Chapter 1

Introduction: Bioassessment and Biocriteria

1.1 Rationale

1.1.1 Water Quality Monitoring

The recognition that chemical water quality analyses do not adequately predict or reflect the condition of all aquatic resources has led to the development of measures of biological integrity expressed by biological criteria. Biological surveys, criteria, and assessments complement physical and chemical assessments of water quality by reflecting the cumulative effects of human activities, and natural disturbances on a water body, including the possible causes of these effects. The biological approach is best used for detecting generalized and non-specific impairments to biological integrity, and for assessing the severity of those impairments. Then, chemical and toxicity tests, and more refined habitat assessments, can be used to identify probable causes and their sources, and to suggest corrective measures.

For the purposes of bioassessment and biocriteria development described here, an estuary is a semi-enclosed water body that has a free connection with the open sea and an inflow of freshwater that mixes with the seawater; including fjords, bays, inlets, lagoons, and tidal rivers. Coastal marine waters are those marine waters adjacent to and receiving estuarine discharges and extending seaward over the continental shelf and/or the edge of the U.S. territorial sea.

1.1.2 Advantages of Bioassessment and Biocriteria

Bioassessment is intended to detect biological responses to pollution and perturbation. Routine water quality monitoring for example, detects effects of nutrient enrichment and chronic acidification, but normally is not designed to detect trace levels of toxicants or contaminants, ephemeral pollution events (e.g., acidic episodes, spills, short-lived toxicants and pesticides, short-term sediment loading), or combined or synergistic impacts. Bioassessment, by monitoring organisms that integrate the effects of environmental changes, may in time detect these effects.

Bioassessment, coupled with habitat assessment; i.e., physical and chemical measurements, helps identify probable causes of impairment not detected by physical and chemical water quality analyses alone, such as nonpoint source pollution and contamination, erosion, or poor land use practices. The detection of water resource impairment, accomplished by comparing biological assessment results to the biological criteria, leads to more definitive chemical testing and investigations which should reveal the cause of the degradation. This, in turn, should prompt regulatory and other management action to alleviate the problem. Continued biological monitoring, with the data collected compared to the criteria, will determine the relative success of the management efforts.

1.2 Legal Origins

1.2.1 Clean Water Act

The CWA, Section 101, requires federal and state governments to "restore and maintain the chemical, physical, and biological integrity of the nation's waters." Thus, the Act mandates the restoration and maintenance of biological integrity in the Nation's waters. The combination of performing biological assessments and comparing the results with established biological criteria is an efficient approach for evaluating the biological integrity of aquatic ecosystems. Other pertinent sections of the CWA are Sections 305(b), 301(h), and 403(c). Table 1-1 outlines suggestions for the application of biological monitoring and biocriteria for estuaries through existing state programs and regulations.

1.2.2 305(b) Reporting

States and the USEPA report on the status and progress of water pollution control efforts in §305(b) reports submitted every two years. Inclusion of biological assessment results in these reports will improve the public understanding of the biological health and integrity of water bodies. Many of the better known and widely reported recoveries from pollution have involved the renewal or reappearance of valued species to systems from which they had nearly disappeared, or the recovery of a viable fishery from contaminants. Examples of such recoveries are the restoration of the lower Potomac River and of shellfish beds in Maine. Incorporation of biological integrity in §305(b) reports will ensure the inclusion of a bioassessment endpoint, and will make the reports more accessible and meaningful to many segments of the public.

1.2.3 301(h) and 403(c) Programs

Two other programs within USEPA that specifically rely on biological monitoring data in coastal marine areas are the §301(h) Waiver Program and the §403(c) Ocean Discharger Program. The §301(h) program allows estuarine and marine dischargers who meet specific criteria set forth by USEPA to defer secondary treatment if they can show that their discharge does not produce adverse effects on resident biological communities. As part of the modified NPDES permit received through this waiver program, the dischargers are required to conduct extensive biological monitoring programs designed to detect detrimental effects to those biological communities.

The §403(c) Ocean Discharge Program requires that all dischargers to marine waters provide an assessment of discharge impact on the biological community in the area of the discharge and on the surrounding biological communities. This program requires extensive biological monitoring for some dischargers. Community bioassessment methods are valuable in this program for trend assessment and, in some cases, refinement into more rigorous and definitive assessments.

1.2.4 304(a) Criteria Methodology

This technical guidance was developed under the §304(a) requirement that, "criteria for water quality accurately reflecting the latest scientific knowledge of the kind and extent of all identifiable effects on health and welfare including, but not limited to, plankton, fish, shellfish, wildlife, plant life, shorelines, beaches, aesthetics, and recreation which may be expected from the presence of pollutants in any body of water . . .” be published and updated as needed.

Under this section, a guidance document must include information on restoration and maintenance of chemical, physical, and biological integrity of navigable and ground waters, waters of the contiguous zone, and the ocean. This also covers information identifying conventional pollutants, such as those classified as biological oxygen demanding,

Table 1-1. Applications of estuarine biological monitoring protocols and biocriteria.

