United States
Environmental Protection
Agency

Office Of Water
(WH-556)

EPA 823-R-92-006
September 1992

EPA

Sediment Classification Methods Compendium

SEDIMENT CLASSIFICATION METHODS COMPENDIUM

Prepared by

U.S. Environmental Protection Agency
Sediment Oversight Technical Committee

EPA Work Assignment Managers

Beverly Baker and Michael Kravitz
Office of Science and Technology
Washington, DC 20460

ACKNOWLEDGMENTS

This document was prepared by the U.S. Environmental Protection Agency Sediment Oversight Technical Committee. The Sediment Oversight Technical Committee, chaired by Dr. Elizabeth Southerland of the Office of Science and Technology, has representation from a number of Program Offices in Headquarters and the Regions.

Appreciation is extended to the authors of each chapter contained in this document. Critical reviews of portions of the document were provided by the following persons: G. Allen Burton, Jr., Tom Chase, Rick Fox, Audrey Massa, George Schupp, and Howard Zar.

Assistance in preparation and production of the Compendium was provided under EPA Contract No. 68-C8-0062.

CONTENTS

Introduction

1.1 BACKGROUND

The problem of contaminated sediments is widespread in freshwater and marine systems throughout the world. Contaminated bottom sediments can have direct adverse impacts on bottom fauna. Contaminated sediments can also be a long-term source of toxic substances to the environment and can impact wildlife and humans through the consumption of food or water or through direct contact. These impacts may be present even though the overlying water meets water quality criteria. As a result, something more than the traditional water and effluent quality-based control and monitoring approaches will be needed to protect and restore the quality of the Nation's rivers, lakes, estuaries, and embayments.

In recognition of the significance of the problem, the U.S. Environmental Protection Agency (EPA) has begun a comprehensive contaminated sediment program. The effort began in 1985, when EPA examined the potential national extent of sediment contamination using existing sediment monitoring data from the EPA Storage and Retrieval System (STORET) database (Bolton *et al.*, 1985). These data were compared to organic carbon-normalized threshold concentrations calculated from existing water quality criteria using the equilibrium partitioning model. In 1986, the EPA formed the Sediment Criteria Technical Advisory Committee to examine possible approaches for deriving regulatory criteria for sediments. In 1988, EPA formed two oversight committees to take a comprehensive look at the whole range of contaminated sediment issues: the Sediment Oversight Steering Committee, which is responsible for overall management of the program, and the Sediment Oversight Technical Committee, which is oriented toward technical issues and is the implementation arm of the Steering Committee. These committees have prepared a draft outline describing EPA's Contaminated Sediment Management Strategy and have formed working groups to focus on specific issues and approaches to sediment management. The committees are also sponsoring a number of activities aimed at providing basic information about contaminated sediment issues to persons within the Agency and to the interested public. This compendium of sediment assessment methods is one of the committees' products.

An important initial step in addressing the contaminated sediments problem is the identification of scientifically sound methods that can be used to assess whether and to what extent sediments are "contaminated" or have the potential for posing a threat to the environment. The Sediment Oversight Technical Committee compiled this compendium of sediment assessment methods through the efforts of the committee members and others who are experienced in the state of the art in sediment assessment.

Many factors can affect the kinds and magnitudes of impacts that contaminated sediments have on the environment. The sediment assessment tools vary in their suitability and sensitivity for detecting these different endpoints and effects. It is, therefore, important to properly match the assessment methods to the site- and program-specific objectives of the study being conducted. The suite of assessment methods presented in this compendium offers a rich repertoire of tools from which to select the most suitable tests for a given situation.

Unfortunately, there simply is no single method that will measure all contaminated sediment impacts at all times and to all biological organisms. This is the result of a number of factors, including environmental heterogeneity and associated sampling problems, variability in the laboratory exposures, analytical variability, differing sensitivities of different organisms to different types of contaminants, the confounding effects caused by the presence of unmeasured contaminants, the synergistic and antagonistic effects of contaminants, and the physical properties of sediments. While one method will suffice for some circumstances, it is often advisable to use several complementary methods rather than a single one. When several of these approaches are used together, they can provide additional insights into the nature and degree of sediment contamination problems. The use of complementary assessment methods can provide a kind of independent verification of the degree of sediment contamination if the conclusions of the different approaches agree. If the conclusions differ, that difference indicates a need for caution in interpreting the data since some unusual site-specific circumstances may be at work. The importance of this type of verification increases with the significance of the decisions that must be made using the information obtained. In fact, the actual decision-making frameworks within which the compendium methods are used often include this verification in the concept of tiered testing.

The assessment methods presented in the compendium are continually being refined and improved. Additional methods are also being developed. As these methods are developed and verified, they will be incorporated into future updates of the compendium.

1.2 OBJECTIVE

This document is a compendium of scientifically valid and accepted methods that can be used to assess sediment quality and predict ecological impacts.

Some regulations require the use of certain types of tests (e.g., the Toxicity Characteristic Leaching Procedure under the Resource Conservation and Recovery Act), criteria (e.g., the limitations in the London Dumping Convention), and procedures (e.g., risk assessment under the Comprehensive Environmental Response, Compensation, and Liability Act). Additional guidance may be issued in the future to provide direction when addressing sediment contamination under particular regulatory programs including these, or other, required tests and approaches. These other test procedures will not be presented in this compendium, however, because the intent here is to provide the most useful overall measures or predictors of ecological impacts currently in use rather than procedures that may have limited application outside of a particular regulatory framework. Nevertheless, many of the methods presented in the compendium can be used as part of regulatory and/or remedial actions.

Guidance on how to use the compendium methods in a decision-making framework will be provided in forthcoming documents and will likely include both chemical and biological methods in a tiered hierarchical framework suitable for testing various hypotheses and endpoints. Currently such a document has been prepared by the Sediment Oversight Technical Committee to summarize existing EPA decision-making processes for managing contaminated sediments (*Managing Contaminated Sediments: EPA Decision-Making Processes*; USEPA, 1990). The information provided in the compendium on the relative strengths and weaknesses of the different assessment methods can provide assistance in selecting the appropriate methods.

1.3 OVERVIEW

The compendium is organized in the following manner. The remainder of this chapter gives a broad overview of the assessment methods in the compendium. The information is presented in tabular form to facilitate comparisons between the different methods. Chapter 2 outlines quality assurance/quality control, sampling, and analytical considerations that apply to all of the methods. Method-specific information is also provided where the procedures differ from the general ones.

The remaining chapters give specific information on each of the sediment assessment methods. The information is organized in a consistent manner for each assessment method so the reader can readily compare the relative strengths, weaknesses, and applicability of each method in order to select the best method(s) for a specific situation. The information provided for each method includes the following:

- How each method is currently used or could be used;
- A detailed description of the method, including types of data, equipment, and sampling procedures needed;
- The applicability of the method to the protection of wildlife and humans;
- The utility of the method to produce numeric sediment quality criteria;
- The method's applicability to making different types of sediment management decisions;
- The method's advantages, limitations, costs, level of acceptance, and accuracy;
- The degree to which the method is actually being used now;
- How well it is validated; and
- Its potential future uses.

Extensive references are provided after each method in case any additional details are required. The names, addresses, and telephone numbers of the authors of the descriptions of each method are provided to facilitate additional follow-up. Given the limited level of detail in the compendium, use of these references is suggested for actual implementation of the methodologies.

The 12 sediment assessment methods described in the compendium are summarized in Table 1-1. The assessment methods can be categorized in many different ways. Differentiation could be made between numeric methods and descriptive methods. Numeric methods are chemical-specific and can be used to generate numerical sediment quality criteria (SQC) on a chemical-by-chemical basis. A potential drawback of descriptive methods is that they are not chemical-specific and cannot be used alone to generate numerical sediment quality criteria for particular chemicals. On the other hand, descriptive methods can be used to directly assess the overall impact of all chemicals that may be present in a sediment, whereas it is difficult to use the chemical-specific methods to predict the combined effects of several chemicals.

Another differentiation that is often made among different sediment assessment methods is whether they are based on the measurement of the concentrations of chemicals of concern or on the measurement of biological impacts. For methods that have ecological validity, this differentiation really applies only to the practical implementation of the methods rather than to their scientific basis since all ecologically valid methods must ultimately be based on an ability to predict or measure biological effects. Many of the assessment methods use both chemical and biological testing or observation.

Yet another differentiating factor is whether the method uses interstitial water (pore water), elutriate, or bulk sediment (whole, including the solids and interstitial water). This difference also relates primarily to implementation rather than to a substantive scientific difference since the chemistry of interstitial water and that of the bulk sediment are closely linked. Except for contaminants that might be transferred directly by ingestion, interstitial water is the medium through which the contaminants in the bulk sediment are transferred to the affected organisms.

Some of the assessment *methods* (which would be more accurately characterized as *approaches*) described in the compendium combine numeric and descriptive measures. For example, the Sediment Quality Triad (Triad) and Apparent Effects Threshold (AET) approaches employ bulk sediment toxicity testing, benthic community structure analysis, and concentrations of sediment contaminants. The Triad is both descriptive and numeric, depending on its use. Typically, the Triad approach has been used in a descriptive manner to identify contaminated sediments. It has also been used, however, to generate criteria for several chemical contaminants. The International Joint Commission (IJC) approach would be more accurately described as an assessment strategy since it employs several of the other sediment assessment methods in a tiered, comprehensive

Table 1-1. Some Characteristics of the Sediment Assessment Methods.

Sediment Method (Chapter Number)	Description
Bulk Sediment Toxicity (3)	Test organisms are exposed to sediments that may contain unknown quantities of potentially toxic chemicals. At the end of a specified time period, the response of the test organisms is examined in relation to a specified biological endpoint.
Spiked-Sediment Toxicity (4)	Dose-response relationships are established by exposing test organisms to sediments that have been spiked with known amounts of chemicals or mixtures of chemicals.
Interstitial Water Toxicity (5)	The toxicity of interstitial water is quantified and identification evaluation procedures are applied to identify and quantify chemical components responsible for sediment toxicity. The procedures are implemented in three phases to characterize interstitial water toxicity, identify the suspected toxicant, and confirm toxicant identification.
Equilibrium Partitioning (6)	A sediment quality value for a given contaminant is determined by calculating the sediment concentration of the contaminant that would correspond to an interstitial water concentration equivalent to the U.S. EPA water quality criterion for the contaminant.
Tissue Residue (7)	Safe sediment concentrations of specific chemicals are established by determining the sediment chemical concentration that will result in acceptable tissue residues. Methods to derive unacceptable tissue residues are based on chronic water quality criteria and bioconcentration factors, chronic dose-response experiments or field correlations, and human health risk levels from the consumption of freshwater fish or seafood.
Freshwater Benthic Macroinvertebrate Community Structure and Function (8)	Environmental degradation is measured by evaluating alterations in freshwater benthic community structure and function.
Marine Benthic Community Structure (9)	Environmental degradation is measured by evaluating alterations in marine benthic community structure.
Sediment Quality Triad (10)	Sediment chemical contamination, sediment toxicity, and benthic infauna community structure are measured in the same sediment. Correspondence between sediment chemistry, toxicity, and biological effects is used to determine sediment concentrations that discriminate conditions of minimal, uncertain, and major biological effects.
Apparent Effects Threshold (11)	An AET is the sediment concentration of a contaminant above which statistically significant biological effects (e.g., amphipod mortality in bioassays, depressions in the abundance of benthic infauna) would always be expected. AET values are empirically derived from paired field data for sediment chemistry and a range of biological effects indicators.
International Joint Commission Sediment Assessment Strategy (12)	Contaminated sediments are assessed in two stages: (1) an initial assessment that is based on macrozoobenthic community structure and concentrations of contaminants in sediments and biological tissues and (2) a detailed assessment that is based on a phased sampling of the physical, chemical, and biological aspects of the sediment, including laboratory toxicity bioassays.
Sediment-Testing Approach Used for Ocean Disposal (13)	A tiered testing strategy consisting of physical, chemical, and biological testing to predict benthic and water column impacts of dredged sediment disposal.
National Status and Trends Program Approach (14)	Three ranges of concentrations are determined for each chemical: the no-effects range, the possible-effects range, and the probable-effects range. These values are arithmetically determined from a database consisting of matching chemical and biological data from laboratory spiked-sediment bioassays, equilibrium-partitioning models, and field studies.

procedure. The Sediment-Testing Approach Used for Ocean Disposal is the tiered, comprehensive testing procedure developed by EPA and the U.S. Army Corps of Engineers (USACE) for determining the suitability of dredged material for disposal at designated disposal sites. The procedure is specified in *Evaluation of Dredged Material Proposed for Ocean Disposal-Testing Manual*, commonly referred to as the 1991 Green Book (USEPA/USACE, 1991).

To facilitate the user's selection of the most suitable sediment assessment method, Tables 1-2 through 1-5 highlight the major characteristics of each method. Information from individual chapters that is useful in management decisions is presented in summary form and includes method descriptions and uses, data and sampling required, ability to generate numerical sediment quality criteria, and outlook for future use. More pointedly, the reader will learn what each method predicts, what it assumes, how much it will cost, and why one might choose a particular method over another for a specific situation.

Regardless of which of the compendium methods one uses, several considerations must be addressed: a sampling program needs to be designed; samples need to be collected, stored, and analyzed; and quality assurance/quality control is needed throughout the process to determine the uncertainty associated with the results of the assessment. Sampling design and QA/QC issues will be discussed in Chapter 2.

1.4 REFERENCES

Bolton, S.H., R.J. Breteler, B.W. Vigon, J.A. Scanlon, and S.L. Clark. 1985. National perspective on sediment quality.

USEPA. 1990. Managing contaminated sediments: EPA decision-making processes. U.S. Environmental Protection Agency, Sediment Oversight Technical Committee. EPA 506/6-90/002.

USEPA/USACE. 1991. Evaluation of dredged material proposed for ocean disposal—Testing manual. U.S. Environmental Protection Agency and U.S. Army Corps of Engineers.

Table 1-2. Summary of Sediment Methods and Applications.

Sediment Method	Current Use	Ability to Generate Numerical SQC	Potential Use	Protects Human Health, Aquatic Life, and/or Wildlife	Type of Sampling Required
Bulk Sediment Toxicity Test	Measures total toxic effect of all contaminants.	No.	Determines toxicity. Can generate SQC in combination with other methods.	Aquatic life.	Field-collected bulk sediments.
Spiked Sediment Toxicity Test	In research state.	Yes.	Can address interactions of chemical mixtures.	Human health, aquatic life, and wildlife.	Field-collected sediments contaminated or uncontaminated.
Interstitial Water Toxicity Test	Several aquatic species, marine and freshwater.	Yes, in combination with TIE procedures.	Provides important toxicity data, particularly in combination with other sediment classification methods.	Aquatic life.	Field-collected bulk sediments.
Equilibrium Partitioning	Regulatory uses of Equilibrium Partitioning (EqP)-based SQC under development.	Yes. Interim SQC for some chemicals have been developed.	EqP-based SQC have a possible major role in the identification, monitoring, and cleanup of contaminated sites.	Human health, aquatic life, and wildlife.	Sediment chemistry, total organic carbon concentrations.
Tissue Residue	Some use in remedial and regulatory actions.	Yes. Most applicable for nonionic organic and organometallic compounds.	Will provide excellent measure of "effective exposure dose."	Human health, aquatic life, and wildlife.	Sediment chemical and physical characteristics. Biota sampling for residue analysis.
Freshwater Benthic Community Structure and Function	A number of uses, including the establishment of criteria and standards.	No.	Will be most successful in concert with sediment chemistry and toxicity results.	Directly applicable to aquatic life and some wildlife, and indirectly to human health and other wildlife.	Sediment collection using a grab sampler.
Marine Benthic Community Structure	Describes reference conditions, baseline conditions, and effects of natural and anthropogenic disturbances.	Not alone. Integral component of AET and Sediment Quality Triad.	Potential for identifying species that are indicative of sediment contaminants at various concentrations.	Directly applicable to aquatic life and some wildlife, and indirectly to human health and other wildlife.	Sediment collection using a grab or core sampler.

Table 1-2. Summary of Sediment Methods and Applications. (Continued)

Sediment Method	Current Use	Ability to Generate Numerical SQC	Potential Use	Protects Human Health, Aquatic Life, and/or Wildlife	Type of Sampling Required
Sediment Quality Triad	Determines extent of pollution-induced degradation. Determines numerical SQC.	Yes. Used for lead, PAHs, and PCBs.	Identifying problem areas, prioritizing and ranking degraded areas, and predicting where degradation will occur.	Aquatic life directly, wildlife and human health directly and indirectly.	Field-collected sediment. Five field replicate benthic samples recommended.
Apparent Effects Threshold	Used by several programs to develop guidelines for protection of aquatic life in Puget Sound.	Yes.	Identifying problem areas, identifying problem chemicals in sediments, and focusing cleanup activities. Screening sediments in regulatory programs.	Aquatic life.	Field-collected sediments from 50 stations or more recommended. Conduct chemical tests for a wide range of chemical classes.
International Joint Commission	Intended as guidance for assessment of contaminated sediments in the Great Lakes.	Yes.	Evaluation of Areas of Concern. Possible use outside of Great Lakes basin.	Directly to aquatic life, and indirectly to wildlife and human health.	Bulk sediment collection, benthic community structure, fish contaminant body burdens, and external abnormalities.
1991 Green Book	Guidance for dredging applicants, scientists, and regulators.	Field-validates SQC currently under development.	Will be applied to dredged material evaluations for the foreseeable future.	Directly to aquatic life, wildlife, and human health.	Comprehensive sampling plan for sediment and water.
National Status and Trends Program Approach	Initially used to develop informal guidelines for use by the NS&T Program.	Yes.	Identify toxic chemicals in sediments; rank and prioritize areas for further study; assess potential ecological hazards of contaminated sediments; design spiked sediment bioassays; describe toxic effects associated with certain chemical concentrations; quantify likelihood of toxicity for range of chemical concentrations.	These guidelines provide an estimate of effects on benthic life. They are not intended to be used for the protection of human life or wildlife.	Minimizes the need for additional sampling through the use of existing data.

Table 1-3. Summary of Sediment Methods and Suitability.

Sediment Method	Summary of Method	Types of Data Required	Suitability					
			Different Sediment Types	Different Chemicals	Predicting Effects on Organisms	In-Place Pollutant Control	Source Control	Disposal Applications
Bulk Sediment Toxicity Test	Exposes test organisms to field-collected sediments.	Physical, chemical, biological.	Any type.	All classes and combinations of chemicals.	In theory, can use any organism.	Can directly monitor in-place pollution.	Can identify suspected sources.	Widely used to determine toxicity prior to disposal.
Spiked Sediment Toxicity Test	Adding ("spiking") sediments with one chemical or a mixture.	Test, reference, and control sediment data. Physical, chemical, and biological data.	Any type.	All classes and combinations of chemicals.	In theory, can use any organism.	Can be used in developing criteria. Can identify extent of problem, monitor trends, and set target cleanup goals.	Can be used in combination with wasteload allocation models to establish maximum allowable effluent concentrations.	Can determine toxicity prior to disposal.
Interstitial Water Toxicity Test	Pore water preparation, toxicity tests, and TIE procedures.	Physical, chemical and biological response, identification of toxic compounds.	Any type from which adequate quantities of pore water can be obtained.	Water-soluble nonionic organics, cationic metals, and ammonia, and their interactions.	Predicts impacts on organisms once toxicant responsible for toxicity is identified.	Identifies sediment toxicants and can design remediation plans. Routine monitoring.	Ideal for point source controls and controllable nonpoint sources.	Can determine toxicity prior to disposal.
Equilibrium Partitioning	Predicts chemical concentration in interstitial water and compares it to chronic water quality criteria.	Bulk sediment analysis and concentration of total organic carbon.	After development of test, expected to apply to wide range of sediment types.	Modifications of method exist for different classes of chemicals.	Can predict toxic effects for a range of representative organisms.	Can identify sources of contamination and identify target cleanup levels.	Predicts concentrations of a chemical above which adverse impacts are likely.	Suitable for addressing aquatic disposal. Unsuitable for addressing upland disposal sites.

Table 1-3. Summary of Sediment Methods and Suitability. (Continued)

Sediment Method	Summary of Method	Types of Data Required	Suitability						
			Different Sediment Types	Different Chemicals	Predicting Effects on Organisms	In-Place Pollutant Control	Source Control	Disposal Applications
Tissue Residue	Links toxic effects to residues, and links chemical residues in organisms to sediment chemistry.	Identification of chemical in sediment through screening of aquatic organisms for residues.	Any type.	Nonionic organics and organometallics that bioaccumulate.	Not limited by organism unless organism residues cannot be obtained.	Provides established assessment method for human health and ecological risks.	Provides strong support for establishing controls in point and nonpoint sources of contamination.	Suitable for use along with partitioning and bioaccumulation models.
Freshwater Benthic Community Structure and Function	Field survey, collection, sorting, and identification of benthic organisms.	Varies from a list of families of taxa present to species-level taxonomy and enumerations.	Any type, but only similar types should be compared.	Many individual chemicals and classes of chemicals.	Facilitates use of benthic macroinvertebrates as indicator organisms.	Can be used to screen for potential sources of contamination.	Extensively used for source characterization and control.	Advised for areas suitable for open-lake disposal.
Marine Benthic Community Structure	Collection, sorting, and identification of benthic organisms.	Number of taxa, abundance of each taxon, and biomass and conventional sediment chemistry variables.	Any type, but only similar types should be compared.	Applies to general categories with exceptions based on level of organic enrichment.	Facilitates use of benthic macroinvertebrates as indicator organisms. Research is needed to predict specific effects on potential predators.	Has not been used to set sediment quality criteria for polluted marine sediments.	Limited value in specific source characterization.	Not required in testing of sediment to be dredged under sections 401 and 404 of the Clean Water Act.
Sediment Quality Triad	Uses sediment chemistry, sediment bioassays, and *in situ* biological variables.	Sediment chemistry, sediment toxicity, and benthic infauna data.	Any type.	All chemicals and classes.	All biological effects data based on a single species.	A comprehensive approach that allows for all potenial interactions between chemical mixtures and the environment.	Comprehensive and complements TIE programs for effluents.	History of regulatory use.

Table 1-3. Summary of Sediment Methods and Suitability. (Continued)

Sediment Method	Summary of Method	Types of Data Required	Suitability: Different Sediment Types	Suitability: Different Chemicals	Suitability: Predicting Effects on Organisms	Suitability: In-Place Pollutant Control	Suitability: Source Control	Suitability: Disposal Applications
Apparent Effects Threshold	Collection of "matched" chemical and biological effects data. Statistically tests significance of adverse biological effects relative to reference conditions.	Statistical analysis of biological effects relative to reference conditions. Generation of AET for each chemical and biological indicator.	Any type.	All chemicals and classes.	Any life stage of any marine or aquatic organism for which a biological response can be determined.	Use as a predictive tool, in the designation of problem areas, and as a database for remedial action.	Well-suited for identifying problem areas and designing source controls.	Generated AET values can predict whether adverse biological effects will occur after disposal of dredged material at aquatic sites.
International Joint Commission	Stage I - *In situ* assessment, physical, chemical and benthic community structure. Stage II - Detailed 4-phase assessment.	Physical characteristics, chemical concentrations in sediments and tissues, benthic community structure, and external abnormalities.	Any type.	Most chemicals in Great Lakes sediments.	Tailored to analysis of indigenous organisms.	Developed specifically for assessment of in-place pollutant problems.	Identifying hot spots and establishing significant differences from background conditions.	Initial disposal decisions.
1991 Green Book	Tiered-testing procedures to characterize dredged material and predict its impact. Include physical and chemical sediment evaluations, toxicity, and bioaccumulation studies.	Physical and chemical sediment data. Bioassay, bioaccumulation, and field species data.	Any type with exception of extremely coarse-or angular-grain sediments.	Wide range of organic and inorganic chemicals.	Considers effects on marine organisms representative of organisms indigenous to Ocean Dredged Material Disposal Sites.	Developed to determine water-column and benthic Limiting Permissible Concentration (LPC) compliance for dredged material.	Not intended purpose but may be useful.	Used in decision-making for ocean-disposal management.

Table 1-3. Summary of Sediment Methods and Suitability. (Continued)

Sediment Method	Summary of Method	Types of Data Required	Suitability					
			Different Sediment Types	Different Chemicals	Predicting Effects on Organisms	In-Place Pollutant Control	Source Control	Disposal Applications
National Status and Trends Program Approach	From a database of matching chemical and biological data, derive 3 ranges of chemical concentrations: the no-effects range, the possible-effects range, and the probable-effects range.	Matching biological and chemical data from laboratory spiked-sediment bioassays, equilibrium partitioning models, and field studies.	Can be applied to any sediment type occurring in freshwater, estuarine, and marine sediments.	Can be applied to a wide variety of chemicals.	Widely applicable to benthic organisms.	Provide a basis for evaluating existing sediment chemical data and ranking areas and chemicals of concern. Can identify the need for further investigations to support regulatory decisions.	Provide a credible and defensible basis for evaluating contaminants in taking source control action.	Can provide an additional assessment tool.

Table 1-4. Summary of Sediment Methods and Ease of Use.

Sediment Method	Ease of Use	Relative Cost	Tendency to Be Conservative	Level of Acceptance	Necessary Level of Effort	Interpretability of Results
Bulk Sediment Toxicity Test	Most are simple. Some require special training.	$150 - $500 per sample replicate.	Tests can be made as sensitive or as conservative as necessary.	Widely accepted.	Relatively small; field sampling and laboratory toxicity tests.	Easily interpretable. Biological data subjected to "pass-fail" or some explanation.
Spiked Sediment Toxicity Test	Most are simple. Some require special training.	$100,000 for chemical and toxicity data to establish SQC for one chemical.	High degree of accuracy. Inherently limited in ability to reflect all ecological processes affected by contaminants.	Widely accepted, with peer review.	Field sampling, laboratory toxicity tests, and calculation of data.	Provides cause-effect relationships.
Interstitial Water Toxicity Test	Straightforward analyses plus highly sensitive instrumentation.	Case-specific.	As sensitive or as conservative as necessary.	Sound theoretical basis.	Field sampling, pore water preparation, toxicity tests, and TIE procedures.	Results easily interpreted.
Equilibrium Partitioning	Calculations are straightforward with necessary data.	Dependent on cost of collecting site-specific chemical data.	Levels of protection of SQC similar to those of water quality criteria -- deemed protective of 95% of organisms.	Wide acceptance.	Varies. No site-specific biological data required.	Requires interpretation but provides pertinent information.
Tissue Residue	Straightforward.	Cost of SQC generally incurs low analytical costs.	Does not tend to be either conservative or liberal.	Accepted as a basis for regulatory decisions.	Varies from none to large.	Varies by number and nature of contamination, complexity of distribution, and regulatory application.
Freshwater Benthic Community Structure and Function	Equipment and materials inexpensive and minimal. Organisms difficult to sort and identify.	$700 per sample site.	High.	Wide acceptance from a historical perspective.	Results can be generated within 1 day.	Data interpretation requires an expert. Results are easily incorporated into a management strategy.

Table 1-4. Summary of Sediment Methods and Ease of Use. (Continued)

Sediment Method	Ease of Use	Relative Cost	Tendency to Be Conservative	Level of Acceptance	Necessary Level of Effort	Interpretability of Results
Marine Benthic Community Structure	Field collection of sediments, extensive laboratory work, data analysis and interpretation.	$400 - $1,000 per replicate.	Moderate.	Wide acceptance from a historical perspective.	Field effort, laboratory work, and analysis may take several months to a full year.	Data interpretation requires an expert.
Sediment Quality Triad	Straightforward, with a high level of expertise required to collect data.	Proper implementation requires substantial resources. Nonetheless, a cost-effective method.	Empirical, field-based nature of test precludes definitive predictions of tendency.	Wide.	Correlated to different levels of results. Fieldwork and synoptic sampling provide most useful results.	Expert judgment required.
Apparent Effects Threshold	Straightforward.	$200 - $1,800 per station.	Highly conservative.	Accepted by several federal and state agencies in the Puget Sound region.	Field-data collection, data entry and verification, and data comparisons.	Test interprets matched biological and chemical data.
International Joint Commission	Tailored to the area under investigation. Intended to be flexible.	Costly. Up to $500,000 for a complete assessment.	Highly conservative.	A combination of widely accepted individual methods.	Relatively high.	Complex analysis and interpretation required by trained investigators.
1991 Green Book	Expertise required. Intended to be flexible.	Difficult to estimate. Tiers I, II, III, and IV ordered by increasing complexity and cost.	Highly conservative.	Widely accepted.	Low in Tier I, relatively high in Tiers III and IV.	Data analysis is complex.
National Status and Trends Program Approach	Approach relies upon existing data. Guidelines can be developed relatively quickly and easily.	Use of the existing database is simple and quick. If the necessary data must be generated, costs could be relatively high.	Predictive capabilities of the guidelines have not been quantified.	Broad acceptance and growing interest.	Level of effort toward the development of original database relatively high. Subsequent use of the database is relatively easy.	Provides the user some flexibility in use and interpretability of the guidelines. All of the data are presented, and the degree of certainty can be assessed by the user.

Table 1-5. Summary of Sediment Methods and Extent of Use.

Sediment Method	Environmental Applicability	Accuracy and Precision	Extent of Use	Extent of Field Validation	Outlook for Future Use
Bulk Sediment Toxicity Test	Wide range of sediment types and environmental conditions.	High.	Wide.	Some field validation; more is necessary.	• Promising for direct measurement of biological effects. • More emphasis needed on measurement of chronic effects. • Methods should be standardized across laboratories. • Central database necessary.
Spiked Sediment Toxicity Test	Wide range of sediment types and environmental conditions.	High.	Under development. In the process of standardization by ASTM's sediment toxicity subcommittee.	Some field validation; more is necessary.	• Promising where direct dose-response data required. • Development and standardization across laboratories necessary. • Central database necessary.
Interstitial Water Toxicity Test	All sediment types and environmental conditions.	High.	Wide acceptance for freshwater and marine applications.	Little field validation; more is necessary.	• Extremely promising; only method that directly includes the identification of compounds responsible for toxicity. • Further development needed.
Equilibrium Partitioning	EqP-based SQC apply to sediments with greater than 0.2% organic carbon and nonionic chemicals for which criteria are available.	Each EqP-based SQC will have associated degree of uncertainty.	Under EPA review for regulatory uses of EqP-based SQC.	Some field validation; more is necessary.	• Only procedure for derivation of SQC that is generic across sediments, accounts for bioavailability, and relates effects to chemicals.
Tissue Residue	All types.	Generally high.	Wide acceptance.	Some field validation for individual chemicals, none for chemical mixtures.	• Can be implemented with minimal effort. • Central database should be developed. • Field validation of residue-based ecological effects predictions essential.

Table 1-5. Summary of Sediment Methods and Extent of Use. (Continued)

Sediment Method	Environmental Applicability	Accuracy and Precision	Extent of Use	Extent of Field Validation	Outlook for Future Use
Freshwater Benthic Community Structure and Function	Lotic and lentic fresh-water ecosystems.	High.	Wide acceptance.	An *in situ* study; therefore, consistently and accurately assesses environmental quality.	• Outlook good because benthic macroinvertebrates provide substantial information that chemical and toxicity data alone cannot provide. • Development most needed in combining benthic community assessments with chemical and toxicological data.
Marine Benthic Community Structure	Direct measure of environmental effects.	High; if necessary replicates are obtained.	Valued tool for several decades.	An *in situ* study; therefore, consistently and accurately assesses environmental quality.	• Outlook bright with continued development toward new data analysis methods to reduce cost or variability within data.
Sediment Quality Triad	Extremely high.	Not quantitatively determined; expected to be high.	Recently developed. Has been used to identify degraded areas.	By its nature an *in situ* study; therefore, automatically field-validated.	• High potential. Provides objective information to judge extent of pollution-induced degradation. • Method development and standardization necessary.
Apparent Effects Threshold	High.	Sensitive and efficient. The number of stations used has a marked effect on AET uncertainty.	Used by Puget Sound agencies for regulatory guidelines. Also widely used by others.	Field-validated for Puget Sound. Further testing desired before application of Puget Sound AETs to other geographic regions.	• High potential for regional use.
International Joint Commission	High.	Not quantitatively determined. Expected to be high.	Published in 1989. Individual methods widely used and accepted.	First field validation in 1989-1991 as part of EPA ARCS program.	• Potential for widespread use in Great Lakes basin and elsewhere.
1991 Green Book	High.	Strongly supports extensive QA program.	Guidance will be applied to all evaluations for dredged material that is proposed for disposal outside of the baseline of the territorial sea.	Large portions were field-validated in the past; additional projects planned.	• EPA and USACE continue to support the guidance nationally and regionally. • Ongoing public and private research and development with concomitant document updates.

Table 1-5. Summary of Sediment Methods and Extent of Use. (Continued)

Sediment Method	Environmental Applicability	Accuracy and Precision	Extent of Use	Extent of Field Validation	Outlook for Future Use
National Status and Trends Program Approach	Highly applicable to the interpretation of environmental data.	Once the minimum number of data sets is determined to develop consistent guidelines, the variability is minimal. Accuracy in predicting toxicity has not been determined.	Has been used by NOAA's National Status and Trends Program, Environment Canada, and the Florida Department of Environmental Regulation. A variation of the approach isbeing developed by the California Water Resources Control Board.	Validations have not yet been quantified.	• Outlook is good. Since the approach relies on existing data, other region-specific guidelines could be easily developed using region-specific data. • Approach can be used to validate criteria determined with other single-method approaches. • Several types of data are needed to further develop the approach.

Quality Assurance/Quality Control, Sampling, and Analytical Considerations

The purpose of this chapter is to provide a brief introduction to some of the most important terms and concepts that are integral to the design of an adequate program for sediment sample collection, handling, and analysis. This chapter is intended only as a general guide to sediment sampling and should not be used as an instruction manual for collecting samples. The subjects mentioned will not be dealt with in an exhaustive manner. The reader is referred to the references cited in this chapter for more complete guidance on the particular techniques.

2.1 ESTABLISHING DATA QUALITY OBJECTIVES

Fundamental to the process of designing a study is the establishment of data quality objectives (DQOs). The most carefully collected and analyzed data are of no use if the data collected are insufficient or of the wrong type. To avoid either of these and other potentially costly errors, EPA has initiated the use of the DQO Process. The DQO Process is a management tool designed to help data users and data collectors design the best sampling strategy to reach their objectives while minimizing resource requirements. It is a multistep, systematic approach to data collection that enables the manager to refine goals and objectives and help answer the question, "How much data is enough?" As the steps of the DQO Process are followed, the decisions made in previous steps should be reviewed to ensure consistency and cohesiveness.

The first step in the process is to specify the problem and identify limitations of time or resources on the data-collection effort. This process allows one to evaluate his or her current knowledge base of the problems and identify all available resources. The next step is to identify what decisions or activities will be made based on the data. The answer to this question is vital to ensure the collection of the right type of data. The decision goals should be as narrow in scope as possible, and considerable effort may be required to define them properly.

The third step involves identifying all variables needed to make a decision. This step focuses on eliminating the potential measurement or collection of data that may not actually be used in the decision-making process. The next step requires the data collector to set or define the boundaries of the study, including the population, which could consist of people, objects, or media, and the boundaries on the population, including space, time, and area.

Developing a decision rule, or how the data will be used and summarized, is the next step in the process. This step involves describing how the study results will be compiled or calculated and defining the decision rule in an "If ..., then ..." format. The statement should incorporate the study results as "If the results are this, then the action should be this." For example, "If PCB levels in fish are greater than 2 ppm, then a fish consumption advisory will be issued." This step, along with the others, helps define the data collection effort by identifying the data needed to fulfill the decision rule.

A very important step in the DQO Process is specifying the limits of uncertainty acceptable in the data. These limits can be expressed as acceptable false-positive and false-negative error rates for the decision. These error rates must be based on careful consideration of the consequences of incorrect conclusions being drawn from the data. The definitions of false-positive and false-negative errors vary with the decision being defined. If a decision to take regulatory action is being made, a possible false-negative error could result in no action being taken because incorrect data results indicated there was no problem. The opposite could also occur, where a false positive error results in regulatory action being taken when no problem exists. It is essential that the potential consequences to economic, health, ecological, political, and social issues be considered when deciding on acceptable false-positive and false-negative error rates. This step may involve the consultation of a qualified statistician.

Finally, all steps in the DQO Process should be reviewed to design the most efficient sampling study. Considerations including cost, time, defined boundaries, the decision rule, and all other factors defined and specified during the DQO Process should be incorporated.

One can refer to "Planning Issues for Superfund Site Remediation" in *Hazardous Material Control* (Ryti and Neptune, 1991) for an excellent example of applying the DQO Process to an actual situation.

Quality assurance and quality control are integral components of every aspect of a program's activities. The collection of reliable data is contingent on the use of and adherence to a good Quality Assurance Project Plan; the development of a sound sampling study is contingent on the use of the DQO Process; and use and implementation of the DQO Process is contingent on a Quality Assurance Program Plan.

2.2 SAMPLING DESIGN

2.2.1 Test, Reference, and Control Sediments

In sediment quality evaluations, there is a substantial precedent for using comparisons between sites rather than comparison of testing results to an independently set numerical benchmark. This is the result of a number of factors including the standard procedures used in biological testing, the paucity of scientifically acceptable numerical sediment quality criteria or standards, and the long-standing "nondegradation" philosophy used in evaluating the acceptability of dredged material for open-water disposal. The degree of sediment contamination in a particular area is often evaluated by comparing the structure of benthic communities, levels of pollutants, or bioassay test results in sediments collected from the area being investigated with those in the surrounding area. The terms used to describe the different sediments in the comparisons are *test* sediments, *control* sediments, and *reference* sediments.

As used in sediment assays and assessments, a *test* sediment is sampled from the area whose quality is being assessed. A *control* sediment is a pristine (or nearly so) sediment, free from localized anthropogenic inputs of pollutants with contamination present only because of inputs from the global spread of pollutants (Lee *et al.*, 1989). A control sediment is fully compatible with the needs of the organisms used in the assay, is known to not cause toxicity, and is used primarily to verify the health of the test organisms and the acceptability of the test conditions (USEPA/USACE, 1991). The control sediment may be artificially prepared in order to achieve sufficient volumes of a known and consistent quality for use in standard testing and for culturing test organisms (ASTM, 1990).

A *reference* sediment, on the other hand, is collected from a location that may contain low to moderate levels of pollutants resulting from both the global inputs and some localized anthropogenic sources, representing the background levels of pollutants in an area (Lee *et al.*, 1989). The reference sediment is to be as similar as possible to the test sediments in grain size, total organic carbon (TOC), and other physical characteristics (Lee *et al.*, 1989; USEPA/USACE, 1991; ASTM, 1990). The physical environment of the reference site should also be as similar as possible to that at the sites where the test sediments will be collected. This is especially significant for benthic community structure comparisons, since community structure can be very significantly affected by water depth, physical transport processes such as waves and currents, sediment grain size, and the presence of organic debris.

As used in dredged material assessment, the results of assays or evaluations on the test sediments are compared to those obtained from reference sediments to determine whether the test sediments are contaminated. In contrast, the results of assays or evaluations using the control sediments are usually compared only to past results using those same control sediments to ensure that the testing was free of some extraneous factors that may have affected the reliability of the test. Depending on the study objectives, however, controls can also be used as a benchmark against which to compare test sediments to determine the relative degree of contamination of sediments collected from different sites (ASTM, 1990).

A clear understanding of the end uses of the data is essential in the establishment of an appropriate sampling program. A cost-effective study for a qualitative overview of potential contaminated sediment impacts will differ markedly from one whose purpose is to make statistically-based numerical comparisons with criteria or indexes, or to reference sites.

Sediment sampling programs are most often undertaken to achieve one or more of the following objectives:

- To fulfill a regulatory testing requirement;
- To determine characteristic ambient levels;
- To monitor trends in contamination levels;
- To identify hot spots of contamination; or
- To screen for potential problems.

These different objectives will lead to different sampling designs. For example, a study for a dredging project may have a specific set of guidelines on sampling frequency, sample site selection methodology, and other parameters already determined by existing specific guidance. The design for a study to determine ambient levels will strive to obtain uniform, random coverage of an area through the collection of samples from a relatively large number of sites. The design for a study to track sediment contamination trends will expend its resources to sample fewer sites but more often. A study to identify hot spots would concentrate efforts on fewer sites within zones most likely to be contaminated, while an initial screening study might take very few, randomly distributed samples for analysis together with some "observation" samples to supplement the analytical results.

Available information about the area to be sampled and its surroundings should be used in determining the final sample design. Knowledge about bottom topography, currents, areas of dredging and the frequency of dredging, locations of point and nonpoint sources of contaminants, distribution of grain sizes, and other factors can provide the basis for determining which of the sampling designs to use (e.g., Are there reasons to expect localized hot spots of contamination?) and where to place sampling locations (e.g., Which parts of the area are likely to be similar enough to group into the same strata?). Preliminary surveys of an area using depth-sounding and sediment-profiling equipment can prove invaluable in delineating vertical and horizontal distributions of sediments (IJC, 1988). This information can be helpful in planning sediment sampling methods (grab samples or core samples) and sample site selection (grouping similar areas into strata, identifying likely locations of hot spots).

The methods most often used for selecting the sample collection sites are haphazard, worst-case, random, stratified random, and exhaustive (Higgins, 1988).

2.2.1.1 Haphazard

The haphazard method, whereby one selects sampling sites based on whim or ease of implementation rather than science or knowledge, really reflects the lack of a design. This method has no validity and should not be used.

2.2.1.2 Worst-Case

The worst-case sampling design is based on knowledge regarding the presence and distribution of potential sources of sediment contamination in an area. It is usually considered cost-effective as long as the study objectives are being met. An inherent problem with this design is that it results in an incomplete characterization of an area and is not statistically robust. However, it can be useful as an initial survey to determine the potential for a contamination problem, which would be followed up with more complete sampling later, if needed. The effectiveness of this technique depends on the availability of reliable historical information on contamination, sources, bathymetry, currents, and other factors.

2.2.1.3 Random

The random sampling design is most useful for cases where little is known about the likely distribution of sediment contamination or sources, or when available information indicates a high degree of homogeneity in an area. The area to be sampled is divided using a grid system. Samples are distributed within the grid randomly, with each location having an equal probability of being sampled. The number of samples is selected statistically based on the requirements of the survey and the acceptability of false-positive or false-negative results. This design yields statistically sound results.

2.2.1.4 Stratified Random

The stratified random design is a variation on the previous two designs. Available information is used to identify different zones that are likely to be similar in degree of contamination or other characteristics. Samples sites are then randomly selected within the different zones. This design also yields statistically reliable results.

2.2.1.5 Exhaustive

In the exhaustive design, an area is subdivided into equal-sized units, each of which is then sampled. This design yields a very complete characterization. However, this design is usually very costly because of the large number of samples that need to be collected.

2.2.2 Numbers of Samples

Statistics can be used to determine the number of samples needed. To use statistics in this way, one needs to decide what comparisons will be made with the resulting data and what will be the desired statistical power of the comparisons (i.e., at what level of confidence will resulting differences be tested). In addition, one needs some information about the inherent environmental variability in the area (i.e., the likelihood that an observed difference is due to an actual difference in contamination rather than just the natural heterogeneity in sediment or benthic population characteristics in the area). There are many different statistical approaches to estimating the number of samples required and to interpreting the resulting test results. Excellent reviews of statistical designs and interpretation are given by Baudo (1990) for sediment physical and chemical testing and by Downing and Rigler (1984) for benthic community structure evaluations.

In practice, constraints on resources often preclude the use of a purely statistical approach to determining the number of samples and some form of a cost-benefit approach is often used to arrive at a reasonable compromise between statistical power and the cost of the study. One of the major advantages of the tiered approaches for testing and assessment is the cost savings that results when information is collected relatively inexpensively initially and additional resources are expended only when the information collected thus far is insufficient to make a decision.

Guidance on how to select a cost-effective approach is usually provided in very general qualitative terms as to the factors that should be considered in arriving at a decision (USEPA/USACE, 1991; Higgins, 1988; Plumb, 1981). Decisions are largely subjective. However, researchers at EPA's Environmental Research Laboratory (ERL)-Narragansett/Newport recently developed a four-step procedure to determine the optimal cost-effective sampling scheme for marine benthic community assessment (USEPA, undated). The procedure begins with an initial limited sampling using two or more sampling schemes at paired sites (test and reference sites). The "costs" in time and money are assessed for each sampling scheme. Next, a statistical power analysis is conducted to calculate the number of replicate samples needed to achieve a desired degree of statistical "power" for each sampling scheme. Finally, the power-cost efficiencies of the alternative sampling schemes are calculated and the optimum scheme is selected as the one with the highest power-cost efficiency.

2.3 QUALITY ASSURANCE/QUALITY CONTROL

Quality assurance and quality control (QA/QC) are essential to the production of environmental monitoring data of known and documented quality in a cost-effective manner. QA/QC should be an integral part of the process of study design, execution, and data evaluation and interpretation.

All EPA data-collection programs have implemented Quality Assurance Program Plans designed and overseen by their management to ensure the quality of all activities for which their organization is responsible. These programs address all quality assurance issues in regard to policy, planning, review, and implementation. QA Project Plans are a vital part of the QA Program Plan. A QA Project Plan is a project-specific guidance compiled to encompass all aspects of the sampling/analytical effort. The preparation of a QA Project Plan is often met with unnecessary trepidation. A QA Project Plan is simply a written record of the plans that must be made and followed in executing a study. A QA Project Plan provides detailed documentation of all facets of how and why a particular study will be undertaken. The Plan also describes the alternative actions that will be taken in the event that things do not go according to the original plans. Once all of the purposes and procedures of the proposed study are recorded in a QA Project Plan, the Plan can be improved or modified, if needed, through reviews by persons knowledgeable about different aspects of the study (e.g., chemical analysis, sampling logistics, navigational positioning, sample preservation techniques).

Because the QA Project Plan is a vital tool for the data-collection process, it is essential that all personnel involved in the project read and understand the Plan and that the Plan be available for reference throughout the project to ensure proper implementation.

QA Project Plans are important for legal as well as scientific reasons. QA Project Plans are required for all EPA-associated projects (EPA Order 5360.1). QA Project Plans become part of contracts that are issued to undertake studies (40 CFR, Part 15). Furthermore, nonadherence to the Plan could result in the data being unusable for court proceedings or regulatory decisions.

The QA Project Plan is just as important after the study is completed and the data are being used to make an evaluation or decision. The Plan provides the information needed to assess the degree of confidence one can place in the data, as well as the comparability of the data collected in a particular study with those from another study. A common problem that managers and scientists have with using existing data is not that the old data are unreliable, but that the data are of unknown reliability.

2.3.1 QA/QC Terminology

A number of important concepts and terms need to be defined to develop an understanding of what makes up an adequate QA/QC program (USEPA, 1983; Delbert and Starks, 1985).

Accuracy is defined as the difference between a measured value and the assumed or expected value. Accuracy in percent is 100 minus the total error, which is composed of bias and random errors.

Bias is the systematic distortion of a measurement process that adversely affects the representativeness of the results. Bias can result from the basic sampling design, the kind of equipment used to collect the samples, the sample-handling procedures, and poor recovery of the analyte. Because bias is systematic, its magnitude can be predicted if proper QA procedures are being used in the field and laboratory.

Comparability is the measure of confidence one has in being able to compare one data set with another. Comparability is increased if similar field and laboratory methods were used and decreased if different or unknown (undocumented) methods were used. Comparability between different laboratories can be evaluated through the use of inter-laboratory comparisons, or "round-robin" studies, wherein standardized samples are analyzed by each of the participating laboratories.

Completeness is the amount of valid data obtained (i.e., that met QA/QC acceptance criteria) compared to the planned amount. Completeness is usually expressed as a percentage.

Data quality refers to the sum of all features and characteristics of the data that determine its capability to satisfy the objectives of the data collection.

Data quality indicators are quantitative statistics and qualitative descriptors that are used to interpret the degree of acceptability or utility of data to the user. Data quality indicators include bias, precision, accuracy, comparability, completeness, and representativeness.

Data quality objectives (DQO) are statements of the overall uncertainty that a decision-maker is willing to accept in results or decisions derived from the data, and they provide the framework for the data-collection effort.

Duplicate samples are two samples taken from and representative of the same population and carried through all the same steps of sampling, storage, and analysis in an identical manner.

Field blank is a clean sample (i.e., distilled water) carried to the sampling site, exposed to sampling conditions, and returned to the laboratory and treated as an environmental sample. Field blanks are used to try to assess contamination problems caused by conditions in the field, including contamination of the sampling device, sample containers, shipping containers, etc.

Measurement error is the difference between the true sample values and the reported values and can occur during analysis, data entry, database manipulation, or other steps.

Method sensitivity/method detection limit defines the lower limits of reliable analysis of a particular parameter inherent in the use of a particular test method. The method detection limit is the minimum concentration of a substance that can be measured with 99 percent confidence that the analyte concentration is greater than zero in a particular medium (40 CFR Part 136, Appendix B).

Precision is the degree of consistency among duplicate/replicate measurements.

Quality assurance is an integrated program for ensuring the reliability of monitoring and measurement data. It includes the well-defined plans and procedures for how to ensure the production of sufficient data of known and documented quality, including monitoring how well QC procedures are actually being implemented.

Quality control is the routine application of procedures for obtaining prescribed standards of performance in the monitoring and measurement process. It is the actual implementation of the QA plan, effected through measurements of data quality through the use of blanks, spikes, etc. Quality control consists of both internal and external checks including repetitive measurements, internal test samples, interchange of technicians and equipment, use of independent methods to verify findings, exchange of samples and standards among laboratories, and use of standard reference materials.

Random error is nonsystematic (and, therefore, unpredictable) error that can occur during any part of the sample collection, handling, and analysis. Hopefully, random errors are normally distributed with a mean of zero so that the overall evaluation will not be affected even though individual measurements will be affected.

Representativeness is the degree to which the data accurately and precisely represent the parameter or condition being sampled. Representativeness is affected by sampling design (e.g., number of samples, method of selecting sampling sites), as well as analytical sampling accuracy and precision.

Sampling error is the difference between the sampled value and the true value, and is a function of natural spatial and temporal variability and sampling design. It also includes error due to improperly selected/collected samples or improperly gathered measurements. Sampling error is more difficult to control than the other type of error, measurement error, and typically accounts for most of the total error.

Uncertainty is the total variability in sampling and analysis including systematic error (bias) and random error.

Duplicates, spikes, and blanks are all used to assess the quality of the data, to identify any systematic problems, and to isolate the sources of such problems.

2.4 SOURCES AND SIGNIFICANCE OF MONITORING ERROR

To increase the accuracy, precision, and representativeness of the data collected in a sediment assessment study, it is important to be aware of and minimize two types of error that can be introduced into sediment contaminant concentration data: bias and scatter. Sources of bias in sediment studies include the actual heterogeneity in the distribution of contaminants in the sediments, the sampling design (number of samples, method for selecting sampling sites), the sampling method, the sample preparation procedures, and the testing methods.

Factors that tend to make sediment contaminants distribute themselves heterogeneously include the differences in the density of the bulk contaminant (e.g., sinking versus floating); differences in the affinity of the contaminant for particles as a function of particle size, organic carbon content, etc.; particle sorting as a function of water currents and particle size; lateral mixing of water and sediments as a function of flow or distance downstream of the sources; resuspension; bioturbation; and biouptake.

The objective of a well-designed sampling program is to minimize the introduction of data artifacts associated with the sampling plan, sample collection, sample preparation, and sample analysis while revealing the actual contaminant concentration profile in space as a function of time. A plan that requires preferential sampling of areas that are devoid of aquatic life will likely be biased toward high toxicant concentrations, resulting in an unrepresentative horizontal spatial sediment contaminant profile. Artifactual variability can be introduced if the number and size of the samples are inappropriate to the scale of the system under investigation, yet the sampling size has to be balanced against cost.

With respect to bias due to sampling method, if certain core samplers are used to quantify the vertical distribution of a sediment contaminant, for example, the actual vertical profile is likely to be distorted because the absolute vertical relationship of contaminant concentrations is lost due to differential compression of the sample during coring. Another example of sampling method bias occurs when a grab sampler is used to collect the surficial sediment sample. The potential disproportionate loss of fine particles from the grab during the drop, closing, and withdrawal phases of sampling can result in an underquantification of the contaminant surficial concentration if the contaminant is preferentially concentrated on the fines.

Regarding sample preparation bias, a sample preparation procedure that transforms, loses, or destroys one member of a homologous series (e.g., PCBs, PCDDs, or PCDFs) will not only result in an underquantification of the total concentration for that toxicant category, but will also misrepresent the relative proportions of the isomers. Analytical method bias can result from the inability to separate complex mixtures into individual constituents (interference), thus resulting in the misidentification or misquantification of a toxicant; from differences in the sensitivity of the detector for a particular pollutant over the range of concentrations encountered in the sediment (nonlinear responses); or from poor or varying recovery of the analyte.

Analytical variability arises primarily from the compounded uncertainty associated with the tolerance on each of the components and steps of the wet or electronic methods of sample preparation (aliquot selection, weighing, drying, grinding, sieving, etc.) and analysis.

2.5 COMPONENTS OF A QUALITY ASSURANCE PROJECT PLAN

As mentioned previously, a QA Project Plan clearly documents the participants' responsibilities; what will be done; why it is being done; the desired accuracy, precision, completeness, and representativeness of the resulting data; who will report what information to whom; and what will be done in the event something goes wrong. Rather than attempting to describe the actual components of a QA Project Plan in any detail here, an example of the table of contents from a recent plan is presented in Figure 2-1. In addition, actual QA Project Plans from projects similar to the one being planned can be extremely useful in suggesting the important issues to consider. For detailed guidance on preparing QA Project Plans, one should refer to *Interim Guidelines and Specifications for Preparing Quality Assurance Project Plans* (USEPA, 1980). Some good examples of actual sediment assessment Quality Assurance Project Plans include Burton (1989), Crecelius (1990), and Valente and Schoenherr (1991).

2.6 SAMPLE COLLECTION AND HANDLING

2.6.1 Sampling for Physical and Chemical Analyses

2.6.1.1 Sample Collection Methods

The most appropriate device for a specific study depends on the study objectives, sampling conditions, parameters to be analyzed, and cost-effectiveness of the sampler. There are basically three types of devices used to collect sediment samples: dredges, grab samplers, and corers (Baudo, 1990).

A ***dredge***[1] is a vessel that is dragged across the bottom of the surface being sampled, collecting a composite of surface sediments and associated benthic fauna. Dredge samplers are more commonly used to sample sediments in marine waters than in fresh water. This type of sampler is primarily used for collecting indigenous benthic fauna rather than samples for analyses or assays. Because the sample is mixed with the overlying water, no pore water studies can be made of dredged samples. Additionally, because the walls of the dredge are typically nets, they act as a sieve and only the coarser material is trapped, resulting in the loss of fine sediments and water-soluble compounds (ASTM, 1990). Results of dredge sampling are considered qualitative in nature since it is difficult to determine the actual surface sampled by the dredge.

Grab samplers have jaws that close by a trigger mechanism upon impact with the bottom surface. Grab samplers offer the advantage of being able to collect a large amount of material in one sample, but they have the disadvantage of giving an unpredictable depth of penetration. Grab samplers are recommended when sampling is being performed for routine dredging projects because the sediments are continually disrupted by marine traffic, homogenizing the sediments that have accumulated since the last dredging (Plumb, 1981).

A ***core*** sampler is basically a tube that is inserted into the sediment by various means to obtain a cylinder or box sample of material at known depths. Corers can be simple, hand-operated devices used by scuba divers, or they can be

[1]Although grab samplers are sometimes referred to as "dredges," in this document grab samplers are distinguished from dredge samplers in that the grab samplers sample a discrete volume of surface sediments in an area defined by the opening size of the sampler's jaws, as opposed to the dredge sampler, which collects a composite of bottom sediments as it is dragged across the bottom.

Cover Page (w/Approval Signatures)
Title Page
Introduction
Table of Contents
List of Tables
List of Figures
List of Appendices
List of Acronyms and Abbreviations
Glossary

1 PROJECT DESCRIPTION
1.1 Introduction
1.2 Project Scope
1.3 Data Quality Objectives
1.4 Sample Network Design and Rationale
1.5 Project Implementation

2 PROJECT ORGANIZATION AND RESPONSIBILITY
2.1 Organization
2.2 Authority and Responsibility
2.2.1 Project Oversight
2.2.2 Field Activities
2.2.3 Laboratory Analyses
2.2.4 Other Regulatory Personnel
2.3 Project Communication

3 QUALITY ASSURANCE OBJECTIVES
3.1 Quality Assurance Documents
3.2 Project Quality Assurance Objectives
3.3 Field Measurement Quality Objectives
3.3.1 Navigation
3.3.2 Sample Collection Parameters
3.3.3 Water Column Measurements
3.4 Laboratory Data Quality Objectives
3.5 Macrobenthic Community Assessment Quality Assurance Objectives
3.6 Computer Model Quality Assurance Objectives

4 SAMPLE COLLECTION AND HANDLING PROCEDURES
4.1 Sample Containers
4.1.1 Volume and Type
4.1.2 Quality Control and Storage
4.2 Sampling Procedures
4.2.1 Selection and Decontamination of Equipment
4.2.2 Sampling Methods
4.2.3 Collection of Sample
4.2.4 Sample Volume, Preservation, and Holding Times
4.2.5 Field-Generated Waste Disposal
4.3 Sample Packaging and Shipment

5 SAMPLE DOCUMENTATION AND CUSTODY
5.1 Field Procedures
5.1.1 Sample Labeling
5.1.2 Field Logbooks
5.1.3 Field Chain of Custody
5.1.4 Transfer of Custody
5.2 Laboratory Procedures
5.2.1 Sample Scheduling and Management
5.2.3 Sample Receipt and Handling
5.2.4 Log Books and Chain of Custody
5.2.5 Sample Disposal
5.3 Final Evidence File
5.3.1 Contents
5.3.2 Custody Procedure

6 CALIBRATION PROCEDURES AND FREQUENCY
6.1 Field Measurements
6.1.1 Records and Traceability of Standards
6.1.2 Initial and Continuing Calibration Procedures
6.1.3 Conditions to Trigger Recalibration
6.2 Physical and Chemical Laboratory Analyses of Sediment
6.2.1 Records and Traceability of Standards
6.2.2 Preparation and Storage of Standards
6.2.3 Initial and Continuing Calibration Procedure
6.2.4 Conditions to Trigger Recalibration
6.3 Biological Effects Tests -- Water Quality Monitoring
6.3.1 Records and Traceability of Standards
6.3.2 Initial and Continuing Calibration Procedure
6.3.3 Conditions to Trigger Recalibration

Figure 2-1. Contents of a Quality Assurance Project Plan

Figure 2-1. Contents of a Quality Assurance Project Plan. (Continued) large, costly, motor-driven mechanisms that can collect samples from great depths. A few types of corers include a gravity corer, which uses weights attached to the head of the sampling tube to push the tube into the sediment; a piston corer, which is similar to a gravity corer but also has a piston inside the tube that remains stationary during sediment penetration and creates a vacuum that helps pull the sampler into the sediment; a vibra-corer, which is like a gravity corer except with a vibrating head attached to enhance penetration; and a multiple corer, which is an array of plastic tubes attached to a frame, allowing for the collection of several samples at the same location. Because gravity corers can compact the sample and distort the vertical profile, a piston corer or vibra-corer is recommended to minimize sample compaction. The corer that disturbs the sediments the least is a box corer. Instead of being cylindrical, it is a large box-shaped sampler that is deployed inside a frame. After the frame is brought to rest on the bottom, heavy weights lower the open-ended box into the sediment. A bottom door then swings shut upon retrieval to prevent sample loss. The advantages of the box corer include its ability to collect a large amount of sample with the center of the sample virtually undisturbed. Corers are not generally recommended for use in sandy sediments since they have difficulty retaining the sample upon withdrawal.

A comparison of the general characteristics of various commonly used sediment-sampling devices for chemical, physical, and biological studies is given in Baudo (1990); Plumb (1981); Downing and Rigler (1984); and ASTM (1990).

2.6.2 Sample Handling, Containers, Preservation, and Holding Times

2.6.2.1 General Requirements

Proper handling of the samples is essential to preserve the sample integrity and the validity of the results. Mishandling of samples at any stage of the sample-collection process could distort analytical results, wasting the effort and expense of the sampling survey. Some of the basic considerations in sediment sample handling include the following (Plumb, 1981):

- It is essential that noncontaminated sampling devices are used and that obvious sources of contamination such as exhaust fumes from the collecting ship, lubricating drilling fluids, and powder from surgical gloves be eliminated.

- Sampling devices should be washed between samples with an appropriate series of cleansers and solvents to prevent cross-contamination from one sample to the next.

- Analysis for different parameters requires different storage containers to ensure noncontamination and to prevent degradation of the sample. Basic rules for containers include using plastic or glass containers for metal analysis, glass containers for organic analysis, and glass or plastic for inorganic analysis. Since no set guidelines have been determined for sediment sampling, a good general rule to follow is to use containers recommended for water testing.

- A reliable and identifiable sample-labeling process should be used.

- Sampling containers should be filled to capacity, allowing only enough air space for possible expansion of the sample resulting from the preservation technique (e.g., freezing) to eliminate or greatly reduce oxidation of the sample (USEPA/-USACE, 1991). Sample containers for volatile organics analyses should be filled completely, allowing no headspace.

Preservation methods are intended to maintain the integrity of the sample by limiting the deterioration or alteration of a specified parameter by hydrolysis, oxidation, and/or biological activity while the sample awaits analysis. Methods are basically limited to pH control, chemical addition or fixation, sample extraction or isolation, or temperature control. Preservation steps should be initiated immediately after collection of the sample since significant alteration of the sample can occur in the first few hours after sampling. Immediately after collection, sediment samples are typically kept on ice or refrigerated. Upon arrival at the laboratory, samples are usually preserved by drying, freezing, or cold storage (ASTM, 1990).

The type of preservation required will depend on the parameters being tested. For example, if the sediment is to be tested for both bulk metals and particle size, either two samples should be collected or the sample should be split, since it is recommended that samples for bulk metal analysis be preserved by dry ice and stored at less than -20°C, whereas samples to be analyzed for particle size should be refrigerated at 4°C (USEPA/USACE, 1991). For this reason, it is essential to know which tests are to be performed, or potentially performed, on the samples in advance to allow for additional sample collection or splitting of samples as needed to comply with differing sampling, handling, and preservation requirements.

Freezing appears to be the generally preferred method for preserving sediment samples for most chemical analysis, although sediments to be used for particle size determination, volatile organics, and toxicity testing should not be frozen (ASTM, 1990).

2.6.2.2 Requirements for Specific Analyses

There are basically four ways to analyze chemical and physical parameters of sediments: bulk analysis, standard elutriate test, fractionation procedures, and physical analysis. Brief descriptions of these types of analyses follow, along with any special sample handling procedures, containers, or preservation techniques needed.

Bulk analysis allows one to evaluate the total concentration of a parameter within a sediment sample or the toxicity of the whole sediment. Most chemical parameters are evaluated by bulk analysis. In general, the collection container and preservation and storage method are dependent on the parameter to be tested. Bulk analysis samples can be stored wet, air-dried, or frozen. If trace organic constituents are to be analyzed, a glass container should be used to store the sample. When preserving and storing samples, one needs to take into consideration that other parameters could change as a result of oxidation, volatilization, or chemical instability (Plumb, 1981).

Elutriate tests indicate the ability of chemical constituents to migrate from the solid phase to the liquid phase. An elutriate sample is prepared by mixing or shaking sediment and water in prescribed proportions for a prescribed period of time and separating the liquid fraction by filtration and/or centrifugation. The liquid fraction, the elutriate, is then analyzed by methods used for analysis of water samples. Sediments to undergo elutriate testing should be stored wet, at 4°C, in airtight containers and should be tested as soon as possible following sample collection. If trace organic analyses are to be performed, glass containers with Teflon lids are required for storage (Plumb, 1981).

Fractionation procedures provide information on the distribution of constituents. The samples are extracted multiple times using a series of extractants and procedures, thereby isolating specific pollutants or classes of pollutants. Pore water extraction is a form of fractionation whereby the interstitial water in the whole sediment sample is extracted by squeezing or centrifugation. The resulting water sample can be used in chemical and biological tests. To date, fractionation has been used primarily for research. As a result, most agencies do not subject their sediment samples to fractionation procedures (Plumb, 1981). However, some fractionation tests, such as the toxicity identification evaluation (TIE), a fractionation procedure to isolate the toxic component of a sample, are beginning to be used to make decisions regarding regulatory actions and remedial approaches since they can be used to assess which pollutants are responsible for the toxicity observed in a sediment. Samples to be analyzed for fractionation should be stored wet, at 4°C, and in airtight containers. Testing procedures should start as soon as possible after sample collection (Plumb, 1981).

Physical analysis provides information on particle size, color, texture, and mineralogical characterization and includes tests for cation exchange capacity, particle size, pH, temperature, salinity, oxidation reduction potential, total volatile solids, and specific gravity. Samples to undergo physical analyses may be stored wet, at 4°C, or frozen, depending on the parameter to be tested. Some of these parameters (e.g., pH) should be analyzed immediately upon collection.

The 1991 Green Book (USEPA/USACE, 1991) suggests the use of a grab sampler or corer for collection of sediment samples and offers the following general guidelines for preservation and handling and sample sizes needed for sediment samples collected for chemical and physical testing:

Bulk metals should be stored in nonreactive containers, such as high-density polyethylene, and analyzed as soon as possible.

Bulk organics, including PCBs, pesticides, and high-molecular-weight hydrocarbons, should be contained in solvent-rinsed glass jars with Teflon lids, preserved by dry ice, and stored at less than -20°C in the dark. The samples can be stored for up to 10 days. Approximately 475 mL of sample should be collected.

Samples to be analyzed for **total organic carbon** (TOC) should be preserved by dry ice and stored at less than -20°C. They can be kept for an undetermined amount of time.

Sediments for **particle size testing** should be kept refrigerated at 4°C in any sealed container and can be kept for an undetermined amount of time.

2.6.3 Minimum Parameters to Be Tested

Sampling efforts are performed with a variety of objectives in mind, and therefore the minimum chemical and physical parameter testing requirements vary between studies or programs. However, some chemical and physical parameters seem to be common to several programs. They include particle or grain size, total organic carbon, heavy metals, acid volatile sulfides, polycyclic aromatic hydrocarbons, polychlorinated biphenyls, and pesticides. Unionized ammonia must also be measured, taking into account its sensitivity to pH and temperature, both of which are affected by sample manipulation. When testing sediment samples from estuarine or marine environments, the analysis methods chosen must address salinity since this can alter the analytical results (USEPA/-USACE, 1991).

Particle or grain size analysis is a physical parameter that determines the distribution of particle sizes. Methods for particle size analysis are suggested in Folk (1968), Buchanan (1984), Plumb (1981), ASTM (1990), and Tetra Tech (1985). Plumb (1981) suggests that analysis will usually require two or more methods, depending on the range of particle sizes encountered. He gives a detailed account of the use of sieves in conjunction with electronic particle counters or sieves and pipet analysis. *Testing and Reporting Requirements for Ocean Disposal of Dredge Material off Southern California under Marine Protection, Research and Sanctuaries Act Section 103 Permits* (Ocean Dredged Material Disposal Program, 1991) recommends the method given in Plumb (1981) for analysis of particle size.

Total organic carbon (TOC) is an important indicator of bioavailability for nonionic hydrophobic organic pollutants. When analyzing for this parameter, it is essential that the sample be stored in a glass or plastic container and that all air bubbles be removed from the sample before it is sealed and stored. The method given in Plumb (1981) is commonly recommended (Tetra Tech, 1985). Plumb (1981) suggests using sample ignition, which uses a hydrochloric acid wash to separate the inorganic and organic carbon, or differential combustion, which uses thermal combustion to separate the two carbons by their different combustion temperature ranges. The 1991 Green Book recommends that the analytical method to test for TOC be based on high-temperature combustion rather than on chemical oxidation. Additionally, it recommends using sulfuric acid rather than hydrochloric acid rinse. *Testing and Reporting Requirements for Ocean Disposal of Dredge Material off Southern California under Marine Protection, Research and Sanctuaries Act Section 103 Permits* recommends EPA Test Method No. 9060 for TOC determinations. The method recommended by EPA for use in applying organic carbon-normalized sediment quality criteria for nonionic hydrophobic organic chemicals uses catalytic combustion and nondispersive infrared detection (Leonard, 1991).

Metals are found naturally occurring in the environment, but an excess of metals can be an indication of anthropogenic contamination. The most commonly used method to analyze sediments for metals is atomic absorption spectrophotometry. Plumb (1981) details the use of the direct-flame atomic absorption method for all metals except arsenic, mercury, and selenium. For these metals, he recommends using arsine generation, cold vapor technique, and digestion/flameless atomic absorption or hydride generation, respectively. The 1991 Green Book points out that the concentration of salt in marine or estuarine samples may cause interference in analysis for metals. Therefore, the approach of an acid digestion followed by atomic absorption spectroscopy should be coupled with an appropriate technique to control this interference. The 1991 Green Book recommends USEPA (1986) for analysis of mercury and EPRI (1986) for the analysis of selenium and arsenic. *Testing and Reporting Requirements for Ocean Disposal of Dredge Material off Southern California under Marine Protection, Research and Sanctuaries Act Section 103 Permits* recommends the following EPA Test Methods: cadmium (Nos. 7130, 7131); hexavalent chromium (Nos. 7190, 7191); copper (No. 7210); lead (Nos. 7420, 7421); mercury (No. 7471); nickel (No. 7520); selenium (Nos. 7740,7741); silver (No. 7760); and zinc (No. 7950).

Acid volatile sulfides (AVS) have been found to be closely related to the toxicity of sediment-associated metals (Di Toro *et al.*, 1990). AVS have been found to be important in binding potentially bioavailable metals, thereby reducing their toxicity. The approved method is given in USEPA (1991).

Polyaromatic hydrocarbons (PAHs) are semivolatile organic priority pollutants, a number of which are potential carcinogens. Plumb (1981) details the methods of methanol extraction/UV analysis and ethanol extraction/UV spectrophotometry to analyze for this parameter. *Testing and Reporting Requirements for Ocean Disposal of Dredge Material off Southern California under Marine Protection, Research and Sanctuaries Act Section 103 Permits* recommends EPA Test Method Nos. 8100, 8250 and 8270 for analysis of PAHs.

Polychlorinated biphenyls (PCB) are chlorinated organic compounds that were once used for numerous purposes including as a dielectric fluid in electrical transformers. Desirable properties of PCBs include low flammability, nonconductivity, and nonreactivity. However, PCBs do not break down readily and they bioaccumulate in the environment. The 1991 Green Book offers gas chromatography/electron-capture detection (GC/ECD) methods as the primary tool for the analysis of PCBs, or the use of GC/MS using selected ion monitoring (SIM). They do not recommend the traditional methods of PCB analysis, which quantify PCBs as arochlor mixtures. *Testing and Reporting Requirements for Ocean Disposal of Dredge Material off Southern California under Marine Protection, Research and Sanctuaries act Section 103 Permits* recommends the use of the methods described in Tetra Tech (1986) and NOAA (1989) for analysis of PCBs.

Pesticides are man-made compounds predominantly used in agriculture to control crop-damaging insects. Some pesticides, especially halogenated compounds, persist in the environment and can contaminate the food chain. Plumb (1981) details the method of hexane extraction in preparation for testing for organophosphorus pesticides. The 1991 Green Book recommends using GC/ECD or GC/MS to analyze for chlorinated pesticides. *Testing and Reporting Requirements for Ocean Disposal of Dredge Material off Southern California under Marine Protection, Research and Sanctuaries Act Section 103 Permits* recommends EPA Test Method No. 8080 to analyze for pesticides.

For analyses of **volatile organic pollutants** and **semivolatile organic pollutants**, the 1991 Green Book recommends the methods described by Tetra Tech (1986), which should always include the use of capillary-column GC or GC/MS techniques. For volatiles, a purge-and-trap method is used, followed by GC/MS analysis according to U.S. EPA Method 624 or U.S. EPA Method 1624, Rev. B, Ref. 3 (Tetra Tech, 1986).

As stated previously, the minimum set of parameters tested in sediments varies and is based on the sampling objectives of the program. Listed below are several examples of minimum data sets required by specific programs.

The 1991 Green Book recommends that all sediment samples be analyzed for TOC, PAHs, grain size, total solids/water content, and specific gravity. The remaining parameters to be sampled are compiled from the priority pollutants list based on historical testing data, potential contaminants due to known industries in the area, and a general knowledge of the area to be sampled.

Testing and Reporting Requirements for Ocean Disposal of Dredge Material off Southern California under Marine Protection, Research and Sanctuaries Act Section 103 Permits has very specific parameters and methods required for materials to be disposed of off the coast. Required analyses for physical parameters include grain size, total solids/water content, and specific gravity. Chemical analyses includes 9 metals, ammonia, arsenic, total sulfides, acid volatized sulfides (AVS), 11 pesticides including total pesticides, 9 organic compounds, all PCB congeners, individual totals of tetra-, penta- and hexa-chlorobiphenyl isomers, and 17 PAHs.

The EPA Environmental Monitoring and Assessment Program-Near Coastal (EMAP-NC) established guidelines identified in its *Near Coastal Program Plan for 1990: Estuaries* (Holland, 1990) for sediment sampling for determination of contaminant levels. They include sample collection by means of a Young-modified Van Veen grab and, initially, analyzing the NOAA Status and Trends suite of contaminants, which include chlorinated pesticides, PCBs, PAHs, major elements, and toxic metals. EMAP-NC, with the assistance of other programs, plans to refine the list of contaminants to include pesticides and herbicides and other toxic chemicals.

2.6.4 Sampling for Benthic Community Structure in Fresh Water

Macrobenthic organisms play an important role in marine, estuarine, and freshwater lotic and lentic ecosystems. As major secondary consumers, they represent an important linkage between primary producers and higher trophic levels for both planktonic and detritus-based food webs. They are a significant food source for juvenile fish and crustaceans and may improve water quality by filter-feeding of particulate matter (Holland, 1990). Benthic populations also represent diverse taxa and can serve as sentinels for environmental stress. Benthic organisms access all aspects of the aquatic habitat with varying feeding strategies, reproductive modes, life history characteristics, and physiological tolerances to environmental conditions. Most benthic organisms have limited mobility and cannot avoid environmental stressors. As a result, the responses of some species serve as indicators of changes in sediment quality (Holland, 1990). This section will detail specific procedures and precautions necessary for proper conduct of benthic sample collection and handling in freshwater, marine, and estuarine ecosystems.

2.6.4.1 Sample Collection Methods

It is helpful to consult *Macroinvertebrate Field and Laboratory Methods for Evaluating the Biological Integrity of Surface Waters* (Klemm *et. al.*, 1990), which thoroughly addresses methodology. State environmental regulatory programs should have a Quality Assurance Program Plan describing the field methods and standard operating procedures for collecting and evaluating benthic macroinvertebrates. This information should be obtained to ensure acceptance and comparability of study results with those obtained by the state agency. If this information is not available, then field methods and standard operating procedures from other existing programs should be used.

In soft freshwater sediments, the most common method used to collect benthos is with a grab sampler such as a Ponar (15 x 15 cm or 23 x 23 cm) or Ekman grab sampler (15 x 15 cm, 23 x 23 cm, or 30 x 30 cm), each of which provides a quantitative sample based on the surface area of the sampler. The smaller of the sampler sizes are most commonly used for freshwater studies because of their relative ease of manipulation.

The Ekman grab sampler is not as effective in areas of vegetative debris but is much lighter than the Ponar and easier to use in softer substrates. Artificial substrates (Hester-Dendy using several 3-inch plates and spacers attached by an eyebolt, or substrate/rock-filled baskets) provide consistent habitat for the benthos to colonize in both soft-bottomed and stony areas. Artificial substrates can be used in almost any water body and have been successfully used to standardize results despite habitat differences (Ohio EPA, 1989; Rosenberg and Resh, 1982; and Resh and Jackson, 1991).

A variety of methods for sampling benthos in hard-bottomed lotic systems are available, including artificial substrates. If quantification by sediment or sampler surface area is needed, a Surber-type square-foot sampler with a Standard #30-mesh (0.589-mm openings) can be used (Klemm *et al.*, 1990). The traveling kick-net (or dip-net) method, also using a #30-mesh net, can be used to quantify the sample collected by the amount of time spent sampling and the approximate surface area sampled (Pollard, 1981; Pollard and Kinney, 1979). The Surber-type and kick methods can each be used to provide consistent, reproducible samples, but both are limited to wadable streams. The Surber sampler's optimal effectiveness is limited to riffles, whereas kick-net or dip-net samplers can be effectively used in all available habitats. Although dip-net samplers have been effectively used to sample riffles and other relatively shallow habitats to determine taxa richness, presence of indicator organisms, relative abundances, similarity between sites, and other information, they do not provide definitive estimates of the number of individuals or biomass per surface area.

2.6.4.2 Sample Handling and Preservation

The following decisions will need to be made once the sample collection method is chosen: (1) whether samples will be picked from debris and sorted in the field, (2) which preservative should be used, (3) whether a stain (rose bengal) or other material will be added to the sample to facilitate separating the organisms from debris, (4) the type of sample containers and labeling of the containers required, and (5) the mode of transportation of the samples to their destination. Many of these decisions are based on professional preference or the required logistics of the study.

Sorting of the benthos from debris and preservation are fully discussed by Klemm *et al.* (1990). American Public Health Association *et al.* (1989) and Klemm *et al.* (1990) defined the benthos by what is retained on a standard #30 sieve. However, some types of Chironomidae and other small benthos pass through a #30-mesh sieve but are retained by a #40-mesh sieve. It has been recommended that samples should first be passed through a #30-mesh sieve. Then the materials washed through should be passed through a #40-mesh sieve, and the materials retained in both sieves should be sorted (Ohio EPA, 1989). Once the material is washed through the sieves, the organisms should be separated from the vegetation and other debris in a white enamel pan. As the materials are separated, the organisms can be placed in different vials for the major taxa. Preservation with either formalin or 70 percent ethanol is common. Although formalin is an excellent fixative, the human health concerns associated with its use require extreme caution and adequate ventilation. Many programs rely on 70 percent ethanol as a fixative and preservative.

A practical technical reference that details procedures for cost-effective biological assessments of lotic systems has been developed. ***Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic Macroinvertebrates and Fish*** (Plafkin *et al.*, 1989) presents three benthic rapid bioassessment protocols (RBPs) and two fish RBPs, with a progressive order of increasing rigor in evaluation within each series for each class of organisms.

The RBPs are based on integrated assessments that compare physical conditions of habitat (e.g., physical structure, flow regime) and biological measures of reference conditions. These reference conditions are derived after systematic monitoring of sites that represent the natural range of variation in water chemistry, habitat, and biological condition.

The functional and structural components evaluated for aquatic communities comprise eight metrics for benthic RBPs and 12 metrics for the fish RBPs. Examples of metrics for benthic communities include the following: taxa richness, the modified Helsenhoff Biotic Index (summarizes overall pollution tolerance of the benthic arthropod community with a single value; this index was modified to include nonarthropod species as well), ratio of scraper and filtering collector functional feeding groups, ratios of the number of organisms in the EPT (Ephemeroptera, Plecoptera, and Trichoptera) to the number of Chironomidae present, and community similarity indexes. The fish protocol is based on the index of biotic integrity (IBI) or a fish community assessment approach developed by Karr *et al.* (1981). As with the approach of metrics in the benthic evaluations, the metrics of the fish protocol represent differing sensitivities.

2.6.5 Sampling for Benthic Community Structure in Marine and Estuarine Waters

Historically, regional monitoring programs have used benthic community studies as an effective indicator of the extent of pollution impacts on marine and estuarine ecosystems, as well as the effectiveness of management actions. In addition, information on changes in benthic population and community parameters due to sediment characteristics can be used to distinguish natural variation from changes due to human activities (Holland, 1990).

2.6.5.1 Sample Collection Methods

Three grab samples are collected for benthic species composition, abundance, and biomass. Additional sediment grabs are collected for chemical analyses and for use in acute toxicity tests. To minimize the possibility of biasing results, benthic biology grabs should not be collected consecutively, but rather interspersed among the chemistry/toxicity grabs. While a biology grab is being processed (sieved), grabs should be collected for chemistry/toxicity (Holland, 1990).

A $^{1}/_{25}$ m^2, stainless steel, Young-modified Van Veen grab sampler may be used to collect sediments for benthic analyses. The sampler is constructed entirely of stainless steel and has been coated with Kynar (similar to Teflon) and is, therefore, appropriate for collecting sediment samples for both biological and chemical analyses. The top of the sampler is hinged to allow for the removal of the top layer of sediment for chemical and toxicity analyses. This gear is relatively easy to operate and requires little specialized training. To minimize the chance of sampling the exact same location twice, the boat should be moved 5 meters downstream after three grabs have been taken, whether successful or not (Holland, 1990).

2.6.5.2 Sample Handling and Preservation

Grab samples to be used in the assessment of macroinvertebrate communities are processed by first extracting a core sample from the sampler. The depth of sediment at the middle of the sampler should be at least 7 cm. Descriptive information on the grab is recorded. The depth to the black layer of sediment within the core, the redox potential discontinuity (RPD), is measured in the field. The sample is then extruded from the core tube to fill a whirl pac bag, labeled, and recorded. The sample should be refrigerated at 4°C, not frozen (Holland, 1990).

The remainder of the grab is processed for benthic community analysis. The sediments are transferred into a basin and then into a 0.5-mm mesh sieve. The sieve is agitated to wash away sediments and leave organisms, detritus, sand particles, and pebbles larger than 0.5 mm. A gentle flow of water over the sample is acceptable, but forceful jets of water should be avoided because they can cause mechanical damage to fauna. The organisms are rinsed and transferred from the sieve into a jar and covered in seawater with MgCl added. This "relaxes" the organisms, reducing damage from addition of the preservative (Holland, 1990). Ten percent buffered formalin is used to fix and preserve samples. After 30 minutes in the relaxant, formalin with a small amount of borax should be added to each sample jar. The jar is filled to the rim with seawater to eliminate any air space, eliminating the problem of organisms sticking to the cap during shipment. Prior to sieving the next sample, the sieve is rinsed and brushed thoroughly to prevent cross-contamination of samples.

2.6.6 Sampling for Bioassays and Toxicity Testing

Environmental impacts on marine ecosystems are primarily assessed and monitored using the tools outlined in the 1991 Green Book. The 1991 Green Book is used to make decisions regarding the suitability of dredged material for ocean dumping. EPA and the USACE have shown that the greatest potential for environmental impact from dredged material disposal is on the benthic environment since benthic organisms burrow into and are exposed to sediments and associated contaminants for extended periods of time. The 1991 Green Book uses whole sediment bioassays to evaluate potential impacts of dredged sediments and, in concert with the identification of contaminants of concern through chemical analysis, serves to determine the extent and type of bioavailability. In addition, sediment toxicity tests can be used to assess spatial and temporal changes in toxicity in contaminated areas, rank sediments based on their toxicity to benthic organisms, and define cleanup goals for contaminated areas. This section will highlight some of the collection and handling methods of sediments for toxicity testing and whole sediment bioassays.

2.6.6.1 Sample Collection, Handling, and Preservation

The sediment environment is composed of many microenvironments, redox gradients, and interacting physicochemical and biological processes. Many of these characteristics influence sediment toxicity and bioavailability to benthic and planktonic organisms, microbial degradation, and chemical sorption. Maintaining the integrity of a sediment sample during its removal, transport, storage, and testing in the laboratory is extremely difficult. Any disruption of this environment complicates interpretations of treatment effects, causative factors, and *in situ* comparisons (ASTM, 1990).

Sample handling, preservation, and storage techniques have to be designed to minimize any changes in composition of the sample by retarding chemical and/or biological activity and by avoiding contamination. Sufficient sample volume must be collected to perform the necessary analyses, partition the samples for respective storage requirements, and archive portions of the sample for possible later analysis. Core sampling is recommended to best maintain the integrity of the sediment for studies of sediment toxicity, interstitial waters, microbiological processes, and chemical fate. Subsampling, compositing, or homogenization of sediment samples may be necessary depending on the study objectives. Subsamples of the inner core area may be taken since this area is more likely to retain its integrity and depth profile and not be contaminated by the sampler. The loss of sediment integrity and depth profile is an important consideration, as are changes in chemical speciation through oxidation and reduction resulting in volatilization, sorption, or desorption; changes in biological activity; completeness of mixing; and sampling container contamination (ASTM, 1990).

Subsamples of the top 1 or 2 cm may be collected with a nonreactive sampling tool (e.g., polytetrafluoroethylene (PTF)-lined calibration scoop). Some studies may require a composite of single sediment samples, which usually consist of three to five grab samples. Subsamples should be collected with a Teflon paddle, placed in a nonreactive bowl or pan, and stirred until the texture and color appear uniform. The sediments should be removed and partitioned for chemical and AVS analysis. Samples should completely fill the storage containers, leaving no airspace. If the sample is to be frozen, just enough air space should be allowed for expansion to take place. The labeling system should be tested prior to use in the field, making sure that labels can withstand soaking, drying, and freezing without becoming detached or illegible (USEPA/USACE, 1991).

Maintaining clean and uncontaminated sampling equipment between samples is necessary. It is important to clean the sampling device, scoop, spatula, and/or mixing bowls between sites. A suggested cleaning procedure includes a soap-and-water wash followed by an organic solvent rinse (ASTM, 1990).

The choice of sample containers for sediment should consider the type of sediment, storage time, chemical sorption, and sample composition. For sediments containing organics, brown borosilicate glass containers with Teflon lid liners are optimal, whereas plastic or polycarbonate containers are recommended for metal-containing sediments. PTF or high-density polyethylene containers are relatively inert and are suggested for use with samples contaminated with multiple chemical types (ASTM, 1990).

Sediment samples for biological testing should be press-sieved through a 1-mm mesh screen to remove all living organisms from the sediment prior to testing. Other matter retained on the screen with the organisms, such as shell fragments, gravel, and debris, should be recorded and discarded. Sediment samples for use in bioassays should be well mixed.

Since the first few hours are the most critical to changes in the sample, preservation steps should be taken immediately upon sediment collection. There is no universal preservation or storage technique, and a technique for one group of analyses may interfere with other analyses. Problems can be overcome by collecting sufficient sample volume to use specific preservation or storage techniques for specific analytes or tests on subsamples. Preservation, whether by refrigeration, freezing, or addition of chemicals, should be accomplished in the field whenever possible. If final preservation techniques cannot be implemented in the field, samples should be temporarily preserved in a manner that retains the integrity of the sample. Sediment samples for biological analysis should be preserved at 4°C, never frozen or dried. Field refrigeration is easily accomplished with coolers and ice; however, samples should be segregated from melting ice or cooling water.

Storage containers can be the same as the transport containers, and where sediments contain volatile compounds, transport and storage should be in airtight PTF or glass containers with PTF-lined screw caps. Exposure of sediments to air should also be prevented in the handling of AVS--containing sediments. AVS is the reactive sulfide pool that can reduce metal toxicity by binding metals in anoxic sediments. Oxidation of these sediments can either increase toxicity by disassociation of the AVS-metal complex and precipitation of the metal species, or reduce toxicity if the AVS-metal complex should volatilize (ASTM, 1990).

It has been found that sediments can be stored at 4°C without significant alterations in toxicity. Completion of testing within a 2-week storage period is recommended, but limits on storage time will depend on sediment and contaminant characteristics (ASTM, 1990).

2.7 REFERENCES

American Public Health Association, American Water Works Association, and the Water Pollution Control Federation. 1989. Standard methods for the examination of water and wastewater. 17th edition. APHA, Washington, DC.

ASTM. 1990. Standard guide for collection, storage, characterization, and manipulation of sediments for toxicological testing. American Society of Testing and Materials. ASTM Designation E 1391-90.

Baudo, R. 1990. Sediment sampling, mapping, and data analysis. pp. 15-60. In: R. Baudo, J.P. Giesy, H. Muntau, Sediments: Chemistry and Toxicity of In-Place Pollutants.

Buchanan, J.B. 1984. Sediment analysis. In: Methods for the Study of Marine/Benthos. IBP Handbook No. 16, 2nd edition. N.A. Holme and A.D. McIntyre (eds.). Blackwell Scientific Publications, Oxford, UK.

Burton, G.A., Jr. 1989. Quality assurance project plan for "A multi-assay/multi-test site evaluation of sediment toxicity." U.S. Environmental Protection Agency, Great Lakes National Program Office.

Crecelius, E. 1990. Quality assurance project plan for "Assessment and remediation of contaminated sediments (ARCS) assistance."

Prepared by Battelle/Marine Sciences Laboratory for the Environmental Protection Agency, Great Lakes National Program Office.

Delbert, S.B, and T.H. Starks. 1985. Project summary sediment sampling quality assurance user's guide. EPA-600/4-85-048. Prepared by Environmental Research Center, University of Nevada, Las Vegas, for U.S. Environmental Protection Agency, Office of Research and Development, Environmental Monitoring Systems Laboratory, Las Vegas, Nevada, May 1985.

Di Toro, D.M., J.D. Mahony, D.J. Hansen, K.J. Scott, W. Burry, M.B. Hicks, S.M. Mayr, and M.S. Redmond. 1990. Toxicity of cadmium in sediments: The role of acid volatile sulfides. In: Environmental Toxicology and Chemistry. In press.

Downing, J.A. and F.H. Rigler, eds. 1984. A manual on methods for the assessment of secondary productivity in fresh waters. Second edition Blackwell Scientific Publications.

EPRI. 1986. Speciation of selenium and arsenic in natural waters and sediments. Vol. 2. Prepared by Battelle Pacific Northwest Laboratories for the Electrical Power Research Institute. EPRI EA-4641.

Folk, R.L. 1968. Petrology of sedimentary rocks. University of Texas, Austin, TX.

Higgins, T.R. 1988. Techniques for reducing the costs of sediment evaluation. Tech. Note EEDP-06-2. U.S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi.

Holland, A.F., ed. 1990. Environmental monitoring and assessment program, near coastal program plan for 1990: Estuaries. Environmental Research Laboratory, U.S. Environmental Protection Agency.

IJC. 1988. Procedures for the assessment of contaminated sediment problems in the great lakes. Report to the Water Quality Board of the International Joint Commission by the Sediment Subcommittee and its Assessment Work Group, International Joint Commission, Windsor, Ontario Canada, December, 1988.

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

Klemm, D.J., P.A. Lewis, F. Fiulk, and J.M. Lazorchak. 1990. Macroinvertebrate fieldand laboratory methods for evaluating the biological integrity of surface waters. U.S. Environmental Protection Agency, Office of Research and Development, EPA/600/4-90/030.

Lee, H. II, B.L. Boese, J. Pellitier, M. Winsor, D.T. Specht, and R.C. Randall. 1989. Guidance manual: Bedded sediment bioaccumulation tests. U.S. Environmental Protection Agency, Pacific Ecosystems Branch, Bioaccumulation Team, Newport, Oregon. EPA-600/x-89-302. ERLN-N111.

Leonard, E. 1991. Standard operating procedures for total organic carbon analysis of sediment samples, U.S. Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratory, Duluth, Minnesota.

NOAA. 1989. Standard analytical procedure of the noaa national analytical facility. 2nd ed. NOAA Tech. Mem. NMFC F/NWC-92, 1985-1986.

Ocean Dredged Material Disposal Program. 1991. Testing and reporting requirements for ocean disposal of dredged material off southern california under marine protection, research and sanctuaries act, section 103 permits.

Ohio EPA. 1989. Biological criteria for the protection of aquatic life: Volume III. Standardized biological field sampling and laboratory methods for assessing fish and macroinvertebrate communities. Division of Water Quality Planning and Assessment, Ecological Assessment Section, Columbus, Ohio.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughs. 1989. Rapid bioassessment protocols for use in streams and rivers: Benthic macroinvertebrates and fish. U.S. Environmental Protection Agency, Office of Water, EPA/444(440)/4-39-001, Washington, DC.

Plumb, R.H., Jr. 1981. Procedure for handling and chemical analysis of sediment and water samples. Tech. Rep. EPA/CE-81-1. Prepared by Great Lakes Laboratory, State University College at Buffalo, NY, for the U.S. Environmental Protection Agency/Corps of Engineers Technical Committee on Criteria for Dredged and Fill Material. Published by the U.S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi.

Pollard, J.E. 1981. Investigator differences associated with a kicking method for sampling macroinvertebrates. J. Freshwater Ecol. 1:215-224.

Pollard, J.E., and W.L. Kinney. 1979. Assessment of macroinvertebrate monitoring techniques in an energy development area: A test of the efficiency of three macroinvertebrate sampling methods in the White River. U.S. Environmental Protection Agency, Office of Research and Development, Las Vegas, NV. EPA-600/7-79/163.

Resh, V.H., and J.K. Jackson. 1991. Rapid assessment approaches to biomonitoring using benthic macroinvertebrates. In: Freshwater Biomonitoring and Benthic Macroinvertebrates. D.M. Rosenberg and V.H. Resh (eds.). Chapman and Hall, New York Press.

Rosenberg, D.M., and V.H. Resh. 1982. The use of artificial substrates in the study of freshwater benthic macroinvertebrates. In: Artificial Substrates. J. Cairns, Jr. (ed.). Ann Arbor Science Publisher, Ann Arbor, MI.

Ryti, R.T., and D. Neptune. 1991. Planning issues for superfund site remediation. *Hazardous Material Control*, November/December, 1991. pp. 47-53.

Tetra Tech. 1985. Summary of U.S. EPA-approved methods, standard methods, and other guidance for 301(h) monitoring variables. Final Report, EPA Contract No. 68-01-6938.

Tetra Tech. 1986. Analytical methods for U.S. EPA priority pollutants and 301(h) pesticides in estuarine and marine sediments. Final Report, EPA Contract No. 68-01-69-38.

USEPA/USACE. 1991. Evaluation of dredged material proposed for ocean disposal-testing manual. U.S. Environmental Protection Agency and U.S. Army Corps of Engineers.

USEPA. 1980. Interim guidelines and specifications for preparing quality assurance project plans, U.S. Environmental Protection Agency, Office of Monitoring Systems and Quality Assurance, Office of Research and Development, Publication Number QAMS-005/80, December 29, 1980.

USEPA. 1983. Guidelines and specifications for preparing quality assurance program plans. U.S. Environmental Protection Agency, Office of Research and Development, Quality Assurance Management Staff.

USEPA. 1986. Test methods for evaluating solid waste. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, DC.

USEPA. 1991. Draft analytical method for determination of acid volatile sulfide (AVS) in sediment, proposed technical basis for establishing sediment quality criteria for nonionic organic chemicals using equilibrium partitioning, U.S. Environmental Protection Agency, Criteria and Standards Division, Washington, DC.

USEPA. undated. Cost-Efficient sampling schemes for marine benthic communities. U.S. Environmental Protection Agency, Environmental Research Laboratory - Narragansett and Environmental Research Laboratory Newport, Publication Number ERLN-N156.

Valente, R., and J. Schoenherr. 1991. Environmental monitoring and assessment program, near coastal Virginian Province, quality assurance project plan. Environmental Research Laboratory, U.S. Environmental Protection Agency.

Bulk Sediment Toxicity Test Approach

Nelson Thomas
U.S. Environmental Protection Agency, Environmental Research Laboratory
6201 Congdon Blvd., Duluth, MN 55804, (218) 720-5702

Janet O. Lamberson and Richard C. Swartz
U.S. Environmental Protection Agency, Pacific Ecosystems Branch, ERL-N
2111 SE Marine Science Dr., Newport, OR 97365-5260, (503) 867-4031

In the bulk sediment toxicity test (BSTT) approach, test organisms are exposed in the laboratory to sediments collected in the field. To measure toxicity, a specific biological endpoint is used to assess the response of the organisms to the sediments. The bulk sediment toxicity approach is a descriptive method and cannot be used by itself to generate sediment quality criteria.

3.1 SPECIFIC APPLICATIONS

3.1.1 Current Use

Sediment toxicity testing has been applied in dredged material disposal permit and other regulatory programs in the following ways (USEPA/USACE, 1991).

- To determine potential biological hazards of dredged material intended for disposal in an aquatic environment;
- To evaluate the effectiveness of various dredged material management actions;
- To indicate the spatial distribution of toxicity in contaminated areas, the relative degree of toxicity, and the changes in toxicity along a gradient of pollution or with respect to distance from pollutant sources (Scott and Redmond, 1989; Swartz *et al.*, 1982, 1985b);
- To reveal temporal changes in toxicity (i.e., by sampling the same locations over time or by assaying layers of buried sediment in core samples) (Swartz *et al.*, 1986, 1991);
- To reveal hot spots of contaminated sediment for further investigation (Chapman, 1986); and
- To rank sediments based on toxicity to benthic organisms and to define cleanup boundaries of small or large problem areas of contaminated sediment.

BSTT integrates interactions among complex mixtures of contaminants that may be present in the field. Many classes of chemical contaminants, including metals, polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), dioxins, and chlorinated pesticides can contribute to toxicity in effluents and sediments (Chapman *et al.*, 1982). The BSTT measures the total toxic effect of all contaminants, regardless of their physical and chemical composition.

3.1.2 Potential Use

By itself, BSTT cannot generate chemical-specific toxic effects data, but it can determine toxicity. Used in conjunction with toxicity identification evaluation procedures (Ankley *et al.*, 1990) such as those described in Chapters 5, 10, and 11, BSTT could help identify causal toxicants. To generate sediment quality criteria, the procedure must be combined with other methods of estimating sediment quality such as the Triad (Chapman, 1986b; Chapman *et al.*, 1987; see Chapter 10) and the Apparent Effects Threshold (AET) approach (Tetra Tech, 1986; PTI, 1988; see Chapter 11). BSTT will be most valuable in verifying other methods used to develop sediment quality criteria.

3.2 DESCRIPTION

3.2.1 Description of Methods

The toxicological approach involves exposing test organisms to sediments. The chemical composition of the sediments, which may be complex, need not be known. At the end of a specified time period, the response of the test organisms is examined in relation to a specified biological endpoint (e.g., mortality, growth, reproduction, cytotoxicity, alterations in development or respiration rate). Results are then statistically compared with control and reference sediment results to estimate sediment toxicity.

3.2.1.1 Objectives and Assumptions

The objective of BSTT is to derive toxicity data that can be used to predict whether the test sediment will be harmful to benthic biota. It is assumed that the behavior of chemicals in test sediments in the laboratory is similar to that in natural *in situ* sediments. The effects of various interactions (e.g., synergism, additivity, antagonism) among chemicals in the field or in dredged materials can be predicted from laboratory results without measuring total or bioavailable concentrations of potentially hundreds of contaminants in the test sediment (Swartz *et al.*, 1989) and without *a priori* knowledge of specific pathways of interaction between sediments and test organisms (Kemp and Swartz, 1989). One of the strengths of this test is to integrate the effects of all contaminants. However, the effect of individual contaminants cannot be determined by BSTT, therefore limiting its use in source control. This method can be used for all classes of sediments and any chemical contaminants, but not to answer cause-and-effect questions.

3.2.1.2 Level of Effort

Implementation of this procedure requires a moderate amount of laboratory effort. A variety of toxicity test procedures (see Methods below) have been developed and are fairly straightforward and well documented.

3.2.1.2.1 Type of Sampling Required

It is recommended that bulk sediments be collected for analysis of total solids, acid volatile sulfide, grain size, and total and dissolved organic carbon (ASTM, 1990a). Bulk and interstitial concentrations of chemicals of interest can be determined in subsamples of the sediment added to the toxicity test chambers to enhance the interpretation of toxicity results. However, methods for sampling interstitial water have not been standardized (ASTM, 1990b). Sediment variables such as pH and Eh should also be monitored.

3.2.1.2.2 Methods

The American Society for Testing and Materials (ASTM) has developed standard guidelines for several BSTTs (ASTM, 1990a, 1991). The most commonly used of these partial life cycle tests feature the marine amphipods *Rhepoxynius abronius*, *Eohaustorius estuarius*, *Ampelisca abdita*, and *Grandidierella japonica* (ASTM, 1990a); the freshwater/estuarine amphipod *Hyalella azteca* (ASTM, 1990c); and the freshwater chironomid species *Chironomus tentans* and *C. riparius* (ASTM, 1990c). Brief generalized descriptions of these tests are given below.

BSTTs with the two freshwater chironomid species are functionally very similar, differing only in the age of the organisms with which the test is initiated and the duration of the test. Both *C. tentans* and *C. riparius* are available from various aquatic toxicology laboratories and commercial sources, and both species are cultured easily in a laboratory setting. Toxicity tests are initiated by adding *C. riparius* <3 days old or *C. tentans* 10-14 days old (second instar) to test chambers that contain bulk sediment with overlying water in various ratios (e.g., 6 water:1

sediment; Giesy *et al.*, 1988). The length of the test also varies with the biological endpoint of interest and the species used. If the biological endpoint of interest is growth and survival of the larvae, the test is terminated after 10-14 days by sieving the *C. riparius* or *C. tentans* from the sediment. It also is possible to conduct the test until the adults emerge, which will occur (depending on temperature) in approximately 30 days for *C. riparius* and 20-25 days for *C. tentans*. Toxicity test procedures with *C. riparius* and *C. tentans* are given in more detail in Adams *et al.* (1985), Nebeker *et al.* (1984), Giesy *et al.* (1988), Ingersoll and Nelson (1989), and ASTM (1991).

Partial life-cycle toxicity tests with the freshwater/estuarine amphipod *H. azteca* and bulk sediments have been conducted in a number of laboratories. *H. azteca* are available from various aquatic toxicology laboratories and commercial sources and can be cultured easily in a laboratory. Toxicity tests are initiated by adding juveniles <7 days old to test chambers that contain bulk sediment with overlying water in various ratios (e.g., 4 water:1 sediment; Ingersoll and Nelson, 1989). The length of the test can range from ≤10 days (short-term partial life-cycle test) to 30 days (long-term partial life-cycle test) (Nebeker *et al.*, 1984; Ingersoll and Nelson, 1989). Depending on the length of the test, biological endpoints include survival, behavior, growth, and reproduction. More detailed descriptions of toxicity test procedures are given by Nebeker *et al.* (1984), Nebeker and Miller (1988), Ingersoll and Nelson (1989), and ASTM (1991).

Partial life-cycle toxicity tests with the marine amphipods *Rhepoxynius abronius*, *Eohaustorius estuarius*, *Ampelisca abdita*, and *Grandidierella japonica* and bulk sediments have been used for some time (Swartz *et al.*, 1985a). Amphipods and bulk sediments generally are collected from the field and acclimated to laboratory conditions for 2-24 days before toxicity testing. The tests are initiated by adding immature or adult amphipods to test chambers that contain bulk sediment with overlying water in various ratios. The length of the test generally is ≥10 days, and the biological responses monitored consist of behavioral effects (e.g., emergence from the sediment, ability to burrow in clean sediment after exposure to test sediment) and mortality. More detailed descriptions of the toxicity test procedures are given by Swartz *et al.* (1985a), DeWitt *et al.* (1989), Nipper *et al.* (1989), Scott and Redmond (1989), ASTM (1990a), and the Puget Sound Estuary Program (1991). Chronic test procedures for marine and estuarine amphipods are under development at several laboratories. Other test procedures for marine and estuarine polychaetes, pelecypods, shrimp, and fish are described in the USEPA/USACE (1991) and Reish and LeMay (1988) manuals for testing dredged materials before disposal.

3.2.1.2.3 Types of Data Required

The physical and chemical data described above under Section 3.2.1.2.1, Type of Sampling Required, are needed to interpret the test results. The required biological data (which vary by test) may include mortality and various sublethal effects (e.g., changes in growth, reproduction, respiration rate, behavior, or development). These data can be compared to control and reference data to determine the occurrence of biological effects (ASTM, 1990a). Dilution experiments in which uncontaminated sediment is added to test sediment collected from the field can be used to calculate LC_{50} values, EC_{50} values, no-effect concentrations, and lowest-observable-effect concentrations (Swartz *et al.*, 1989).

3.2.1.2.4 Necessary Hardware and Skills

In general, only readily available and inexpensive field and laboratory equipment is needed, procedures are fairly simple and straightforward, and a minimum of training is necessary to detect endpoints through toxicity tests. Interpretation of the toxicity data (chemical and biological) requires a higher degree of skill and training. Chemical sampling methods are generally simple and routine, although analysis of chemical samples requires specialized training and equipment. Some biological effects tests also require specialized training, handling, and facilities.

3.2.1.3 Adequacy of Documentation

Various sediment toxicity test procedures have been developed and well documented for testing field sediments (ASTM 1990a; Chapman 1986a, 1988; Lamberson and Swartz, 1988; Melzian, 1990; Puget Sound Estuary Program (PSEP) 1991; Swartz, 1987; Thompson *et al.*, 1989; USEPA/ USACE, 1991). Although standardization of methodology is progressing, intercalibration among laboratories and better field validation are needed.

3.2.2 Applicability of Method to Human Health, Aquatic Life, or Wildlife Protection

The BSTT approach is suitable only for protection of aquatic life. Sediment toxicity test procedures incorporate a direct measure of sediment biological effects and can be used to predict biological effects of contaminated sediments before approval of state and federal disposal permits. These procedures can be used to assess the toxicity of sediments in the natural environment and to predict the effects of these sediments on resident aquatic life. Combined with other approaches such as the AET and the Triad approaches (Chapman, 1986b), BSTTs can be used to establish sediment quality criteria. Use of the most sensitive species within a benthic community as a test organism will serve to protect the structure and function of the entire ecosystem (Becker *et al.*, 1990).

3.2.3 Ability of Method to Generate Numerical Criteria for Specific Chemicals

The BSTT approach cannot be used by itself to generate sediment quality criteria. Instead it must be combined with chemical measurements and other data to generate information on the effects of individual contaminants. Both the Triad and the AET approaches rely on bulk sediment toxicity data to derive numerical criteria. BSTTs in conjunction with sediment quality criteria derived from equilibrium partitioning (USEPA, 1980; Swartz *et al.*, 1990) can also be used in assessments of potentially contaminated sediments (see Chapter 6, Equilibrium Partitioning Approach).

3.3 USEFULNESS

3.3.1 Environmental Applicability

3.3.1.1 Suitability for Different Sediment Types

The sediment toxicity test approach is suitable for any type of sediment. In some cases, the physical or chemical properties of the test sediment, such as salinity or grain size, may limit the selection of organisms that can be used for testing (Ott, 1986; DeWitt *et al.*, 1989). Appropriate controls or statistical models (DeWitt *et al.*, 1988) for sediment properties may be necessary to discriminate chemical toxicity from conventional effects. In establishing sediment quality criteria, the effects of features of the sediment itself, such as grain size, must be recognized (DeWitt *et al.*, 1988). Data can be normalized to such factors as organic carbon or acid volatile sulfide (DiToro *et al.*, 1990, 1991; Nebeker *et al.*, 1989) and thus can be applied to any sediment. However, normalization techniques are in the developmental stage (see Chapter 6, Equilibrium Partitioning Approach).

3.3.1.2 Suitability for Different Chemicals or Classes of Chemicals

BSTT is the only currently available approach that directly measures the biological effects of all classes of chemicals, including the combined interactive (additive, synergistic, antagonistic) toxic effects among individual chemicals in mixtures of contaminants usually found in field sediments (Plesha *et al.*, 1988; Swartz *et al.*, 1989). Bioaccumulative chemicals can be evaluated if the length of the test is extended to ensure adequate exposure of the test organism.

3.3.1.3 Suitability for Predicting Effects on Different Organisms

Theoretically, any organism can be used in sediment toxicity testing. To protect a biological community and to predict the effects of contaminated sediments on different organisms, test organisms should be selected on the basis of their sensitivity to contaminants, their ability to withstand laboratory handling, and their ability to survive in control and reference treatments (DeWitt *et al.*, 1989; Reish and LeMay, 1988; Shuba *et al.*, 1981). In tests to determine the effects of contaminated sediments on a particular biological community, the test species selected should be among the most sensitive found in the community of interest, or should be comparably sensitive. Test species should include more than one type of organism to ensure a range of sensitivity to various types of contaminants (Becker *et al.*, 1990).

3.3.1.4 Suitability for In-Place Pollutant Control

Sediment toxicity testing can be used directly to monitor in-place pollution. As discussed in Section 3.2.1.1, sediment toxicity testing can be used to determine the extent of the problem area, monitor temporal and spatial trends, detect the presence of unsuspected hot spots, assess the need for remedial actions, and monitor changes in toxicity after remediation. Such tests can also be used as a cost-effective and rapid screening tool for *in situ* pollutant reconnaissance surveys and in *a priori* simulations of proposed remedial actions to test the effectiveness of capping or other remedial alternatives.

3.3.1.5 Suitability for Source Control

Bulk field sediment toxicity testing can be used to identify suspected sources of sediment pollution. Field reconnaissance surveys can reveal hot spots near contaminant sources, and a map showing contours of sediment toxicity values can reveal gradients that identify point and nonpoint sources (Swartz *et al.*, 1982). Toxicity testing cannot be used by itself to verify reductions in the mass loading of chemicals that might be expected as a result of source control. However, the biological effects of source control can be represented through the use of BSTT.

3.3.1.6 Suitability for Disposal Applications

BSTT has been used widely in regulatory programs to determine the toxicity of material before disposal (Reish and LeMay, 1988; USEPA/USACE, 1991). The potential hazard to benthic organisms at the disposal site (which is determined by making comparisons with the "reference" sediments collected near the disposal site) can be predicted from laboratory toxicity test results. Sediment toxicity tests also can be used to monitor conditions at the disposal site both before and after a disposal operation.

3.3.2 General Advantages and Limitations

3.3.2.1 Ease of Use

Most sediment toxicity test procedures are simple to use, requiring limited expertise and standard inexpensive laboratory equipment (PSEP, 1991). Only a few sublethal effects tests require specialized training. Field sampling requires only readily available equipment and standard procedures (ASTM, 1990b).

3.3.2.2 Relative Cost

Individual laboratory toxicity tests and field sampling are cost-effective because they require limited expertise and inexpensive equipment. Such costs generally range from $150 to $500 per sampling replicate. Laboratory sediment toxicity testing is a comparatively inexpensive and cost-effective method of monitoring the field distribution of sediment toxicity because it integrates the effects of all toxic contaminants, does not require individual chemical measurements, and does not require time-consuming analysis of benthic community structure.

3.3.2.3 Tendency to Be Conservative

Sediment toxicity tests can be made as sensitive or as conservative (i.e., environmentally protective) as necessary through selection of biological endpoints and species of test organism. Reliance on mortality as an endpoint may be underprotective, while some sublethal endpoints (e.g., enzyme inhibition) may be overprotective.

3.3.2.4 Level of Acceptance

BSTT is widely accepted by the scientific and regulatory communities and has been tested and contested in court. Field sediment toxicity test results have been published widely in peer-reviewed journals and incorporated into other measures of sediment quality such as the AET and the Triad approaches. Standard guides for sediment toxicity testing continue to be developed by ASTM (1990a, 1990b, 1991), and field sediment toxicity testing is incorporated into most dredged material disposal regulatory programs (PSEP, 1991; Reish and LeMay, 1988; USEPA/USACE, 1991). Toxicity testing in general has long been the basis for water quality criteria, dredged material testing, effluent testing, and discharge monitoring.

3.3.2.5 Ability to Be Implemented by Laboratories with Typical Equipment and Handling Facilities

Sediment toxicity test methods are easily implemented by laboratories with typical equipment using inexpensive glassware and procedures requiring little specialized training, although the interpretation of some sublethal biological endpoints may require some degree of training and experience. Field sediment sample collection procedures are routine.

3.3.2.6 Level of Effort Required to Generate Results

This procedure consists of field sampling and a laboratory toxicity test. Compared to an extensive survey of chemical concentrations or benthic community structure analysis, the level of effort is relatively small.

3.3.2.7 Degree to Which Results Lend Themselves to Interpretation

Biological responses to toxic sediment can be easily interpreted. Generally, data fit "pass-fail" criteria (i.e., the result is either above or below a predetermined acceptance level) or the result is compared statistically to control and reference results to determine whether there is a toxic effect. Little expert guidance is required for interpretation of mortality data although chronic or sublethal effects might require some explanation.

3.3.2.8 Degree of Environmental Applicability

As noted in Section 3.3.1.1, the sediment toxicity test approach applies to a wide range of environmental conditions and sediment types. The effects of various sediment properties such as grain size and organic content can be addressed experimentally with appropriate uncontaminated controls.

3.3.2.9 Degree of Accuracy and Precision

Because the sediment toxicity test is a laboratory-controlled experiment, its results have a high degree of accuracy, precision, and repeatability.

3.4 STATUS

3.4.1 Extent of Use

Sediment toxicity tests are widely used in research and regulatory programs in both marine and freshwater systems (ASTM, 1990a, 1991), as described in Section 3.2.1.1. Sediment toxicity tests also are incorporated into the evaluation of applications for dredged material disposal permits and are used to assess the toxicity of sediments subject to regulatory decisions. BSTTs are used to investigate the mechanisms of sediment toxicity to benthic organisms (Kemp and Swartz, 1989; Swartz *et al.*, 1988).

3.4.2 Extent to Which Approach Has Been Field-Validated

Field validation of BSTT includes several publications in peer-reviewed literature (Chapman, 1986b; Plesha *et al.*, 1988; Swartz *et al.*, 1982, 1986, 1989). As more data become available, results can be compared with available information on contaminant concentrations in sediment in areas where biological effects have been observed. The effects of interactions among contaminants, as well as the effects of nonchemical sediment variables, must be taken into consideration when attempts are made at field validation (DeWitt *et al.*, 1988; Swartz *et al.*, 1989). As noted in Section 3.2.1.3, better field validation of predicted effects is needed.

3.4.3 Reasons for Limited Use

BSTT has been widely used in research and regulatory programs (see Section 3.4.1, Extent of Use).

3.4.4 Outlook for Future Use and Amount of Development Yet Needed

The outlook for future use of sediment toxicity tests is promising where direct measurement of biological effects of toxicants in sediments is desired, especially where the effects of chemical interactions are of interest. Development and standardization of biological testing methods should continue, especially for tests using species locally available in geographic areas that have not been represented such as tropical and arctic regions. More emphasis should be placed on the development of procedures to measure chronic effects. Methods should be compared and standardized among laboratories, and results should be field-validated to establish their ability to predict biological effects on populations and communities in the field. As more toxicity tests are conducted and the results subject to a quality assurance review, results should be compiled in a central database so that comparisons can be made among species, methods, and laboratories.

3.5 REFERENCES

Adams, W.J., R.A. Kimerle, and R.G. Mosher. 1985. Aquatic safety assessment of chemicals sorbed to sediments. pp. 429-453. In: Aquatic Toxicology and Hazard Assessment: Proceedings of the Seventh Annual Symposium. ASTM STP 854. R.D. Cardwell, R. Purdy and R.C. Bahner (eds.). American Society for Testing and Materials, Philadelphia, PA.

Ankley, G.T., A. Katko, and J.W. Arthur. 1990. Identification of ammonia as an important sediment-associated toxicant on the lower Fox River and Green Bay, Wisconsin. Environ. Toxicol. Chem. 9:313-322.

ASTM. 1990a. E 1367-90. Guide for conducting 10-day static sediment toxicity tests with marine and estuarine amphipods. In: Annual Book of ASTM Standards, Water and Environmental Technology, Vol. 11.04. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1990b. E 1391-90. Guide for collection, storage, characterization and manipulation of sediments for toxicological testing. In: Annual Book of ASTM Standards, Water and Environmental Technology, Vol. 11.04. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1990c. E 1383-90. Guide for conducting sediment toxicity tests with freshwater invertebrates. In: Annual Book of ASTM Standards, Water and Environmental Technology, Vol. 11.04. American Society for Testing and Materials, Philadelphia, PA.

Becker, D.S., G.R. Bilyard, and T.C. Ginn. 1990. Comparisons between sediment bioassays and alterations of benthic macroinvertebrate assemblages at a marine Superfund site: Commencement Bay, Washington. Environ. Toxicol. Chem. 9: 669-685.

Chapman, P.M. 1986a. Sediment bioassay tests provide data necessary for assessment and regulation. In: Proceedings of the Eleventh Annual Aquatic Toxicology Workshop; Technical Report 1480. Green, G.H. and K.L. Woodward (eds.). Fish. Aquat. Sci., pp. 178-197.

Chapman, P.M. 1986b. Sediment quality criteria from the sediment quality triad: an example. Environ. Toxicol. Chem. 5: 957-964.

Chapman, P.M. 1988. Marine sediment toxicity tests. In: Chemical and Biological Characterization of Sludges, Sediments, Dredge Spoils, and Drilling Muds, ASTM STP 976, J.J. Lichtenberg, F.A. Winter, C.I. Weber, and L. Fradkin, (eds.). American Society for Testing and Materials, Philadelphia, PA. pp. 391-402.

Chapman, P.M., G.A. Vigers, M.A. Farrell, R.N. Dexter, E.A. Quinlan, R.M. Kocan, and M. Landolt. 1982. Survey of biological effects of toxicants upon Puget Sound biota. 1. Broad-scale toxicity survey. NOAA Technical Memorandum OMPA-25, Boulder CO.

Chapman, P.M., R.N. Dexter, and E. R. Long. 1987. Synoptic measures of sediment contamination, toxicity and infaunal community composition (the sediment quality triad) in San Francisco Bay. Mar. Ecol. Prog. Ser. 37: 75-96.

DeWitt, T.H., G.R. Ditsworth, and R.C. Swartz. 1988. Effects of natural sediment features on the phoxocephalid amphipod, *Rhepoxynius abronius*: Implications for sediment toxicity bioassays. Mar. Environ. Res. 25: 99-124.

DeWitt, T.H., R.C. Swartz, and J.O. Lamberson. 1989. Measuring the toxicity of estuarine sediments. Environ. Toxicol. Chem. 8: 1035-1048.

DiToro, D.M., J.D. Mahony, D.J. Hansen, K.J. Scott, M.B. Hicks, S.M. Mayr, and M.S. Redmond. 1990. Toxicity of cadmium in sediments: the role of acid volatile sulfide. Environ. Toxicol. Chem. 9: 1487-1502.

DiToro, D.M., J.D. Mahony, D.J. Hansen, K.J. Scott, A.R. Carlson, and G.T. Ankley. 1991. Acid volatile sulfide predicts the acute toxicity of cadmium and nickel in sediments. Environmental Science and Technology.

Giesy, J.P., R.L. Graney, J.L. Newsted, C.J. Rosiu, A. Benda, R.G. Kreis, and F.J. Horvath. 1988. Comparison of three sediment bioassay methods using Detroit River sediments. Environ. Toxicol. Chem. 7: 483-498.

Ingersoll, C.G., and M.K. Nelson. 1989. Solid-phase sediment toxicity testing with the freshwater invertebrates: *Hyalella azteca* (Amphipoda) and *Chironomus riparius* (Diptera). In: Aquatic Toxicology Risk Assessment: Proceedings of the Thirteenth Annual Symposium, ASTM STP; American Society for Testing and Materials, Philadelphia, PA.

Kemp, P.F., and R.C. Swartz. 1989. Acute toxicity of interstitial and particle-bound cadmium to a marine infaunal amphipod. Marine Environ. Res. 26: 135-153.

Lamberson, J.O., and R.C. Swartz. 1988. Use of bioassays in determining the toxicity of sediment to benthic organisms, Chapter 13, In: Toxic Contaminants and Ecosystem Health: A Great Lakes Focus. Evans, M.S. (ed.). John Wiley and Sons, New York, NY. pp. 257-279.

Melzian, B.D. 1990. Toxicity assessment of dredged materials: acute and chronic toxicity as determined by bioassays and bioaccumulation tests. In: Proceedings of the International Seminar on the Environmental Aspects of Dredging Activities, Goubault Impremeur, Nantes; France, pp. 49-64.

Nebeker, A.V., M.A. Cairns, J.H. Gakstatter, K.W. Maleug, G.S. Schuytema, and D.F. Krawczyk. 1984. Biological methods for determining toxicity of contaminated freshwater sediments to invertebrates. Environ. Toxicol. Chem. 3: 617-630.

Nebeker, A.V. and C.E. Miller. 1988. Use of the amphipod crustacean *Hyalella azteca* in freshwater and estuarine sediment toxicity tests. Environ. Toxicol. Chem. 7: 1027-1034.

Nebeker, A.V., G.S. Schuytema, W.L. Griffis, J.A. Barbitta, and L.A. Carey. 1989. Effect of sediment organic carbon on survival of *Hyalella azteca* exposed to DDT and endrin. Environ. Toxicol. and Chem. 8: 705-718.

Nipper, M.G., D.J. Greenstein, and S.M. Bay. 1989. Short- and long-term sediment toxicity test methods with the amphipod *Grandidierella japonica*. Environ. Toxicol. and Chem. 8:1191-1200.

Ott, F.S. 1986. Amphipod sediment bioassays: Effect of grain size, cadmium, methodology, and variations in animal sensitivity on interpretation of experimental data. Ph.D. dissertation. University of Washington, Seattle, WA.

Plesha, P.D., J.E. Stein, M.H. Schiewe, B.B. McCain, and U. Varanasi. 1988. Toxicity of marine sediments supplemented with mixtures of selected chlorinated and aromatic hydrocarbons to the infaunal amphipod, *Rhepoxynius abronius*. Mar. Environ. Res. 25: 85-97.

Puget Sound Estuary Program. 1991. Recommended guidelines for conducting laboratory bioassays on Puget Sound sediments. Draft report prepared for U.S. Environmental Protection Agency, Region 10, Office of Puget Sound, Seattle, WA.

PTI Environmental Services. 1988. Sediment quality values refinement: Tasks 3 and 5 - 1988 update and evaluation of the Puget Sound AET. PTI Environmental Services, Bellevue, WA.

Reish, D.J., and J.A. Lemay. 1988. Bioassay manual for dredged materials. Contract DACW-09-83R-005. U.S. Army Corps of Engineers, Los Angeles District, Los Angeles, CA.

Scott, K.J., and M.S. Redmond. 1989. The effects of a contaminated dredged material on laboratory populations of the tubicolous amphipod, *Ampelisca abdita*. In: Aquatic Toxicology and Hazard Assessment: Vol 12. U. M. Cowgill and L. R. Williams (eds.). ASTM STP 1027. American Society for Testing and Materials, Philadelphia, PA.

Shuba, P.J., S.R. Petrocelli, and R.E. Bentley. 1981. Considerations in selecting bioassay organisms for determining the potential environmental impact of dredged material. Technical Report EL-81-8. U.S. Army Engineer Waterways Experimental Station, Vicksburg, MS.

Swartz, R. C. 1987. Toxicological methods for determining the effects of contaminated sediment on marine organisms. pp. 183-198. In: Fate and Effects of Sediment Bound Chemicals in Aquatic Systems. K. L. Dickson, A.W. Maki, and W. A. Brungs (eds.). Pergamon Press, New York.

Swartz, R.C., W.A. DeBen, K.A. Sercu, and J.O. Lamberson. 1982. Sediment toxicity and the distribution of amphipods in Commencement Bay, Washington, USA. Mar. Poll. Bull. 13: 359-364.

Swartz, R.C., W.A. DeBen, J.K.P. Jones, J.O. Lamberson, and F.A. Cole. 1985a. Phoxocephalid amphipod bioassay for marine sediment toxicity. In: Aquatic Toxicology and Hazard Assessment. R.D. Cardwell, R. Purdy and R.C. Bahner (eds.). ASTM STP 854, pp. 284-307. American Society for Testing and Materials, Philadelphia, PA.

Swartz, R.C., D.W. Schults, G.R. Ditsworth, W.A. DeBen, and F.A. Cole. 1985b. Sediment toxicity, contamination, and macrobenthic communities near a large sewage outfall. pp. 152-175. In: Validation and Predictability of Laboratory Methods for Assessing the Fate and Effects of Contaminants in Aquatic Ecosystems. T.P. Boyle (ed.). ASTM STP 865. American Society for Testing and Materials, Philadelphia, PA.

Swartz. R.C., F.A. Cole, D.W. Schults, and W.A. DeBen. 1986. Ecological changes on the Palos Verdes Shelf near a large sewage outfall: 1980-1983. Mar. Ecol. Prog. Ser. 31: 1-13.

Swartz, R.C., P.F. Kemp, D.W. Schults, and J.O. Lamberson. 1988. Effects of mixtures of sediment contaminants on the marine infaunal amphipod, *Rhepoxynius abronius*. Environ. Toxicol. Chem. 7: 1013-1020.

Swartz, R.C., P.F. Kemp, D.W. Schults, G.R. Ditsworth, and R.O. Ozretich. 1989. Acute toxicity of sediment from Eagle Harbor, Washington, to the infaunal amphipod *Rhepoxynius abronius*. Environ. Toxicol. and Chem. 8: 215-222.

Swartz, R.C., D.W. Schults, T.H. DeWitt, G.R. Ditsworth, and J.O. Lamberson. 1990. Toxicity of fluoranthene in sediment to marine amphipods: a test of the equilibrium partitioning approach to sediment quality criteria. Environ. Toxicol. and Chem. 9: 1071-1080.

Swartz, R.C., D.W. Schults, J.O. Lamberson, R.J. Ozretich, and J.K. Stull. 1991. A toxicological record in cores of contaminated sediment. Mar. Environ. Res.

Tetra Tech, Inc. 1986. Eagle Harbor preliminary investigation. Final Report EGHB-2, TC-3025-003. Tetra Tech, Inc., Bellevue WA.

Thompson, B.E., S.M. Bay, J.W. Anderson, J.D. Laughlin, D.J. Greenstein, and D.T. Tsukada.

1989. Chronic effects of contami-nated sediments on the urchin *Lytichinus pictus*. Environ. Toxicol. and Chem. 8: 629-637.

USEPA. 1980. Water quality criteria for fluoranthene. U.S. Environmental Protection Agency, Washington, DC.

USEPA/USACE. 1991. Evaluation of dredged material proposed for ocean disposal-testing manual. EPA-503-8-91/001. U.S. Environmental Protection Agency and U.S. Army Corps of Engineers, Washington, DC.

Spiked-Sediment Toxicity Test Approach

Janet O. Lamberson and Richard C. Swartz
U.S. Environmental Protection Agency, Pacific Ecosystems Branch, ERL-N
2111 Southeast Marine Science Dr., Newport, OR 97365-5260
(503) 867-4031

The toxicological approach to generating sediment quality criteria uses concentration-response data from sediments spiked in the laboratory with known concentrations of contaminants. Sediments are spiked to establish cause-and-effect relationships between chemicals and adverse biological responses (e.g., mortality, reduction in growth or reproduction, physiological changes). Individual chemicals or other potentially toxic substances can be tested alone or in combination to determine toxic concentrations of contaminants in sediment. This approach can be used to generate sediment quality criteria or to validate sediment quality criteria generated by other approaches.

4.1 SPECIFIC APPLICATIONS

4.1.1 Current Use

The spiked-sediment toxicity test (SSTT) approach is in the research stage. Although the procedures used resemble those used to generate water quality criteria, the influence of the variable properties of sediment makes generating quality criteria values much more complex.

Where LC_{50} values and chronic effects data are available for chemicals in sediments (see Section 4.3.2.3), they can be used to identify concentrations of chemicals in sediment that are protective of aquatic life. The predictive value of sediment quality criteria generated by this approach should be tested by comparing them with field data on chemical concentrations in natural sediments and observed biological effects. However, interim laboratory-derived criteria can be implemented before field validation.

4.1.2 Potential Use

This method can be used to address empirically the problem of interactions among complex mixtures of contaminants that are almost always present in the field (Swartz *et al.*, 1988, 1989). Chemical-specific data can be generated for a wide variety of classes of chemical contaminants, including metals, PAHs, PCBs, dioxins, and chlorinated pesticides. Both acute and chronic criteria can be established, and the approach is applicable to both marine and freshwater systems (Tetra Tech, 1986; Battelle, 1988). However, unless the sediment factor that normalizes for bioavailability is known, this procedure must be applied to every sediment (i.e., a value derived for one sediment may not be applied with predictable results to another sediment with different properties).

4.2 DESCRIPTION

4.2.1 Description of Method

The toxicological approach involves exposing test organisms to sediments that have been spiked with known quantities of potentially toxic chemicals or mixtures of compounds. At the end of a specified time period, the response of the test organism is examined in relation to a biological endpoint (e.g., mortality, growth, reproduction, cytotoxicity, alterations in development or respiration rate). Results are then statistically compared with results from control or reference sediments to identify toxic concentrations of the test chemical.

4.2.1.1 Objectives and Assumptions

The objective of this approach is to derive in the laboratory concentration-response values that can be used to predict the concentrations of specific chemicals harmful to resident biota under field conditions. The effects of the interactions—synergism, additivity, antagonism—among chemicals in the field can be predicted from laboratory results with sediments spiked with combinations of chemicals. This method can be used for all classes of sediments and any chemical contaminant. The bioavailable component of contaminants in sediment can be determined by this method, and an *a priori* knowledge of specific pathways of interaction between sediments and test organisms is not necessary. Any method of expressing the bioavailability of contaminants in sediment can be used with sediment toxicity tests, including the "free" interstitial concentration and normalization to organic carbon, acid volatile sulfide, and other sediment properties.

Data generated by this method may be difficult to interpret if the normalizing factor for bioavailability is unknown. If the normalization factor is known, this method can be used to validate sediment quality criteria generated by other approaches. It is assumed that laboratory results for a given sediment and overlying water represent biological effects of similar sediments in the field, and that the behavior of chemicals in spiked sediments is similar to that in natural, *in situ* sediments.

4.2.1.2 Level of Effort

Implementation of this procedure requires a moderate to considerable amount of laboratory effort. The various toxicity test procedures that have been developed are generally straightforward and well documented (Lamberson and Swartz, 1988; Melzian, 1990; Nebeker *et al.*, 1984; Swartz *et al.*, 1989; PSEP, 1991). However, many individual tests would be required to generate an extensive database of sediment quality values for a large number of chemicals, chemical combinations, and sediment types.

4.2.1.2.1 Type of Sampling Required

Collection of sediments from the field is required. Depending on the particular study objectives, the sediments may be clean (uncontaminated) sediments from a control area, uncontaminated reference sediments for comparison with similarly contaminated sediments, or contaminated sediments to be spiked with known concentrations of chemicals in a test for interactions among contaminants. Sufficient sediment must be collected to provide samples for chemical analysis, spiking, and reference or controls (i.e., sediment for statistical comparison with spiked sediment). Depending on the experimental design, the following controls might be required: sediment from the collection site for test animals (or culture sediment for laboratory-cultured animals), positive controls with a reference toxicant, carrier controls, and reference sediment controls for natural sediment features that may affect test animals, such as grain size distribution (DeWitt *et al.*, 1988).

4.2.1.2.2 Methods

Various methods of adding chemicals to sediment (spiking sediments) have been used. In general, the chemical is either added to the sediment and mixed in (Birge *et al.*, 1987; Ditsworth *et al.*, 1990; Francis *et al.*, 1984) or added to the overlying water (Hansen and Tagatz, 1980; Kemp and Swartz, 1988) or to a sediment slurry (Landrum, 1989; Oliver, 1984; Schuytema *et al.*, 1984) and allowed to equilibrate with the sediment. Sediments are spiked with a range of concentrations to generate LC_{50} data or to determine a minimum concentration at which biological effects are observed.

The effect of sediment contaminants on benthic biota is determined either by exposing known numbers of individual benthic test organisms to the sediment for a specific length of time (Swartz *et al.*, 1985) or by exposing larvae of benthic species to the sediment in flowing natural waters (Hansen and Tagatz, 1980). Biological responses are determined at the end of the test period using response criteria that include mortality, changes in growth or reproduction, behavioral or physiological alterations, or differences in numbers and species of larvae in contaminated versus control sediments.

4.2.1.2.3 Types of Data Required

Spiked sediments, as well as reference or control sediments, must be analyzed for total solids, grain size, and total and dissolved organic carbon. The concentrations of toxicants added to sediment must be determined in stock solutions as well as in the test sediment. Bulk and interstitial levels of the spiked chemicals in the test sediment must be determined throughout a concentration range at least at the beginning and at the end of the toxicity test. However, methods for sampling interstitial water have not been standardized. If sediment properties that control availability, such as acid volatile sulfides or dissolved or total organic carbon, change during exposure, measurements must be taken before, during and at the end of the exposure period. In addition, these changes must be taken into account in interpreting the data. Sediment parameters such as pH and Eh should also be monitored.

Biological and chemical data are compared statistically with control or reference data to determine the occurrence of biological effects, and can be used to calculate LC_{50} values, EC_{50} values, no-effect concentrations, or lowest-observable-effect concentrations. Establishment of the maximum acceptable toxicant concentration requires data from a chronic or life-cycle test.

Data correlating observed biological effects with chemical concentrations in spiked sediment can be used to calculate probit curves for derivation of biological effect level values (e.g., EC_{50}). Data from several species of test organisms can be ranked, and the lowest contaminant concentrations that affect the most sensitive species can be used to establish sediment quality criteria that will protect the entire benthic community and associated aquatic ecosystem. This approach has regulatory and scientific precedence in the development of water quality criteria.

4.2.1.2.4 Necessary Hardware and Skills

Most toxicity test procedures require a minimum of specialized hardware and level of skill. In general, only readily available and inexpensive laboratory equipment is needed, procedures are fairly simple and straightforward, and a minimum of training is necessary to detect and interpret biological endpoints. Although analysis of chemical samples requires specialized training and equipment, the chemical sampling methods for spiked-sediment toxicity are generally simple and routine. Some biological effects tests also require specialized training and experience, especially to interpret the results.

4.2.1.3 Adequacy of Documentation

Various acute sediment toxicity test procedures have been developed and are well documented for testing freshwater and marine field sediments (Chapman, 1986, 1988; Lamberson and Swartz, 1988; Melzian, 1989; Swartz, 1987). Although only a few of these procedures have been used with laboratory-spiked sediments, most of the established methods could be used with laboratory-prepared sediments as well as with field sediments.

In contrast to acute tests, there are relatively few procedures for testing the chronic effects of contaminated sediments on benthic invertebrates. Life-cycle test methodology has been presented for the amphipods *Ampelisca abdita* (Scott and Redmond, 1989), *Hyalella azteca* (ASTM, 1990c; Borgmann and Munawar, 1989), and *Grandidierella lutosa* and *G. lignorum* (Connell and Airey, 1982); the polychaetes *Neanthes arenaceodentata* (Pesch, 1979) and *Capitella capitata* (Chapman and Fink, 1984); freshwater oligochaetes (Wiederholm *et al.*, 1987); and species of *Daphnia* and *Chironomus* (ASTM, 1991; Nebeker *et al.*, 1988). Chronic exposures to most sensitive life stages are also inherent in the benthic recolonization procedure (Hansen and Tagatz, 1980). Further research is needed to develop and validate methodology for other species.

4.2.2 Applicability of Method to Human Health, Aquatic Life, or Wildlife Protection

Spiked-sediment toxicity tests incorporate a direct measure of sediment biological effects. This approach is the only method that can quantify the interactive effects of combinations of contaminants directly.

When chemical concentrations in tested biota are measured after a spiked-sediment toxicity test, uptake of contaminants by benthic organisms (bioaccumulation) can be predicted. As an important component of food webs in aquatic ecosystems, benthic organisms can contribute toxicants accumulated from contaminated sediments to higher levels of the aquatic food web and ultimately affect human health. Sediment quality criteria and bioaccumulation studies using sediment toxicity test methods can help to set limits on the disposal of toxic sediments and predict uptake of toxicants into food webs. If this approach is combined with chemical analysis of sediment samples and BSTT, these limits can be used to define areas from which food species should not be harvested or consumed or where direct contact with contaminated sediments can be hazardous to human health.

Bioaccumulation studies and sediment quality criteria established using data from SSTT with several benthic species can also be used to protect benthic communities and aquatic species that feed on the benthos. Assuming that a sufficient mix of taxonomic groups is used, a sediment quality criterion based on the responses of the most sensitive species within a benthic community can be developed. This criterion can then be employed to protect the structure and function of the entire ecosystem (Hansen and Tagatz, 1980).

4.2.3 Ability of Method to Generate Numerical Criteria for Specific Chemicals

Laboratory tests with the SSTT approach can be used to measure the effects of specific chemicals in various types of sediments directly and to establish unequivocal analysis of causal effects. Test conditions allow this method to determine the effects of individual chemicals or mixtures of chemicals on benthic biota (Plesha *et al.*, 1988; Swartz *et al.*, 1988, 1989), establish pathways of toxicity, and provide specific effects concentrations (e.g. LC_{50}, EC_{50}, no-effect concentration). The influence of various physical characteristics of the sediment on chemical toxicity also can be determined (DeWitt *et al.*, 1988; Ott 1986). The available data represent concentrations at which toxicity occurs rather than numerical sediment quality criteria. Recent spiked sediment studies have provided data that can be useful in setting preliminary sediment criteria levels based on equilibrium partitioning models and water quality values (Swartz *et al.*, 1990).

Concentration-response data have been generated using SSTT for a variety of chemicals, including metals and organic compounds. Specific data are available for phenanthrene, fluoranthene, zinc, mercury, copper, cadmium, hexachlorobenzene, pentachlorophenol, Aroclor 1242 and 1254, chlordane, DDE, DDT, dieldrin, endosulfan, endrin, sevin, creosote, and kepone (Adams *et al.*, 1985; Cairns *et al.*, 1984; DeWitt *et al.*, 1989; Kemp and Swartz, 1989; McLeese and Metcalfe, 1980; McLeese *et al.*, 1982; Nebeker *et al.*, 1989; Swartz *et al.*, 1986, 1988, 1989; Tagatz *et al.*, 1977, 1979, 1983; Word *et al.*, 1987). Concentrations of nonionic organic compounds are usually normalized to sediment organic carbon or acid volatile sulfide (DiToro *et al.*, 1990, 1991; Nebeker *et al.*, 1989). Normalizing factors for other compounds in sediment currently are being researched.

4.3 USEFULNESS

4.3.1 Environmental Applicability

4.3.1.1 Suitability for Different Sediment Types

The SSTT approach is suitable for any type of sediment. This approach also can be used to establish the bioavailable component of the sediment responsible for the observed toxicity. The effects of various physical properties of the sediment on chemical toxicity can be determined experimentally. In some cases, the physical or chemical properties of the test sediment such as salinity or grain size may limit the species of organisms that can be used for testing, and a substitute species must be used (DeWitt *et al.*, 1988, 1989). When establishing sediment quality criteria, the effects of adverse physical or chemical properties of the sediment itself must be reflected. When factors controlling bioavailability (e.g., organic carbon, acid volatile sulfide) are known, data can be normalized to such factors, and the approach applied to any sediment type.

4.3.1.2 Suitability for Different Chemicals or Classes of Chemicals

A major advantage of the SSTT method is that it is suitable for all classes of chemicals. In addition, it is the only approach currently available that can empirically determine the interactive effects among individual chemicals in mixtures of contaminants usually found in real-world sediments (Swartz *et al.*, 1988, 1989). This approach also can be used to provide experimental validation of sediment quality criteria generated by other approaches.

4.3.1.3 Suitability for Predicting Effects on Different Organisms

Theoretically, any organism can be used in SSTT. To protect a biological community and to predict the effects of a toxicant on different organisms, test organisms should be selected based on the following criteria: (1) their sensitivity to contaminants, (2) their ability to withstand laboratory handling, and (3) their ability to survive in control treatments. Tests to determine the effects of toxicants on a particular biological community should use the most sensitive species found in the community or a species with comparable sensitivity.

4.3.1.4 Suitability for In-Place Pollutant Control

SSTT can be used to develop sediment quality criteria, which will then be used to determine the extent of the problem area. It also can be used to monitor temporal and spatial trends and to assess the need for remedial action. Criteria can be used in setting target cleanup levels and in post-cleanup monitoring of actual contaminant levels.

4.3.1.5 Suitability for Source Control

SSTT can be combined with wasteload allocation models and used in source control to establish maximum allowable effluent concentrations or mass loadings of single chemicals and mixtures of chemicals.

4.3.1.6 Suitability for Disposal Applications

SSTT can be used to predict the biological effects of contaminants before approval of dredged material disposal or sewage outfall permits.

4.3.2 General Advantages and Limitations

4.3.2.1 Ease of Use

Most sediment toxicity test procedures are simple to use, require limited expertise, and use standard laboratory equipment. Some of the sublethal-effects tests require specialized training.

4.3.2.2 Relative Cost

The cost of individual toxicity tests is relatively low because such tests require limited expertise and inexpensive equipment. (See Chapter 3, Bulk Sediment Toxicity Approach.) The costs to implement this approach as a regulatory tool would be comparatively high because SSTT requires the collection of sediment chemistry data for comparison to data established by the sediment toxicity test method. The cost of developing a large toxicological database would be relatively high because of the large number of individual chemicals and sediments that would have to be tested. Generating the chemical and toxicological data necessary to establish a sediment quality criterion for one chemical by this method is estimated to cost $100,000.

4.3.2.3 Tendency to Be Conservative

Laboratory-controlled SSTT experiments provide a high degree of accuracy. The tests are controlled sufficiently to give an estimate of the toxicity of individual chemicals in sediment. Laboratory bioassays, especially acute toxicity tests, are inherently limited in their ability to reflect all of the ecological processes through which sediment contaminants may affect benthic ecosystems in the field.

4.3.2.4 Level of Acceptance

SSTT methods, which follow the procedures and rationale used to develop water quality criteria, are easily interpreted, technically acceptable, and legally defensible. The procedures and resulting data have been accepted and published in peer-reviewed journal articles, and some procedures have been incorporated into standard guidelines by ASTM's subcommittee on sediment toxicology (ASTM, 1990a, 1990c).

4.3.2.5 Ability to Be Implemented by Laboratories with Typical Equipment and Handling Facilities

SSTT methods are implemented easily by laboratories with typical equipment, requiring inexpensive glassware and little specialized training. Spiking sediments may require special handling facilities for preparing stock solutions of highly toxic substances, and the interpretation of some sublethal biological endpoints may require some degree of training and experience.

4.3.2.6 Level of Effort Required to Generate Results

This procedure consists of a laboratory toxicity test and requires a moderate amount of effort to begin and end an experiment. The data generated must be compiled, and some calculations must be made to derive concentration-response relationships. The generation of chemical and biological data required for a large database of sediment quality values based on this approach would require a relatively high level of effort.

4.3.2.7 Degree to Which Results Lend Themselves to Interpretation

Sediment toxicity tests applied to spiked sediments provide an unequivocal analysis of cause-and-effect relationships between toxic chemicals and biological responses. Because the procedures follow the rationale used in the development of water quality criteria, the methods are legally defensible. Toxicity tests have long been accepted by both the public and the scientific community as a basis for water quality criteria and dredged material testing.

4.3.2.8 Degree of Environmental Applicability

The SSTT approach is applicable to a wide range of environmental conditions and sediment types. The confounding effects of sediment variables such as grain size and organic content can be addressed experimentally by using toxicity test methods or can be addressed by using normalization equations (DeWitt *et al.* 1988). A major advantage of SSTT is the ability to predict interactive effects of chemical mixtures such as those found in field sediments.

4.3.2.9 Degree of Accuracy and Precision

Because the SSTT is a laboratory-controlled experiment, results have a high degree of accuracy and precision. The procedure produces a direct dose-response data set for individual chemicals in sediment. Sediment criteria generated by this approach must be field-validated.

4.4 STATUS

4.4.1 Extent of Use

SSTT procedures are under development in several laboratories. Spiking procedures, as well as biological test procedures, are currently being standardized by ASTM's sediment toxicology subcommittee (ASTM, 1990b).

4.4.2 Extent to Which Approach Has Been Field-Validated

Although some results have been published, spiked-sediment toxicity test values have not been well validated in the field, (Plesha *et al.*, 1988; Swartz *et al.*, 1989). As more data and criteria values become available, they can be compared with existing information on contaminant levels in sediment in areas where biological effects have been observed. The effects of interactions among contaminants, as well as the effects of nonchemical sediment variables, must be considered during field validation (DeWitt *et al.*, 1988; Swartz *et al.*, 1989).

4.4.3 Reasons for Limited Use

Although some data have been generated and compared to field conditions, the approach is still in the developmental stage in several laboratories, and a relatively large expenditure of effort will be needed to generate a large database. To date, there have been few comparisons of methods and species sensitivity, and few chronic toxicity tests have been developed.

4.4.4 Outlook for Future Use and Amount of Development Yet Needed

The outlook for future use of SSTTs or other sediment toxicity tests is promising where accurate, direct dose-response data are desired, or where the effects of chemical interactions need to be examined. Development of sediment-spiking and biological-testing methods should continue, methods should be compared and standardized among laboratories, and results should be field-validated to establish their ability to predict biological effects in sediments. As more toxicity tests are conducted, results should be compiled in a central database so that comparisons can be made among species, methods, and laboratories and so that sediment quality criteria can be developed.

4.5 REFERENCES

Adams, W.J., R.A. Kimerle, and R.G. Mosher. 1985. Aquatic safety assessment of chemicals absorbed to sediments. pp. 429-453. In: Aquatic Toxicology and Hazard Assessment: Seventh Symposium, R.D. Cardwell, R. Purdy, and R. C. Bahner (eds.). ASTM STP 854. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1990a. E 1367-90. Guide for conducting 10-day static sediment toxicity tests with marine and estuarine amphipods. In: Annual Book of ASTM Standards, Water and Environmental Technology, Vol. 11.04. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1990b. E 1391-90. Guide for collection, storage, characterization and manipulation of sediments for toxicological testing. In: Annual Book of ASTM Standards, Water and Environmental Technology, Vol. 11.04. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1990c. E 1383-90. Guide for conducting sediment toxicity tests with freshwater invertebrates. In: Annual Book of ASTM Standards, Water and Environmental Technology, Vol. 11.04. American Society for Testing and Materials, Philadelphia, PA.

Battelle. 1988. Overview of methods for assessing and managing sediment quality. Report prepared for U.S. Environmental Protection Agency, Office of Water, Office of Marine and Estuarine Protection, Washington, D.C., Battelle Ocean Sciences, Duxbury, MA.

Birge, W.J., J. Black, S. Westerman, and P. Francis. 1987. Toxicity of sediment-associated metals to freshwater organisms: biomonitoring procedures. pp. 199-218. In: Fate and Effects of Sediment Bound Chemicals in Aquatic Systems. K. L. Dickson, A.W. Maki, and W. A. Brungs (eds.). Pergamon Press, New York.

Borgmann, U., and M. Munawar. 1989. A new standardized sediment bioassay protocol using the amphipod *Hyalella azteca* (Saussure). Hydrobiologia 188/189: 425-431.

Cairns, M.A., A.V. Nebeker, J.H. Gakstatter, and W.L. Griffis. 1984. Toxicity of copper-spiked sediments to freshwater invertebrates. Environ. Toxicol. Chem. 3: 435-445.

Chapman, P.M. 1986. Sediment bioassay tests provide data necessary for assessment and regulation. In: Proceedings of the Eleventh Annual Aquatic Toxicology Workshop. G.H. Green and K.L. Woodward (eds.). Technical

Report 1480. Fish. Aquat. Sci., pp. 178-197.
Chapman, P.M. 1988. Marine sediment toxicity tests. In: Chemical and Biological Characterization of Sludges, Sediments, Dredge Spoils, and Drilling Muds, J.J. Lichtenberg, F.A. Winter, C.I. Weber, and L. Fradkin (eds.). ASTM STP 976, American Society for Testing and Materiáls, Philadelphia, PA. pp. 391-402.
Chapman, P.M., and R. Fink. 1984. Effects of Puget Sound sediments and their elutriates on the life cycle of *Capitella capitata*. Bull. Environ. Contamination Toxicol. 33: 451-459.
Connell, A.D., and D.D. Airey. 1982. The chronic effects of fluoride on the estuarine amphipods *Grandidierella lutosa* and *G. lignorum*. Water Res. 16: 1313-1317.
DeWitt, T.H., G.R. Ditsworth, and R.C. Swartz. 1988. Effects of natural sediment features on the phoxocephalid amphipod, *Rhepoxynius abronius:* Implications for sediment toxicity bioassays. Mar. Environ. Res. 25: 99-124.
DeWitt, T.H., R.C. Swartz, and J.O. Lamberson. 1989. Measuring the toxicity of estuarine sediments. Environ. Toxicol. Chem. 8: 1035-1048.
DiToro, D.M., J.D. Mahony, D.J. Hansen, K.J. Scott, M.B. Hicks, S.M. Mayr, and M.S. Redmond. 1990. Toxicity of cadmium in sediments: the role of acid volatile sulfide. Environmental Toxicology and Chemistry 9: 1487-1502.
DiToro, D.M., J.D. Mahony, D.J. Hansen, K.J. Scott, A.R. Carlson, and G.T. Ankley. 1991. Acid volatile sulfide predicts the acute toxicity of cadmium and nickel in sediments. In press. Environmental Science and Technology.
Ditsworth, G.R., D.W. Schults, and J.K.P. Jones. 1990. Preparation of benthic substrates for sediment toxicity testing. Environ. Toxicol. Chem. 9: 1523-1529.
Francis, P.C., W.J Birge, and J.A. Black. 1984. Effects of cadmium-enriched sediment on fish and amphibian embryo-larval stages. Ecotoxicol. and Environ. Safety 8: 378-387.
Hansen, D.J., and M.E. Tagatz. 1980. A laboratory test for assessing impacts of substances on developing communities of benthic estuarine organisms. pp 40-57. In: Aquatic Toxicology. J.G. Eaton, P.R. Parrish and A.C. Hendricks (eds.). ASTM STP 707. American Society for Testing and Materials, Philadelphia, PA.
Kemp, P.F., and R.C. Swartz. 1989. Acute toxicity of interstitial and particle-bound cadmium to a marine infaunal amphipod. Mar. Environ. Res. 26: 135-153.
Lamberson, J.O., and R.C. Swartz. 1988. Use of bioassays in determining the toxicity of sediment to benthic organisms. pp. 257-279. In: Toxic Contaminants and Ecosystem Health; Evans, M.S., (ed.) A Great Lakes Focus. John Wiley and Sons, New York, NY.
Landrum, P.F. 1989. Bioavailability and toxicokinetics of polycyclic aromatic hydrocarbons absorbed to sediments for the amphipod *Pontoporeia hoyi*. Environ. Sci. Technol. 23: 588-595.
McLeese, D.W., and C.D. Metcalfe. 1980. Toxicities of eight organochlorine compounds in sediment and seawater to *Crangon septemspinosa*. Bull. Environ. Contam. Toxicol. 25: 921-928.
McLeese, D.W., L.E. Burridge, and J. Van Dinter. 1982. Toxicities of five organochlorine compounds in water and sediment to *Nereis virens*. Bull. Environ. Contam. Toxicol. 28: 216-220.
Melzian, B.D. 1990. Toxicity assessment of dredged materials: acute and chronic toxicity as determined by bioassays and bioaccumulation tests. pp. 49-64. In: Proceedings of the International Seminar on the Environmental Aspects of Dredging Activities, Goubault Impremeur, Nantes, France.
Nebeker, A.V., M.A. Cairns, J.H. Gakstatter, K.W. Malueg, G.S. Schuytema, and D.F. Krawczyk. 1984. Biological methods for determining toxicity of contaminated freshwater sediments to invertebrates. Environ. Toxicol. Chem. 3: 617-630.
Nebeker, A.V., S.T. Onjukka, and M.A. Cairns. 1988. Chronic effects of contaminated sediment on *Daphnia magna* and *Chironomus tentans*. Bull. Environ. Contam. Toxicol. 41: 574-581.
Nebeker, A.V., G.S. Schuytema, W.L. Griffis, J.A. Barbitta, and L.A. Carey. 1989. Effect of

sediment organic carbon on survival of *Hyalella azteca* exposed to DDT and endrin. Environ. Toxicol. Chem. 8: 705-718.

Oliver, B.G. 1984. Biouptake of chlorinated hydrocarbons from laboratory-spiked and field sediments by oligochaete worms. Environ. Sci. and Technol. 21: 785-790.

Ott, F.S. 1986. Amphipod sediment bioassays: Effect of grain size, cadmium, methodology, and variations in animal sensitivity on interpretation of experimental data. Ph.D. dissertation, University of Washington, Seattle, WA.

Pesch, C.E. 1979. Influence of three sediment types on copper toxicity to the polychaete *Neanthes arenaceodentata*. Marine Biol. 52: 237-245.

Plesha, P.D., J.E. Stein, M.H. Schiewe, B.B. McCain, and U. Varanasi. 1988. Toxicity of marine sediments supplemented with mixtures of selected chlorinated and aromatic hydrocarbons to the infaunal amphipod, *Rhepoxynius abronius*. Mar. Environ. Res. 25: 85-97.

Puget Sound Estuary Program 1991. Recommended guidelines for conducting laboratory bioassays on Puget Sound sediments. Draft report prepared for U.S. Environmental Protection Agency, Region X, Office of Puget Sound, Seattle, WA.

Schuytema, G.S., P.O. Nelson, K.W. Malueg, A.V. Nebeker, D.F. Krawczyk, A.K. Ratcliff, and J.H. Gakstatter. 1984. Toxicity of cadmium in water and sediment slurries to *Daphnia magna*. Environ. Toxicol. and Chem. 3: 293-308.

Scott, K.J., and M.S. Redmond. 1989. The effects of a contaminated dredged material on laboratory populations of the tubicolous amphipod, *Ampelisca abdita*. In: Aquatic Toxicology and Hazard Assessment: Vol. 12 ASTM STP 1027. U.M. Cowgill and L.R. Williams, (eds.). American Society for Testing and Materials, Philadelphia, PA.

Swartz, R.C. 1987. Toxicological methods for determining the effects of contaminated sediment on marine organisms pp. 183-193. In: Fate and Effects of Sediment Bound Chemicals in Aquatic Systems. K.L. Dickson, A.W. Maki, and W.A. Brungs, (eds.). Pergamon Press, New York.

Swartz, R.C., D.W. Schults, G.R. Ditsworth, W.A. DeBen, and F.A. Cole. 1985. Phoxocephalid amphipod bioassay for marine sediment toxicity. pp 284-307. In: Aquatic Toxicology and Hazard Assessment: Proceedings of the Seventh Annual Symposium. R.D. Cardwell, R. Purdy and R.C. Bahner (eds.). ASTM STP 854, American Society for Testing and Materials, Philadephia, PA.

Swartz, R.C., G.R. Ditsworth, D.W. Schults, and J.O. Lamberson.. 1986. Sediment toxicity to a marine infaunal amphipod: cadmium and its interaction with sewage sludge. Mar. Environ. Res. 18: 133-153.

Swartz, R.C., P.F. Kemp, D.W. Schults, and J.O. Lamberson. 1988. Effects of mixtures of sediment contaminants on the marine infaunal amphipod, *Rhepoxynius abronius*. Environ. Toxicol. Chem. 7: 1013-1020.

Swartz, R.C., P.F. Kemp, D.W. Schults, G.R. Ditsworth, and R.J. Ozretich. 1989. Toxicity of sediment from Eagle Harbor, Washington to the infaunal amphipod, *Rhepoxynius abronius*. Environ. Toxicol. Chem. 8: 215-222.

Swartz, R.C., D.W. Schults, T.H. DeWitt, G.R. Ditsworth, and J.O. Lamberson. 1990. Toxicity of fluoranthene in sediment to marine amphipods: a test of the equilibrium partitioning approach to sediment quality criteria. Environ. Toxicol. Chem. 9: 1071-1080.

Tagatz, M.E., J.M. Ivey, and H.K. Lehman. 1979. Effects of sevin on development of experimental estuarine communities. J. Toxicol. and Environ. Health 5: 643-651.

Tagatz, M.E., J.M. Ivey, J.C. Moore, and M. Tobia. 1977. Effects of pentachlorophenol on the development of estuarine communities. J. Toxicol. and Environ. Health 3: 501-506.

Tagatz, M.E., G.R. Plaia, C.H. Deans, and E.M. Lores. 1983. Toxicity of creosote-contaminated sediment to field- and laboratory-colonized estuarine benthic communities. Environ. Toxicol. Chem. 2: 441-450.

Tetra Tech, Inc. 1986. Development of sediment quality values for Puget Sound. Task 6 Final Report. Tetra Tech, Inc., Bellevue, WA.

Wiederholm, T., A.M. Wiederholm, and G.

Milbrink. 1987. Bulk sediment bioassays with five species of fresh-water oligochaetes. Water, Air and Soil Pollut. 36: 131-154.

Word, J.Q., J.A. Ward, L.M. Franklin, V.I. Cullinan and S.L. Kiesser. 1987. Evaluation of the equilibrium partition theory for estimating the toxicity of the nonpolar organic compound DDT to the sediment dwelling organism *Rhepoxynius abronius*. Report prepared for U.S. Environmental Protection Agency, Criteria and Standards Division, Battelle Washington Environmental Program Office, Washington, DC.

Interstitial Water Toxicity Identification Evaluation Approach

Gerald Ankley and Nelson Thomas
U.S. Environmental Protection Agency, Environmental Research Laboratory
6201 Congdon Boulevard, Duluth, MN 55804
(218) 720-5702

The interstitial water toxicity approach is a multiphase procedure for assessing sediment toxicity using interstitial (pore) water. The use of pore water for sediment toxicity assessment is based on the strong correlations between contaminant concentrations in pore water and observed exposure of benthic macroinvertebrates to sediment-associated contaminants (Adams *et al.*, 1985; Swartz *et al.*, 1985; 1988; 1990; Connell *et al.*, 1988; Knezovich and Harrison, 1988; USEPA, 1989a; DiToro *et al.*, 1990), as well as correlations between the actual toxicity of pore water and bulk sediments to epibenthic or benthic species (Ankley *et al.*, 1991a). The approach combines the quantification of pore water toxicity with toxicity identification evaluation (TIE) procedures to identify and quantify chemical components responsible for sediment toxicity (U.S. Environmental Protection Agency, 1988; 1989b; 1989c, 1991a). TIE involves the use of toxicity-based fractionation procedures to identify toxic compounds in aqueous samples containing mixtures of chemicals (Burkhard and Ankley, 1989). In the interstitial water toxicity method, TIE procedures are implemented in three phases to characterize the nature of the pore water toxicant(s), identify the suspect toxicant(s), and confirm identification of the suspect toxicant(s).

5.1 SPECIFIC APPLICATIONS

5.1.1 Current Use

The TIE procedures described herein were developed over the last 4 years using municipal and industrial effluents from more than 50 locations, as well as sediment samples from more than 10 different sites. They have been used with several aquatic species including cladocerans, fishes, and epibenthic macroinvertebrates. Although the methods were developed largely with freshwater species, they are generally applicable to, and are currently being used with, marine organisms as well. The procedures have proven to be successful in identifying acutely toxic substances in more than 90 percent of the samples to which they have been applied (e.g., Ankley *et al.*, 1990a, 1991b; Kuehl *et al.*, 1990; Amato *et al*, 1991; Norberg-King *et al.*, 1991; Schubauer-Berigan and Ankley, 1991; Ankley and Burkhard, 1992).

5.1.2 Potential Uses

The use of pore water as a fraction to assess sediment toxicity, in conjunction with associated TIE procedures, can provide data concerning specific compounds responsible for toxicity of contaminated sediments. These data could be critical to the success of remediation of toxic sediments, including the control of inputs of contaminants.

In spite of existing uncertainties in preparing and using pore water to assess sediment toxicity, the ability to identify specific toxicants responsible for acute toxicity in contaminated sediments makes pore water an important test fraction. Thus this method, in conjunction with other sediment classification methods, could prove to be extremely valuable.

5.2 DESCRIPTION

5.2.1 Description of Method

The interstitial water toxicity method involves three major steps:

- Isolation of pore water from sediment samples;

- Performance of toxicity tests on pore waters; and
- Application of TIE procedures to pore water fractions.

Pore water can be isolated from sediment samples by compression (squeezing) techniques, displacement of water from sediment via the use of inert gases, centrifugation, extraction via dialysis, and micro-syringe sampling (Knezovich *et al.*, 1987; Knezovich and Harrison, 1988; Sly, 1988; USEPA, 1991b). The most representative pore water samples may be obtained using the latter two procedures. However, the resulting sample volumes are too small to be useful for toxicity tests and associated TIE work. Centrifugation has been used in a number of studies evaluating the toxicity of sediment pore water (Giesy *et al.*, 1988; Swartz *et al.*, 1989; Hoke *et al.*, 1990; Ankley *et al.*, 1990a; Schubauer-Berigan and Ankley, 1991) and comparative studies at Duluth, as well as other laboratories, indicate that centrifugation is a reasonable technique for pore water preparation (Schults *et al.*, 1991; U.S. Environmental Protection Agency, 1991b). Regardless of the techniques chosen for pore water isolation, the method should *not* involve filtration either during or after preparation (Schubauer-Berigan and Ankley, 1991; USEPA, 1991b).

After preparation of pore water, toxicity tests can be performed using either standard test species (e.g., USEPA, 1985a, 1985b) or various types of epibenthic or benthic organisms amenable to toxicity testing in aqueous samples (Ankley *et al.*, 1991a; USEPA, 1991b). In samples exhibiting acute toxicity, it is then possible to directly apply the TIE procedures described below in Section 5.2.1.2.2.

5.2.1.1 Objectives and Assumptions

The objective of the interstitial water toxicity method is to derive chemical-specific toxicity data in the laboratory that can be used to assess sediment toxicity in field situations. With this approach, it is possible to quantify toxicity in a sample and potentially to identify chemical components responsible for toxicity. The major assumption in using this method is that the compounds that are toxic to test organisms in the pore water, as it is isolated in the laboratory, are the same compounds that cause toxicity in sediments *in situ*.

5.2.1.2 Level of Effort

Implementation of this method requires a moderate amount of laboratory effort, both to perform toxicity tests and to conduct TIE studies. The effort expended in the TIE studies will be proportional to the complexity of analyses required for the identification of suspected toxicants.

5.2.1.2.1 Type of Sampling Required

Bulk sediment must be obtained and pore water prepared from the sediments. Routine measurement of certain chemical components of the pore water should be conducted. These measurements should include (but are not limited to) pH, hardness, alkalinity, salinity (where appropriate), dissolved oxygen, sulfides, and ammonia. Certain of these variables, in particular pH, also should be monitored in the bulk sediment.

5.2.1.2.2 Methods

The framework for existing TIE procedures is summarized below. Greater detail (e.g., with respect to all possible results that could be generated) is available elsewhere (USEPA, 1988, 1989b, 1989c), as are specific methods for performing sediment TIEs (USEPA, 1991b).

Toxic sediment samples can potentially contain thousands of chemicals, and usually only a handful are responsible for the observed toxicity. The goal of the TIE method is to identify quickly and cheaply the chemicals causing toxicity. However, components causing toxicity can vary widely, and potential toxicants include cationic metals, polar and nonpolar organics, and anionic inorganics, as well as ammonia or hydrogen sulfide. In addition, when multiple toxicants are present, it must be possible to determine the proportion of the overall toxicity due to each toxicant.

After preparation of pore water and performance of initial toxicity tests, the first step in the TIE is to separate toxic from nontoxic components in the pore water sample. To isolate the toxicants, sample manipulations and subsequent fractionation techniques are used in combination with toxicity tests (toxicity tracking). Each fractionation step consists of manipulations to identify the physical/chemical properties of the sample toxicants, thereby enabling selection of the "correct" analytical technique for detecting, identifying, and quantifying the toxicants in the manipulated samples. Because there may be significantly fewer chemical components in the manipulated samples than in the original sample, the task of deciding which component is causing the toxicity is much easier. The toxicity-based TIE approach enables direct relationships to be established between toxicants and measured analytical data because toxicants are tracked through all sample fractionations, using the most relevant detector available, the organism. Establishing this relationship ultimately results in highly efficient TIEs.

With the toxicity-based TIE approach, detection of synergistic and antagonistic interactions, as well as matrix effects, for the toxicants is possible via toxicity tracking. *A priori* knowledge of the toxicants' behavior in the aqueous phase is not required.

The TIE approach is divided into three phases. Phase I consists of methods to identify the physical/chemical nature of the constituents causing acute toxicity. Phase II describes fractionation schemes and analytical methods to identify the toxicants, and Phase III presents procedures to confirm that the suspected toxicants are the cause of toxicity.

Phase I: Toxicant Characterization—In Phase I, the physical/chemical properties of toxicants are characterized by performing manipulations to alter or render biologically unavailable generic classes of compounds with similar properties. Toxicity tests, performed in conjunction with the manipulations, provide information on the nature of the toxicants. Successful completion of Phase I occurs when both the nature of the components causing toxicity, as well as their consistent presence in a number of samples, can be established. After Phase I, the toxicants can be tentatively categorized as having chemical characteristics of cationic metals, nonpolar organics, polar organics, volatiles, oxidants, and/or substances whose toxicity is pH-dependent.

An overview of the sample manipulations employed in Phase I is shown in Figure 5-1. Not shown in Figure 5-1, but performed on all samples, are routine water chemistry measurements including pH, hardness, conductivity, and dissolved oxygen. These routine measurements are needed for designing sample manipulations and interpreting test data. The manipulations shown in Figure 5-1 are usually sufficient to characterize toxicity caused by a single chemical. When multiple toxicants are present, various combinations of the Phase I manipulations will most likely be required for toxicant characterization.

Many of the manipulations in Phase I require samples that have been pH-adjusted. The adjustment of pH is a powerful tool for detecting cationic and anionic toxicants since their behavior is strongly influenced by pH. By changing pH, the ratio of ionized to un-ionized species in solution for a chemical is changed significantly. The ionized and un-ionized species have different physical/chemical properties as well as toxicities. In Phase I, pH manipulations are used to examine two different questions:

- Is the toxicity different at various pHs?
- Does changing the pH, performing a sample manipulation, and then readjusting to ambient pH affect toxicity?

The graduated pH test examines the first question, and the pH adjustment, aeration, filtration, and solid phase extraction (SPE) manipulations examine the second.

In the graduated pH test, the pH of a sample is adjusted within a physiologically tolerable range (e.g., pH 6.0, 7.0, and 8.0) before toxicity testing. In many instances, the un-ionized form of a toxicant is able to cross biological membranes more readily than the ionized form and thus is more toxic. This test is designed primarily for

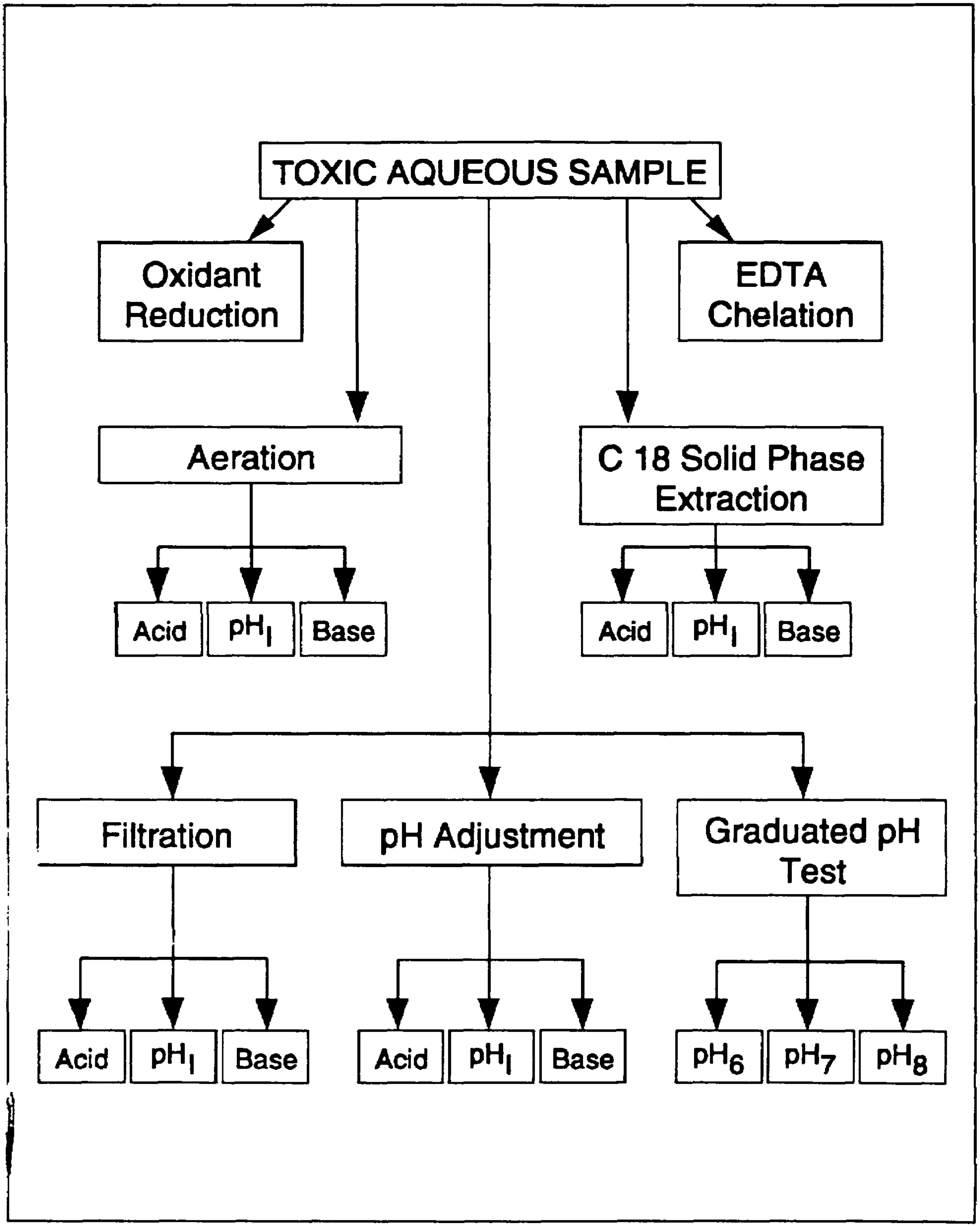


Figure 5-1. Overview of the Phase 1 Toxicity Characterization Process.
The ambient pH of the sample is indicated as pH$_I$.

ammonia, a relatively common toxicant whose toxicity is extremely pH-dependent (USEPA, 1985c). However, different pH values can strongly affect the toxicity of many common ionizable pesticides, and also may influence the bioavailability and toxicity of certain heavy metals and surfactants (Campbell and Stokes, 1985; Doe *et al.*, 1988).

Aeration tests are designed to determine whether toxicity is attributable to volatile, oxidizable, or sublatable compounds. Samples at pH_i (ambient pH), pH 3, and pH 11 are sparged with air for 1 h, readjusted to pH_i and tested for toxicity. The different pH values affect the ionization state of polar toxicants, thus making them more or less volatile, and also affect the redox potential of the system. If toxicity is reduced by air sparging at any of the pH values, the presence of volatile or oxidizable compounds may be suggested. To distinguish the former from the latter situation, further experiments are performed using nitrogen rather than air to sparge the samples. If toxicity remains the same, oxidizable materials are implicated; if toxicity is again reduced, volatile compounds are suspect. The pH at which toxicity is reduced is important. If nitrogen sparging decreases toxicity at pH_i, neutral volatiles are present; if toxicity decreases at pH 11.0 or pH 3.0, basic and acidic volatiles, respectively, are implicated. An additional process through which aeration can remove sample toxicants is sublation, which is the movement of compounds through aqueous solutions at the surface of the air bubbles, often followed by deposition on the aeration glassware. Compounds that exhibit this behavior include resin acids and surfactants; in some instances it may be possible to implicate the presence of sublatable compounds by rinsing the aeration glassware with clean laboratory dilution water and testing this fraction (Ankley *et al.*, 1990b).

Filtration provides information concerning the amount of toxicity associated with filterable components. In this test, samples at pH_i, pH 3.0, and pH 11.0 are passed through 1-μm glass fiber filters, readjusted to pH_i, and tested for toxicity. Reductions in toxicity due to filtration could be related to factors such as decreased physical toxicity, rather than chemical toxicity (Chapman *et al.*, 1987), or removal of particle-bound toxicants, which could be important, particularly if filter-feeding organisms such as cladocerans are the test species.

Reversed-phase, solid-phase extraction (SPE) is designed to determine the extent of toxicity due to compounds that are relatively nonpolar at pH_i, pH 3.0, or pH 9.0. This test, in conjunction with associated Phase II analytical procedures, is an extremely powerful TIE tool. In this procedure, filtered sample aliquots at pH_i, pH 3.0, and pH 9.0 are passed through small columns packed with an octadecyl (C_{18}) sorbent. At pH_i, the C_{18} sorbent will remove neutral compounds such as certain pesticides (Junk and Richard, 1988). By shifting ionization equilibria at the low and high pH values, the SPE column also can be used to extract organic acids and bases (Wells and Michael, 1987). During extraction, the resulting post-column effluent is collected and tested for toxicity to determine whether the manipulation removed toxicity and/or whether the capacity of the column was exceeded. Following this, the column is eluted with solvents, such as methanol, which then can be tested for recovery of toxicity. If sample toxicity is decreased and subsequently recovered in solvent elutions, a nonpolar toxicant would be suspected.

The presence of toxicity due to cationic metals is tested through additions of ethylenediaminetetraacetic acid (EDTA), a strong chelating agent that produces nontoxic complexes with many metals. As with SPE chromatography, the specificity of the EDTA test for a class of ubiquitous toxicants makes it a powerful TIE tool. Cations chelated by EDTA include certain forms of aluminum, barium, cadmium, cobalt, copper, iron, lead, manganese, nickel, strontium, and zinc (Stumm and Morgan, 1981). EDTA does not complex anionic forms of metals, and only weakly chelates certain cationic metals, such as silver, chromium, and thallium (Stumm and Morgan, 1981).

The oxidant reduction test is designed to determine the degree of toxicity associated with chemicals reduced, or in some instances chelated, by sodium thiosulfate. The toxicity of oxidants such as chlorine, bromine, iodine, and manganous ions is neutralized by sodium thiosulfate, and metals such as copper, cadmium, and silver are chelated and rendered biologically unavailable (Hockett and Mount, 1990). Because sodium thiosulfate, like EDTA, has low toxicity to most aquatic organisms, a relatively wide range of concentrations can be tested.

Phase II: Toxicant Identification—Initial laboratory work in Phase II focuses on isolation of the toxicants using chemical fractionation techniques with toxicity tracking. The ideal isolation process would create a subsample that contains one chemical, the toxicant. Depending on the nature of the toxicants, wide differences in the techniques, as well as in the amount of effort required for fractionation, will occur.

In general, after fractionation, instrumental analyses are performed on the toxic subsamples, and lists of identified chemicals are assembled for each subsample. For each chemical in a list, an LC_{50} value is obtained, usually from the literature or occasionally from structure activity models (Institute for Biological and Chemical Process Analyses, 1986). By comparing concentrations of the identified chemicals to their LC_{50} values, a list of suspect toxicants is made. This list is then refined by actually determining LC_{50} values for the suspects using the TIE test species. If only one toxicant is present, it should be easily identified. For samples with multiple toxicants, identification becomes significantly more protracted since interactions among toxicants may need to be examined. If none of the suspected toxicants appears to explain the toxicity, the true toxicants were probably not detected during instrumental analysis. Usually, additional separation and associated concentration steps are required to increase the analytical sensitivity for toxicant identification.

The information obtained in Phase I provides the analytical roadmarks for performing the fractionation and identification tasks in Phase II. To illustrate the relationship between Phase I data and the analytical approaches employed in Phase II, results for two typical Phase I TIE evaluations are presented in Table 5-1. The Phase II methods and approaches appropriate for these examples are discussed below.

In the first sample in Table 5-1, SPE reduced toxicity. In Phase II, the SPE column is eluted with graded, increasingly nonpolar methanol/-water solutions, and toxicity testing is performed on each fraction (Burkhard *et al.*, 1990). Although solvents other than methanol are routinely used in analytical work with C_{18} chromatography columns, the low toxicity of methanol to aquatic organisms (e.g., LC_{50} ≥1.5 percent for cladocerans) makes it a solvent of choice for toxicity tracking in the fractions. If no toxicity occurs in the fractions, the toxicants have been lost and further characterization (Phase 1) work is required. If toxicity occurs in the fractions, Phase II methods feature concentration of the toxic methanol/water fractions; high performance liquid chromatography fractionation of the concentrate (again with a C_{18}/methanol/water solvent system), with concurrent toxicity testing of the fractions; and, ultimately, identification of suspected toxicants in the toxic fractions via gas chromatography/mass spectroscopy. For pore water TIE, toxicity caused by high log k_{ow} nonpolar organics is often not elutable with methanol. In these cases, it is useful to elute the SPE column with a less polar solvent (i.e., methylene chloride) (Schubauer-Berigan and Ankley, 1991).

In the second sample, both EDTA additions and SPE reduced toxicity. The reduction of toxicity by EDTA strongly suggests the presence of toxic levels of cationic metals. Thus, Phase II procedures would include both metal analyses and the concentration, fractionation, and identification techniques described for nonpolar organics in the first example. If analyses identify specific metals at concentrations high enough to cause toxicity, various mass balance procedures can be used to define the portion of the sample toxicity due to the suspected metals and the portion of the toxicity due to the suspect nonpolar compounds.

Only a very small subset of possible Phase I results is shown in Table 5-1, particularly when one considers that three of the tests (aeration, filtration, SPE) are conducted at three different pH values. A complete discussion of the types of Phase I results that may be encountered and

Table 5-1. Phase I Characterization Results and Suspect Toxicant Classification for Two Samples.

Phase I Test	Sample	
	One	Two
Oxidant reduction	NR[a]	NR
EDTA addition	NR	R
Graduated pH test	NR	NR
pH adjustment		NR
Filtration	NR	NR
Aeration	NR	NR
SPE	R[b]	R
Methanol fractions	T[c]	T
Suspected toxicant classification	Nonpolar organics	Nonpolar organics/heavy metals

[a]NR = No reduction in toxicity.
[b]R = Reduction in toxicity.
[c]T = Toxicity recovered.

subsequent Phase II strategies that could be implemented is beyond the scope of this review.

Phase III: Toxicant Confirmation—After Phase II identification procedures implicate suspected toxicants, Phase III is initiated to confirm that the suspects are indeed the true toxicants. Confirmation is perhaps the most critical step of the TIE because procedures used in Phases I and II may create artifacts that could lead to erroneous conclusions about the toxicants. Furthermore, there is a possibility that substances causing toxicity are different from sample to sample within a supposedly homogeneous geographic region. Phase III enables both situations to be addressed. The tools used in Phase III include correlation, relative species sensitivity, observation of symptoms, spiking, and mass balance techniques. In most cases, no single Phase III test is adequate to confirm suspects as the true toxicants; it is necessary to use multiple confirmation procedures.

In the correlation approach, observed toxicity is regressed against expected toxicity due to measured concentrations of the suspected toxicants in samples collected over time or from several sites within a location. For the correlation approach to succeed, temporal or spatial variation has to be wide enough to provide a range of values adequate for meaningful analyses. To use the correlation approach effectively when there are multiple suspect toxicants, it is necessary to generate data concerning the additive, antagonistic, and synergistic effects of the toxicants in ratios similar to those found in the samples. These data also are needed for the spiking and mass balance techniques described below.

The relative sensitivity of different test species can be used to evaluate suspected toxicants. If two or more species exhibit markedly different sensitivities to a suspected toxicant in standard reference tests, and the same patterns in sensitivity are seen with the toxic pore water sample, this provides evidence for the validity of the suspect being the true toxicant.

Another Phase III procedure is observation of symptoms (e.g., time to mortality) in poisoned animals. Although this approach does not necessarily provide support for a given suspect, it can be used to provide evidence against a suspected toxicant. If the symptoms observed in a standard reference test with a suspected toxicant differ greatly from those observed with the pore water sample (which contains similar concentrations of the suspected toxicant), this is strong evidence for a misidentification.

Confirmatory evidence also can be obtained by spiking samples with the suspect toxicants. While the results may not be conclusive, an increase in toxicity by the same proportion as the increase in concentration of the suspect toxicant in the sample suggests that the suspect is correct. To obtain a proportional increase in toxicity from the addition of a suspect toxicant when in fact it is not the true toxicant, both the true and suspect toxicants would have to have very similar toxicity levels and their effects would also have to be additive.

Mass balance calculations can be used as confirmation steps when toxicity can be at least partially removed from the pore water sample, and subsequently recovered. This approach can be useful in instances when SPE removes toxicity. The methanol fractions eluted from the SPE column are evaluated individually for toxicity; these toxicities are summed and then compared to the total amount of toxicity lost from the sample.

Other techniques, including alteration of water quality characteristics (e.g., pH, salinity) in a manner designed to affect the toxicity of specific compounds, and analysis of body burdens of suspected toxicants in exposed animals, also can be useful confirmation steps.

5.2.1.2.3 Types of Data Required

In addition to the routine measurements described above, biological response data, either acute or chronic, will be obtained. Specific data collected will depend on the choice of test organism and endpoints. If the TIE process is initiated, the researcher will first obtain data concerning the physical/chemical characteristics of the toxicants in the pore water, followed by actual identification of toxic compounds, and standard determination of their concentrations in the toxic samples (see Section 5.2.1.2.2 above).

5.2.1.2.4 Necessary Hardware and Skills

Pore water preparation and toxicity test procedures are fairly straightforward and require commonly available equipment and facilities. Many of the TIE procedures also require only routine facilities. However, certain TIE techniques require some degree of advanced analytical capability (e.g., atomic absorption spectroscopy, gas chromatography/mass spectroscopy). Similarly, although many of the routine toxicity tests require relatively little training, certain of the TIE procedures, in particular some of the chemical analyses, require advanced technical expertise and experience.

5.2.1.3 Adequacy of Documentation

The theoretical basis for using pore water to assess toxicity appears to be scientifically sound, and pore water has been used for sediment toxicity evaluation (Adams *et al.*, 1985; Swartz *et al.*, 1985, 1988, 1990; Knezovich and Harrison, 1988; Connell *et al.*, 1988; Giesy *et al.*, 1988; USEPA, 1989a; Ankley *et al.*, 1990a, 1991a, 1991b; Hoke *et al.*, 1990; Schubauer-Berigan and Ankley, 1991). Toxicity tests that can be used are in many instances well-documented, standard procedures (U.S. EPA, 1985a; 1985b). The TIE techniques involved, including those specifically for sediments, have been documented (USEPA, 1988, 1989b, 1989c, 1991a, 1991b). Also, sediment TIEs with pore water have been successfully demonstrated (Ankley *et al.*, 1990a, 1991b; Schubauer-Berigan *et al.*, 1990; Schubauer-Berigan and Ankley, 1991).

5.2.2 Applicability of Method to Human Health, Aquatic Life, or Wildlife Protection

This method can be used to predict acute and chronic (i.e., growth or reproductive) effects of toxic sediment on aquatic organisms and can identify toxicants responsible for observed effects. The data generated thus can be used to design sediment remediation programs that would have an optimal likelihood of success. These procedures are not suitable, however, for evaluating human health effects or protecting wildlife, and they cannot be used to address bioconcentratable toxicants.

5.2.3 Ability of Method to Generate Numerical Criteria for Specific Chemicals

Pore water toxicity assessment, in conjunction with successful TIE procedures, can be used to generate numerical criteria for toxic compounds in sediment pore water because the toxicants are actually identified. However, it must be established that compounds identified as being toxic to test organisms in the laboratory are the same compounds (both in form and concentration) responsible for toxicity to organisms in field situations. This relationship can be evaluated both through biosurveys (possibly in conjunction with analysis of contaminant residues in organisms collected from the field), and laboratory toxicity tests in which benthic organisms perceived to be affected in contaminated sediments *in situ* are exposed to toxicants identified in the pore water. Both types of data also would be required for any sediment classification method based on toxicity or chemical analyses.

5.3 USEFULNESS

5.3.1 Environmental Applicability

5.3.1.1 Suitability for Different Sediment Types

The pore water toxicity assessment approach is suitable for any sediment from which adequate quantities of pore water can be isolated. In typical sediments, 20-50 percent of the volume of the bulk sediment sample is pore water. For a complete Phase I characterization with a test species of relatively small body size (e.g., cladocerans, larval fishes), approximately 1.5 L of pore water is required. This translates into a bulk sediment requirement of 3-8 L. Bulk sediment volumes needed for Phase II identification will, of course, be dependent on the toxicants present in the pore water, but typical volumes required would be expected to range from 1 to 20 L.

5.3.1.2 Suitability for Different Chemicals or Classes of Chemicals

This approach appears to be suitable for various nonpolar organics, cationic metals, and ammonia (Adams *et al.*, 1985; Swartz *et al.*, 1985, 1988, 1990; Knezovich and Harrison, 1988; Connell *et al.*, 1988; USEPA, 1989a; Ankley *et al.*, 1990a, 1991b; DiToro *et al.*, 1990). The applicability of the approach to toxicants such as polar organics or extremely lipophilic compounds has yet to be established. Also, the TIE procedures enable the evaluation of interactive (additive, synergistic, antagonistic) effects among various toxicants present in pore water samples.

5.3.1.3 Suitability for Predicting Effects on Different Organisms

If the TIE procedures successfully identify specific toxicants responsible for sediment toxicity, the impacts of these toxicants on various species of concern can be easily predicted, provided that there are data concerning the toxicity of the identified compounds to these species. Although toxicity data may not be available for certain benthic species, once suspect toxicants are identified, it would be possible to generate toxicity data for specific species of concern.

5.3.1.4 Suitability for In-Place Pollutant Control

The pore water toxicity assessment method and associated TIE procedures could prove to be a powerful tool for in-place pollutant control. Because sediment toxicants are actually identified, it is possible to design remediation plans for toxicants from point sources or controllable nonpoint sources, and to routinely monitor the success of these plans through continued assessment of pore water for toxicity and specific chemical toxicants.

5.3.1.5 Suitability for Source Control

Because the potential exists for identifying specific sediment toxicants, this method is ideal for point source control, as well as controllable nonpoint source inputs.

5.3.1.6 Suitability for Disposal Applications

As stated above, because specific sediment toxicants can be identified, it would be possible to identify potential hazards of contaminated sediments to aquatic organisms before disposal operations, such as those associated with dredging (Ankley *et al.*, 1991c).

5.3.2 General Advantages and Limitations

5.3.2.1 Ease of Use

Pore water preparation, routine chemical analyses, toxicity tests, and certain of the TIE procedures are reasonably straightforward and require relatively little technical expertise or extensive laboratory facilities. Because it is possible to work with aqueous samples, many of the standard toxicity tests developed for toxicity assessment of surface waters and effluents can be used, in addition to tests with various benthic species (e.g., USEPA, 1985a, 1985b). However, interpretation of results of certain of the TIE procedures, as well as analytical support for the TIE work, requires advanced training and experience. Also, several TIE analyses require highly sensitive analytical instrumentation for procedures, such as atomic absorption spectroscopy and gas chromatography/mass spectroscopy.

5.3.2.2 Relative Cost

Cost of the actual toxicity test procedures is relatively low. Cost of the TIE procedures will vary depending on the nature of the toxic compounds; certain toxicants (e.g., pesticides) are more costly to identify and quantify than others (e.g., ammonia). Also, identification and determination of the effects of multiple toxicants in samples costs more than the identification of single toxicants. Thus, cost analysis for the TIE portion of the toxicity assessment is case-specific.

5.3.2.3 Tendency to Be Conservative

Depending on the species used and the endpoint evaluated, pore water toxicity tests can be as conservative as desired. However, acute pore water toxicity tests described for sediment TIE are not meant to represent chronic or bioaccumulation endpoints.

5.3.2.4 Level of Acceptance

The theoretical basis of pore water toxicity assessment is sound (Adams *et al.*, 1985; Swartz *et al.* 1985, 1988, 1990; Knezovich and Harrison, 1988; Connell *et al.*, 1988; USEPA, 1989a; DiToro *et al.*, 1990; Ankley *et al.*, 1991a). The most important advantage of using pore water as a sediment test fraction, however, is the fact that it enables the application of recently developed TIE procedures for the identification of toxic compounds in aqueous samples containing complex mixtures of chemicals (USEPA, 1988, 1989b, 1989c, 1991a, 1991b). TIE procedures have proven to be extremely powerful tools for work with both complex effluents and sediment pore water (Ankley *et al.*, 1990a, 1991b; Kuehl *et al.*, 1991; Amato *et al.*, 1991; Norberg-King *et al.*, 1991; Schubauer-Berigan and Ankley, 1991; Ankley and Burkhard, 1992). The ability to identify specific compounds responsible for the toxicity of contaminated sediments clearly could be critical to the success of remediation.

5.3.2.5 Ability to Be Implemented by Laboratories with Typical Equipment and Handling Facilities

Pore water preparation, toxicity test procedures, and certain of the TIE methods are easily implemented by laboratories with typical equipment and a moderate degree of expertise. Interpretation of some TIE results requires additional technical training and experience, and certain of the analytical procedures associated with TIE work require both specialized training and analytical instrumentation.

5.3.2.6 Level of Effort Required to Generate Results

This procedure consists of field sampling, preparation of pore water, toxicity tests, and various TIE procedures. Depending on the results of the TIE work, the level of effort expended to obtain potentially important data can be relatively small.

5.3.2.7 Degree to Which Results Lend Themselves to Interpretation

Biological responses (i.e., toxicity) can be easily interpreted, and when properly performed, the results of the TIE procedures can be straightforward and easily interpreted; however, this is dependent on the complexity of the sample and the number of compounds contributing to sample toxicity.

5.3.2.8 Degree of Environmental Applicability

Pore water toxicity assessment and TIE procedures are applicable to virtually all environmental conditions and sediment types. Moreover, a wide variety of test organisms can be evaluated with this approach. However, although data indicate that the toxicity and/or bioaccumulation of a variety of contaminants are correlated with their pore water concentrations, there is no guarantee that this relationship exists for all types of contaminants. For example, a potentially important route of exposure for highly lipophilic compounds is thought to be via ingestion of contaminated particles. This route is not addressed using pore water exposures. Finally, existing TIE procedures are applicable for acutely toxic samples, and thus generally would not be useful for identifying chronically toxic sediment contaminants.

5.3.2.9 Degree of Accuracy and Precision

Because the procedures consist of laboratory-controlled experiments, results obtained are statistically accurate and precise.

5.4 STATUS

5.4.1 Extent of Use

Various toxicity tests have been widely applied to the evaluation of both freshwater and marine sediments, and pore water is merely one of the possible fractions that can be tested. Theoretically, pore water appears to be appropriate for sediment toxicity assessment and there have been many examples of its use for this purpose (Adams *et al.*, 1985; Swartz *et al.*, 1985, 1988, 1990; Giesy *et al.*, 1988; Knezovich and Harrison, 1988; Connell *et al.*, 1988; USEPA, 1989a; Ankley, 1990a, 1991a, 1991b; DiToro *et al.*, 1990; Hoke *et al.*, 1990; Schubauer-Berigan and Ankley, 1991). The TIE procedures (USEPA, 1988, 1989b, 1989c, 1991a, 1991b) although developed only relatively recently, already are widely used in both research and regulatory programs.

5.4.2 Extent to Which Approach Has Been Field-Validated

Because the procedure is relatively new, there has been little field validation. This area requires research, not only for the pore water TIE methods described herein, but for virtually any other sediment method involving toxicity tests or chemical analyses.

5.4.3 Reasons for Limited Use

Various sediment toxicity tests have been widely used; however, relatively few studies have evaluated pore water toxicity. This is primarily because the theoretical basis for using pore water has only recently been critically evaluated. For this reason, there are no standard methods for pore water preparation. Systematic TIE procedures for toxic aqueous samples have only recently been developed and thus have not yet been widely applied to the area of sediment toxicity assessment. Because current TIE procedures cannot be used with bulk sediment samples, pore water appears to be the best fraction with which to attempt to identify specific sediment contaminants responsible for acute toxicity.

5.4.4 Outlook for Future Use and Amount of Development Yet Needed

The outlook for this approach is extremely promising because it is the only method currently available that enables the identification of specific compounds responsible for sediment toxicity with some degree of certainty. This information could be critical to the success of remediation. However, as with all of the existing sediment methods, further development is needed, particularly in the following areas:

- The development of standard and scientifically sound techniques for pore water isolation;
- Further characterization of relationships between sediment toxicity *in situ* and the toxicity of sediment pore water in the laboratory for different classes of compounds; and
- The development of TIE procedures to identify chronically toxic compounds in aqueous samples.

Research in all these areas is ongoing at ERL-Duluth.

For more information please contact:

Gerald Ankley and Nelson Thomas
U.S. Environmental Protection Agency
Environmental Research Laboratory
6201 Congdon Boulevard
Duluth, MN 55804
(218) 720-5603

Mary K. Schubauer-Berigan
AScI Corporation
6201 Congdon Boulevard
Duluth, MN 55804
(218) 720-5619

5.5 REFERENCES

Adams, W.J., R.A. Kimerle, and R.G. Mosher. 1985. Aquatic safety assessment of chemicals sorbed to sediments. pp. 429-453. In: Aquatic Toxicology and Hazard Assessment: Seventh Symposium. R.D. Cardwell, R. Purdy, and R.C. Bahner (eds.). ASTM STP 854. American Society for Testing and Materials, Philadelphia, PA.

Amato, J.R., D.I. Mount, E.J. Durhan, M.T. Lukasewycz, G.T. Ankley, and E.D. Robert. 1991. An example of the identification of diazinon as a primary toxicant in an effluent. Environ. Toxicol. Chem. In press.

Ankley, G.T. and L.P. Burkhard. 1992. Identification of surfactants as toxicants in a primary effluent. Environ. Toxicol. Chem. Submitted.

Ankley, G.T., A. Katko, and J.W. Arthur. 1990a. Identification of ammonia as an important sediment-associated toxicant in the lower Fox River and Green Bay, Wisconsin. Environ. Toxicol. Chem. 9:313-322.

Ankley, G.T., M.T. Lukasewycz, G.S. Peterson, and D.A. Jenson. 1990b. Behavior of surfactants in toxicity identification evaluations. Chemosphere. 21:3-12.

Ankley, G.T., M.K. Schubauer-Berigan, and J.R. Dierkes. 1991a. Predicting the toxicity of bulk sediments to aquatic organisms with aqueous test fractions: Pore water versus elutriate. Environ. Toxicol. Chem. In press.

Ankley, G.T., G.L. Phipps, P.A. Kosian, D.J. Hansen, J.D. Mahony, A.M. Cotter, E.N. Leonard, J.R. Dierkes, D.A. Benoit, and V.R. Mattson. 1991b. Acid volatile sulfide as a factor mediating cadmium and nickel bioavailability in contaminated sediments. Environ. Toxicol. Chem. In press.

Ankley, G.T., M.K. Schubauer-Berigan, and R.A. Hoke. 1991c. Use of toxicity identification evaluation techniques to identify dredged material disposal options: A proposed approach. Environ. Management. In press.

Burkhard, L.P., and G.T. Ankley. 1989. NETAC's toxicity-based approach to identify toxicants. Environ. Sci. Technol. 23:1438-1443.

Burkhard, L.P., E.J. Durhan, and M.T. Lukasewycz. 1990. Identification of nonpolar toxicants in effluent using toxicity-based fractionation with gas chromatography/mass spectrometry. Anal. Chem. 63:277-283.

Campbell, P.G.C., and P.M. Stokes. 1985. Acidification and toxicity of metals to aquatic biota. Can. J. Fish. Aq. Sci. 42:2034-2049.

Chapman, P.M., J.D. Popham, J. Griffin, D. Leslie, and J. Michaelson. 1987. Differentiation of physical from chemical toxicity in solid waste fish bioassays. Water Air Soil Pollut. 33:295-308.

Connell, D.W., M. Bowman, and D.W. Hawker. 1988. Bioconcentration of chlorinated hydrocarbons from sediment by oligochaetes. Ecotoxicol. Environ. Safety 16:293-302.

DiToro, D.M., J.D. Mahony, D.J. Hansen, K.J. Scott, M.B. Hicks, S.M. Mays, and M.S. Redmond. 1990. Toxicity of cadmium in sediments: the role of acid volatile sulfide. Environ. Toxicol. Chem. 9:1489-1504.

Doe, K.G. W.R. Ernst, W.R. Parker, G.R.J. Julien, and P.A. Hennigar. 1988. Influence of pH on the acute lethality of fenitrothion, 2,4-D and aminocarb and some pH-altered sublethal effects of aminocarb on rainbow trout (*Salmo gairdneri*). Can. J. Fish. Aq. Sci. 45:287-293.

Giesy, J.P., R.L. Graney, J.L. Newsted, C.J. Rosiu, A. Benda, R.G. Kreis, and F.J. Horvath. 1988. Comparison of three sediment bioassay methods using Detroit River sediments. Environ. Toxicol. Chem. 7:483-498.

Hockett, J.R., and D.R. Mount. 1990. Use of metal chelating agents to differentiate among sources of toxicity. Eleventh Annual Meeting of the Society of Environmental Toxicology and Chemistry, Abstract, p. 162.

Hoke, R.A., J.P. Giesy, G.T. Ankely, J.L. Newsted, and J.R. Adams. 1990. Toxicity of sediments from western Lake Erie and Maumee River at Toledo, Ohio, 1987: Implication for current dredged material disposal practices. J. Great Lakes Res. 16:457-470.

Institute for Biological and Chemical Process Analyses. 1986. User manual for QSAR system. Montana State University, Bozeman, MT.

Junk, G.A., and J.J. Richard. 1988. Organics in water: Solid phase extraction on a small scale. Anal. Chem. 60:451-454.

Knezovich, J.P., and F.L. Harrison. 1988. The bioavailability of sediment sorbed chlorobenzenes to larvae of the midge *Chironomus decorus*. Ecotoxicol. Environ. Safety 15:226-241.

Knezovich, J.P., F.L. Henderson, and R.G. Wilhelm. 1987. The bioavailability of sediment--sorbed organic chemicals: A review. Water Air Soil Pollut. 32:233-245.

Kuehl, D.W., G.T. Ankley, L.P. Burkhard, and D.A. Jensen. 1990. