I. INTRODUCTION

The purpose of this workshop was to assess the role of **biocriteria*** and information generated by **ambient biological sampling** in the variety of surface water programs. The workshop was convened in response to (1) the 1987 Office of Water report, Surface Water Monitoring: A Framework For Change, calling for increased ambient biomonitoring and an improved framework for monitoring, assessment and reporting, and (2) the sections of the Water Quality Act (WQA) of 1987, which amended the Clean Water Act (CWA), referring directly to biological assessments, requiring:

- o restoration and maintenance of **biological integrity**, Sec. 101;
- o U.S. EPA to develop criteria based on biological assessment methods when numerical criteria for toxicity are not established, Sec. 303(c)(2)(B); and
- o guidance and criteria based on biological monitoring and assessment methods, Sec. 304(a)(8).

There are several other sections of the CWA whose implementation would be improved by biocriteria and **biosurvey** data, requiring:

- o development of improved measures of the effects of pollutants on biological integrity, Sec. 105;
- o guidelines for evaluating Nonpoint Sources (NPS), Sec. 304(f);
- o lists of waters unable to support balanced biological communities, Sec 304(l);
- o biennial reports of the extent to which waters support balanced aquatic communities, Sec. 305(b);
- o assessments of lake trophic states and trends, Sec. 314;
- o lists of waters that cannot attain designated uses without additional NPS controls, Sec. 319; and
- o prohibitions against dredge and fill disposal adversely affecting balanced wetland communities, Sec. 404.

In addition, there are several U.S. EPA program activities that address biological assessments; including, the 305(b) reporting guidance, the Surface Water Monitoring Strategy, the Water Quality-based Program Policy (49 FR 9016), the Rapid Bioassessment Protocol, and the Water Quality Standards Framework being developed by the Office of Water Regulations and Standards.

This report summarizes the recommendations of the workshop to illustrate to regulatory agencies that biosurveys are an important monitoring and evaluation tool, and that biocriteria can provide, in a quantifiable regulatory context, a measure of the interim goals of Section 101(a)(2) of the CWA.

The workshop participants represented 18 States, nine U.S. EPA Regional offices, three laboratories and three headquarters Offices and Divisions, as well as other organizations (TVA, USGS, Environment Canada, etc.) and the academic community. This illustrates the regulatory

---

*Boldface terms "defined" in the Terminology Section p. iv.

community's interest in ambient biological assessments and the twin needs for increased use of biocriteria to protect the Nation's waters and for guidance and support in developing these tools. A list of workshop participants is in Appendix C. The workshop consisted of two parts. On the first day, two university researchers related the current state of aquatic ecological knowledge to biocriteria and biosurvey issues, two speakers from EPA-Office of Water (OW) discussed conceptual and practical issues facing regulators, a speaker from EPA-Office of Research and Development (ORD) presented ecological regions as a framework for stratifying variability to solve regulatory problems, and, most importantly, representatives from ten States presented their development, methods, and uses of biocriteria (See Appendix B for a list of speakers).

For the remainder of the workshop, attendees participated in one of five discussion groups. These groups were responsible for defining issues of concern, areas of consensus, and areas requiring further effort, including research questions as well as the guidance and support needed to use ambient biological sampling and biocriteria. The discussion topics were: water quality-based effluent limitation (permits), water quality standards, sampling and data evaluation for fish and for benthic macroinvertebrates, and use of ecoregions.

An Overview of Issues

The WQA of 1987 seeks to restore and maintain the chemical, physical and biological integrity of the Nation's waters. The U.S. EPA and the States have developed a variety of standards to achieve these goals, by specifying beneficial uses for waterbodies and criteria to evaluate whether the uses are being attained (and thus, whether the standards are being met). The traditional approach to criteria has been to develop chemical-specific limits deemed to be protective of the designated uses. These limits were usually established by performing single chemical toxicity tests on test organisms. More recently whole effluent testing has been added to the array of assessment tools.

This approach to water quality has produced significant improvement in the Nation's waters. These laboratory derived criteria are particularly appropriate for designated uses related to human activities, e.g., drinking water, irrigation water, human contact (swimming, water skiing), etc. Chemical-specific criteria and subsequent water quality sampling are essential to establish effluent limitations in the National Pollution Discharge Elimination System (NPDES) permits program. The chemical-specific criteria approach also has been applied to protecting designated aquatic life uses. However, an increasing body of evidence indicates that regulatory agencies cannot rely solely on these criteria without also sampling the ambient biological communities to verify the results of pollution control measures.

The existing criteria are insufficient to protect aquatic life uses for two main reasons. Primarily, they are laboratory-based simulations that cannot address all factors affecting resident aquatic communities (e.g. habitat limitations, additive impacts of multiple dischargers, etc.). Thus they are only surrogates for actually evaluating for achievement of the desired results. There are several related limitations to single chemical toxicity tests: they have been completed on a minority of the suspected toxicants; they cannot assess bioaccumulation and indirect (food chain, competition, etc.) effects; the test organisms often are not appropriate to the waters in question; and chemical toxicity changes as receiving water chemistry changes. Whole effluent toxicity testing compensates for some of these deficiencies, but must be evaluated on a site-specific basis and is thus relatively expensive. Chemical concentrations and toxicity must be monitored continuously to effectively track changes in processes, spills, and flows; this is an expensive (and rare) undertaking for regulators and dischargers.

In addition, chemical-specific criteria cannot always address potential impacts from habitat modifications and many "nontoxic" pollutants, particularly those from nonpoint sources (e.g. sediments and nutrients) which affect a vast majority of river miles. These issues are discussed

in more detail in the 1987 Office of Water - Office of Processes, Planning and Evaluation report, Surface Water Monitoring: A Framework For Change.

In the past, the U.S. EPA and States have generally been discouraged by actual and perceived problems in sampling and evaluating the health of the ambient "aquatic life" and using this information in regulation. Natural communities were thought to be too variable and complex to be precisely and consistently measured and therefore could not be used as the basis for standards and criteria in the way "hard" data such as chemical concentrations were used. Indices (such as species diversity) intended to reduce this complexity proved unreliable. There was considerable debate about sampling methods (what kinds, how extensive/intensive, and where to sample) and which organisms to sample. Finally, the costs of sampling and evaluating ambient biological conditions were considered too high for the level of information obtained.

