.

- *Number of benthic species* (LE, L)

This metric is thought to primarily respond to environmental disturbance from excess sedimentation and secondarily to toxicity and low oxygen levels. It comprises darters, sculpins, and madtoms. Other benthic species of generally greater environmental tolerance, such as bullheads and suckers, were excluded in order to maintain sensitivity.

- *Number of sunfish species* (LE, L)

This metric includes sunfish and species of the genera *Pomoxis* and *Micropterus*.

- *Number of cyprinid species* (L)

Cyprinid species were historically a prominent community component that could be found in all lacustuary habitats, and several highly sensitive species (now apparently extirpated in Ohio) were primarily associated with Lake Erie near shore areas. This metric can accommodate future changes in the ecosystem if environmental conditions improve to the point that locally extirpated species become reestablished.

- *Number of phytophilic species* (LE)

Variations in this metric are associated with increases in submerged aquatic vascular plants (especially *Potamogeton* and *Vallisneria*) which are found in high quality, clear, low-polluted waters which is an ecological parameter of substantial historical prominence.

- *Percent lake individuals* (LE)

This metric reflects a species guild that has proven to be sensitive to environmental disturbances in Lake Erie. Because sufficient numbers of lake-associated species still exist and much room for improvement is possible, this metric is ideal for measuring the long term trends of Lake Erie fish communities.

- *Percent phytophilic individuals* (L)

As with percent lake species, this metric is highly sensitive to slight environmental change. Historically, lacustuaries exhibited high numbers of phytophilic species and very high numbers of individuals. Though numerous phytophilic species have disappeared from Lake Erie's lacustuaries, many species still subsist at very low numbers in almost all areas. As even the most polluted sites generally have the same phytophilic species, we decided to use the number of individuals as sites of higher environmental quality exhibited much higher abundances than degraded sites. This allows discrimination between the very bad sites and fair sites. If lacustuary habitats should improve in the future, this metric may to be converted to a number of species metric.

- *Percent top carnivores* (LE, L)

Only those species that at an adult size feed on fish or crayfish more than 80% of the time are considered top carnivores. Species such as crappie and channel catfish that have a more plastic feeding behavior and can convert to other forms of food resources under sub-optimal conditions are excluded.

- *Percent non-indigenous species* (LE, L)

Non-indigenous species have been found, in this study, to increase in areas of higher disturbance, especially that associated with extensive urban development. Only species which were present in the system originally (pre 1700s) are considered indigenous.

- *Percent diseased individuals* (LE, L)

This metric is a measure of the percent of individuals that have externally observable deformities, eroded fins, lesions or tumors.

Scoring Considerations and Attainment Criteria

Setting the Ninety-fifth Percentile Line

Because the fish community of Lake Erie has experienced pervasive negative impacts (Hartman 1972, Regier and Hartman 1973, Trautman 1981, Van Meter and Trautman 1970, and White et al.1975), the selection of reference sites and 95 percent lines is problematic. If one sets expectations at levels thought to be equivalent to the historic potential of Lake Erie, all sites would score so low that it would not be possible to distinguish highly, moderately, and slightly polluted areas. Alternatively, if a straightforward ninety-fifth percentile line is employed it becomes possible that sites will score in the exceptional range. This prospect is unacceptable in light of the present condition of Lake Erie. The intent of the IBI is to measure integrity and Lake Erie presently exhibits very little integrity. A score of exceptional would be construed as an indication that Lake Erie is approaching full recovery, which it is not.

The approach employed in this IBI effort has been to use a modification of the ninety-fifth percentile methodology. When drawing the ninety-fifth percentile line, the line was always drawn between the ninety-fifth percent value and the next value point. This acknowledges the fact that if greater integrity existed the ninety-fifth percent value would be more stringent while keeping scoring criteria at a level that allows discrimination of the present conditions. Using this methodology, none of the sites sampled in this study have scored fifty or higher.

Karr (in press) proposes the use of ecological dose-response curves to devise scoring criteria for IBI metrics. Such an approach may prove to be the best methodology to score Lake Erie's fish communities because of the extensive disturbances experienced and the lack of reference conditions. Future work on the Lake Erie and lacustuary IBI will examine ecological dose-response curves.

<u>Integrity classifications</u>

Integrity ranges of exceptional (>50), good (>42), fair (>31), poor (>17), and very poor (<=17) have been set for Lake Erie and its lacustuaries. The predicament of setting specific integrity ranges for Lake Erie and its lacustuaries is difficult because all sites sampled have been affected to some degree by dramatic ecological changes. One approach has been to use the IBI value that occurs at the 25 percentile of the reference sites selected as representative of a habitat type as the level at which the "good" classification begins. It is incumbent in the 25 percentile approach that the reference site data base is composed of sites that very nearly approach biological integrity. In the Huron-Erie Lake Plan (HELP) ecoregion, where most sites have been impacted and do not display ecological integrity, the Ohio EPA elected to use the 90th percentile of all sites sampled to derive attainment criteria. Because the Lake Erie system displays pervasive negative environmental effects, an approach like the HELP ecoregion strategy is desirable. This work differs from the previous HELP effort by using only the least impacted sites to set the 90 percentile instead of all sites. Use of a 25 percentile in Lake Erie waters would result in a criteria that accepts environmental degradation while the 90 percentile of least impacted sets a goal that the data have demonstrated can be attained in a reasonable time frame with some environmental amelioration (even in light of pervasive impacts). Once the good attainment point was set, exceptional, fair, poor and very poor integrity ranges were set based on an understanding of species composition at differing IBI levels.

The potential for this scoring system to change is great, as Lake Erie is currently in a state of dynamic flux. New non-indigenous species are invading at increasing rates (Mills et al. 1993) and phosphorus levels are decreasing (Bertram 1993, Makarewicz and Bertram 1991) and the two are interacting in unpredictable ways to create considerable uncertainty. Continued monitoring will be required to track changing community conditions, and attainment criteria will need to be reviewed in light of future changes.

Application examples:

<u>Lacustuary assessments</u>

Four examples are provided that demonstrate the effectiveness of the IBI to identify areas with no improvement, improving conditions and gradients of impact, which can be related to site-specific anthropogenic activities. Multiple examples exist in each type of situation.

Index of biotic integrity scores from 4 years of biological monitoring in the Ottawa River have consistently remained in the poor to very poor range (Figure 4.7.1). Numerous combined sewer overflows, urban runoff, leaking landfills and contaminated sediments combine to suppress communities to extreme low levels. Over the 10 year period of monitoring, none of the impact sources have been addressed and consequently no changes are detectable in fish communities. Restoration potential for this lacustuary is good because depths are still shallow enough to allow reestablishment of aquatic macrophyte communities, a factor critical to fish community integrity.

The Black River lacustuary was sampled in 1992 and 1994. Scores for the IBI were consistently poor in 1982 and poor to mostly fair in 1992 (Figure 4.7.2). Community improvements over the 10 year period were due to upgrades at the upstream Elyria waste water treatment plant that reduced loading to the Black River and its lacustuary. Removal of contaminated sediments after 1992 probably will lead to further fish community improvements. Presently the lacustuary is limited by nutrient enrichment primarily from upstream nonpoint pollution, both urban and rural. Very little submerged aquatic vegetation exists in the lacustuary although habitat structure is suitable. With further reductions of pollutant loads and a resurgence of plant life, fish communities in the Black River lacustuary should recover and attain exceptional conditions.

Seven sites have been sampled since 1989 in the 2.5 mile (4 km) length of the Ashtabula River lacustuary (Figure 4.7.3). Downstream of river mile (RM) 2.3, much of the waterway was lined with sheet piling and boat docks. A ship channel extended from the river mouth to RM 0.7. Fields Brook joins the Ashtabula River at RM 1.6. Sediment contamination has been documented downstream of Fields Brook. In 1989, fish community sampling was conducted to evaluate the degree of impact associated with chemical degradation originating from Fields Brook and habitat alteration of the lacustuary. It was concluded that shoreline development was the principal factor impacting fish communities in the lower Ashtabula River with a lesser effect from chemical pollutants. In general, IBIs were good to fair in upper reaches, fair to poor near Fields Brook, and fair to very poor in the ship channel area.

The Conneaut Creek lacustuary extends for 2.2 miles (3.5 km) upstream from the mouth. A total of 6 sites have been sampled since 1989 (Figure 4.7.4). Very little environmental deterioration was seen in the lotic portions of the system and extensive areas of the basin are wooded. The lower 0.5 mile of the stream was a ship channel with deep sheet piling lined banks while upstream from RM 0.5, the channel was shallower and at least partially vegetated along the banks; most of this reach was relatively narrow with moderate accumulations of silt and sediment. An area of thick silt and sediment with a large expanse of emergent and submergent vegetation was present at RM 1.0. Upstream from the ship channel, in the area of vegetation, IBI scores were in the good range while ship channel sites (RMs 1.3 and 0.6) had IBI scores in the poor to fair range. The dichotomy of good and poor community conditions found in this lacustuary illustrate the strong effect that habitat alterations can have on biological conditions even in areas where no impacts from water column chemistry exist.

Lake Erie

In Lake Erie, three factors affect fish community structure; lake-wide trophic changes as a result of nutrient enrichment, habitat loss primarily in the form of wetland destruction or diking and shoreline modifications, and localized environmental impacts from industrial and municipal discharges. Of principal significance is the predominant effect of lake-wide trophic changes and associated species losses. These changes have resulted in most sites scoring as fair with few good and no exceptional values attained (Figure 4.7.5). Four of the nine sites that clearly fall into the good range are from the shorelines of the Lake Erie Islands. Island sites score better, in part, due to their distance from lacustuaries and associated impacts. Habitat was also an important factor for island sites. The principal habitat type encountered around the islands was boulder - rubble strewn shorelines with high levels of substrate texture. It was observed in this study that the greater the habitat texture the greater the relative abundance and number of species. Breakwater sites, at the mouths of lacustuaries, had habitat textures similar to island sites, but failed to reach the levels attained at island sites. This was due to lacustuaries experiencing environmental stress from higher loads of pollutants. Beaches were the area of lowest substrate texture and tended to score lower than other habitat types (in the absence of other environmental stresses). Examples of localized pollution impacts were found in the Maumee Bay and Cuyahoga River at Cleveland, areas where in spite of the fact that habitats were highly textured breakwaters, IBI values remained in the poor range. The only site in this study that fell in the very poor classification was just east of the Maumee Bay area. This site was a riprapped beach in an area where extensive settling of organic debris and urban waste was occurring. The dominant species at this site was goldfish, a highly tolerant fish.

General

None of the lake or lacustuary sites attained an integrity level of exceptional and only a few attained the good level. This was reflective of the widespread and pervasive nature of environmental impacts in the region. Many species were missing (Trautman 1981, Hartman 1972) and trophic dynamics were radically changed (Regier and Hartman 1973, Stoermer et al. 1987). Five of the 20 most abundant species were non-indigenous species. Ninety three species were recorded and the average relative abundance of individuals (number per kilometer) was 687.

At the good-fair integrity interface, similarities between Lake Erie and its lacustuaries begin to diverge. In the lake proper, environmental impacts are more widely dispersed and less intense, whereas in lacustuaries they can be very intense and are always more concentrated. In the lake, only 73 species were recorded and the average relative number of individuals (number per kilometer) was 934. Integrity levels of fair dominated the lake results (59%), poor to very poor (24%) comprised the next largest classification, and good (17%) the least. In the lacustuaries 87 species were recorded and the average relative number of individuals (number per kilometer) was 552. Poor to very poor IBI scores dominated the results (71%) while fair comprised 23% and good equaled only 6%.

References

Bertram, P.E. 1993. Total Phosphorous and Dissolved Oxygen Trends in the Central Basin of Lake Erie, 1970-1991. J. Great Lakes Res. 19:224-236.

Brant, R.A. and C.E. Herdendorf. 1972. Delineation of Great Lakes Estuaries. Proceedings, 15th Conference, International Association for Great Lakes Research 15:710-718.

Hartman, W.L. 1972. Lake Erie: Effects of Exploitation, Environmental Changes and New Species on the Fishery Resource. J. Fish. Res. Board Can. 29:899-912.

Karr, J.R., (in press). Rivers as Sentinels: Using the Biology of Rivers to Guide Landscape Management.

Makarewicz, J.C. and P. Bertram. 1991. Evidence for the Restoration of the Lake Erie Ecosystem. BioScience vol. 41, no. 4.

Mills, E.L., J.H. Leach, J.T. Carlton, and C.L. Secor. 1993. Exotic species in the Great Lakes: a history of biotic crises and anthropogenic introductions. J. Great Lakes Res. 19:1-54.

Regier, H. A. and W. L. Hartman. 1973. Lake Erie's Fish Community: 150 Years of Cultural Stress. Science 180: 1248-1255.

Stoermer, E. F., J. P. Kociolek, C. L. Schelske, and D. J. Conley. 1987. Qualitative Analysis of Siliceous Microfossils in the Sediments of Lake Erie's Central Basin. Diatom Research 2:113-134.

Trautman, M.B. 1981. The Fishes of Ohio. Ohio State Univ. Press. Columbus, OH, 782 pp.

Van Meter, H.D. and M.B. Trautman. 1970. An annotated list of the fishes of Lake Erie and its tributary waters exclusive of the Detroit River. Ohio J. Sci. 70(2):65-78.

White, A.M., M.B. Trautman, E.J. Foell, M.P. Kelty, and R. Gaby. 1975. Water quality baseline assessment for the Cleveland area-Lake Erie. V 2. The fishes of the Cleveland metropolitan area including the Lake Erie shoreline. U.S. E PA, Chicago, Ill. 181pp.

Table 4.7.1 Metrics used in Ohio EPA's two IBIs developed to evaluate Lake Erie nearshore ecosystems and lacustuaries.

Lake Erie Metrics	Lacustuary Metrics
Species number metrics	
# Species	# Species
# Sunfish species	# Sunfish species
# Phytophilic species	#Cyprinid species
# Benthic species	# Benthic species
Behavior/trophic guild metrics	
% Lake assoc. individuals	% Phytophilic individuals
% Top carnivores	% Top carnivores
# Intolerant species	# Intolerant species
% Omnivore individuals	% Omnivore individuals
% Non-indigenous ind.	% Non-indigenous ind.
% Tolerant individuals	% Tolerant individuals
Community health metrics	
% DELT*	% DELT*
Relative numbers**	Relative numbers**

* Externally observable deformities, eroded fins, lesions, and tumors.

** Includes non-indigenous species and excludes gizzard shad.

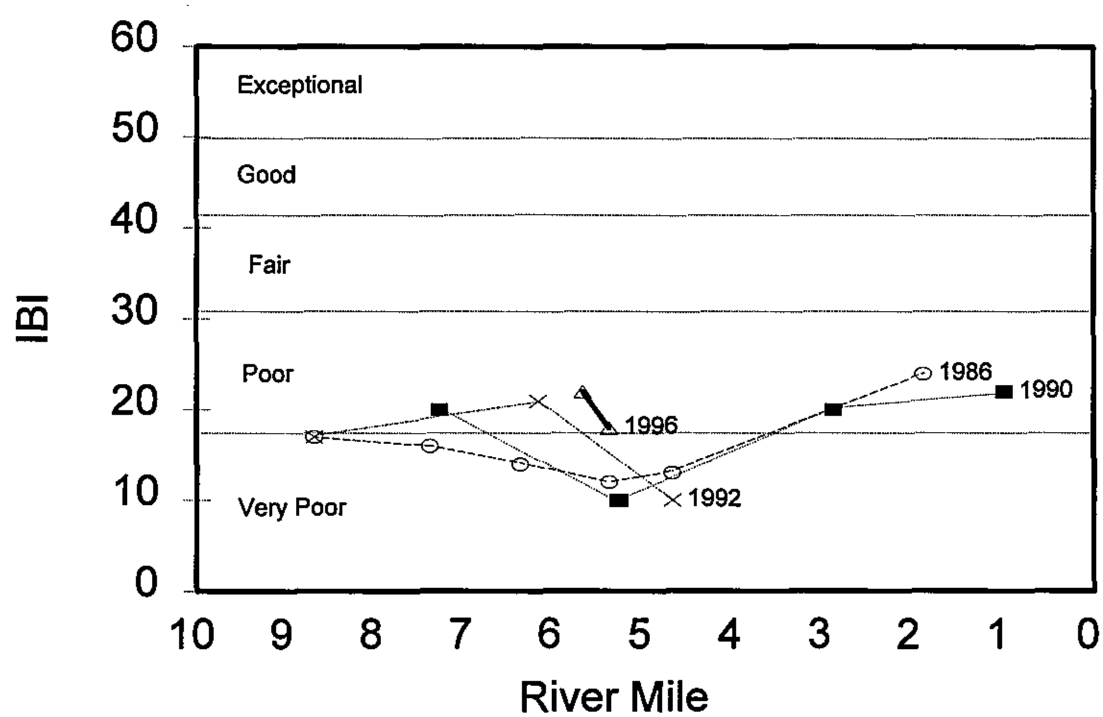


Figure 4.7.1 Ottawa River IBI scores for 1986, 1990, 1992, and 1996. Exceptional, good, fair, poor, and very poor classifications are delimited by dashed lines.

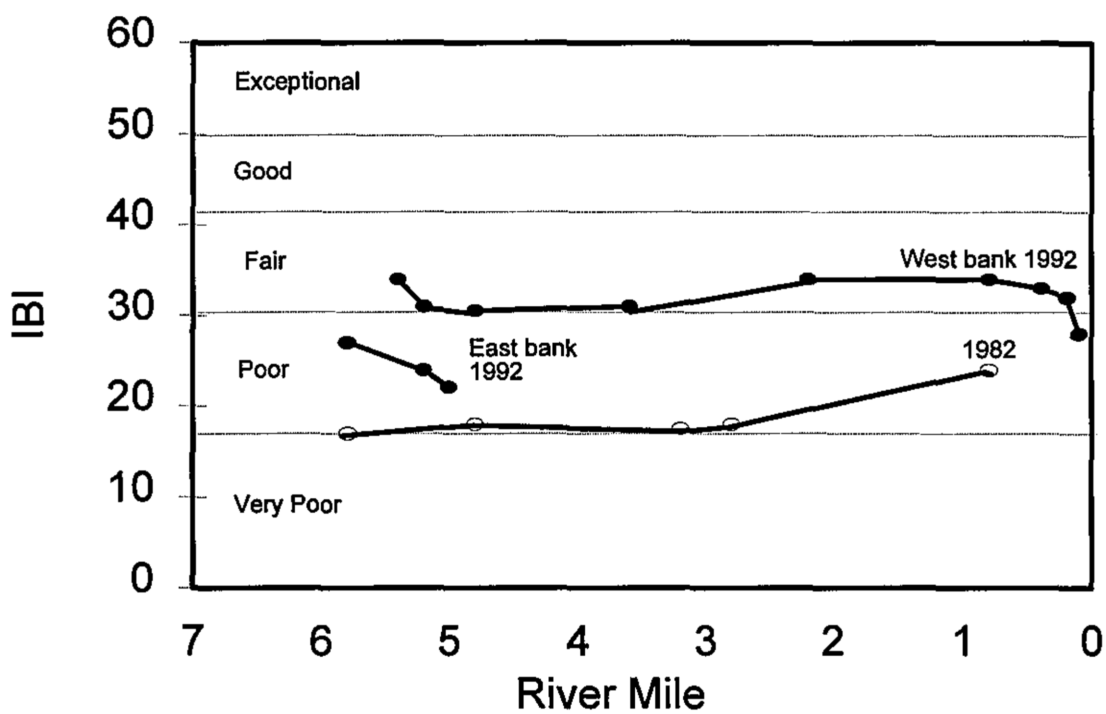


Figure 4.7.2 Black River IBI scores for 1982 and 1992. Exceptional, good, fair, poor, and very poor classifications are delimited by dashed lines.

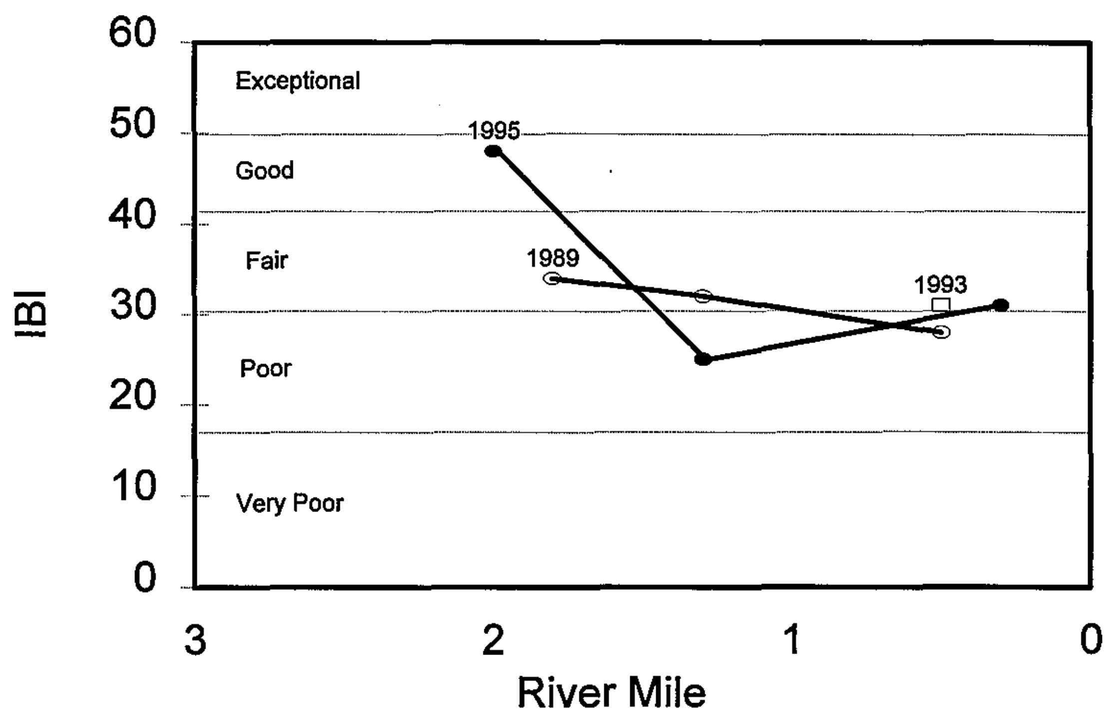


Figure 4.7.3 Ashtabula River IBI scores for 1989, 1993, and 1995. Exceptional, good, fair, poor, and very poor classifications are delimited by dashed lines.

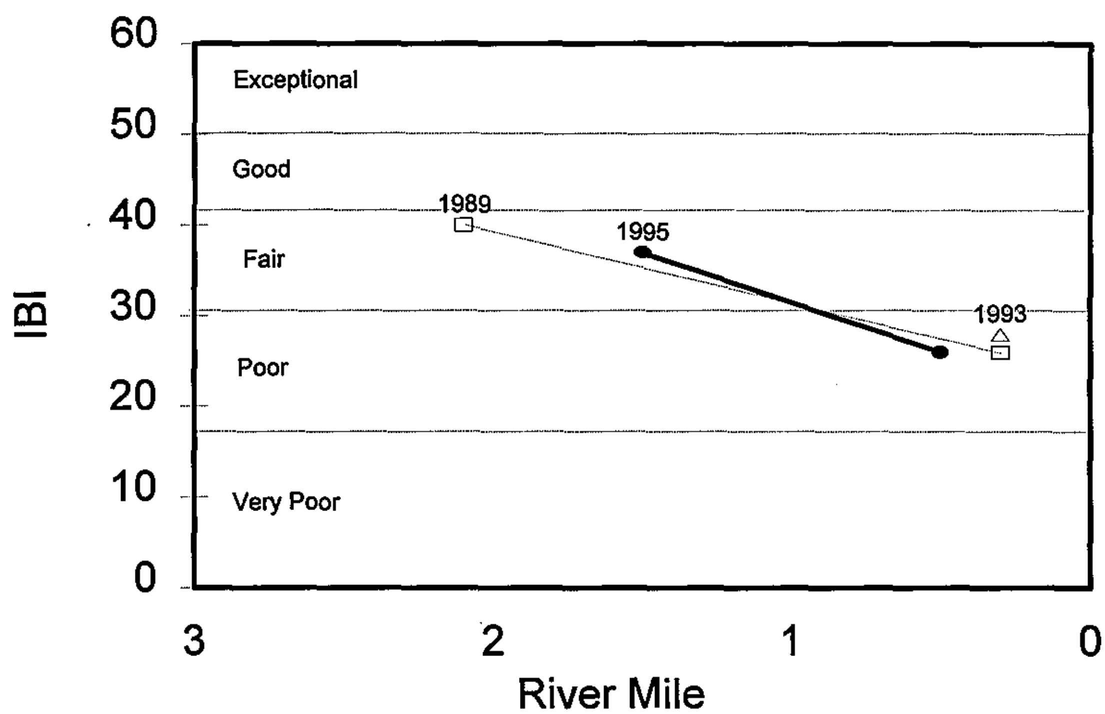


Figure 4.7.4 Conneaut Creek IBI scores for 1989, 1993, and 1995. Exceptional, good, fair, poor, and very poor classifications are delimited by dashed lines.

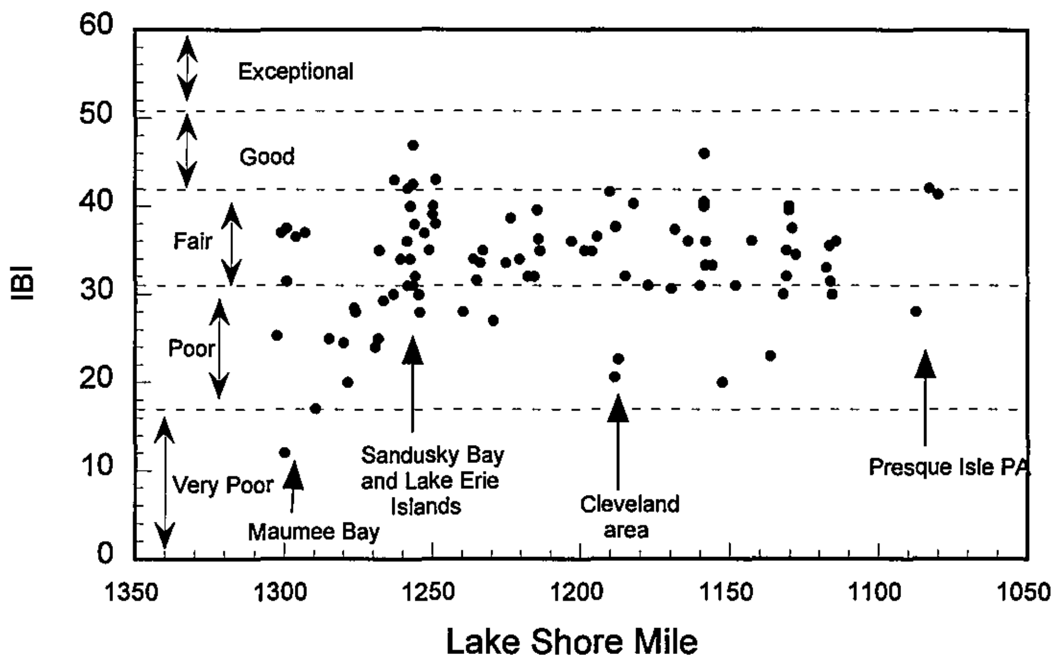


Figure 4.7.5 IBI scores for all Lake Erie sites. Habitats include Sandusky Bay, Bass islands area, and miscellaneous shore types (rocky and sandy beaches). Lake shore miles are measured from east to west. Exceptional, good, fair, poor, and very poor categories are delimited by dashed lines.

4.8 AN INDEX OF BENTHIC CONDITION TO DETERMINE THE MAGNITUDE OF ENVIRONMENTAL STRESS

Kevin Summers and Virginia Engle

US EPA, Gulf Ecology Div., Gulf Breeze, FL

The Environmental Monitoring and Assessment Program for Estuaries (EMAP-E) in the Louisianian Province has collected data from 644 stations in four years (1991-1994). One of the objectives of EMAP is to develop and test indicators of environmental quality and to use these indicators to determine the status of, and trends in environmental condition over large geographical areas. A core response indicator that has been developed for EMAP-E is the benthic index. The benthic index is a useful and valid indicator of estuarine condition that is intended to provide environmental managers with a simple tool for assessing the health of benthic macroinvertebrate communities. It represents the response of the benthic macroinvertebrate community to environmental stressors.