Program	Biological Monitoring and Assessment	Biological Criteria
Section 305(b)/ Reporting	• Improving data for beneficial use assessment. • Improving water quality reporting.	• Identifying waters that are not achieving their aquatic life use support. • Defining an understandable endpoint in terms of “biological health” or “biological integrity” of waterbodies."
National Estuary Program (NEP)	• Assessing status of biological components of estuarine systems. • Develop monitoring objectives and performance criteria. • Establish testable hypothesis and select statistical methods. • Assessing estuarine trophic status and trends, and assessing biological trends. • Select analytical methods & alternative sampling designs. • Evaluate expected monitoring study performance. • Implement monitoring study & data analysis. [Monitoring and sampling needs vary for each estuary]	• Identifying estuaries that are not attaining designated use (including aquatic life use) support. • Defining estuarine biological integrity based on a reference condition. • Identifying impairments due to toxic substances, eutrophication, and habitat modification.
Section 319/Nonpoint Source Program	• Evaluating nonpoint source impacts and sources. • Measuring site-specific ecosystem response to remediation or mitigation activities. • Assessing biological resource trends within watersheds.	• Determining effectiveness of nonpoint source controls.
Watershed Protection Approach	• Assessing biological resource trends within watersheds.	• Setting goals for watershed and regional planning.
TMDLs	• Identifying biological assemblage and habitat impairments that indicate nonattainment of water quality standards. • Priority ranking waterbodies. • Documenting ecological/water quality response as a result of TMDL implementation.	• Identifying water quality-limited waters that require TMDLs. • Establishing endpoints for TMDL development, i.e., measuring success.

Table 1-1 (cont'd). Applications of estuarine biological monitoring protocols and biocriteria.

Program	Biological Monitoring and Assessment	Biological Criteria
NPDES Permitting	• Measuring improvement or lack of improvement of mitigation efforts. • Developing protocols that demonstrate the relationship of biological metrics to effluent characteristics.	• Performing aquatic life use compliance monitoring. • Helping to verify that NPDES permit limits are resulting in achievement of state water quality standard.
State Monitoring Programs	• Improving water quality reporting. • Documenting improvement or lack of improvement of mitigation efforts including estuary clean-up efforts, TMDL application, NPDES efforts, nonpoint source pollution controls, etc. • Problem identification and trend assessment. • Prioritizing waterbodies.	• Providing a benchmark for measuring effectiveness of controls and performing watershed/regional planning.
Risk Assessment	• Providing data needed to estimate ecological risk to assessment endpoints.	• Providing an assessment or measurement endpoint.
Water Quality Criteria and Standards	• Developing data bases for estuarine phytoplankton, macroinvertebrates, fish, plants, and other assemblages. • Developing indices that assess estuarine biota compared to a reference. • Providing data for aquatic life use classifications.	• Providing benchmark for identifying waterbodies that are not attaining aquatic life use classification. • Developing site-specific standards.
Section 301(h)/ Waiver Program	• Allows marine discharges who meet USEPA criteria to defer secondary treatment if discharge does not produce adverse effects on resident biological communities.	• Providing threshold against which to measure detrimental effects on biological communities.
Section 403(c)/Ocean Discharge Program	• Requires marine dischargers to provide an assessment of discharge impact on biological community in discharge area as well as surrounding communities.	• Providing threshold against which to measure discharger impacts on biological communities.
Section 304(a)/ Criteria Methodology	• Provides information on restoration and maintenance of chemical, physical, and biological integrity of waters. • Identifies conventional pollutants, their concentrations and effects on surrounding communities.	• Providing the benchmark for measuring the effects of pollutants on the biological community.

suspended solids, fecal coliform, and pH. Section 304(a)(8) authorizes USEPA to develop and publish methods for establishing and measuring water quality criteria for toxic pollution, on other bases than a pollutant by pollutant approach. This includes biological monitoring and assessment methods. Specific states have the authority to enforce more stringent regulations as necessary.

1.2.5 Biocriteria

A major purpose of developing biological assessment methods is to establish biological criteria for surface waters. Biological criteria are guidelines or benchmarks adopted by states to evaluate the relative biological integrity of surface waters. The criteria are defined as "narrative expressions or numerical values that describe the biological integrity of aquatic communities inhabiting waters of a given designated aquatic life use" (USEPA 1990). Biological criteria are, in effect, a practical approach to establishing management goals designed to protect or restore biological integrity. Biocriteria can be adopted by a State into their water quality standards, along with chemical, physical and toxicity criteria to better protect aquatic life uses of waterbodies.

Biocriteria can be developed from reasonable expectations for the locality based on: historical data; reference conditions; empirical models; and the consensus judgment of regional experts (Section 1.4.2). The reference condition component of biocriteria requires minimally impaired reference sites against which the study area may be compared. Minimally impaired sites are not necessarily pristine; they must, however, exhibit minimal influence by human activities relative to the overall region of study (USEPA 1996a). In some instances, "minimally impaired" sites are not available because the entire area has been degraded. Biocriteria are then based on historical data, empirical models if appropriate, and expert judgement to set a condition better than present sites. Restoration of the degraded area must therefore be accomplished before any such reference sites can be established.

Biological criteria typically include the condition of aquatic communities at designated reference sites as an important component. The conditions of aquatic life found at these sites are used to help detect both the causes and levels of risk to biological integrity at other sites of that type in a region. In keeping with the policy of not degrading the resource, the reference conditions—like the criteria they help define—are expected to be upgraded with each improvement to the water resource. It is important that biological criteria not be based on data derived from degraded reference sites. In fact, a concerted effort should be made by States and other jurisdictions to preserve the quality of designated reference sites by setting those areas aside in preserves or parks or by inclusion in use protection programs so that continuity of the biocriteria data base can be maintained. Biocriteria supported by bioassessment surveys serve several purposes in surface water programs, discussed in the following section.

1.3 Uses of Biocriteria

The biocriteria-bioassessment process helps resource managers identify impairment of designated beneficial uses. It expands and improves designated beneficial use classifications and their associated water quality standards. It detects problems other survey methods may miss or underestimate. It is a process which helps the resource manager set program priorities. It can also be used to evaluate management and regulatory efforts. For example, the information summarized in Table 1-2 indicates that wastewater outfalls are a controlling factor of soft bottom benthic communities and that there is a moderate scientific understanding of the effects of these outfalls specifically in the Southern California Bight (USEPA 1992).

1.3.1 The Use of Bioassessment Data to Establish Biocriteria Appropriate to Designated Beneficial Uses

The hypothetical information presented in Figure 1-1 represents data collected for a given class of similar estuarine or coastal reaches (e.g., similar sediments, depths, and salinities) within the same geographic region. For these areas some high level of resource quality can be conceived which represents a pristine condition, essentially the optimum potential or integrity of those waters. A completely unimpaired (no negative human impacts upon the organisms of the natural system) estuary or coastal marine area is referred to as having biological integrity. The approximation of this ideal quality at the top of a continuum can be expressed by a variety of environmental measures of the biota indicated on the vertical axis of the graph. The determined **ideal** level of biological measurements at the maximum score is shown by the upper horizontal line (equivalent to biological integrity). A second horizontal line somewhat below this is the level set as the reference condition, the **attainable** level of integrity derived from actual measurements from among the highest quality areas in the class. All information on this axis is expected to be objectively derived through the scientific process and usually is presented in a comprehensive index of many biological characteristics such as an IBI or the EMAP benthic index (Chapter 11).

The horizontal axis represents a progression of socially determined **use designations**; i.e., those predominant uses the State has concluded are appropriate for a particular estuary, region or area within the class. These hypothetical designated uses are arranged on the graph from those usually associated with relatively low water resource quality on the left, to those associated with very high, relatively natural, resource quality on the far right.

The potentially optimal array of **biological criteria** for this class of waters, then, are scores between the reference condition and the level of biological integrity; i.e., between that which is achievable and that which is ideal. The narrower this area, the higher the quality of the waters throughout the class, and the less restoration management is required. The objective, then, is to protect these resources.

On the same horizontal axis, a class of high quality regional uses are further described by a subset of **aquatic life uses**. These are the designated uses for which management goals are also described by desirable characteristics of the aquatic biota to be especially protected, such as “protection of the health and diversity, undiminished, of all indigenous species of fish and invertebrates” for those designated as exceptional natural waters. Resource managers need to apply their first, concerted efforts to those uses because it is usually more cost-effective and resource-conservative to protect existing high quality areas than it is to restore degraded ones.

Table 1-2. Impacts on the marine environment of the Southern California Bight. Modified from Bernstein et al. 1991.

Sources of Perturbation	Valued Ecosystem Components														
	Intertidal	Phytoplankton	Zooplankton	Soft Bottom Benthos	Hard Bottom Benthos	Kelp Beds	Wetlands & Estuaries	Commercial Shellfish	Pelagic Fish	Demersal Fish	Fish Eggs & Larvae	Marine Mammals	Marine Birds	Human Health	All
Storms	△▨			▲	●	☆▨	○▨							○	
El Niños	✚	★	★	△	○	☆▨	△		★	▲		○	▲		
California Current		★	★						★	★	☆				
Upwelling		★	★			☆▨			★		★		●		
Blooms/Invasions				△▨		○▨			○▨	○▨					Net effect of each source on all components
Ecol. Interactions	★	★	★	●		★	○	▲	✙		△		✙		
Power Plants							●		●		○				
Wastewater Outfalls		○		☆▨	●	○		○		✚		○	✙▨	○	
Dredging				○▨		○	✙▨								
Rivers/Storm Runoff	△▨					✙▨	☆▨	○		○	●	●	✙▨		
Commercial Fishing						△▨		★	△	✚		○▨	○▨	○	
Sport Fishing								✚	✚	✚		○		✚	
Habitat Loss/Mod.	☆▨					✙▨	☆▨	○		○	●	●	✙▨		
Oil Spills	☆▨						▲						✙▨	○	
All				Net effect of each source on all components											

KEY			
Potential Importance		Understanding	
☆ Controlling	✙ Moderate	■	High
△ Major	○ Some	▨	Moderate
		□	Low

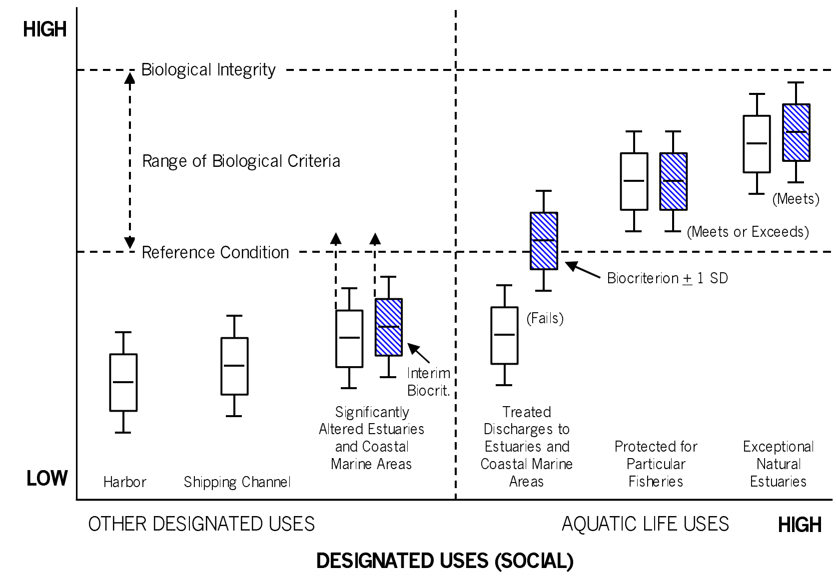


Figure 1-1
Biocriteria for given classifications of estuaries and coastal marine areas. Shaded boxes represent the appropriate biocriterion range for selected classes. Unshaded boxes represent the range of measurement results for test sites in given classes. The vertical arrows above the boxes for the "significantly altered estuaries and coastal marine areas" class indicate the goal of raising the biocriterion for these waters over time in response to restoration efforts.

Selected biocriteria with an acceptable range of variation, perhaps one standard deviation, are shown as cross hatched boxes appropriately located for each designated use. Test results for a given area in any use classification ("box and whisker" plots showin g the full range of measurements including variation for that area) can then be compared graphically to the biocriterion for that designated use. Three interpretations of an estuarine or coastal marine area meets its criterion, meets or perhaps even exceeds its criterion, and fails to meet the criterion are illustrated.

A fourth possible result is the marginal condition of significantly altered systems such as urban harbors or shipping channels. The original condition of these areas may very well have been within the optimal range of biotic health and diversity for the region, but intense development has significantly altered them so that as a group they no longer meet the minimum reference condition for similar areas of the region. An interim biocriterion for these areas may be set with the intention of progressively raising the criterion when sequential restoration efforts are accomplished through a long range management effort.

The "other desi gnated uses" to the left of the bifurcation line may still be surveyed to assist management decision makers; however, they fail to meet the criteria, and there are no designated aquatic life uses which apply.

The designated uses, aquatic life uses and biocriteria are all hypothetical in this illustration, but the interrelationships of societal and scientific elements of decision making should be evident. They are independent processes linked by an environmental ethic and the USEPA policy of antidegradation of water resource quality (the reference condition "bottom line" so to speak). A rational decision can be made which balances that which is ideal with that which is achievable measured by the objective processes of science.

1.3.2 Expansion and Improvement of Water Quality Standards

When a State adopts biological criteria in their water quality standards to protect aquatic life uses, the criteria become benchmarks for decision making, and may form the basis for requirements in NPDES permits and other regulatory programs.

1.3.3 Detection of Problems Other Methods May Miss or Underestimate

In the process of establishing biocriteria, more data and information is inevitably developed than was previously available. The review of this new information often reveals problems not evident before or provides expanded insight into existing concerns and issues. Armed with this information, a water resources manager is better able to examine issues and make decisions.

1.3.4 Helping the Water Resource Manager Set Priorities

In light of the new information described above, the schedule of activities, allocation of funds, and uses of personnel and equipment may be more appropriately prioritized according to the urgency or magnitude of the problems identified.

With the expanded available biological information augmenting chemical and physical information, managers can apply a triage approach to water resource projects based on the actual condition of the biota affected. This is much like a physician evaluating multiple emergency medical patients. Essentially, areas that are critically impaired, those that are moderately impaired, and those in good condition for which protection rather than remediation is required, can all be identified. Rational decisions can then be made about how to apply limited resources for the best results in accordance with the needs and priorities of the state.

1.3.5 Use of Biosurveys and Biocriteria to Evaluate the Success or Failure of Management Initiatives or Regulations

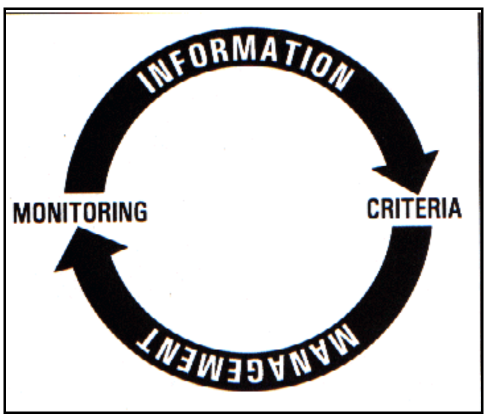
The manager may design a biosurvey to collect data before and after a permit, regulation or other management effort has been implemented, perhaps augmented by spatially distributed nearfield/farfield sampling as well. With this information and the biocriteria decision making benchmark, it is possible to clearly evaluate the environmental response of the system to the methods applied. This is useful in the NPDES permit review process as a way to help determine the effectiveness of permit controls. Typically, biocriteria are not used directly in NPDES permits as effluent limitations. Biomonitoring above and below a permit site when compared to the established biocriteria will reveal the adequacy of the permit to achieve its intended purpose.

If the biota are unimpaired or recovering, it may be wise to leave the permit, management practice or regulation as is. If the biota are impaired or declining, the review recommendation may be to change the permit, management technique or regulation accordingly. With NPDES permits, the five year review cycle allows sufficient time for extensive biological information to be developed so this determination can be made with reasonable confidence.

1.4 Program Interdependence

It should be readily evident from the applications described above that physical, chemical, and biological surveys and monitoring (repetitive surveys of the same area) and biological criteria are interrelated in the water resource management process. Figure 1-2 illustrates this interrelationship, often referred to as "adaptive management." In this continually cycling process, monitoring provides the information necessary to identify problems and to establish biocriteria for the decision making, management planning, and implementation necessary to respond appropriately. Continued monitoring then reveals the relative success of the effort by comparing the new results to those criteria again. At this point the criteria or the management plan may be adjusted as needed and the cycle repeats. Ideally, the estuarine or coastal waters improve with each cycle.

Figure 1-2
Program Interdependence



1.5 Implementing Biological Criteria

Implementing biocriteria requires an established and standardized methodology for biological assessment adjusted to regional or state conditions. Hence, guidance for state and regional development of biocriteria has two elements which are described in the biological criteria technical guidance documents such as this one:

- **Bioassessment Protocols** are methods used to assess the status and trends of water bodies. Guidance documents for bioassessment contain suggested methods and protocols for establishing monitoring programs that use biological assessment.

- **Biocriteria Guidance** assists states in establishing biological criteria for water bodies. Biocriteria are a series of ambient water resource quality values or statements of condition that relate to the desired biological integrity for that class of waters. When established they can be used to evaluate similar water bodies in that region. Implementation of biocriteria requires use of bioassessment protocols and a state or regional biomonitoring database. The National Program Guidance for biocriteria describes issues related to development and implementation (USEPA 1990). The first biocriteria technical guidance issued was for streams and small rivers (USEPA 1996a). It incorporated both biosurvey techniques and biocriteria development methods. It was followed by the Lakes and Reservoir Bioassessment and Biocriteria Guidance (USEPA 1998). Each of these documents incorporated biosurvey techniques and the same approach is being followed in similar documents for rivers, wetlands, and coral reefs in addition to this present technical guidance for estuaries and coastal marine waters.

1.6 Characteristics of Effective Biocriteria

Generally, effective biocriteria share several common characteristics:

- Provide for scientifically sound, cost-effective evaluations;
- Protect sensitive biological values;
- Protect healthy, natural aquatic communities;
- Support and strive for protection of chemical, physical, and biological integrity;
- May include specific characteristics required for attainment of designated use;
- Are clearly written and easily understood;
- Adhere to the philosophy and policy of nondegradation of water resource quality;
- Are defensible in a court of law.

In addition, effective biocriteria are set at levels sensitive to anthropogenic impacts; they are not set so high that sites that have reached their full potential are considered as failing to meet the criterion, nor so low that unacceptably impaired sites are rated as meeting them, which defeats the purpose of the CWA. The establishment of formal biocriteria warrants careful consideration of planning, management, and regulatory goals and the best attainable condition at a site. Balanced biocriteria will allow multiple uses to be considered so that any conflicting uses are evaluated at the outset. The best balance is achieved by developing biocriteria that closely represent the natural biota, protect against further degradation, and stimulate restoration of degraded sites.

Developing and implementing biological criteria occurs in three steps (USEPA 1996a):

1. Planning the biocriteria development program, including:
   - definition of program objectives;
   - establishment of interagency cooperation;
   - identifying acceptable levels of uncertainty for decisions made on the basis of biocriteria;
   - establishing data quality objectives.
2. Characterizing reference conditions for biocriteria and identifying candidate reference sites, which may require a biological survey.
3. Establishing biocriteria based, in part, on characterized reference conditions and designated use classes of the state.

1.7 Conceptual Framework

The central principle of biological assessment is comparison of the biological resources of a water body to a biological criterion based, in part, on a reference condition. Impairment of the water body is judged by its departure from the biocriteria. This approach presumes that the purpose of management is to prevent and repair anthropogenic; i.e., human-induced, damage to natural resources. Biological assessment of water bodies is predicated on our ability to define, measure, and compare biological integrity between similar systems. This requires an operational definition of biological integrity as follows:

> "...the condition of the aquatic community inhabiting unimpaired water bodies of a specified habitat as measured by community structure and function (USEPA 1990)."

The functional definition also requires definitions of "unimpaired" and "community structure and function", and the habitat must be specified. Community structure and function is operationally defined by the biological measures chosen for bioassessment, consisting primarily of measures of species richness, trophic diversity (relative numbers of herbivores and top carnivores), and indicator species. In addition to biological community structure and function, chemical (DO, salinity, contaminants, dissolved TOC, inorganic nitrogen, etc.) and physical (sediment composition) attributes are measured to define an unimpaired site. The combined attributes form the basis for defining reference conditions for biological criteria. When unimpaired water bodies do not exist within a region, an operational definition of unimpaired can be developed from a combination of minimally impaired estuaries and coastal waters, historical information, and professional judgment (Section 1.7.2). Figure 1-3 shows a simplified framework for progressing from an estuarine classification to assessing the health of the estuary.

1.7.1 Indicators of Biological Integrity and Survey Protocols

Several analytical approaches have been developed to assess the biological condition of waterbodies within the framework of comparison to reference, ranging in complexity from simple comparison of indicator values, to development of multivariate models:

- Comparison of indicator values — Indicator of metric values can be compared directly to the reference condition, without development of an index. This has been used most often for paleoecological comparison, where biological indicators are limited to certain indicator species, deposition rates, organic carbon loss, etc. (Turner and Rabalais 1994, Sen Gupta et al. 1996, Cooper and Brush 1991, Latimer et al. 1997).

- Multimetric index — The multimetric approach is to define an array of metrics or measures that individually provide limited information on biological status, but when integrated, function as an overall indicator of biological condition. Metrics incorporate information from individual, population, and community levels into a single, ecologically-based index of water resource quality (Gray 1989, Plafkin et al. 1989, Karr 1991). The index is typically a sum or an average of standardized scores of its component metrics (Barbour et al. 1999). Developed initially for streams, the multimetric approach has increasingly been applied to estuaries (Weisberg 1997, Hyland et al. 1998).

- Discriminant analysis to develop an index from metric values — In this approach, metrics (calculated as above) are used to develop a multivariate discriminant analysis model to distinguish reference sites from impaired sites. The calibrated model is then applied to assessment sites to determine whether they are impaired. This approach was used in EMAP-Near Coastal for the

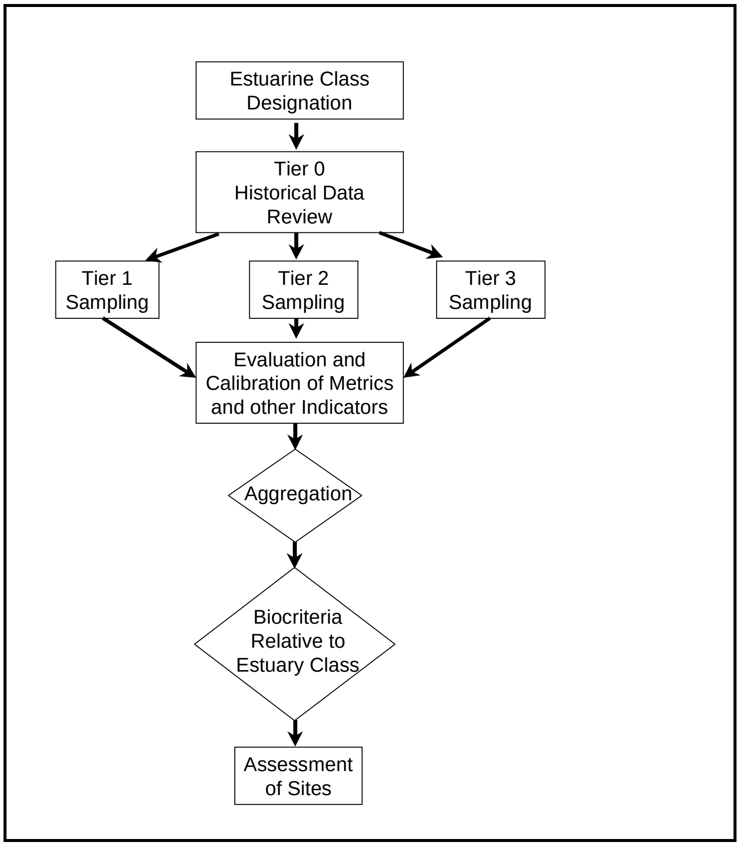


Figure 1-3

The process for progressing from the classification of an estuary to assessing the health of the estuary. Adapted from Paulsen et al. 1991.

Virginian and Gulf provinces (Paul et al. 1999, Engle et al. 1999).

- Multivariate ordination approaches — Several approaches have been developed using multivariate ordination to examine differences in species composition between reference and impaired sites. The purpose of ordination analysis is to reduce the complexity of many variables (for example, abundances of over 100 species from many estuarine sites), by re-ordering the information into fewer variables. These approaches have been used to show the effects of oil drilling in the North Sea (Warwick and Clarke 1991), and to develop an index of benthic quality in California (Smith et al. 2000).

While all of these approaches are appropriate to biocriteria development when properly applied, the multimetric approach is highlighted in this guidance. This is because it is the best developed and most extensively used method to date. Investigators should carefully consider what is most appropriate for their specific program. Time and experience will ultimately determine the best approach or combination for each state to use. Chapter 11 goes into further detail about methods of classification and assessment using all three approaches.

The multimetric concept came to fruition with the fish Index of Biotic Integrity (IBI) first conceived by Karr (1981). The IBI aggregates various elements and surrogate measures of process into a single assessment of biological condition. Karr (1981) and Karr et al. (1986) demonstrated that combinations of these attributes or metrics provide valuable synthetic assessments of the status of water resources.

A metric is a calculated term or enumeration representing some aspect of biological assemblage structure, function, or other measurable characteristic. Similarly, each of the assemblages (e.g., fish, benthic macroinvertebrates composing the aquatic community) measured would be expected to have a response range to perturbation events or degraded conditions. Thus, biosurveys targeting multiple species and assemblages; i.e., multimetric, will likely provide detection capability over a broad range of impacts, and the biocriteria derived from their results could provide protection to a large segment of the ecosystem.

Metrics can be expressed numerically as integers or ratios. Consistent routines in normalizing individual metric values provide a means of combining metric scores which initially consisted of dissimilar numerical expressions. However, final decisions on impact/no impact or management actions are not made on the single, aggregated value alone. Rather, if comparisons to established reference values indicate an impairment in biological condition, component parameters (or metrics) are examined for their individual effects on the aggregated value and for indications of potential causes.

Assessment of biological integrity using this multimetric approach typically focuses on four broad classes of community properties. Ecological systems respond to anthropogenic impacts with changes in one or more of these classes of properties (e.g., Karr et al. 1986, Schindler 1988, Plafkin et al. 1989, Schindler et al. 1989, Karr 1991, Barbour et al. 1992). The four properties are:

- **Health** of populations, typically expressed as number of individuals per m^2 or as biomass, reflecting possible stress from anthropogenic sources;

- **Community structure and composition**, or the number and kinds of species in an assemblage. Exotic species are typically undesirable, and high diversity is usually desirable. Species structure metrics include diversity and evenness indexes as well as presence of indicator species, counts of tolerant or intolerant species, the percentage of individual taxa in comparison to the total number sampled, and abundance proportions of taxonomic groups (e.g. crustaceans, mollusks, polychaetes), or comparisons of infauna vs. epifauna;

- **Trophic structure**, or the relative proportion of different trophic levels and functional feeding groups (e.g., Barbour et al. 1992). In estuaries, abundant, diverse, and relatively large top carnivores (e.g., piscivorous fish) are typically desirable as representative of a broad, stable, and substantial trophic network;

- **System function**, or the productivity and material cycling of the system or its components (trophic levels, assemblages, species). Measures of system function include primary production and standing stock biomass.

Since biological integrity is defined as an indicator of undisturbed conditions, it too must be measured relative to those conditions. The requirement of the biological criteria process for a reference by which to measure biological integrity makes it a practical tool (*sensu* Peters 1991) for managing society's impact on the natural environment.

Monitoring and assessment programs typically do not have the resources to measure all ecological attributes of concern to the public and to managers, and assessment tools must be cost-effective. Ideally, metrics selected for monitoring must be scientifically valid; should not require large amounts of expensive equipment; and should be relatively rapid in the field. The selected variables must be:

- **Related to Biological Integrity** In general, almost any biological measurement is related to biological integrity, but some are more clearly tied to the properties of biotic systems of concern to society (e.g., native species, fish production, diverse trophic structure) (Suter 1993);

- **Responsive to Environmental Stresses** Biological measurements and the metrics developed from them must respond to environmental stress. Metrics that are not monotonic; i.e., they do not consistently exhibit low values in response to one end of a stressor continuum and high values in response to the opposite end, or that respond oppositely to different stresses, are difficult to interpret in practice;

- **Measurable with Low Error** Variability and measurement error should be controllable so that a reasonable sampling effort yields sufficient precision. Index period sampling; i.e., sampling during specific time periods in the annual cycle, is one way to reduce seasonal variability. However, there are costs in terms of information derived which may be prohibitive (see later discussion on seasonality);

- **Cost-effective** Cost of a metric should be proportional to the value of the information obtained. Usually, the simplest approach is most cost-effective and should be selected so long as results are sufficient to the agency's objectives;

- **Environmentally Benign to Measure** Sampling methods that significantly disturb or alter habitats and biota should be avoided.

1.7.2 Comparison to a Reference

As noted earlier, establishing biocriteria includes determining the reference condition. The reference condition establishes the basis for making comparisons and for detecting use impairment. Because absolutely pristine estuarine and coastal marine habitats probably do not exist, resource managers must decide on acceptable levels of minimum impacts that exist or that are achievable in a given region. Acceptable reference conditions will differ among geographic regions and states because estuarine salinity gradients, trophic state, bottom sediment types, morphology and biological communities differ between regions.

Reference conditions can be established in a variety of ways. It is important to recognize that the reference condition is best developed from a population of sites, not from a single site. However, in some instances, particularly coastal environments and sites influenced by controversial land uses, the use of site-specific nearfield/farfield stations may be necessary and appropriate to augment the reference condition. They should include information derived from:

- **Historical Data** are usually available that describe biological conditions in the estuary or coastal marine region over some period of time in the past. Careful review and evaluation of these data provide insight about the communities that once existed and/or those that may be reestablished. Review of the literature and existing data is an important initial phase in the biocriteria development process. However, if data have not been collected for this specific purpose, they need to be carefully reviewed before being applied;

- **Reference Sites** are minimally impaired locations in the same or similar water bodies and habitat types at which data are collected for comparison with test sites. Reference sites could include sites that are away from point sources or concentrated nonpoint loadings; sites in sub-estuaries; sites occurring along impact gradients (nearfield/farfield); and regional reference sites that may be applied to a variety of test sites in a given area;

- **Models** include mathematical models (logical constructs following from first principles and assumptions), statistical models (built from observed relationships between variables), or a combination of the two. Paleobiological reconstructions of historic or prehistoric conditions are typically statistical or empirical models (Latimer et al. 1997, Alve 1991, Dixit et al. 1992). The degree of complexity of mathematical models to predict reference conditions is potentially unlimited with attendant increased costs and loss of predictive ability as complexity increases (Peters 1991). Mathematical models that predict biological reference conditions should only be used with great caution, because they are complex and often untestable hypotheses (Oreskes et al. 1994, Peters 1991);

- **Expert Opinion/Consensus** A consensus of qualified experts is always needed for assessing all of the above information; establishing the reference condition; and helping develop the biocriteria. This is especially the case in impaired locales where no candidate reference sites are acceptable and models are deemed unreliable. In these cases, expert consensus is a workable alternative used to establish reference "expectations". Under such circumstances, the reference condition may be defined using a consensus of expert opinion based on sound ecological principles applicable to a region of interest. The procedures for these determinations and decisions should be well documented for the record.

1.7.3 Assessment Tiers

Biological surveys of estuaries and coastal marine waters can be implemented in several tiers, ranging from a simple and inexpensive screening to detailed field sampling, analysis, and assessment. The tiered approach gives agencies one suggested approach for planning, organizing, and implementing biological surveys. Other approaches may also be available. Agencies should consider the approach that would work best to meet their program objectives. The tiers are intended to be implemented cumulatively, that is, each tier should incorporate the elements in the preceding tier as appropriate for the waters in which they are applied. Each integrated tier includes both biological and habitat components. Higher tiers require successively more effort and yield more detailed information on specific biotic assemblages and potential stresses on the system. Higher tiers reflect higher quality information and reduced uncertainty in the final assessment (Costanza et al. 1992). A desktop screening and three field survey tiers are described in this document. Figure 1-4 provides a summary of the requirements for each tier.

Tier 0 is a desktop screening assessment that consists of compiling documented information for the estuary or coastal marine areas of concern through a literature search and sending survey questionnaires to local experts. No field observations are made at this assessment level. Desktop screening should precede any of the three subsequent tiers. Its purpose is to support the planning for monitoring and more detailed assessments. Information to be compiled in Tier 0 includes: area and geomorphometric classification, habitat type, watershed land use, population density, NPDES discharges, water quality data (salinity, temperature, DO, pH, turbidity), biological assemblage data, and water column and bottom characteristics.

Tier 1 is the least complex of the survey approaches. It consists of a one-time visit to sites during a suitable, predetermined index period to collect biological and habitat data using standardized methods. The focus of this tier is on developing screening or survey information. These variables include a rudimentary identification of organisms (benthos, fish, macrophytes, or phytoplankon), water column characteristics (salinity, temperature, DO, pH, Secchi depth, water depth), and bottom characteristics (grain size, RPD layer depth, total volatile solids, and sediment toxicity). States may choose some variation of this list depending on regional characteristics and resources. Evaluation of the data collected, as well as historical data for the area, leads to an initial classification of sites and identification of candidate reference sites.

Tier 2 is somewhat more complex. A higher level of detail is incorporated into the standardized biological methods and multiple visits to the site are made to address temporal variability and/or seasonality. Another assemblage (epibenthos) could be selected in addition to those listed above. Water column nutrient measurements are added to the Tier 1 water column characteristics. A tactile categorization of grain size, plus total organic carbon, are added to the bottom characteristics. The data collected in this tier will allow the development of preliminary biological criteria.

Tier 3 is the most rigorous survey tier. Three or more assemblages are sampled here, through multiple site visits to account for seasonal variations in the selected estuarine and coastal marine biological assemblages and should incorporate supplemental studies which might be necessary for diagnostic assessment of the potential causes of observed impairments. This tier adds water column pesticides and metals measurements, plus full grain size characterization (sieving to determine percent grain size composition), acid volatile sulfides, and sediment contaminants. This tier also allows the resource agency to develop a database sufficient to support resource management activities to reduce the identified impairments and to develop and refine biocriteria.

Biological Assessment

The procedure of biological assessment is to sample two or more biological assemblages and record data such as abundance, condition, biomass, and

Figure 1-4
General comparison of Tiered Approach. Tiers are intended to be implemented cumulatively. Each tier should incorporate the elements in the preceding tier as appropriate for the waters in which they are applied, as necessary for specific programs.

Tier 0
-No field observations
-Desktop screening
-Literature search
-Questionnaires to local experts
-Support planning for monitoring and more detailed assessments

Tier 1
-One time visit to sites during suitable, predetermined index period
-Least complex survey approach
-Develop screening/survey information
-States choose variation of variables (assemblages + water column & bottom characteristics) according to regional characteristics & resources
-Leads to initial classification & ID of candidate reference sites

Tier 2
-2 or more visits to site
-More complex
-Possible to add another assemblage
-Add to water column & bottom characteristics samples
-Allows for development of preliminary biological criteria

Tier 3
-4 or more visits to sites
-Most rigorous
-3 or more assemblages
-Incorporate supplemental studies
-Additions to water column & bottom characteristics
-Develop database to support resource management activities to reduce impairments & define/refine biocriteria

other characteristics of each species. These data are then used to calculate metrics, such as taxa richness, percent dominance, number of intolerant species, and percent abundance of tolerant species. Each metric is compared to its expected value under reference conditions, and rated good (similar to reference), fair (different from reference), or poor (substantially different from reference). Numeric scores are assigned to the ratings, and the scores of all metrics of an assemblage are summed for a total score for the assemblage. The total score is again compared to the expected total score under reference conditions, and the assemblage as a whole is assigned an ordinal rating of good, fair, or poor. This second comparison to reference conditions is necessary because not all metrics are expected to score "good" at all times even in pristine conditions; the final assemblage score thus takes into account natural variability in metric values.

Once these values are satisfactorily established they can be incorporated in the development of a biocriterion for a particular estuarine or coastal marine class. "Biological assessment" at this point becomes a comparison of monitoring scores to the biocriteria for management decision making. The following several chapters describe the processes necessary to the development of suitable metrics and finally their incorporation in biological criteria for water resource management decision making.