The past ten years of aquatic ecosystem research and assessments support a change in direction. There is an increased awareness of the limitations of the chemical-specific, toxicological approach, as outlined above. At the same time the tools for assessing ambient biological conditions (sampling protocols, sampling gear, ecoregions, new indices, analytical methods, etc.) have been markedly improved, tested and refined, thus addressing many of the issues that limited the establishment of biocriteria and use of biosurveys. Most importantly, advances have been made in the development of more tractable definitions of biological integrity, and the analytical tools to assess the integrity. These developments have allowed for standardization of sampling and analysis, thus making workable biocriteria possible.

One result of these improvements is that expensive, intensive biological surveys of all organisms over long periods are not necessary for many regulatory purposes. Research in field methods has shown that reliable data, useful for monitoring, can be collected without excessive expense, usually a few hundred dollars annually per site. This makes biosurvey data especially valuable as a screening tool, while reserving more expensive investigatory techniques for more complex situations, e.g., where there is a high probability of litigation.

Biosurveys provide a valuable set of tools to more directly assess attainment of the objectives of Section 101 of the 1987 WQA; to restore and maintain the biological integrity of the Nation's waters. Biological communities integrate impact effects over a variety of spatial and temporal scales. Water chemistry sampling usually provides only instantaneous "snapshots" of conditions. Limited biological sampling often suffices to demonstrate ongoing impacts (e.g., "midnight dumpers", chronic sublethal toxic levels, etc.) that may be missed even by monthly chemical samples. Meanwhile, equipment costs and increasingly complex chemical analyses are making chemical-specific regulation relatively more expensive.

On the other hand, biosurvey data indicate effects; they also may indicate the category of pollution causing an effect on the biological community (e.g., enrichment, siltation, toxicity). As such, biosurvey data are generally used in a "reactive" manner, e.g., to detect where current controls are inadequate. Chemical-specific criteria are a "proactive" tool, e.g., to set limits for discharge permits. These proactive and reactive functions form an important balance in the regulatory process. Biosurvey data and biocriteria are not meant to replace chemical water quality assessments and chemical-specific criteria, but should be integrated into the surface water programs to more fully evaluate and protect water quality.

II. BIOCRITERIA IN STATE PROGRAMS

Several States have been acutely aware that a reliance solely on chemical criteria and toxicity tests may not be fully protective of aquatic life as mandated in the CWA and State legislation. These States have begun to develop biocriteria to directly address the protection of aquatic life. At the workshop, ten States presented their development, methods, and uses of biocriteria. The States have used four general approaches to develop biological criteria: (1) State law explicitly defines an appropriate biological use classification system, or a particular biological measure or set of measures (Maine, Vermont, Florida and Ohio have such legislative mandates). (2) The State regulatory agency develops detailed lists of characteristic and key organisms expected in healthy waterbodies for each ecological region (Arkansas uses this approach). (3) The State regulatory agency develops quantitative numerical criteria based on several biological indices and sets regional expectations for minimally-impacted waterbodies (e.g., Ohio's numerical biocriteria). (4) The State regulatory agency uses biosurvey data as decision-support information. Many States currently use this less structured approach.

Experiences in Common Among States Developing Biocriteria

While no two States use exactly the same set of ambient biological assessment tools, several common themes were evident from the presentations. This section discusses these themes and presents an overview of the States' development of biocriteria, the kinds of problems being addressed, and the sampling methods.

- States in different parts of the country face a variety of water quality problems and natural conditions. Despite this variety of regulatory issues, several States have been able to successfully develop appropriate biocriteria tools to address the problems of protecting aquatic life.

- There is considerable potential for applying similar biosurvey methods among neighboring states that share similar ecological regions. For example, the southeast Atlantic Seaboard States (except Florida) all share portions of four distinct ecoregions and should be able to conserve fiscal resources to develop the needed biosurvey tools.

- Within each State there exists extensive knowledge about the biological condition of the waterbodies among State regulatory personnel and academic researchers. These experiences and databases from biological surveys can provide the initial framework for establishing biocriteria. That is, no State needs to start from a "no information" position.

- In the ten States presenting talks at the workshop, professional aquatic biologists have been active participants in the process of establishing water quality standards and biocriteria and in monitoring the environmental results of pollution and pollution controls. This is an important point, because development of meaningful biocriteria and biosurvey methods requires knowledge of both aquatic biological communities and the regulatory process.

- All ten States use community-based biological data (as opposed to indicator species) from ambient biosurveys. These data range from a fairly simple diversity index of macroinvertebrate communities, to a combination of several more complex measures of macroinvertebrate and fish community health. All ten States have methodologies to define benchmarks of attainability to which they compare this community-based data. They have established levels of deviation from attainable conditions that, when exceeded, warrant remedial action.

- Biosurvey data have been used to establish and define aquatic life uses, to plan programs and as a corrective feedback device in monitoring activities. Biocriteria have been integrated into existing State programs and have not replaced ongoing chemical monitoring programs.

State Applications of Biosurvey Data and Biocriteria

From a practical standpoint, there are three interrelated issues that bear directly on performing ambient biological sampling and establishing biocriteria. They are, in broad terms, defining health or integrity of aquatic communities, developing measurements that address the definition(s), and applying the information from biosurveys to the legalities of monitoring, standards and permits. A healthy aquatic community is usually defined as one that possesses a diversity, an abundance of organisms, a trophic structure, and a species composition closely matching those found in relatively unimpacted or natural conditions. This definition implies a spatial and ecological framework for comparing sites. Measuring the health of a system has two components, sampling (data collection) and analysis of data. States must also decide how to apply this information in their water quality programs. States implement their findings by determining which waters will be subject to which measures, which impacts can be assessed biologically, and how to establish reasonable and adequate standards. A summary of the States' approaches to these three issues follows.

Data Collection -- Data collection represents a compromise between the resource expended and the information gained. This compromise is one of the factors that has led to a fairly wide variety of State biosurvey programs. However, all ten States presenting talks at the workshop have developed standard field methods and quality assurance documents to insure consistent and defensible results.

All ten States sample macroinvertebrates in their biosurveys. There was no clear preference between artificial substrate samplers or sampling natural substrates. The major expense of macroinvertebrate sampling is in the identification of specimens. This has led to varying levels of resolution among State programs (e.g., field identification to the family level as in North Carolina, subsampling and identifying 100 organisms per sample as in New York's rapid bioassessment protocol, and identification to the species level and counting all organisms in a large subsample as with, Ohio and Vermont). The suitability of these levels depends on the impact being examined, the level of certainty needed, the resolution needed to detect changes, the need for rapid turnaround and the probability of litigation.

Five of these States also sample fish. Fish sampling requires a larger field effort than for macroinvertebrates, but the specimens are typically identified to the species level and counted in the field. Thus, the data are quickly available for analysis. In addition, fish have a greater public appeal. Four of the States use electrofishing techniques for most of their collecting. Arkansas uses a combination of electrofishers and rotenone, depending on the waterbody being sampled.

The States also reported collecting physical habitat data. The physical condition of the waterbody greatly affects the health of the biological communities and the ability of a waterbody to attain a given use. A separate assessment of the habitat helps to determine whether any measured impairment to the biota results from water quality or habitat limitations and whether they are controllable. These data begin to address the issues of nonchemical and nonpoint degradation. They also help refine the comparability of sites. There is a wide range of uses of this information among States.

Reference Sites -- To assess the level of impact at a site it is necessary to refer to one or more similar sites that are relatively unimpaired. These are often upstream from a point or nonpoint source of pollution. All States use this approach for most of their assessments. Many States assess

a point source impact by surveying several upstream and downstream recovery sites to better evaluate ambient variability.

When conditions upstream are also impacted, the States typically choose reference sites in nearby streams (or lakes). Four States are beginning to use ecoregions to define attainable conditions. This allows assessment without paired samples. Ohio developed explicit ecoregional standards for all stream classes, for fish and macroinvertebrates. Arkansas uses ecoregional species signatures (profiles) of "characteristic" and "key" fish to define the expected communities in unimpacted streams. Nebraska and North Carolina are in the process of developing regional criteria.

Interpretation of Field Data -- For regulatory purposes there are two parts to the analysis of field data: defining quantifiable features of the biological community that describe its health, and comparing those results to some standard usually derived from unimpacted or minimally impacted communities. The simplest measures of biotic health are species richness and diversity. Florida defines an impact as reducing the macroinvertebrate species diversity index to less than 75% of established background levels. Because opportunistic species can colonize a site and inflate the index, Florida's state biologists have proposed removing these species from their analysis. Other States calculate species diversity among their analyses, but only Florida's rules have specific reference to that index.

Another feature of a healthy biotic community is the presence and abundance of species intolerant of degraded conditions. A growing body of research continues to improve our knowledge of species' tolerances. Traditionally, three Orders of insects (Stoneflies, Mayflies, and Caddisflies) have been considered intolerant of a variety of pollution impacts. This relatively coarse level of resolution (initially developed by North Carolina) is one of the metrics used by New York, Ohio, North Carolina and others. Maine, Ohio, Vermont and others identify individual species tolerances to environmental disturbances. Florida uses Beck's Biotic Index of macroinvertebrate tolerances to pollution and New Mexico uses Winget and Mangum's Biotic Condition Index that was developed for the Rockies. Fish species tolerances are also used by some States.

There is growing interest in and use of more complex assessments of biotic health. These evaluations use multiple features of community structure e.g., diversity, abundance, trophic structure, and tolerance to degradation. The advantage of this type of evaluation is that it provides a more complete, ecologically sound assessment of the health of the aquatic community. One approach examines each component of biological integrity serially through a decision matrix. Maine and Vermont use this approach with macroinvertebrate community data.

The other approach sums the values of individual metrics into one index value. For fish, the Index of Biotic Integrity (IBI) is gaining wider acceptance. Currently, the IBI is most applicable to small streams and rivers, but research is continuing to expand it for use in major rivers, lakes and estuaries. Ohio and Illinois are using the IBI and Vermont is modifying it for use in its streams. Several similar indices are being developed for macroinvertebrate communities. Ohio has developed an Invertebrate Community Index (ICI) modeled closely on the IBI.

Application of biosurvey data to criteria -- There is currently a wide range in the uses of ambient biological data by the States. New Mexico, Nebraska and North Carolina do not have explicit biological language in their standards but are considering its inclusion. New York, Illinois and Ohio have proposed explicit narrative and numerical criteria. Vermont, Maine, Florida and Arkansas have biocriteria in place. Vermont and New Mexico use biosurvey data primarily to measure the level of impact of recreational development on mountain streams. Florida uses ambient biological data to assess impacts from point sources and dredge/fill operations. Maine has developed a detailed decision matrix to assess whether a waterbody is attaining its designated uses. Ohio's proposed criteria use ecoregional values for several indices of biotic health to assess point and nonpoint source, and habitat impacts and as an arbiter of aquatic life use attainment.

Thus, despite diverse natural settings and legislative mandates, at least ten States have taken a community oriented approach to analyzing biosurvey data and to applying the results to regulatory decision making and planning.

III. WORK GROUP SUMMARIES AND RECOMMENDATIONS

To facilitate effective consideration of the diverse technical, and program/policy related issues specific discussion groups were established to address each major topic. This section summarizes the discussions of the five workgroups. Some recommendations were presented by more than one group. These repetitions emphasize the importance of those issues, and concurrence among groups working independently, while showing their relationship to each topic. The recommendations of the two sampling and data evaluation workgroups (for fish and for benthic macroinvertebrates) are combined to present an overview of biological sampling and data issues related to the surface water programs.

WATER QUALITY-BASED EFFLUENT LIMITATION (PERMITS) WORKGROUP

Technical Support Document for Biocriteria and Biosurvey

Standardized, documented methods and procedures are necessary to integrate biocriteria and biosurvey information into the water programs. Further, development and documentation of these procedures are particularly critical to their integration into the National Pollutant Discharge Elimination System (NPDES) permits program.

To promote and guide development of these procedures, it is recommended that U.S. EPA produce a Technical Support Document that describes the use of ambient biosurveys and development of biocriteria. This document should draw upon the experiences and knowledge of the existing State programs while encouraging States to tailor biosurvey programs to their specific needs. The document should present guidance on a broad range of subjects to include:

- The necessity and description of standard operating procedures. Carefully conceived and documented field and laboratory procedures are essential for effective evaluation of impacts and for enforcement of standards;
- Quality assurance/quality control plans. These are an integral part of standardized operating procedures;
- The use of specific indices, i.e., definitions of each index, what each measures, the range resolution each provides, and appropriate uses for each category of regulatory concern;
- The ecoregion concept and regionalization of criteria, both biological and chemical;
- Consideration of the appropriate uses (and limitations) of Rapid Bioassessment Protocols for fish and macroinvertebrate community evaluation.

A Biocriteria "Policy" Document

The Region V draft statement "Regarding the Use of Ambient Biosurvey Data in Implementing the Objectives, Goals and Policies of the Clean Water Act" served as the focal point for the workgroup's discussions regarding integration of biocriteria and biosurvey data into the regulatory permits program. The Region V statement was subsequently redrafted, based on recommendations contained in this report. The concepts contained in the statement are broadly applicable and this

workgroup recommends the statement (Appendix A) as a guide for the Office of Water and other Regions and States. The following are specific recommendations from this workgroup.
Most importantly, biosurvey information should be used, where applicable in all surface water programs, to include broad use in monitoring, planning, water quality standards, permitting, hazardous waste, wetlands and other programs.

To facilitate integration of biosurvey data specifically into the regulatory permits program:

- The U.S. EPA should actively encourage the States to use ambient biological sampling as an additional monitoring tool for toxicity screening and environmental problem discovery.

- In general, site-specific biosurvey data should be considered the optimum assessment of attainment of designated aquatic life uses. However, if chemical-specific, toxicity, and biosurvey methods yield apparently contradictory indications, none of the three types of evaluation should be assumed, a priori, to be superior to the others. Rather, the quality of the data and analysis used in each of the three approaches should determine the appropriate course of action.

- An integrated approach should be taken in the development of NPDES permit limitations, using a weight of evidence evaluation of bioassay, biosurvey, and chemical-specific information at a level of complexity dictated by site-specific concerns. The necessity of any particular piece of information should be evaluated on a case-by-case basis.

- Although many population effects are revealed, it should be recognized that biosurvey information may not address potential wildlife or human health concerns, allocations necessary to prevent the cumulative impacts of long-term low level discharges to lakes and wetlands, or potential accumulation of pollutants to deleterious levels in sediment or tissue.

- Discussion of the use of biosurvey data in the water programs should be included in an update of the Regional Water Monitoring Strategy.

- Personnel performing ambient biological sampling should be actively involved in formulating and approving water quality based permitting requirements in State and Regional programs.

WATER QUALITY STANDARDS WORKGROUP

The discussion focused on two main topics. First, does the water quality standards program need criteria that are based on assessments of the resident biological community? The States have limited resources and are justifiably reluctant to develop new types of criteria and assessment procedures, unless there is a clear and convincing need for them. Second, how can those States that are interested in biocriteria incorporate them into their standards and what needs to be done next, by both EPA and the States? The experiences of the States that have developed biocriteria were discussed. A summary of the uses of biosurvey data and biocriteria in several States is presented in Section II

Biosurvey Data and Biocriteria Fill a Critical Gap in the Standards Program

Ambient biological sampling provides a direct assessment of the biological community in the waterbody being evaluated, and therefore, a measurement of aquatic life use attainment that cannot be obtained by laboratory-based criteria, (chemical-specific, toxicity or whole effluent testing). Biocriteria provide a basis for evaluating the effectiveness of existing pollution controls and for identifying previously unknown sources of impairment. They also can be used to identify where site-specific criteria modifications are needed. Currently, there is no ecologically based mechanism within the standards program for evaluating the appropriateness of existing criteria for protecting aquatic life uses. It is necessary to recognize the value of biosurvey data or the adoption of biocriteria in State standards to provide a "reality check" on State and Federal chemical-specific criteria.

There are no criteria to address many pollutants from nonpoint sources, particularly sedimentation and its effect on aquatic habitats. Nonpoint sources are the major causes of surface water use impairment (Judy, et al. 1984). The Agency policy on nonpoint source control (updated SAM-32, of 1987) recommends that biological assessments and other site-specific evaluations (e.g., for physical habitat degradation) be used to determine the effectiveness of Best Management Practices (BMP's).

Biosurvey data and biocriteria can also be applied to use attainability analyses. Biocriteria, coupled with regional use classifications, such as those that employ the ecoregion concept, provide a more quantitative basis for establishing aquatic life uses. This specificity is often necessary to identify use impairment. These tools provide an effective mechanism for characterizing the ecological value of high quality waters, i.e., those waters that possess exceptional water quality or that support sensitive and/or endangered aquatic species. This evaluation of high quality waters is a necessary component of State antidegradation policies and State standards.

Biocriteria and biosurvey data provide a screening tool for other programs such as monitoring and permits. These tools along with all other relevant information such as chemical-specific and whole effluent criteria, should be used in the permit setting process. Conflicts between these assessment approaches should be resolved using a "weight of evidence" approach, which considers the "power" of each type of assessment to address the issues and the level of resolution of the data. The lack of evidence of biological impairment in the receiving water should not preclude establishing permit limits based on other evidence. Biosurvey data maybe useful in modifying the application of chemical-specific and bioassay based criteria.

Characterization of a waterbody based on biosurvey data and in terms of biological criteria is more understandable to the general public and encourages their involvement in the regulatory process.

Biocriteria Development in State Programs

Many of the applications discussed above do not require any changes in State water quality standards statutes. They include the following:

- Use attainability analyses
- Identifying areas where site-specific criteria (both biological and chemical) are needed
- Characterizing the value of high quality waters as required by antidegradation statutes
- Determining the need for additional controls
- Identifying previously unknown sources of pollution
- Use attainment to achieve the swimmable/fishable goals of the CWA.

Some minor modifications or clarifications of State statutes may be warranted to better describe the specific applications of biosurvey data (and biocriteria where they exist) in the standards.

Existing aquatic life and fishery use designations generally do not adequately describe the types of organisms that should be found in the waterbody. Uses are generally distinguished only as warmwater and coldwater aquatic life (usually meaning salmonid and nonsalmonid waters). This has been adequate where one numeric criterion is developed for a particular contaminant; e.g., national criteria recommendations under Section 304(a) of the Clean Water Act. However, there are several examples that show that national criteria are not applicable in all cases; e.g., natural conditions (regional patterns in Arkansas dissolved oxygen), species composition responses to impacts in warmwater habitats, etc.

There is evidence that the need for site-specific criteria (biological, chemical and physical) may follow regional patterns similar to those developed by Omernik (i.e., ecoregions). The biosurvey information used to define these patterns can be used to develop more descriptive designated uses. The Ohio and Arkansas standards are examples of ecoregional use designations that describe the types of species that should be found in waterbodies within each class. These more descriptive use classes provide a basis for site-specific criteria, conducting use attainability analyses, long-term monitoring, problem discovery, etc. In addition, designated uses based on species composition allow for greater public involvement in a range of surface water programs.

Biocriteria can be numeric, narrative, or both. Where a narrative statement is used, a process should be in place that describes how biological information will be collected and evaluated to determine compliance. This is similar to the implementation procedure required for the narrative criteria currently in all State standards. Biocriteria should initially be narrative until sufficient data have been collected to develop numeric criteria. Narrative criteria can take the form of a general statement that the waterbody exhibit an unimpaired biological community. This must be supported by procedures to evaluate the level of impairment. Maine's standards are an example of appropriate use of narrative criteria.

Numeric criteria may be based on individual measurements or indices such as species richness, diversity, trophic composition, abundance or biomass. Preferably, they should also integrate several indices to provide an overall measure of aquatic health, e.g., Karr's Index of Biotic Integrity (IBI), Gammon's Index of Well-Being (IWB), or the Invertebrate Community Index (ICI). Ohio's standards, for example, set regional numeric biocriteria for all three of these indices.

Implicit in any "standard" is the issue of comparability. In the case of biocriteria, the standards are biological conditions in relatively unimpacted waterbodies. When assessing suspected impacts on the biological community, and appropriate nearby unimpacted stations exist upstream, the

biological characteristics of upstream and downstream reaches can be directly compared. This would be most applicable in high quality headwater streams. Where such stations do not exist, (e.g. heavily industrialized waterbodies, waters impaired by nonpoint sources or large waterbodies such as lakes and estuaries) the biological quality must be compared to similar (usually nearby) unimpacted or less impaired waterbodies. Another approach uses data from a set of minimally impacted reference sites that are considered to be similar ecologically. This approach does not require paired sample comparisons and is often used in conjunction with the ecoregion methodology for defining relatively homogeneous areas and a series of minimally impacted reference reaches. State standards should allow both types of methods for using biosurvey data and when establishing biocriteria.

It is recommended that:

- States begin developing biocriteria. Biocriteria and ambient biological surveys fill important gaps in the water quality standards program.

- EPA develop national guidance on the use of biocriteria and biosurvey data in the water quality-based program; similar to the 1985 "Technical Support Document for Water Quality-based Toxics Control". This should include guidance on physical habitat assessments.

- EPA develop a national policy on the inclusion of biocriteria in State standards and clarify how this information will be used in the surface water programs. This policy should clarify the legal basis for application of biocriteria and biosurvey data.

- Further evaluations be made of how biosurvey information compares with chemical-specific or whole effluent assessments.

SAMPLING AND DATA EVALUATION: FISH AND BENTHIC MACROINVERTEBRATES WORKGROUPS

Two aquatic life discussion groups were established, one for fish and the other for benthic macroinvertebrates. The intent of separating these workgroups was not to discourage exchange between them but to consider the particular attributes of each aquatic life groups in greater detail. The experiences and knowledge of the participants convinced us that evaluation of both aquatic life groups is clearly optimal and often necessary. We do not advocate the routine use of one group over the other on a programmatic basis. The knowledge and professional experience of a qualified field biologist should be applied to decide how best to biologically assess a particular environmental project.

Surface Water Program Issues

Ambient biological sampling and evaluation of these data should be integrated into a variety of State and Federal surface water programs to more completely assess and monitor the health of the Nation's waters. Biosurveys improve our ability to assess attainment and nonattainment of aquatic life uses, as well as enhance use attainability studies. These assessments should be included among the measures of the success of pollution abatement programs.

Biosurvey data should be used for surveillance, monitoring, and enforcement of water quality standards for point source discharges. These data enable evaluation of multiple point sources, and can be used to identify additive impacts of multiple dischargers. Even when designated uses are being attained, small incremental impacts can be identified by quantitative and qualitative shifts in biological community structure and function. A single sampling session is adequate for screening purposes, to indicate where further sampling is needed, or to initiate some other regulatory action. However, the need for more complete information increases as the complexity of the issues increases.

Whereas there are a variety of surrogate methods (e.g., chemical monitoring and bioassay) for estimating point source impacts to the biota, surrogates are unsatisfactory for estimating many kinds of nonpoint source (NPS) impacts. Biosurveys should be used to screen for location, severity and extent of suspected (or known) areas of NPS impacts and to evaluate long-term trends. Results of previous biological surveys should be incorporated into subsequent 305(b) reports to identify lakes and stream segments with NPS impacts, and to demonstrate trends.

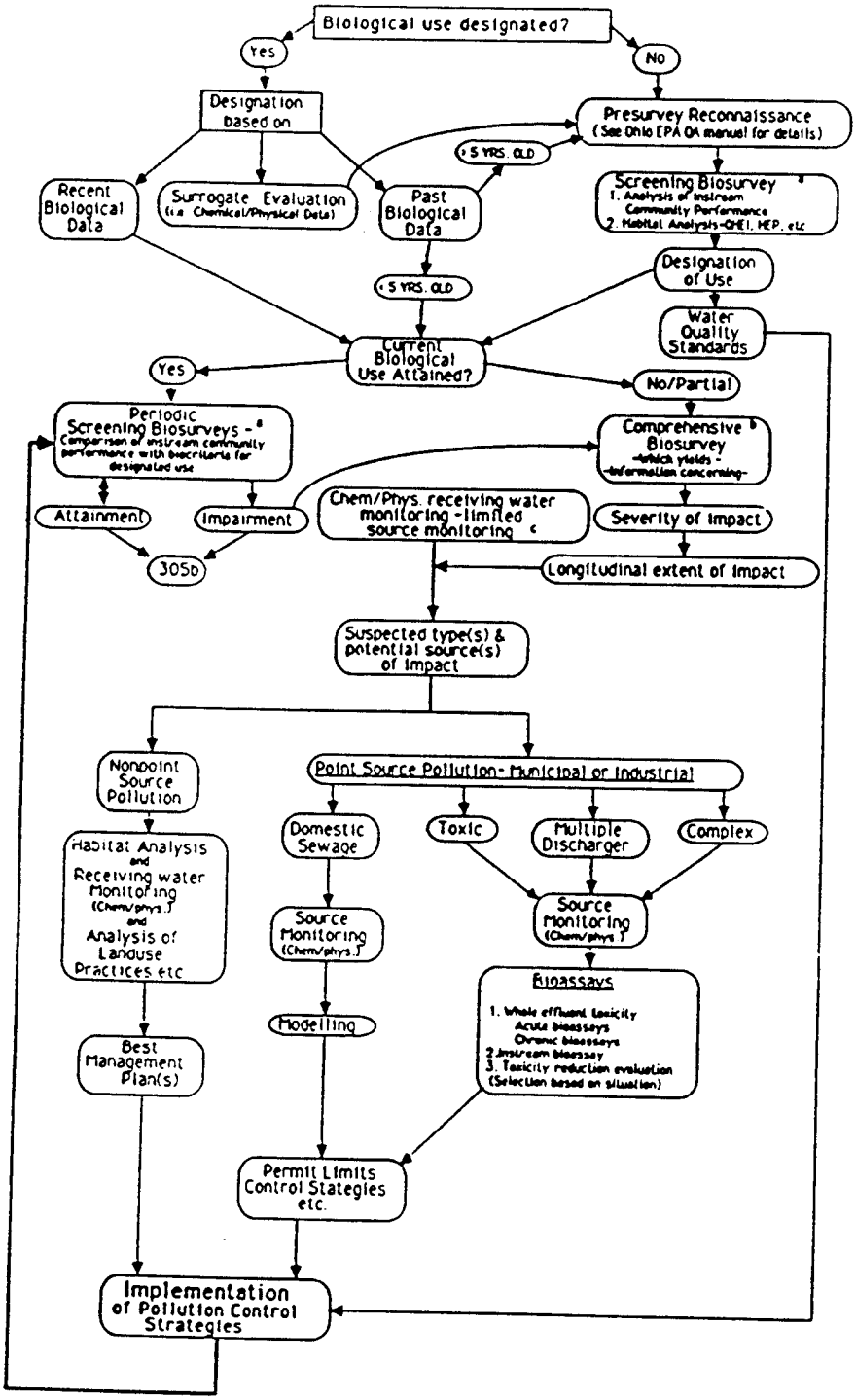
A conceptual framework (Figure 1) for using ambient biological sampling data in surface water programs has been proposed by some of the staff of the Ohio EPA. Most importantly, this framework shows that when decisions are made on aquatic life use attainability, attainment and nonattainment, those decisions should be based largely on a direct assessment of the aquatic community. This framework does not represent a policy statement by either U.S. EPA or Ohio EPA, rather it should be a starting point for further discussions.

Professional aquatic biologists (both fisheries and benthic) should be active participants in all phases of the biosurvey programs; from planning, sampling site selection, data collection, specimen identification, analysis and interpretation of data, to report and permit writing and review.

A technical support or guidance document should be developed by U.S. EPA to identify how biocriteria and ambient biological assessments can be implemented in the surface water programs. This document should provide:

- Clear definitions of terms to eliminate the confusion that currently exits about usage and meaning. These workgroups recommend the definitions used in this report;
- Guidance to State and Federal personnel in surface water programs about the appropriate uses and advantages of ambient biological sampling; and,

Figure 1. A conceptual framework for assessing ambient biological performance and the success of implemented pollution control strategies.



[a] Screening biosurveys use widely spaced sampling sites to obtain general water quality information about a watershed.

[b] Comprehensive biosurveys require numerous sampling sites to document severity and extent of impacts. Sampling occurs at reference sites, sites up and downstream from point sources, and areas of expected changes in water quality.

[c] Biological communities react predictably to various water quality problems (e.g., toxicity, D.O. sags, siltation). As a result, biosurveys indicate the type of impact, and when coupled with receiving water monitoring provide greater resolution of the source(s) and type(s) of impacts.

- An evaluation of the relative strengths and weaknesses of the various chemical, physical and biological assessment methods to address different regulatory issues. Table 1 is one such evaluation and is presented here only as an example and a suggestion for additional calculations. The cost effectiveness of biological sampling as an assessment tool has been demonstrated by the Ohio EPA (Ohio EPA 1987). Other States and regulatory agencies should document the relative costs of these methods.

Sampling Issues

It is essential that biosurveys be performed with the same degree of procedural rigor used in chemical monitoring and toxicity testing. Clear data quality objectives (DQOs) need to be developed and implemented before ambient biological sampling is done. States need to develop (with guidance from U.S. EPA) and document standard field and data evaluation methods to achieve the data quality objectives. The rationale for using particular assessment procedures to address a given regulatory issue should be included in such a document. However, it is understood that States may need to use the results of ambient biological sampling without having completely documented their methods or DQOs,(e.g., while they are in the process of testing and selecting methods). The data evaluation techniques currently used by many of the States should serve as models or beginning points for other State programs.

Both fish and benthic macroinvertebrate (benthos) communities should be used in any ambient biological sampling program. The experiences of the State biologists at the workshop indicate that frequently results of ambient fish and benthos assessments corroborate each other. However, these two groups are not equally sensitive to all perturbations in all circumstances. An understanding of these differences is important in selecting assessment methods and evaluating data.

The following are some considerations in selecting assessment methods. These considerations do not imply a superiority of either biological group for regulatory assessment purposes and are not reasons to use only one group exclusively.

Fish

- Fish populations and individual fish tend to remain in the same area during summer, when most sampling occurs. Fish communities are persistent and recover rapidly from droughts and floods. Thus, large community fluctuations are unlikely for purely natural reasons. Comparable results can be expected from an unperturbed site at various times during a summer.

- Fish are a highly visible component of the aquatic community to the general public. Aquatic life uses and regulatory language are typically characterized in terms of fish (e.g., coldwater, warmwater, sport, etc.).

- Most fish species have long life spans (3-10+ years) and can reflect both long term and current environmental quality. The sampling frequency required for trend estimates is less than for shorter-lived organisms.

- Fish have larger ranges and are less affected by natural microhabitat differences than smaller organisms. This makes fish especially useful for assessing regional and macrohabitat differences.

- Fish continuously inhabit the receiving waters and integrate the chemical, physical, and biological histories of the waters that are not directly measured by chemical or short-term bioassays alone.

Table 1. The comparative ability and "power" of various chemical, physical, and biological assessment techniques to measure or indicate key components of factors affecting biological integrity of surface waters (D - directly measures; I - indirectly measures; S - strongly reflects; C - casual relationship). Modified from Ohio EPA (1987).

Factors/Components	Level 1&2 Exposure Assess.[1]	Level 3 Exposure Assess.[2]	Toxicity (acute)	Toxicity (chronic)	Physical Assessment	Ambient Biological Evaluation
I. CHEMICAL WATER QUALITY						
Conventional substances	D	D	I	I	-	S
Heavy metals	D	D	I	I	-	S
Toxic organics	-	D	I	I	-	S
Static interactions	S	S	I	I	-	N/A
Dynamic interactions	-	I	-	I	-	S
II. ENERGY DYNAMICS						
1° and 2° dynamics	C	I	-	-	-	I
Nutrient cycling	C	I	-	-	-	I
Organic inputs	-	C	-	-	-	I
III. HABITAT QUALITY						
Substrate	-	-	-	-	D	S
Water velocity	-	-	-	-	D	S
Instream cover	-	-	-	-	D	S
Channel integrity	-	-	-	-	D	S
Riparian buffer	-	-	-	-	D	S
Habitat diversity	-	-	-	-	D	S
IV. FLOW REGIME						
Low Extremes	I	I	-	-	-	S
High Extremes	-	-	-	-	-	S
Temporal cycles	-	C	-	-	C	S
Volume	D	D	-	-	D	S
V. BIOTIC RESPONSES						
Acute effects	I	I	D	D	-	S
Chronic effects	I	I	I	S	-	S
Abundance, biomass	-	-	-	-	-	D
Structural	-	-	-	-	-	D
Functional	-	-	-	-	-	D
Disease, etc.	-	-	C	C	-	D
Tolerances	-	-	-	-	-	D
Competition	-	-	-	-	-	S
Reproduction	-	-	-	S	-	S
Predation	-	-	-	-	-	S
Growth	-	C	-	S	-	D

1 primarily models for oxygen demanding substances and simple mass-balance dilution calculations for other substances; steady-state conditions assumed.

2 applications ranging from probabalistic dilution to dynamic fate-assessment models.

- The taxonomy of fish is well established, allowing experienced biologists to identify most species in the field, reducing data generation time.

- Distributions, life histories and tolerances to environmental stresses of many fish species are documented in the scientific literature. Fishes are collected annually by a variety of agencies and organizations. Thus, the relative abundances of tolerant and intolerant species and the absence of expected, but generally intolerant species, provide evidence about a site's quality.

Benthic Macroinvertebrates

- Most benthos are sessile or have limited migration patterns, making them well-suited for assessing site-specific impacts (upstream/downstream studies).

- Benthic communities integrate the effects of short-term impacts since most species have complex life cycles of one to two years or less. The sensitive life stages respond quickly to stress.

- Degraded stream conditions may be detected with only a cursory examination of the benthos since many sensitive taxa are relatively easy to identify to the family level in the field.

- Sampling for benthos is easy, requiring few biologists with inexpensive equipment, and has no detrimental effect on the resident biota.

- Benthos are a primary food source for important recreational and commercial fish. Abundant food is a primary requirement for healthy fish populations.

- Many small streams (1st and 2nd order) naturally support a diverse macroinvertebrate fauna and can provide ample data for many rigorous statistical analyses and tests. Likewise, tributaries to lakes and large rivers are best evaluated using benthos due to the ability of fish to freely migrate in and out of the larger waterbody.

- Many State regulatory agencies routinely collect benthos and thus have extensive data available.

For evaluating specific impacts, a series of reference sampling sites should be selected from upstream and downstream recovery areas wherever possible. If the upstream or recovery segments are impacted, a series of reference sites may be chosen in adjacent streams of similar drainage areas and morphology, within the same ecoregion. These data should serve as the basis for evaluating data from suspected impact areas.

Long-term biomonitoring should be conducted at regional reference sites. Data from these sites may be used to evaluate the effectiveness of the overall water quality control program and to detect any long-term trends. The range of data collected from regional reference sites should be the basis for developing regional biocriteria. For either long-term or site-specific purposes, sampling and data evaluation should be done within an ecoregional framework.

Selection of actual sampling locations and specific habitats within those locations should be made by the field biologists, within State defined guidelines. The choice of habitats to sample (e.g., riffles, pools, etc) will be dictated by their availability at the site, other site conditions and whether they are representative of that ecoregion. Likewise, the use of artificial or natural substrate sampling for benthos should depend on site conditions, the data needs of a particular program and past experience. The rationale for site and substrate choices should be included in State guidelines and the choice documented for each sample.

Larval fish, or young-of-the-year, although not used in most indices, are sensitive to particular kinds of stresses. Their presence should be included in a narrative discussion of the survey results.

The issues of seasonality and low-flow sampling continue to generate much discussion. Although sampling is not limited by season, for purposes of assessing environmental impact, fish should be sampled during the low- to moderate-flow periods of summer and early fall and benthos should be sampled during stable base flows conditions. Sampling during these periods produces the most consistent and ecologically meaningful results.

It is not necessary to sample during 7Q10 conditions because not only do the biological communities integrate ecosystem effects over a long time period, but biological populations are most stable and easily collected at base-flow. The longer term (annual, seasonal) impact of chronic toxicity can be evaluated. Low-level chronic toxicity occurring only during the 7Q10 may not be detected; however, the ecological significance of this toxicity is also questionable. The ecological significance of toxicity occurring only during extreme low flow conditions can be highly variable, depending in part on the flow regime of the stream. It is recommended that, whenever possible, flow data be used to help interpret results of biosurveys.

Benthic surveys are useful for instances when potential toxicity is predicted from effluent toxicity tests. If the biosurvey data does not support the toxicity evaluation, the information may be reevaluated, where feasible using a "weight of evidence" approach. If biosurveys indicate a potential toxicity problem with a discharge, an effluent toxicity test is recommended.

An evaluation of physical habitat quality should be conducted in conjunction with the biological assessment. The health of biological communities is as dependent on the physical habitat as on the chemical water quality. Many NPS impacts affect the physical condition of the waterbody. A habitat evaluation is necessary to delineate these effects and to account for natural ecoregional and site-specific differences. Different evaluations should be made for fish and benthos. Benthos habitat assessments should collect information on the riparian zone.

State and regional training is needed on the use of the Rapid Bioassessment Protocols (RBP). The RBP may be used to augment, but does not necessarily replace, any more intensive State biosurvey methods. In addition, the U.S. EPA biological field methods manual should be updated to address all of these issues and to provide information about advances in sampling techniques.

Data Evaluation Issues

The State and Federal surface water programs need to develop guidance regarding the quantity, quality and kinds of data (from biosurveys, chemical sampling, bioassays, etc.) required to address each regulatory issue. There needs to be a mechanism (e.g., a decision tree) to determine the level of sampling and data effort and resolution of information that is sufficient for various purposes.

Perturbations in aquatic communities should be demonstrated by measuring shifts in the structural and functional composition from conditions expected in unimpaired or minimally impacted situations. For example, increases in the number of species tolerant of pollution and decreases in the number of carnivore species indicate structural and functional shifts. This approach should be considered for parallel studies with in situ bioassays and toxicity testing.

To effectively assess the health (or integrity) of a biological community, it is necessary to use a multimetric approach. A reliance on one or two measures (e.g., species richness and diversity) is often inadequate. Data evaluation methods now are available to more completely assess biotic integrity. One approach combines several community measures, e.g., species richness and composition, trophic composition, condition of individuals, etc. into one index, while making the original measures available for further analyses or evaluation. Karr's Index of Biotic Integrity (IBI) is an example of this approach. Another method uses several community measures placed in a "decision matrix" or "decision tree" to evaluate the level of impact on a biological community.

Again, these measures are used to compare results with those expected of communities in minimally impacted sites.

There should be increased use of computer programs and databases, including (but not limited to) the U.S. EPA BIOS system for storage, retrieval and analyses of biological data. Rigorous data quality checks, including automated spelling, locational controls and duplicate data entry, are needed to ensure data quality. The U.S. EPA should encourage exchange of data and analytical tools among all agencies (Federal and State) involved with surface water programs. Ideally, there should be a mechanism to evaluate existing data and sampling programs from among various agencies and to coordinate efforts by neighboring States to evaluate cross-boundary streams and watersheds.

ECOREGIONS WORKGROUP

Ecological regions (rather than political or hydrographic regions) are a useful geographic framework for developing biological criteria. Ecoregions reduce (but do not eliminate) overall variability among sites. Thus they provide a method to determine a range of attainable conditions and allow development of protective yet reasonable standards on a regional basis. This framework allows a means to develop biological expectations that is far less expensive than site specific approaches and much more accurate than national criteria.

Relatively unimpacted reference sites are essential for evaluating known or suspected impacts. The ecoregion methodology provides alternatives to, or can augment, the currently used upstream/downstream site selection methods. This is especially important for monitoring and reporting an entire State's waterbodies (305b and 319 reports) and where upstream sites are disturbed or naturally different.

Regional reference sites function as: checks on upstream reference sites and downstream recovery sites; tools to evaluate nonpoint source problems; a means to develop regional biological criteria; and benchmarks for evaluating use attainment. In addition, the process of establishing regional reference sites provides the States with a list of their best waters. These sites may be considered for special protection from degradation.

The U.S. EPA and the States should invest in the up-front cost of locating and evaluating regional reference sites. These sites would then form the framework for a long term monitoring network that has enormous potential in long term cost savings and status and trend assessment.

The States and U.S. EPA should evaluate the usefulness of alternative ecoregion maps, based on their needs and applications. For example, in Minnesota, Ohio, and Oregon, Omernik's ecoregions showed better correspondence with patterns in fish communities and/or water chemistry than did Bailey's ecoregions, river basins or physiographic regions. In Arkansas, Omernik's ecoregions and physiographic regions both corresponded better to aquatic ecosystem patterns than did Bailey's ecoregion or river basins. However, the USDA map of Major Land Resource Areas may be more applicable to chemical patterns in Nebraska and Iowa. These ecoregion evaluations must include data from minimally impacted regional sites. Using a mix of disturbed and "clean" sites blurs any regional patterns and defeats the purpose of establishing benchmarks of attainability.

Regional biological criteria should initially be narrative, until sufficient data exist to develop numeric criteria. Narrative criteria should be in the form of key, characteristic and dominant species expected in minimally impacted sites, as well as those species that dominate disturbed sites on a regional basis. Numeric criteria should be based on expected regional values for: species richness and composition, diversity, trophic composition, abundance and/or biomass, and condition. As many of these measures as possible should be analyzed to evaluate the overall health of the biotic communities. Eventually the individual metrics should be incorporated into a multimetric

analysis of aquatic community integrity; either as a "decision matrix" such as Maine's, or as an index of integrity such as Karr's IBI.

Ecoregions can be used to organize data for 305(b) reporting on an ecological, rather than a statewide or site specific, basis. Regions can stimulate a broader ecosystem perspective on water issues, helping managers to prioritize problems and locate outliers. Ecoregions do not yield biological numbers for permits. Ecoregions are not "magic bullets" with static, precisely defined boundaries. Instead, they require a fair amount of thought and biogeographical knowledge to use properly.

EPA and the States should:

- Use ecoregions as a geographic framework for stratifying biological variability;
- Use minimally-impacted ecoregional reference sites to set regional biological expectations (narrative or numerical);
- Locate and monitor these ecoregional reference sites;
- Test the relative usefulness of various ecoregion maps based on data from these reference sites;
- Support a technology transfer program to assist States in the development of regional biological criteria.