The benthic index was developed by first choosing a set of test sites that represent extreme degraded and reference conditions based on *a priori* guidelines for dissolved oxygen, sediment toxicity, and sediment contamination. These test stations were also chosen to represent both the range of natural habitat conditions found in the province and the entire geographic area included in the province. We compiled a suite of parameters that represent indicators of benthic community health including species richness and diversity, overall abundance, and the proportional abundance of major taxonomic and trophic groups of benthos. Parameters that showed a high degree of correlation with natural habitat conditions (e.g., salinity or sediment grain-size) were adjusted accordingly. Stepwise and canonical discriminant analyses were used to determine which subset of the benthic parameters best discriminated between the degraded and reference test sites and to assign coefficients or weighting factors to each of the parameters.

We originally developed a benthic index using data from the 1991 demonstration project in the Louisianian Province. That benthic index combined the Shannon-Wiener index (adjusted for salinity) and the percentages of total abundance represented by tubificids (Family: *Tubificidae*) and bivalves (Class: *Bivalvia*). This original index successfully discriminated between reference sites and sites that were degraded with respect to sediment contaminants, sediment toxicity, and hypoxia. However, when this benthic index was applied to an independent set of data from the Louisianian Province (EMAP's 1992 sampling of 159 new sites), validation of the index was unsuccessful. This was partly the result of 1992 sites that had benthic conditions that were substantially more degraded than the original test sites used to develop the index. A new, revised benthic index was developed using test sites from 1991 and 1992 that represented a broader set of environmental conditions.

The revised benthic index that was developed for EMAP-E in the Louisianian Province is a linear combination of 1) the proportion of expected diversity, 2) mean abundance of tubificid oligochaetes, 3) the percent of total abundance represented by capitellid polychaetes, 4) percent bivalves, and 5) percent amphipods. The weights on each of the independent variables were determined empirically based on the data. This benthic index successfully delineates benthic communities that have characteristics similar to those found in areas known to be degraded, from benthic communities that are similar to those found in known, reference areas. The difference in benthic community structure indicated by our benthic index is more likely to be due to anthropogenic stress than to natural habitat variability.

Validation of the benthic index was accomplished by using an independent set of data from two subsequent years, 1993 and 1994, as well as data from special study sites representing between-year and within-year replicates. Validation of the benthic index consisted of three steps: assessment of the correct classification by the index of an independent set of degraded and reference sites, comparison of the cumulative distribution function of the index among four years, and correct classification of replicate sites by the index. The revised benthic index was validated successfully using the independent data and was then retrospectively applied to all of the data collected from Gulf of Mexico estuaries during 1991-1994.

The benthic index is intended to be used as an indicator of the ecological health of estuaries by ranking and classifying the conditions of benthic invertebrate communities over large geographical areas. It can also be used successfully to classify specific areas of a single estuary as degraded or reference with respect to benthos. We can then try to identify what possible stressors may exist only in the degraded areas. This provides a clue to what environmental impacts may be affecting the benthic communities at the degraded areas.

Monitoring ecological indicators of condition on a regional scale can produce information that is useful to resource managers. EMAP's probabilistic sample design and standardized methodologies allowed for the collection of data that can be used in preforming assessments across the region with a quantifiable level of confidence. Benthic index estimates for the estuaries of the Gulf of Mexico based on the 1991-1994 monitoring indicate that $23 \pm 6\%$ of the estuarine area in the Louisianian Province had degraded benthic resources based on low benthic index scores.

Using the benthic index as an indicator of benthic condition, we explored the spatial distribution of degraded benthic communities in individual estuaries, Pensacola Bay, FL and Mobile Bay, AL. These estuaries were sampled as part of a regional EMAP effort to characterize ecological conditions on a smaller geographic scale. We also investigated statistical associations between various environmental indicators and the benthic index in these estuaries.

Pensacola Bay, an estuary in northwest Florida, has a history of sedimentation problems due to poor flushing and locally high inputs of suspended sediments which are generally retained within the system. However, the sediment and biological quality of Pensacola Bay have deteriorated since the 1950s and recovery is improbable without substantial intervention. The benthic index identified 12 degraded sites that were located primarily in the mainstem of Pensacola Bay and in the three bayous proximal to the city of Pensacola (Bayous Chico, Grande, and Texas). Pensacola Bay has severely contaminated sediments with as many as 40 chemicals at concentrations greater than ER-L guidelines, especially in the bayous. The benthic community is impoverished throughout the bay, but severely so in the areas with low sediment quality.

The benthic communities of Mobile Bay are more affected by hypoxia and nutrient enrichment than by toxic sediments. Although hypoxia in Mobile Bay is primarily driven by salinity stratification and the timing and duration of wind events, the severity and extent of hypoxic bottom waters may be exacerbated by nutrient enrichment. In this case the dominant benthic taxa at degraded sites are small, tube-dwelling polychaetes indicative of a stressed environment.

We have successfully synthesized benthic community information into a benthic index of ecological condition that provides environmental managers with an easy way to assess the status of the health of benthic communities over large geographical areas. A response indicator like the benthic index provides a numerical quantification of the response of the benthic communities to environmental stresses. Because the benthic index is scalable and the criteria for determining the classification of degraded or reference are numeric, the application of the benthic index to other estuaries is straightforward. The application of the benthic index to data from an independent sampling program in Pensacola Bay illustrates this point.

4.9 A BENTHIC INDEX FOR ESTUARIES OF THE SOUTHEASTERN UNITES STATES

Robert F. Van Dolah[1], Jeffrey L. Hyland[2], A. Frederick Holland[1], Jeffrey S. Rosen[3] and Timothy R. Snoots[1]

[1] SC Marine Resources Division, P.O. Box 12559, Charleston, SC
[2] NOAA Carolinian Province Office, P.O. Box 12559, Charleston, SC
[3] TPMC, Mill Wharf Plaza, Suite 208, Scituate, MA

Introduction

We have developed and validated a benthic index for southeastern estuaries using data from the joint EPA-NOAA Environmental Monitoring and Assessment Program (EMAP) in the Carolinian Province (Cape Henry, VA–St. Lucie Inlet, FL). Our approach follows methods developed by Weisberg et al. (1997) to characterize the condition of infaunal assemblages in Chesapeake Bay. This approach differs from the one used in previous EMAP estuarine surveys of the Virginian Province (Weisberg et al. 1992) and Louisianian Province (Engle et al. 1994), which produces an index derived from multivariate stepwise and canonical discriminant analysis. The approach we have adopted here is a modification of the Index of Biotic Integrity (IBI) developed originally for freshwater systems (Karr et al. 1986, Karr 1991). Though there are similarities to the latter IBI approach, one major difference is the way in which scoring criteria for selected biological attributes were established.

Our goal was to develop an index that characterizes the quality of estuarine habitats based on the condition of resident benthic infaunal assemblages. Additionally, the index should be:

1. suitable for use throughout the region;
2. applicable to all habitat types;
3. easy to understand and interpret; and
4. effective in discriminating between degraded and undegraded habitats.

Methods

Results of the EMAP survey completed in 1994 indicated that several natural abiotic factors (salinity, latitude, silt-clay, and TOC) had strong influences on infaunal variables (Hyland et al. 1996). The approach used here attempted to produce an index of integrated benthic response variables independent of these abiotic factors. The basic steps used to develop the index involved:

1. selecting a test data set (75 stations sampled in the summer 1994 from the NC/VA

border to the southern end of Indian River Lagoon, FL);

2. defining major habitat types based on classification analysis of the benthic species test data and evaluation of the physical attributes associated with the resulting site groups;

3. comparing various candidate benthic attributes between reference sites and degraded sites for each of the major habitat types;

4. selecting the attributes that best discriminated between reference and degraded sites for inclusion in the index (key criteria considered were whether differences were in the right direction and statistically significant);

5. establishing scoring criteria (thresholds) for the selected attributes based on the distribution of attribute values at reference sites;

6. deriving a combined index value for each sample by assigning an individual score for each attribute, based on the scoring criteria, and then averaging the individual scores; and

7. validating the index with an independent data set (96 stations sampled during the summer 1993 and 1995).

Several criteria were used to classify stations as degraded or undegraded on the basis of chemistry and toxicity data. Stations were considered to be degraded if:

1. sediments were contaminated (i.e., three or more contaminants in excess of lower, threshold ER-L/TEL sediment bioeffect guidelines, or one or more contaminants in excess of higher ER-M/PEL probable effect guidelines);

2. laboratory sediment bioassays showed toxicity ($\geq$ 2 hits using amphipods, seed clams, and/or Microtox®); or

3. there was low dissolved oxygen observed in the water column ($<$ 0.3 mg/L for any observation, $<$ 2.0 mg/L for 20% or more of observations, or $<$ 5.0 mg/L for all observations over a 24-hr time series). ER-L and ER-M values are from Long et al. (1995) and Long and Morgan (1990); TEL and PEL values are from MacDonald (1994).

Forty benthic infaunal attributes were considered and statistically compared within each of four habitat groups. These groups were oligohaline–mesohaline stations ($\leq$ 18 ppt) from all latitudes, polyhaline–euhaline stations ($>$ 18 ppt) from northern latitudes ($>$ 34.5° N), polyhaline–euhaline stations from middle latitudes (30–34.5° N) and polyhaline-euhaline stations from southern latitudes ($<$ 30° N). The initial list of attributes included various measures of diversity, abundance, dominance, and presence of indicator species (e.g., pollution sensitive vs. tolerant species, surface vs. subsurface feeders). A subset of six candidate metrics that best discriminated between reference and degraded sites was identified for possible inclusion in the index. Scoring criteria for each of these metrics were developed based on the distribution of values at undegraded sites (score of 1, if value of metric for sample being evaluated was in the lower 10th percentile of corresponding reference-site values; score of 3, if value of metric for sample was in the 10th–50th percentile of reference-site values; or score of 5, if value of metric for sample was in the upper 50th percentile of reference-site values). Scoring criteria were determined separately for each metric and habitat type.

Forty different combinations of the six candidate benthic metrics were further evaluated to determine which represented the best combined index. For each, a combined index value was calculated by assigning a score for each component metric (based on the individual scoring criteria) and then averaging the individual scores. A combined score < 3 was used to suggest the presence of a degraded benthic assemblage (very unhealthy to some apparent level of stress). The metric combination that produced the highest percentage of correct classifications (i.e., agreement with predictions of sediment bioeffects based on the various exposure measures) was then selected to represent the final index.

Results

The final index was the average score of four metrics: total abundance, number of species, 100% - % abundance of the two most dominant taxa, and % abundance of pollution-sensitive taxa. Percent pollution-sensitive taxa consisted of the percent of total faunal abundance represented by Ampeliscidae + Haustoriidae + Hesionidae + Tellinidae + Lucinidae + Cirratulidae + *Cyathura polita* and *C. burbanki*.

This combined benthic index correctly classified stations 93% of the time in the developmental data set and 75% of the time in the independent validation data set (Table 4.9.1 and Figure 4.9.1). Figure 4.9.1 further illustrates that stations with index values below 3 (suggestive of some apparent stress to highly degraded conditions) usually coincided with sites considered to be degraded based on a combination of chemistry and toxicity data, and that stations with scores of 3 or higher usually coincided with undegraded sites. Agreement was the highest at the two ends of the scale. Thus, the evaluation of sediment quality based on the benthic index appears to agree reasonably well with predictions of sediment bioeffects based on the combined exposure data. Additional comparisons revealed that the benthic index detected a higher percentage of samples where bioeffects were expected (based on contaminant bioeffect exceedances) than did any of the four individual sediment bioassays (Figure 4.9.2).

References

Engle, V.D., J.K. Summers, and G.R. Gaston. 1994. A benthic index of environmental condition of Gulf of Mexico estuaries. *Estuaries*, 17 (2): 372-384.

Hyland, J.L., T.J. Herrlinger, T.R. Snoots, A.H. Ringwood, R.F. Van Dolah, C.T. Hackney, G.A. Nelson, J.S. Rosen, and S.A. Kokkinakis. 1996. Environmental Quality of Estuaries of the Carolinian Province: 1994. Annual Statistical Summary for the 1994 EMAP-Estuaries Demonstration Project in the Carolinian Province. NOAA Technical Memorandum NOS ORCA 97. NOAA/NOS, Office of Ocean Resources Conservation and Assessment, Silver Spring, MD.

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

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

Long, E.R., D.D. MacDonald, S.L. Smith, and F. D. Calder. 1995. Incidence of adverse biological effects within ranges of chemical concentrations in marine and estuarine sediments. *Envir. Man.*, 19: 81-97.

Long, E.R. and L. G. Morgan. 1990. The potential for biological effects of sediment-sorbed contaminants tested in the National Status and Trends Program. NOAA Technical Memorandum NOS OMA 52. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Rockville, MD.

MacDonald, D.D. 1994. Approach to the assessment of sediment quality in Florida coastal waters. Vols. I-IV. Report prepared for Florida Department of Environmental Protection.

Weisberg, S.B., J.B. Frithsen, A.F. Holland, J.F. Paul, K.J. Scott, J.K. Summers, H.T. Wilson, R. Valente, D.G. Heimbuch, J. Gerritsen, S.C. Schimmel, and R.W. Latimer. 1992. EMAP-Estuaries Virginian Province 1990 demonstration project report. U.S. EPA Environmental Research Laboratory, Narragansett, R.I. EPA/600/R-92/100.

Weisberg, S.B., J. A. Ranasinghe, D.M. Dauer, L.C. Schaffner, and J.B. Frithsen. 1997. An estuarine benthic index of biotic integrity (B-IBI) for Chesapeake Bay. *Estuaries*, 20 (1): 149-158.

Table 4.9.1 Classification efficiencies of the Carolinian Province benthic index.

	1994 "development" data		1993/95 "validation" data	
Habitat	# Sites	% Correct Classifications	# Sites	% Correct Classifications
Oligo. - Mesohaline, All Latitudes	20	90	46	78
Poly. - Euhaline, Northern Latitudes	24	92	13	85
Poly. - Euhaline, Middle Latitudes	22	95	27	74
Poly. - Euhaline, Southern Latitudes	9	100	10	50
Overall	75	93	96	75

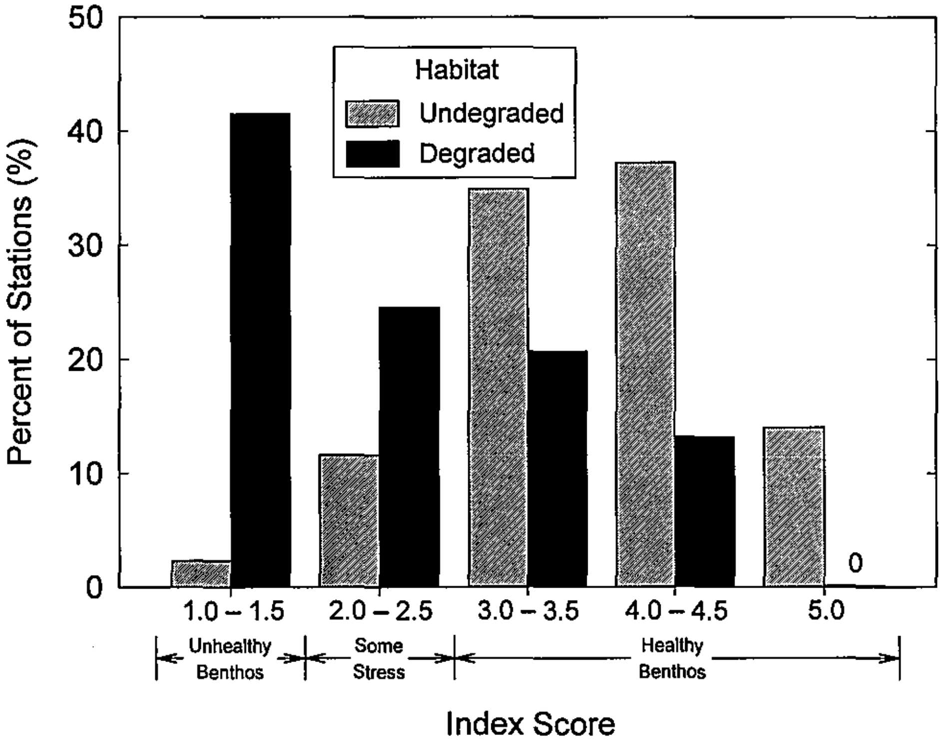


Figure 4.9.1. Frequency distribution of index scores for undegraded vs. degraded stations in 1993/1995 "validation" data set.

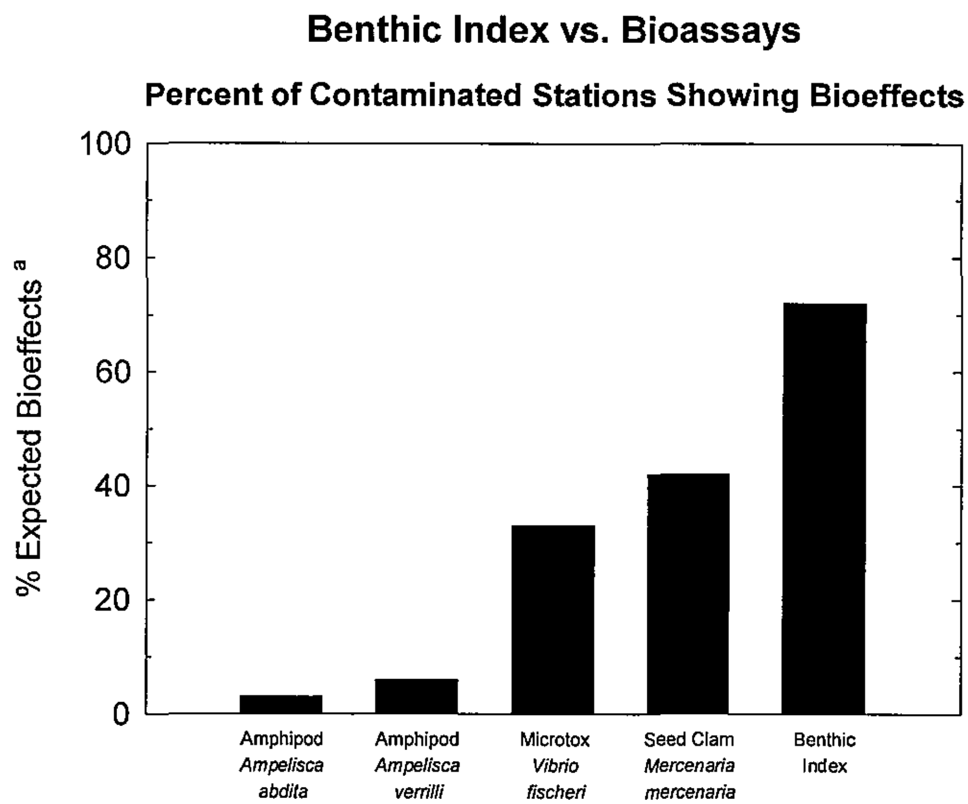


Figure 4.9.2. Comparison of the percent of expected bioeffects detected with the benthic index vs. sediment bioassays. [a] Percent expected bioeffects = # stations (1995 core & supplemental) where an effect was detected / # stations with ≥ 1 ER-M/PEL or ≥ 3 ER-L/TEL exceedance.

4.10 CHESAPEAKE BAY BENTHIC COMMUNITY RESTORATION GOALS

Ananda Ranasinghe[1], Stephen Weisberg[2], Daniel Dauer[3], Linda Schaffner[4], Robert Diaz[4] and Jeffrey Frithsen[1]

[1]Versar Inc. Columbia, MD
[2]Southern California Coastal Water Research Project, Westminster, CA
[3]Old Dominion Univ., Norfolk, VA
[4]Virginia Inst. of Marine Science, Gloucester Point, VA

Benthic macroinvertebrate assemblages have been an integral part of the Chesapeake Bay monitoring program since its inception due to their ecological importance and their value as biological indicators. The condition of benthic assemblages reflects an integration of temporally variable environmental conditions and the effects of multiple types of environmental stresses. As such, benthic assemblages provide a useful complement to more temporally variable chemical and water quality monitoring measures.

While assessments using benthic monitoring data have been useful for characterizing changes in environmental conditions at individual sites over time, and for relating the condition of sites to pollution loadings and sources, the full potential of these assessments for addressing larger management questions, such as "What is the overall condition of the Bay?" or "How does the condition of various tributaries compare?" has not yet been realized. Regional-scale assessments of ecological status and trends using benthic assemblages are limited by the fact that benthic assemblages are strongly influenced by naturally varying habitat elements, such as salinity, sediment type, and depth. Such natural variability confounds interpretation of differences in the benthic community as simple responses to anthropogenic environmental perturbations. An additional limitation is that different sampling methodologies used in various programs often constrain the extent to which the benthic data can be integrated for a unified assessment.

The objective of this project was to develop a practical and conceptually sound framework for assessing benthic environmental conditions in Chesapeake Bay that would address the general constraints and limitations just described. This was accomplished by standardizing benthic data from several different monitoring programs to allow their integration into a single, coherent data base. From that data base a set of measures (Chesapeake Bay Benthic Restoration Goals) was developed to describe characteristics of benthic assemblages expected at sites having little evidence of environmental stress or disturbance (CBP 1994, Weisberg et al. 1997). Using these goals, benthic data from any part of the Bay could be compared to determine whether conditions at that site met, were above, or were below expectations defined for reference sites in similar habitats.

The approach used to develop these restoration goals was similar to that used by Karr et al. (1986) to develop an index of biological integrity for freshwater fish. A set of candidate attributes believed to have properties that differentiate high and low quality assemblages were first identified, and reference sites believed to be "minimally impacted" were designated. Properties of the biotic assemblages at these sites were then compared to assemblage properties at all other sites. Properties that differed significantly between these two groups of sites were selected as metrics to be included in the restoration goals. An index was developed to assist managers in identifying the extent to which these restoration goals were being achieved. The Restoration Goals Index (RGI) is calculated as the average score of metrics, after each metric is scored as 5, 3, or 1, depending on whether its value at an individual site approximated, deviated slightly, or deviated strongly from its value at the best reference sites.

The restoration goals were developed based on available data from seven benthic survey projects: the Maryland and Virginia Chesapeake Bay Benthic Monitoring Programs, U.S. EPA's Environmental Monitoring and Assessment Program (Holland et al. 1990), the Maryland and Virginia Biogenics studies, a James River study, and a study in the Wolf Trap area of the Chesapeake Bay. These seven projects were selected for several reasons: each provided data readily available on electronic media; collectively they provided sample representation in all salinity habitats of Chesapeake Bay; and all used a 0.5 mm sieve in sample processing, which was a critical aspect of the study, since the numbers and types of organisms collected depend on the mesh size used to sieve the sediment.

The attributes incorporated into the restoration goals included metrics from each of the following five categories:

1. benthic biodiversity measures

2. measures of assemblage abundance and biomass

3. life history strategy measures

4. measures of activity beneath the sediment surface

5. feeding guild measures

Restoration goals were developed independently for eight habitat classes defined by salinity and sediment type to ensure that natural differences in benthic communities related to these habitat factors did not confound interpretation of the indices. The eight habitat classes were determined by cluster analysis of the composite data set.

Restoration goals were developed using data from only the summer period, July 15th through September 30th. This restriction avoided seasonal variation that would confound interpretation of benthic community responses to environmental degradation. The summer sampling period was common to six of the seven benthic survey projects. Using data from a different season would have reduced the data available because the various programs differed substantially in the extent of sampling during other seasons of the year. An index developed for summer was desirable because benthic communities are expected to show the greatest response to pollution stress during the summer.

Three approaches were used to validate the goals and the accompanying index. First, the RGI was computed for all samples taken from each reference site to test whether expectations of RGI values greater than three were met. This test indicated a high degree of correct classification; classification efficiency was more than 95% in five of the seven habitat classes. The lowest correct classification efficiency for reference sites was 92.3% in the high mesohaline mud habitat class. Second, RGI values were computed for all samples taken from degraded habitats to test whether expectations of RGI values less than three were met. This test used data that had been excluded from development of the RGI; therefore, it was an independent validation test. A high level of classification efficiency was observed in this test; classification efficiency was 85 % or better for degraded sites in five of the six habitat classes in which data from degraded sites were available. The one habitat class that did not validate as well was tidal freshwater. For the third validation test, sites that were sampled more than once during the summer of any year were identified, and the RGI was computed for each visit. RGI values at each site were evaluated for differences in status between visits within each year to ascertain the stability of the index. Instability of the index would indicate an unacceptable signal-to-noise ratio in the attributes. The results indicated that the RGI index was relatively stable. The correlation between RGI values for the first and second visits exceeded 80% for all habitats.

The validation results indicate that these preliminary restoration goals are effective for distinguishing between sites of high quality and those of lower quality in six of the seven habitats for which data were available for goal development. The only habitat class for which the restoration goals did not validate well was tidal freshwater. Although restoration goals validated well, additional analysis and development of goals appears to be appropriate before the goals are applied rigorously for environmental management purposes. Steps for further goal development are recommended.

References

Holland, A.F. (ed.). 1990. Near coastal program plan for 1990: Estuaries. EPA 600/4-90/033. U.S. EPA, ERL, ORD, Narragansett, RI.

Karr, J.R., K.D Fausch, P.L. Angermeier, P.R. Yant and I.J. Schlosser. 1986. Assessing biological integrity in running waters: A method and it's rationale. Ill. Nat. Hist. Surv., Pub #5. Champaign, Ill.

Weisberg, S.B., J.A. Ranasinghe, D.M. Dauer, L.C. Schaffner, R.J. Diaz and J.B. Firthsen. 1997. An estuarine benthic index of biotic integrity (B_IBI) for Chesapeake Bay. Estuaries, 20: 149-158.

4.11 A PRELIMINARY STUDY OF THE USE OF MARINE BIOCRITERIA SURVEY TECHNIQUES TO EVALUATE THE EFFECTS OF OCEAN SEWAGE OUTFALLS IN THE MID-ATLANTIC BIGHT

George Gibson

US EPA, Office of Water, Office of Science and Technology, Washington, D.C.

Objectives

This project investigates the practical, low cost application of marine biological community measurements and the nearfield/farfield survey technique, for use by coastal States as a water resource quality management tool. The methods applied here were derived from work reported by Pearson and Rosenberg (1978) and Mearns and Word (1982) with modifications.

Study Methods

The study area was a 10 mi coastal reach between Bethany Beach, Delaware and Ocean City, Maryland (Figure 4.11.1). These are nearly adjacent resort communities on the Mid-Atlantic seaboard between Delaware Bay and Chesapeake Bay. Each has a secondary treatment municipal sewage discharge site about 1.5 nautical miles (nm) offshore. Discharge in both cases is through a diffuser at a water depth of approximately 40 ft (12 m). The Bethany Beach sewage treatment plant discharges about 14 mgd and Ocean City about 32 mgd.

A series of nine north-south trending stations were installed parallel to the coast at intervals of about 1 nm, in about 40 ft depth of water and over medium to fine sandy bottoms to obtain as similar a habitat as possible. The stations were labeled "A" through "I", with station "C" at the Bethany Beach outfall and station "G" at the Ocean City outfall.

The variables measured were benthic fish and macroinvertebrate communities as reflected in indexes and metrics incorporating number of taxa and number of individuals per taxa. Fish surveys were made with a 20 ft (16 ft effective opening), 1 inch mesh otter trawl. Tows were made parallel to the shoreline at 2 knots over 0.5 nm with the station coordinates located at the mid-point of the tow. Benthic macroinvertebrate samples were collected with a 0.1 m^2 Smith-McIntyre grab or with a 0.1 m^2 Young grab, and three replicates were taken for each sample at each station site.

Sampling surveys have been conducted twice a year in July-September and January-February since 1993 to determine if multiple season indexing is necessary or appropriate. While the mid-Atlantic area is considered to have four discrete seasons, benthic communities are expected to be in flux during spring and fall and to be most stable in summer and winter (U.S. EPA 1994).

To make comparisons between the sample sites, habitat control in the survey design was maintained as well as possible by attention to four major variables; 1) sediment grain size,

2) water depth, 3) water quality (conductivity, temperature, depth, dissolved oxygen, pH, transmissivity, and 4) salinity. At the beginning of the project, sediment samples were collected from all nine stations and analyzed for heavy metals and a for a standard array of toxic contaminants. All results were insignificant, suggesting no other sources of biotoxicity or impairment indigenous to the immediate area.

In keeping with the objective of low cost, applications of standard, but robust taxonomic indexes were applied to the biological community data for impact detection, The underlying premise for the indices is that once the raw data for species and numbers of individuals per species are compiled, the investigator's primary question is whether or not there is a detectible impact. More refined indices and indicators can later be applied or developed as needed. In this regard, the treatments selected for this project were: total number of individuals, total number of taxa (species), evenness index, Simpson's dominance index, Margalef's taxa richness index, and the Shannon index of general diversity.

Results

Fish Survey Data

Analysis of the fish data showed no significant differences in trawl data between the stations in summer or winter collections for either number of taxa or numbers of individuals. Qualitatively, taxa and number of individuals overall shifted considerably between summer and winter surveys at the nine stations. Greater numbers of both species and individuals (excepting winter runs of striped anchovy, *Anchoa hepsetus*) occur in the summer surveys.

Benthic Macroinvertebrate Data

Benthic macroinvertebrate results have been much more promising, but the same seasonal trend observed with fish for number of taxa and number of individuals has prevailed. Summer measurements were much more indicative of the condition of the benthic macroinvertebrate assemblages. The data in this instance was for three replicates at each station twice a year for three years. Significant differences are evident between each of the outfall sites and the other stations in the summer data. The number of individuals show a gradient from high to low, proceeding from north to south, with an increase in the vicinity of the Ocean City outfall station. This suggests enhanced and or enriched conditions perhaps from the Delaware Bay discharge, and at the Ocean City outfall site.

When numbers of species were compared, a negative trend in outfall impact was evident, especially for the Bethany Beach outfall station (Figure 4.11.2). A similar pattern occurred at Ocean City, but was not as strong. Ludwig and Reynolds (1988) state that a simple count of the number of species present, for samples of equal size, avoids some of the problems of using indices which combine and may confound a number of variables that characterize community structure. However, in this instance, it appears that at least some indices enhance the measurement of outfall perturbations. Box plots of Margalef richness index (Figure 4.11.3) over the three years of summer data provide strong indications of the negative effect of both discharges on the benthic macroinvertebrate community. Simpson's dominance index and the Shannon index of general diversity reveal a similar effect.

Discussion and Conclusions

The nearfield/farfield survey design for biological surveys, together with basic indices of community structure, appears to work equally well on the west coast and in mid-Atlantic coast open water environments. Summer benthic macroinvertebrate data from stations "A" and "C" were significantly different, lending confidence to the conclusion that the wastewater discharges were having a measurable impact on the coastal marine environment. This is of particular interest because routine water quality and sediment investigations at the sites failed to detect change between the outfalls and the surrounding stations. The standard indices such as Margalef's richness index, Simpson's dominance index, and Shannon's diversity index are robust and were entirely appropriate for this survey.

For biocriteria development and site monitoring, it is important to account for seasonality. For the mid-Atlantic Bight, late June to early September appears to be a time of relatively high, stable community productivity and an optimal index period if once a year sampling is preferred. Because Bethany Beach and Ocean City are summer resort communities, their populations increase at least ten-fold in warm weather (Bethany Beach, DE, and Ocean City, MD Chambers of Commerce, personal communication, 1990). Their lower winter discharge rates, together with a natural cyclic depletion of the marine community, may account for the failure of our data to reveal sewage impacts in this season. This may not be the case with a year-round municipality of fairly large size. In any case, if the responsible agency can afford to sample at least occasionally in winter, that baseline biological data may prove invaluable in the event of oil spills or other marine accidents.

References

Ludwig, J.A. and J.F. Reynolds. 1988. Statistical Ecology. J.A. Wiley & Sons, NY, NY. 337pp.

Mearns, A.J. and J.Q. Word. 1982. Forecasting the effects of sewage solids on marine benthic communities. In, Ecological Stress and the New York Bight: Science and Management. G.F. Mayer (ed.). Estuarine Research Federation, Columbia, SC. p 495-512.

Pearson, T.H. and R. Rosenberg. 1978. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. In, Oceanography and Marine Biology, an Annual Review, 16:229-311.

U.S. EPA. 1994. Chesapeake Bay Benthic Community Restoration Goals. U.S. EPA, Chesapeake Bay Program, Annapolis, MD. CBP/TRS 107/94. 88pp.

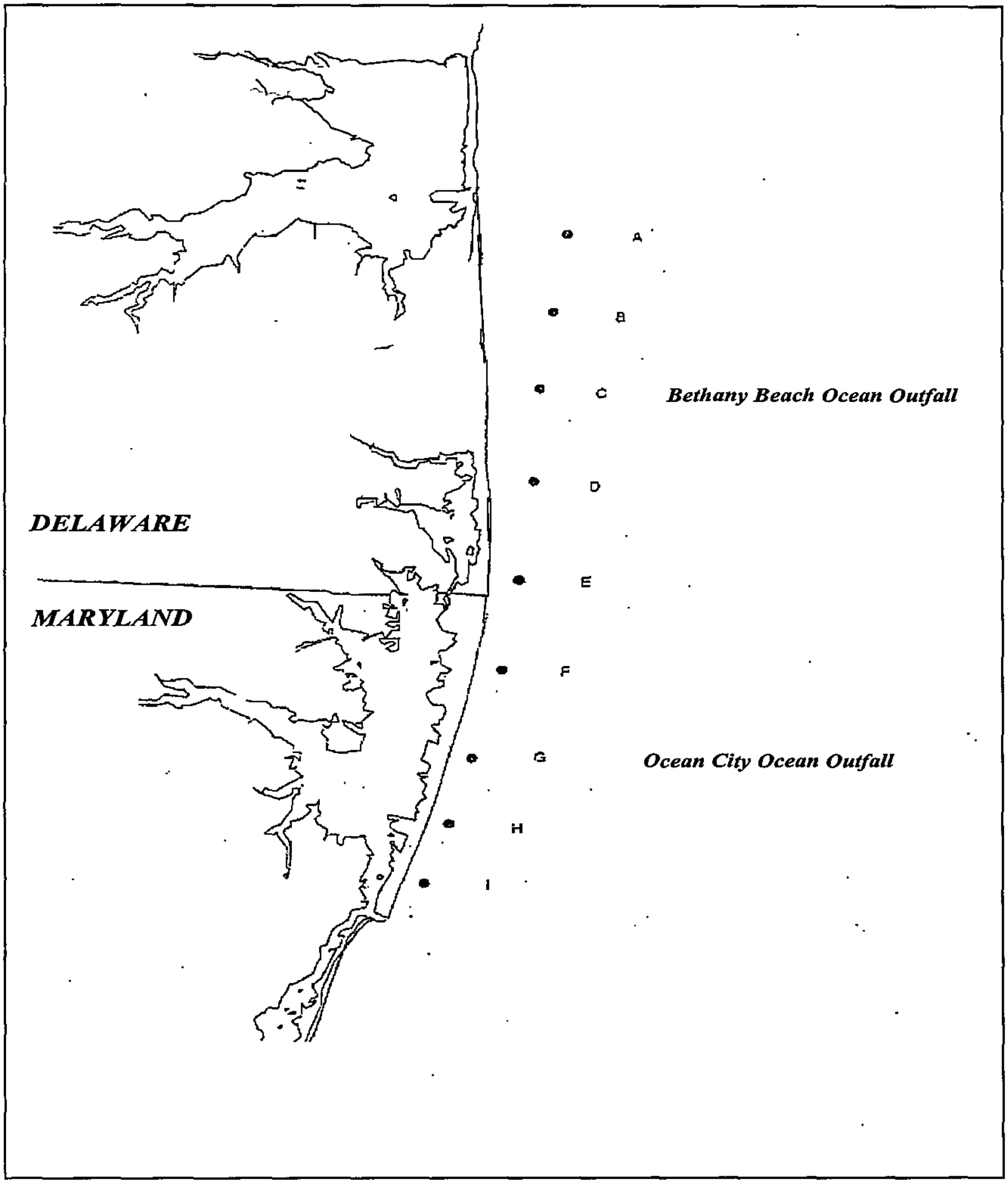


Figure 4.11.1 Offshore sampling locations off the coasts of Delaware and Maryland during the summers of 1992-1994.

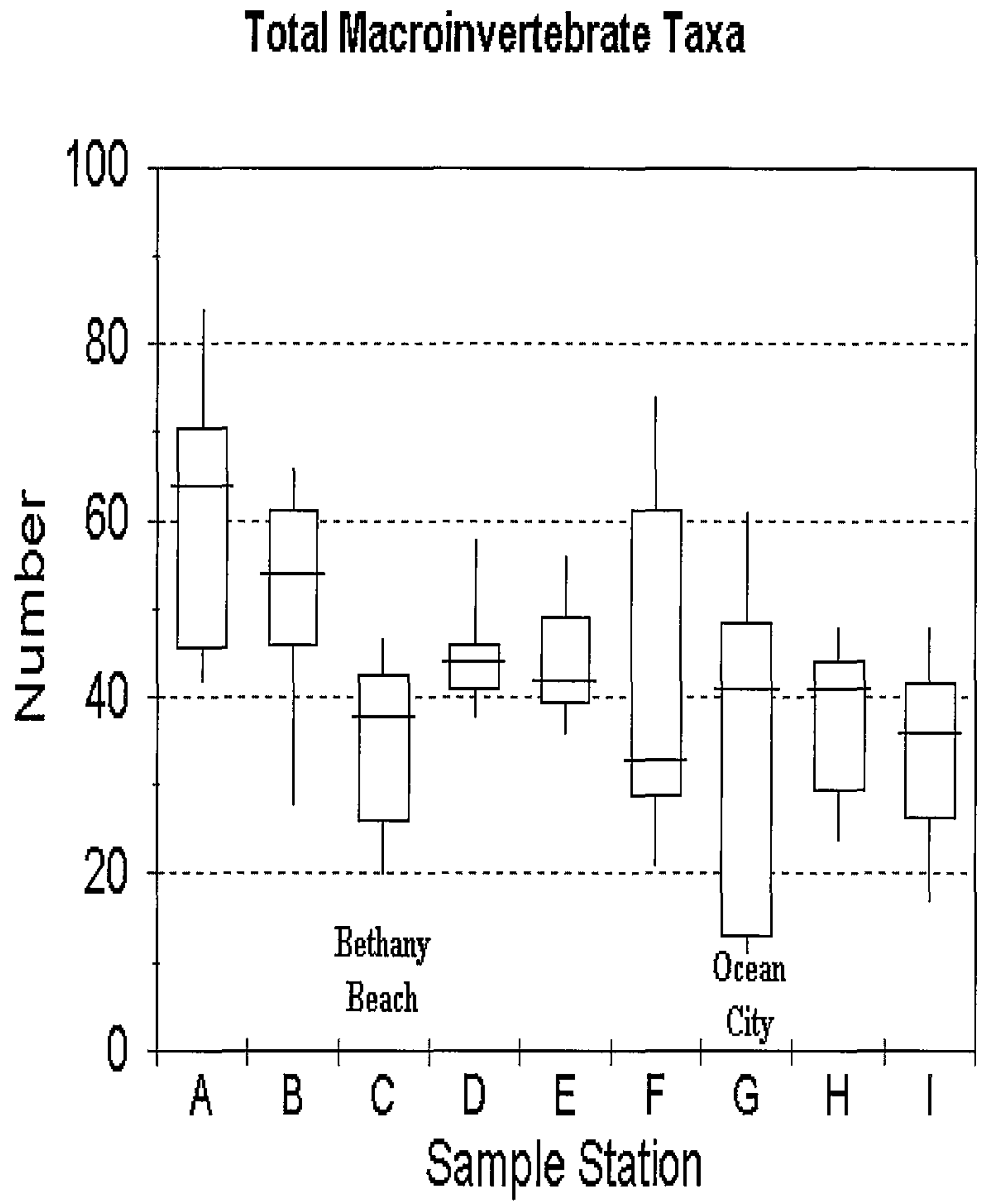


Figure 4.11.2 Number of taxa of macroinvertebrates collected at nine stations off the coasts of Delaware and Maryland during summers of 1992-1994. Lines and bars show maximum, minimum, 75%, 25% and median values.

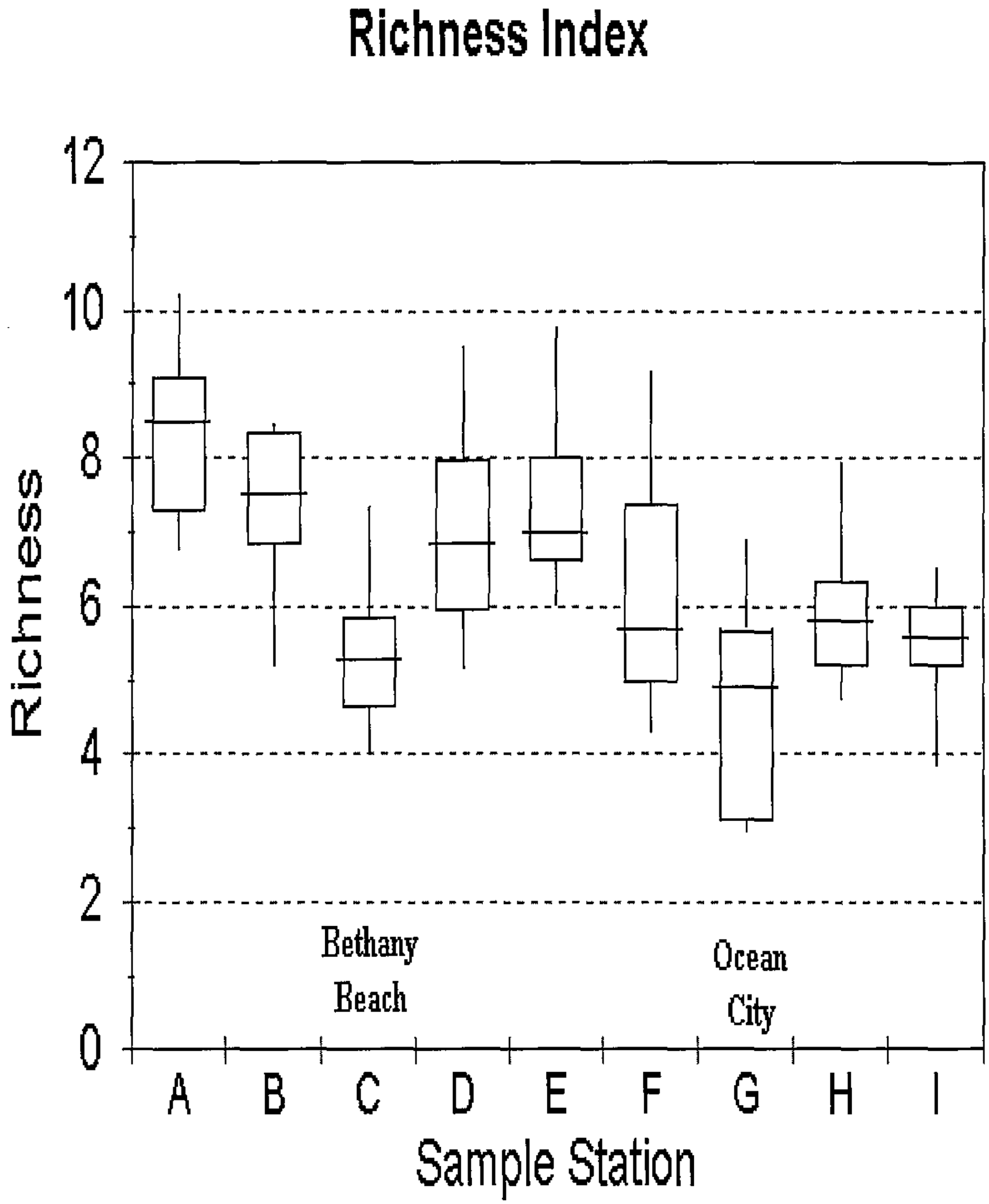


Figure 4.11.3 Taxa richness of macroinvertebrate species collected at nine stations off the coasts of Delaware and Maryland during summers of 1992-1994. Lines and bars show maximum, minimum, 75%, 25% and median values.

4.12 THE BENTHIC RESPONSE INDEX: MEASURING IMPACTS ON BENTHIC ASSEMBLAGES IN SOUTHERN CALIFORNIA

Mary Bergen[1], Bob Smith[2], Steve Weisberg[1], Don Cadien[3], Ann Dalkey[4], Dave Montagne[3] , Jan Stull[3], and Ron Velarde[5]

[1]Southern California Coastal Water Research Project, Westminster, CA
[2]EcoAnalysis, Inc., Ojai, CA
[3] Technical Services Division, County Sanitation Districts of Los Angeles Co., Whittier, CA
[4]Environmental Monitoring Division, City of Los Angeles, El Segundo, CA
[5]Metropolitan Wastewater Department, City of San Diego, San Diego, CA

To determine the best course of action, managers need to know if a resource has been impacted and how much it has been impacted. They also need to know if the condition of the resource is improving or degrading. Ideally, the manager should be able to evaluate the status of the resource using an objective, quantitative measure. This measure should clearly discriminate impacted from reference sites, be easily quantified and communicated, and be insensitive to differences in habitat, seasons or other sources of natural variability. In addition, there should be breakpoints or thresholds for the measure that indicate meaningful points of change in the resource, such as the limit of reference condition or the initiation of severe degradation in the resource. In southern California, we have been able to develop a benthic index which meets these criteria.

To develop the index, we followed an approach used in other programs. We assembled a database including information from known reference and impacted sites. We identified 26 measures, including measures of diversity, abundance, biomass, species composition and mode of feeding, and tested the ability of the measures to discriminate impacted and reference sites. Most measures were not informative. However, two measures based on species composition, the Infaunal Trophic Index (ITI) and an ordination score, provided meaningful discrimination.

While either ITI or ordination could be used as an index, each has limitations. For this reason, we decided to create a new index that would capture the information in the measures but that would be easy to calculate and communicate. The Benthic Response Index (BRI) is the abundance-weighted average of the pollution tolerance of the species in the samples:

$$BRI = \frac{\sum P_i \left(\sqrt[3]{No_i}\right)}{\sum \sqrt[3]{No_i}}$$

where P is the average position of the species on an impact gradient.

To determine P for the species, the impact gradient was defined in an ordination space. The P value was calculated from the distribution of the abundance of the species on the impact gradient. P values were calculated for 537 species in three depth zones: 10-35m, 25-120m, and 110-324 m.

Thresholds were defined for reference and four response levels: (1) marginal deviation, (2) loss in biodiversity, (3) loss in community function and (4) defaunation. The threshold for the reference was the 90th percentile of the index value in samples from reference areas. The endpoints of the distributions of species on the impact gradient were used to define response levels. The threshold for loss in biodiversity was exclusion of 25% of the species pool. Loss in community function was exclusion of 75 and 90% of the arthropods and echinoderms, respectively. Defaunation was exclusion of 90% of the species pool.

The index was validated with data from monitoring programs that had not been used in index calibration. Validation was based on the ability of the index to reproduce known spatial and temporal patterns as well as the ability of the index to discriminate impacted from reference sites across habitat types. The index was validated in every test case.

Based on the results, we believe that we have developed an effective index for soft-bottom assemblages from Point Conception to the Mexican border between 10 and 250 m. Our approach to index development could be used in other geographic areas and with other assemblages, particularly in areas where there is clear separation of the impact gradient from natural gradients.

5.0 Workshop Discussion Group Summaries

5.1 VEGETATED HABITATS

The vegetated habitat category included submerged aquatic vegetation (vascular plants and algae), emergent wetlands, mangrove and kelp habitats. The group established several basic concepts as ground rules in evaluating which biological attributes are most appropriate for the habitat type under consideration. For the purposes of using biological indicators as measures of the condition of habitat designated as EFH, a healthy habitat implies a healthy and sustainable fish population. This links an ecosystem approach (as called for by the Magnuson Stevens Act) to the management of individual populations. A second consideration was that there must be some way to connect human impact effects to all chosen attributes. In addition, reference habitats result from evolutionary and biogeographic processes. With extensive discussion the working group came up with the following attributes that should be considered when developing indices of biological integrity for submerged aquatic vegetation.

Submerged Aquatic Vegetation (SAV)
Plant Attributes:

- Diversity:
    - Vascular plants vs. Macroalgae
    - Genetic diversity
- Abundance:
    - Shoot density, SAVs
    - Patchiness, SAVs
    - Plant exotics
    - Macroalgal biomass
- Function:
    - Blade width/length
    - Epiphyte biomass
- Population/processes:
    - Plant age structure
    - Perennial vs. annual species
    - Runners vs. tuft rhizomes
    - Rhizome density
    - Number flowering
- Tolerance:
    - Frequency of wasting disease
- Physical:
    - Total Organic Carbon in sediments
    - Light
    - Temperature

Animal Communities in SAV Communities:

Diversity:

Infauna; epifauna; fish

Abundance:

Presences of large bivalves/ invertebrates

Avoid biomass

Function:

Number of nursery fish species

Number of resident species

Number of spawning species

Number of shellfish species (as bioindicators of SAVs but varies geographically)

Tolerance:

Fish lesions

Number of tolerant species

In the time available the working group was able only to outline the attributes for submerged aquatic vegetation habitats. However, general discussions provided several observations on the other types of vegetated habitats. For example, emergent wetlands are usually characterized by greater plant species richness and diversity and this attribute should be properly captured. One might also want to emphasize the importance of exotic species such as *Phragmites*. On the other hand, kelp beds and mangroves represent much more limited plant diversity communities and are more "monocultures" like SAV communities.

5.2 BENTHIC HABITATS

The open water benthic habitat category included soft bottom, hard bottom, and live bottom substrates. Soft bottom habitats were defined as having unconsolidated sediment, including anything from silt-clay deposits to coarse sand or gravel. Hard bottom habitats included cobble, consolidated rock and other solid surfaces to which benthic organisms can attach, primarily in the near shore and intertidal zone or estuaries. Live bottom was considered to be those habitats in which the physical structure of the habitat was composed of, or built by, sessile organisms, and included oyster bars, coral reefs and offshore benthic assemblages with significant three-dimensional relief. For the purpose of discussion, a distinction was made between estuarine (including submerged and intertidal), coastal shore zone, and offshore, in addition to substrate type. The coast was also subdivided into regions based on large scale oceanographic and geological features.

The following regions were delineated in the Atlantic and Gulf of Mexico coasts:

1. Canadian border to Cape Cod;

2. Cape Cod to Cape Hatteras;

3. a combination of the area from Cape Hatteras to Cape Canaveral and from Tampa Bay to the Rio Grande; and

4. southern Florida and the Caribbean

On the Pacific coast the regions were:

1. Mexico to Pt. Conception;

2. Pt. Conception to the Columbia River;

3. Columbia River to Canada and Alaska's Pacific coast; and

4. Alaska's Bearing Sea and Arctic Ocean coasts.

The major determinants of regions on the Atlantic and Gulf coasts were climate, ocean circulation patterns, and geology. A primary determinant for regional definition on the Pacific coast was temperature, driven by climate and circulation patterns. Habitats were included out to the limit of the continental shelf and/or the Exclusive Economic Zone (EEZ). Inshore, habitats were considered up to the limit of the tidal fresh zone, or upstream to salmon spawning grounds, where that was relevant.

For each of the regions, the value and practicality of a series of parameters were discussed as measures of the health of the benthic community. The discussions were limited to biological measurements only. Measures of physical and chemical parameters of the habitat were excluded. As discussions proceeded, it became clear that the types of proposed measurement categories and metrics were basically similar for all estuarine environments, regardless of region, but this was not always the case for near shore or deep ocean habitats. The categories of parameters listed below applied to virtually all habitat types. Additional selected habitat or regional specifics were identified.

- Infauna- community structure, composition, number of organisms and biomass by taxa;
- Shellfish, epibenthic fish, benthic foraging fish- community structure, composition, number of organisms and biomass by taxa;
- Percent spatial extent of 3-dimensional refugia- SAV, mangrove, reef etc.;
- Percent spatial extent of living refugia vs. total refugia;
- Dominance by selected species (opportunistic vs equilibrium);
- Changes in dominance; and
- Biomass of fish food.

In addition, parameters which apply specifically to estuaries included measures of resident vs. migratory species, and functional parameters of selected species (e.g., filtering capacity). Parameters which reflected contaminant impact such as body burdens, incidence of disease or the dominance of pollution tolerant species, were considered to be useful on a site-specific basis and are applicable to all habitats. Contamination was not considered to be an issue in the offshore habitat of the southeast, unlike some other regions. In shoreline and offshore habitats, the benthic epifaunal community was considered more appropriate than the infaunal community, depending upon the bottom type. The age structure of selected species, particularly in live bottom areas, was included as a measure of physical disturbance and/or chemical impact. In some locations, inclusion of shorebirds in the community metrics may be appropriate.

Eight general contrasts between degraded and healthy biological communities were enumerated. These were considered to be the functional ecological consequences of habitat degradation, and would be quantified by measurement of the specific parameters reviewed above. These are applicable to any habitat type including benthic, water column and vegetated habitats although some are targeted toward contaminant impacts and may not apply to all site specific locations. These contrasts were considered to represent the extremes of a continuum between healthy and degraded habitats. It is important to recognize that this continuum may not be linear, and may contain threshold points at which a small change in habitat integrity results in a large change in signal.

Degraded	Healthy
low diversity	high diversity
high dominance by selected species	low dominance
high proportion of immature individuals	stable age structure
high proportion of tolerant species	low proportion of tolerant species
high proportion of r selection species	high proportion of K selection species
high chemical body burdens	low chemical body burdens
high disease/lesion incidence	low disease/lesion incidence
low coverage by biological refugia*	high coverage by biological refugia*

*emergent or submerged vegetation, coral reef, live bottom, oyster bar

5.3 WATER COLUMN HABITATS

This session included the open water column habitats of freshwater streams, estuaries, near shore and coastal waters. The session opened with a discussion on the use of IBIs for the water column. Should an IBI be developed for the EFH of managed species under the Magnuson-Stevens Fishery Conservation and Management Act? Should separate IBIs be developed for each principle habitat type -- fresh, estuarine, coastal? The objective is to protect ecological units required to support a sustainable fishery and the managed species' contribution to a healthy ecosystem. Are the right components present? Are the organisms which are present healthy? Is the habitat capable of supporting transient organisms? The discussion progressed to what types of organisms might provide information for developing water column IBIs and the potential value of water column indicators. The group considered general classes of organisms, plankton and pelagic nekton.

The group proceeded with enumeration of potential water column IBI metric measurements that might be involved. These should include metrics which are indicative of community or population impact(s) from chemical, physical, and biological parameters. Examples of chemical attributes included contaminants, excess nutrients, and paralytic shellfish poisons adversely affecting the habitat or species of concern. Examples of physical impacts included low dissolved oxygen, stressful temperatures or salinities, inadequate light penetration, and altered stream flow or tidal circulation. Biological attributes might include the presence of pathogens, exotic species, or species compositions indicating degraded conditions (e.g., harmful algal blooms). Any taxa and/or trophic levels could be included as a biologically mediated stressor.

The group returned to discussing the approach to assessment of habitat quality with the use of IBI sampling. The first step might be to develop a rapid, relatively inexpensive screening IBI to determine whether or not a problem exists with a particular area or region as a diagnostic tool. After potential problem areas have been identified, a follow-up assessment protocol, which would be more site specific, could be developed, based upon investigation of potential stressors causing the habitat degradation. This might be accomplished by examining extant and historical data (e.g., NOAA's Eutrophication Survey, Status and Trends data, EMAP data, state agency information, land use practices, etc.). After potential factors have been identified, a selection of the most important variables would be possible, and an IBI should be developed to address them. Ultimately, it would be desirable to focus on IBI metrics that would address habitat quality of populations managed under Fishery Management Plans. Also, IBI metrics that would be indicative of, or sensitive to, impacts on habitat conditions before they affect managed populations to a significant degree is the objective. Thus, sensitive species, growth indicators, disease condition factors and abnormalities may be desirable to include. Some of these characteristics may only be exhibited in adults. It would be desirable to develop some characteristics specific to juveniles as well. Other water quality metrics were discussed, including the presence of fecal coliforms and parasites, biological oxygen demand, and phytoplankton phaeophytin/chlorophyll ratios might be a useful combination of chemical-biological indicators.

For plankton, the group considered both phytoplankton and zooplankton. For example, as waters degrade (i.e., move toward eutrophic conditions) phytoplankton composition can change from larger cell sized diatoms easily grazed by fish to smaller celled, less nutritious green algae often difficult or impossible for many species of fish to feed upon. Presence of brown tides (i.e., algal blooms) or toxic blooms of algae (e.g., *Pfiesteria*) would represent a degraded water column habitat. However, the group was concerned that water column IBIs might not provide the best bang for the buck. The water column may be too variable, too dynamic, or too transient in quality. Concern was expressed over seeing a signal within appropriate temporal and spatial scales. The group concluded that a holistic approach to IBIs should be developed and that the inclusion of more than just the water column might be desirable (i.e., water column and benthic for coastal, estuarine, and fresh waters).

5.4 SYNTHESIS

In habitats that have a high degree of structural complexity, such as reefs and vegetated areas, all components of the various communities should be assessed in a single index. In some cases, this may include the physical characteristics of the biota (e.g., canopy density, areal extent of live bottom). This makes for a more complex index because it will involve both plants and animals, and benthic and free swimming animals. However, the degree of ecological integration between them in these types of habitats is more intimate than in other habitat types. The functional connection between benthic habitats and the overlying water column habitat in open water and/or offshore habitats is less direct, except for bottom feeding fish which may come and go. In these types of habitats, the benthic and pelagic communities should be assessed independently. Assessment parameters in soft bottom habitats in estuaries were consistent regardless of location. Regional differences in assessment metrics and/or the appropriate community to sample were identified in coastal and hard bottom environments. Strictly pelagic communities will be difficult to assess with the IBI approach, due to high spatial and temporal variability. However, planktonic communities are more easily sampled in a quantitative fashion than nekton. Plankton also includes many trophic guilds (algae, zooplankton, larval stages of larger organisms) which respond to differing types of stressors. The process and activities required to develop bioindicators will be separate from the process of application of bioindicator measurement for monitoring and assessment purposes.

Table 5.1 summarizes the recommendations of potential metrics from the three breakout groups. Many potential metrics are common to all or multiple habitat types. While there is no need for commonality of metrics between different habitats, it is logical that there would be similar types of measures of ecological condition regardless of habitat. There was considerable overlap in the metrics in the diversity, abundance and condition categories. It is instructive that there is very little overlap in functional metrics. Functional roles of a species in a habitat is much more site specific than other parameters. In the tables, tolerant and intolerant refer to pollution indicative and sensitive, respectively.

Tables 5.2 - 5.4 summarize metrics currently in use in the development programs for benthos and fish discussed in the technical presentations of this workshop. These correspond to benthic estuarine, and water column estuarine or vegetated habitats in Table 5.1. The tables also contain metrics from additional programs which were not represented at the workshop, but are summarized here in collaboration with the investigators. Overall, the metrics which have been found to work in the field, do not cover as wide a range of metrics as presented in the recommendation table. Clearly, our knowledge of marine biological communities on the ground does not match our expectations based on ecological theory. Either the theoretical paradigms need to be tested or refined in the marine environment, or sampling techniques need to be refined. On the other hand, except for species-specific metrics of abundance and/or function, which are largely site-specific, virtually all the metrics which are in use are addressed in Table 5.1. The range and specificity of metrics utilized in fish IBI projects are greater than those used in benthic invertebrate projects. This probably reflects a greater relationships and simpler community structure of fish assemblages relative to benthic invertebrate communities. Several of the metrics could be placed in more than one category as they could represent more than one attribute (e.g., %# tubificid oligochaetes reflects abundance and function).

Table 5.1 Summary of recommendations of potential metrics for marine IBI development.

CATEGORY	VEGETATED: PLANTS	VEGETATED: ANIMALS	WATER COLUMN: ESTUARINE	WATER COLUMN: COASTAL	BENTHIC: ESTUARINE	BENTHIC: COASTAL/LIVE
DIVERSITY	diversity (marsh, macroalgae)	diversity (infauna, epifauna, fish)	diversity (phytoplankton, zooplankton, fish)	temporal diversity (phytoplankton zooplankton)	diversity (infauna, epifauna- incl. fish)	diversity (infauna - soft bottom, epifauna- incl. fish)
	# species (marsh only)	# species	# species	# species	# species	#species
	dominance (marsh only)	dominance	dominance	dominance	dominance	dominance
ABUNDANCE	shoot density (SAV)	abundance (#)	μ plankton abundance (temporal)	μ plankton abundance (temporal)	abundance	abundance
	biomass	biomass/taxa	μ biomass	μ biomass	biomass/taxa (≈age structure)	biomass/taxa (≈age structure)
	chlorophyll		phytoplankton chlorophyll	phytoplankton chlorophyll		
	patchiness (SAV)					
		large bivalves				
			fish metrics			
FUNCTION			phytoplankton cell size	phytoplankton cell size		
			zooplankton size	zooplankton size		
			%zooplankton by trophic guild	%zooplankton by trophic guild		
			% larvae by trophic guild	% larvae by trophic guild		

Table 5.1 (Cont.)

CATEGORY	VEGETATED		WATER COLUMN		BENTHIC	
FUNCTION (cont.)			% diatoms, green, blue green	% diatoms, green, blue green		
					% 3-D refugia	% 3-D refugia
					fish food biomass	fish food biomass
		# shellfish species				
		# benthic species				
		# nursery species				
		# resident species				
		# spawning species				
	blade area					
	epiphyte biomass					
CONDITION	disease	disease	disease	disease	disease	disease
	indicator species (r/K, tolerant, opport/equilib, exotic)	indicator species (r/K, tolerant, opport/equilib, exotic)	indicator species (r/K, tolerant, opport/equilib, exotic)	indicator species (r/K, tolerant, opport/equilib, exotic)	indicator species (r/K, tolerant, opport/equilib, exotic)	indicator species (r/K, tolerant, opport/equilib, exotic)
	age structure				age structure of live refugia	age structure of live refugia
	tissue burdens				tissue burdens	tissue burdens
			freq, duration, timing, extent of blooms	freq, duration, timing, extent of blooms		
			freq, duration, timing, extent of toxic blooms	freq, duration, timing, extent of toxic blooms		

Table 5.1 (Cont.)

CATEGORY	VEGETATED		WATER COLUMN		BENTHIC	
			coliform count	coliform count		
			freq, duration, timing, extent of fish kills	freq, duration, timing, extent of fish kills		
					% live vs dead refugia area	% live vs dead refugia area
CONDITION	rhizome density					
(cont.)	# flowering					
	runners vs tuft rhizomes					

Table 5.2 Metrics utilized in current benthic estuarine and coastal IBI development projects.

CATEGORY	GoM Estuary	S.E. Atl. Estuary	Chesapeake Bay	S. Calif. Coastal	EMAP Virginian Prov. †	NY/NJ Harbors ‡
DIVERSITY	Shannon-Wiener		Shannon-Wiener		Gleason's D*	
		# species				# species
		%dominance (1/# top 2 species)				
ABUNDANCE		abundance (#)	abundance (#)			abundance (#)
			biomass			biomass
			%taxa >5cm			
	%# tubificid oligochaetes				# tubificid oligochaetes*	
					# spinoids*	
	%# bivalves					
FUNCTION			% biomass >5cm deep*			
			%carnivores & omnivores*			% # carnivores & omnivores
			% deep deposit feeders*			
CONDITION		% sensitive taxa	% tolerant taxa			% tolerant taxa
			% intolerant taxa*			% intolerant taxa
				cumulative taxa-specific tolerance		

*Salinity specific/normalized

† Strobel et.al. 1995. Statistical Summary: EMAP-Estuaries Virginian Province, 1990-1993. EPA/620/R-94/026.

‡Ranasinghe et.al. (in review). A benthic index of biotic integrity for the New York/New Jersey Harbor. J. Aq. Ecosystem Stress and Recovery.

Table 5.3 Metrics utilized in current vegetated habitat IBI development projects.

CATEGORY	New England† SAV	Chesapeake B.† SAV
DIVERSITY	# species	# species
	dominance (# species =90%)	
ABUNDANCE	# individuals (or biomass)	# individuals (or biomass)
	# estuarine spawning sp.	# estuarine spawning sp.
FUNCTION	# resident sp.	# resident sp.
	% benthic (# or biomass)	% benthic (# or biomass)
	# nursery sp.	# nursery sp.
		# benthic sp.
		# invertevore sp.
CONDITION	% diseased	

†Deegan et al. (in preparation),

Table 5.4 Metrics utilized in current water column habitat IBI development projects.

CATEGORY	Chesapeake B. water column	Lake Erie	Texas NRCC † (seine*,trawl+)	Chesapeake B. Plankton‡
DIVERSITY	# species	# species	# taxa*+	
	dominance (# species=90%)		dominance*	
	# trawl species			
		# sunfish species		
		#phytophilic species		
		# benthic species		
				Margalef (zooplankton)
ABUNDANCE	# individuals W/O menhaden	# individuals W/O gizzard shad	# individuals*	mesozooplankton $\#/m^2$
				microzooplankton $\#/m^2$
	# anadromous species			
				mesozooplankton biomass
				microzooplankton biomass
			% penaeid*+	
			% shad or anchovy*+	
FUNCTION	# est. spawners = residents			
	% benthivores			
	% carnivores	% top carnivores		
	%planktivores			
		% omnivores		

Table 5.4 (Cont.)

CATEGORY	Chesapeake Bay water column	Lake Erie	Texas NRCC† (seine*, trawl+)	Chesapeake Bay Plankton‡
FUNCTION (cont.)		% lake assoc. individuals		
				%microzooplankton: mesozooplankton
				%calanoid: cyclopoid + cladocerans
				% *Bosmina sp.*
CONDITION		# intolerant species		
		% tolerant individuals		
		% exotic individuals		
		% diseased		Δ%mean abundance

†Guillen, G.J. 1996. Development of a Rapid Bioassessment Method and Index of Biotic Integrity for Tidal Streams and Bayous located along the Northwest Gulf of Mexico. 1996. TNRCC Special Report. Houston, Texas.
‡ Alden et al. (in preparation) Long-Term Trends in Zooplankton Indices of Environmental Health in the Chesapeake Bay and its Tributaries, Draft Report, Ches. Bay Prog.

6.0 Workshop Consensus and Conclusions

Based on knowledge gained from preliminary studies, the IBI approach will be useful for assessing habitat quality in two primary ways. It brings together multimetric information to describe the quality of the biological community in simple, yet quantitative terms, and can be used for technical ecological assessment or to formulate research hypotheses for testing. The approach was specifically designed to assess environmental harm resulting from anthropogenic stressors. It provides a more site-specific assessment of target habitats than EMAP, which provides a probabilistic assessment over an entire region. This will be essential for application to EFH quality assessment. The IBI approach addresses a broader range of habitats and stressors than the Status and Trends or Mussel Watch approach, which are specifically geared toward contaminant exposure.

In addition to the regulatory need for site specific biological measurements, it is useful to be able to represent the condition of complex ecosystems concisely, by means of composite indices or simple graphics, so that managers and non-specialists can readily evaluate and compare information, establish goals, and set priorities for remediation or protection. This requires the use of succinct, understandable statistics that are also meaningful, reproducible and technically valid. Indicators are essential for:

1. determining management priorities;

2. measuring the effectiveness of management actions;

3. measuring progress towards restoration goals;

4. developing the capability to predict environmental consequences of management options; and

5. communicating to the general public.

Technical formulation and testing of an IBI for a specific habitat requires a logical accumulation of data on parameters specifically selected because they are considered to be symptomatic of the ecological consequences of anthropogenic degradation. Point and non-point source runoff, toxic contamination, hydrologic alteration of watersheds and overharvest all affect biological communities. However, direct, quantitative cause-and-effect relationships between specific activities and ecological consequences are difficult to assess due to the complex interactions between ecosystems and anthropogenic stressors. No single parameter such as diversity, richness, indicator taxa or abundance, has been identified which can reliably distinguish between degraded and undegraded habitats in disparate environments, or in response to different stressors. The underlying ecological principle of the IBI is that a healthy community requires a diversity of intact ecosystem functions and processes to persist. Confirmation of deleterious effects at the community level is an inherent confirmation that population level effects are occurring within that community. IBIs also provide a mechanism to support contrasts of similar habitats in different regions. Data gaps and method deficiencies will become readily apparent in the process. While the cumulative index may be a ranking based on a number of metrics, the quantitative behavior of those metrics in relation to each other, and our ability to assess them in relation to anthropogenic impacts is instructive. The detailed information from individual metrics is not lost in the process. The IBI approach provides a framework for assessing habitat quality with a consistent, technically defensible method. It has a demonstrated utility in fresh water environments as a technical assessment method and as a management tool.

One difficulty with the application of IBI techniques to complex marine systems has been the relative lack of intimate knowledge of the ecological roles and interactions of specific species and/or functional guilds, compared to fresh water systems. Therefore, a basic element of any future IBI development work is simple taxonomic and natural history documentation of the species selected for use as markers of stress. Data gaps in life histories of critical species, including the degree of natural variation, must be identified and resolved. While it is preferable that metrics can be related to known functional aspects of an ecosystem, this has not always been the case. For example, if the presence of a specific taxonomic group or trophic guild of organisms is shown to be sensitive to habitat degradation, a measure of their abundance may be used as a meaningful metric. The presence of species with demonstrated sensitivity to heavy metals contamination is one example. (Indicator species for specific stresses must be selected with great care because of the potential for differential sensitivity to different stressors, and/or in different environmental settings.) Alternatively, the presence of a species or guild which correlates with some known measure of habitat degradation, without any specific knowledge of cause and effect, has been used successfully. The important aspects are that the metric must be demonstrated to be correlated with habitat degradation, and that this relationship can be quantified. Formulation of management options is difficult unless this 'dose-response' relationship can be demonstrated. To overcome this problem, statistical methods to select metrics from an array of potential measures have been successfully utilized to correlate ecological responses with anthropogenic stressors. These methods have included discriminant analysis, correlation coefficients, cluster analysis and multiple regression techniques. A comparative assessment of these methods has not been performed.

A related problem is the definition of what constitutes a reference condition. A-priori selection of 'reference' sites based upon one set of parameters (e.g., contaminants) have not been tested for efficacy in habitats which may have been impacted by other stressors (e.g., eutrophication). Also, the delineation of a pristine reference site is problematic in regions where anthropogenic impacts are ubiquitous.

It is assumed there is a continuum of biological response between degraded and undegraded conditions, although the biological response may not be linear, and there are probably thresholds beyond which dramatic ecological changes occur. Some metrics have been shown to be bimodal. Negative and positive signals must be selected carefully. For example, diatom blooms in New York bight might be considered as indicative of healthy conditions. However, if early diatom blooms, resulting from eutrophication, cause Si limitation, this may result in a subsequent *Nanachloris* bloom. Because this is a smaller species, menhaden cannot feed on them. Thus, in this case, an early diatom bloom results in a deleterious dominance shift in the phytoplankton community.

It is not necessary to sample all subunits of an ecosystem. This would not be possible in any case, as all gear is selective to some degree. Assuming the ecosystem is integrated at some level, assessment of specific habitat types and/or locations within a system should be adequate. The level of effort for a given location will have to be determined on a site-specific basis. Locations for follow-up programs will build upon existing efforts and will further develop methods in a consistent approach. In addition, specific index periods will need to be established for specific habitats. This will vary in different parts of the country for the same habitat type but should consider a time of year when environmental conditions and the action of stressors are relatively stable. The process will also need to include consideration of sampling efficiency, expense, safety and complexity.

7.0 Follow-up Actions

Under the provisions of the Magnuson-Stevens Fishery Conservation and Management Act, NMFS must describe and map essential fish habitat. This process will also involve characterizing habitat quality. NMFS should move forward to identify appropriate attributes that would constitute biological indicators of habitat quality for the following habitat types: SAV, riparian, estuarine benthic/water column, coastal benthic, and coral reef habitats. Ongoing activities around the nation that are involved in developing and applying biological indicators should be inventoried (Figure 7.1). Estuarine fish bioindicators have been, or are being developed, in Massachusetts, Chesapeake Bay, North Carolina, Florida and Texas. Investigations on the transferability of fish community bioindicator metrics for submerged aquatic vegetation (SAV) developed for Cape Cod estuaries and tested in Chesapeake Bay, and from Chesapeake Bay pelagic habitats to coastal embayments have been instructive. The degree of modification to the metrics which was necessary to adapt the systems to different regions was relatively straight forward. Benthic indicator development projects have employed complex mathematical schemes to develop metrics, due to the more complex and less understood biological communities associated with benthic invertebrate communities. Current development projects in the New York/New Jersey harbors, the Virginian province, Chesapeake Bay, SE Atlantic, and Gulf of Mexico rely heavily on EMAP data. Chemical contamination data has been used extensively to guide definition of reference sites, and therefore metric selection. Coastal benthic efforts on the Atlantic and Pacific continental shelves have taken divergent approaches from estuarine studies due to the more diffuse nature of impacts in those habitats. However, gradients of habitat degradation can be identified and quantified.

NMFS must develop partnerships with other Federal, state, university and private research institutions that are involved, or interested in developing and applying indices of biological integrity. Maximum use of ongoing programs should be made. A coordinated IBI development effort for South Atlantic estuarine benthic habitats is already well under way between the South Carolina Department of Natural Resources, NOAA/NOS, and U.S. EPA. NMFS should participate in that effort, and consider adopting the existing system. A benthic invertebrate IBI for northwest Pacific salmon spawning streams has been proposed, based on data gathered in Washington and Oregon streams. That assessment approach should also be evaluated for utility in EFH quality assessment. In addition U.S. EPA is developing a program for a mid-Atlantic integrated assessment of water and benthic habitats in estuaries and streams, which NMFS should become involved in.

A program should be devised to initiate the necessary data collection efforts addressing both evaluation of existing data, and the design of a systematic sampling and research program to fill data gaps, and generate new data to develop fish community IBI metrics. The major steps in the process should include:

1. Prioritization of habitat types and identification of habitat delineation parameters;

2. An inventory of existing data residing in state and Federal agencies which would be appropriate for metric development in selected habitats;

3. Development of a research program to derive quantitative relationships between potential metrics and anthropogenic stress, including contaminants, eutrophication, and physical habitat alteration. The response pattern and mechanism(s) of cause and effect are particularly important if the IBIs are to have direct management utility. In addition, conduct research, where necessary, to illuminate the ecological role of potential indicator species, including their natural variation and limiting factors;

4. Identification of spatial and temporal gradients, and gradients between habitat types that will influence the distributional pattern of biota in a natural setting. This step, together with item #3 will be essential for designation of reference sites and derivation of regional vs. site-specific IBIs;

5. Initiate sampling and/or assessment of existing data sets to identify potential metrics, gear requirements, sample requirements and statistical treatment;

6. Metric evaluation in independent validation sites;

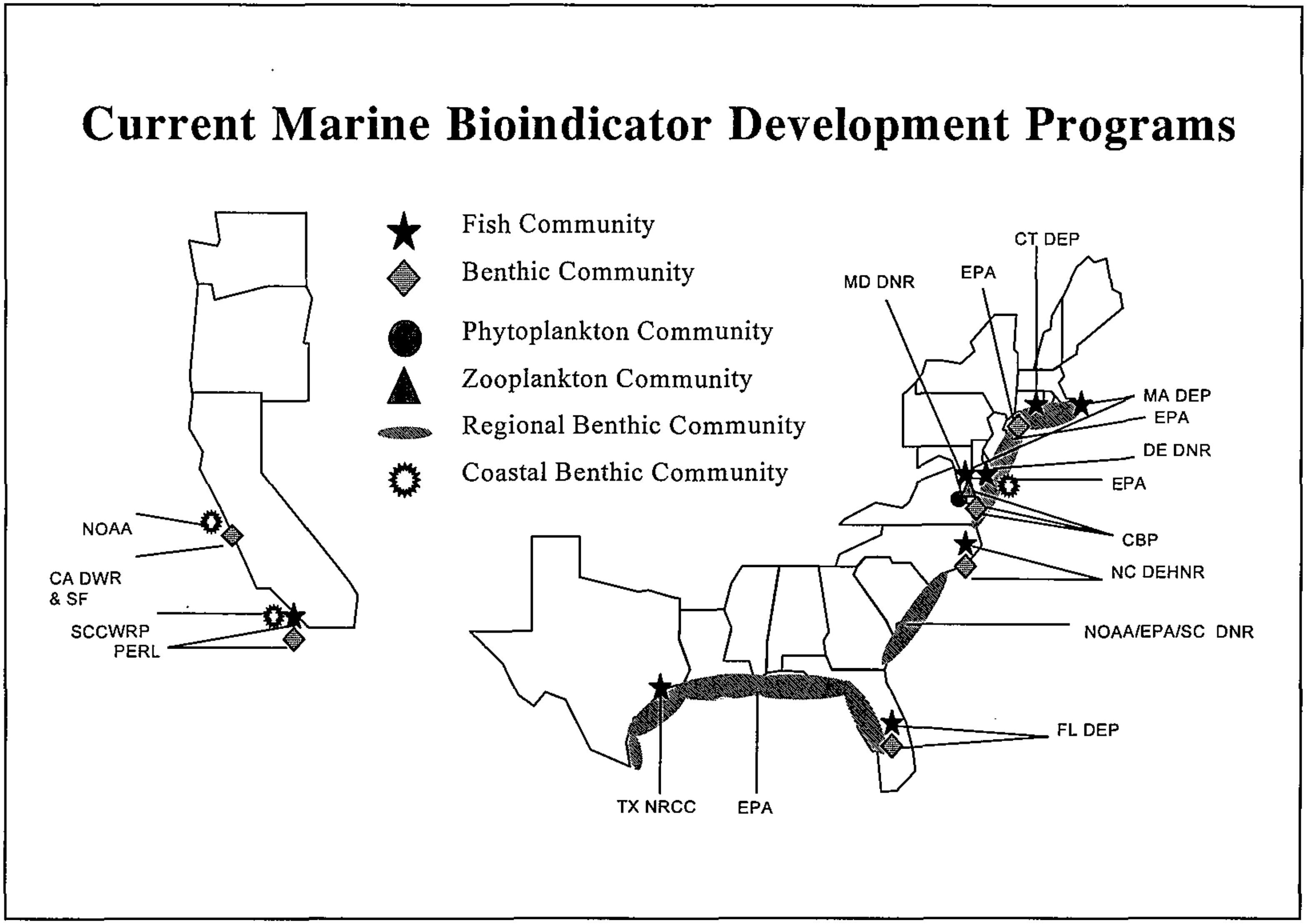


Figure 7.1 Map of locations of site-specific and regional marine IBI development programs.

Appendix 1

AGENDA
July 14-15, 1997

Day 1- Introduction and Presentations of Existing Approaches and Applications

I. Hartwell/D. Brown (NMFS) - Welcome and introduction

J. Karr (Univ. Wash) - Attaining environmental goals

C. Linder (MD DNR) - An estuarine IBI for Chesapeake Bay

M. Weaver (Woods Hole Mar. Biol. Lab.) - Estuarine biotic integrity index

K. Summers (US EPA) - An index of benthic condition

M. Bergen (S. Cal. Coastal Res. Proj.) - The benthic response index

A. Ranasinghe (VERSAR Inc.) - Chesapeake Bay benthic restoration goals

G. Gibson (US EPA) - Marine biocriteria survey techniques

J. Hyland (NOS) - A benthic index for estuaries of the S.E. US

R. Thoma (OH EPA) - Ohio biological monitoring program
- Lake Erie and lacustuary monitoring program

F. Holland (SC DNR)- Assessment of watershed development on tidal creeks

M. Monaco/P. Orlando (NOS) - Spatial framework for EFH data collection

Day 2 Assessment of Metrics and Index Derivation Methods

A) Morning breakout groups to assess matrix of regional metrics types by system

Starting point.

For a given habitat group, what are the parameters, or types of parameters, necessary to assess habitat condition? How would those parameters be measured? How would you combine the measurements to arrive at a conclusion?

Group 1 - Vegetated {SAV (vascular plants & seaweed), emergent marsh, mangrove, kelp}

Group 2 - Open water benthic {soft bottom, hard bottom, live bottom (oyster bar, coral reef, offshore benthic assemblage), deep vs shallow & intertidal}

Group 3 - Water column {tidal-fresh, estuarine, near-shore, coastal}

B) Reconvene to compare and critique metric parameters and approaches

C) Afternoon breakout groups to continue discussions about parameters based on morning sessions.

D) Final session to formulate a consensus statement on a bioindicator framework for EFH habitat quality assessment, identify research needs and, potential pilot program locations/data bases.

Appendix 2

PARTICIPANTS

Robert Boyles (803) 727-2078 boylesrh@musc.edu
S.C. Sea Grant
Univ. of South Carolina
287 Meeting St.
Charleston, SC 29401

Mary Bergen (714) 894-2222 bergenm@sccwrp.org
Southern California Coastal Water Research Project
7171 Fenwick Lane
Westminster, CA 92683

Dail Brown (301) 713-2325 dail.brown@noaa.gov
NOAA/NMFS F/HC
1315 East West Hwy
Silver Spring, Md 20910

Tracey Collier (206) 860-3312 tracey.k.collier@noaa.gov
NOAA/NMFS NWFSC
2725 Montlake Blvd. E.
Seattle, WA 98112

David Dow (508) 495-2249 david.dow@noaa.gov
NOAA/NMFS NEFSC
166 Water St.
Woods Hole, MA 02543-1026

David. W. Evans (919) 728-8752 dave.evans @noaa.gov
NOAA/NMFS SEFSC
101 Pivers Island Rd
Beaufort, NC 28516

Gary Fitzhugh (904) 234-6541 fitzhugh@bio.fsu.edu
NOAA/NMFS SEFSC
3500 Delwood Beach Rd.
Panama City, FL 32408

Michael Fulton (803) 762-8576 mike.fulton@noaa.gov
NOAA/NMFS SEFSC
219 Ft. Johnson Rd.
Charleston, SC 29412

George Gibson (410) 573-2618 gibson.george@epamail.epa.gov
US EPA
839 Bestgate Rd
Annapolis, Md 21401

Ian Hartwell (301) 713-2325 ian.hartwell@noaa.gov.
NOAA/NMFS F/HC
1315 East West Hwy
Silver Spring, Md 20910

Fred Holland (803) 762-5107 holland@mrd.dnr.state.sc.us
S.C. Dept. of Natural Resources
P.O. Box 12559
Charleston, SC 29422-2559

Jeff Hyland (803) 762-5415 jeff.hyland@noaa.gov
NOAA Carolinian Province Office
217 Ft. Johnson Rd.
P.O. Box 12559
Charleston, SC 29422-2559

James R. Karr (206) 685-4784 jrkarr@u.washington.edu
Univ. of Washington
Box 357980
Seattle, WA 98195

Pete Key (803) 762-8596 pete.key@noaa.gov
NOAA/NMFS SEFSC
219 Ft. Johnson Rd
Charleston, SC 29412

Cecilia Linder (302)645-4384 clinder@udel.edu
Univ. Delaware
700 Pilottown Rd.
Lewes, DE 19958

Mark Monaco (301) 913-3005 mark.monaco@noaa.gov
NOAA/NOS ORCA
1305 East-West Hwy
Silver Spring, Md 20910

Paul Orlando (301) 713-3000 porlando@seamail.nos.noaa.gov.
NOAA/NOS ORCA
1305 East-West Hwy
Silver Spring, MD 20910

Paul Pennington (803) 762-8620 paul.pennington@noaa.gov.
NOAA/NMFS SEFSC
219 Fort Johnson Rd.
Charleston, SC 29412

Ananda Ranasinghe (410) 740-6085 ranasinghana@versar.com
Versar Inc.
9200 Rumsey Rd.
Columbia, MD 21045

Casimere Rice (425) 743-3307 crice@sci.nwfsc.noaa.gov
NOAA/NMFS NWFSC
10 Park Ave.
Mukilteo, WA 98275

Robert Stickney (409) 845-3854 stickney@unix.tamu.edu
Texas Sea Grant
Texas A&M Univ.
1716 Briarcrest Dr., Suite 702
Bryan, TX 77802

Kevin Summers (904) 934-9244 summers.kevin@epamail.epa.gov
US EPA, Gulf Ecology Division,
Sabine Island Dr.,
Gulf Breeze, FL 32561

Roger F. Thoma (216) 963-1141 roger.thoma@epa.state.oh.us
Ohio EPA
2110 E. Auroa Rd
Twinsburg, OH 44087

Jim Thomas (301) 713-2325 james.thomas@noaa.gov
NOAA/NMFS F/HC
1315 East-West Hwy.
Silver Spring, Md 20910

Robert VanDolah (803)762-5048 vandolah@mrd.dnr.state.sc.us
S.C. Dept. of Natural Resources
Marine Resources Division
P.O. Box 12559
Charleston, SC 29422

Melissa Weaver (423) 974-3065 weaver@tiem.utk.edu
Dept. of Ecology & Evolutionary Biol.
Univ. of Tennesee. 569
Dabney Hall
Knoxville, TN 37996-1610

☆ U.S. GOVERNMENT PRINTING OFFICE: 1998—693-362 / 62088 Region No. 10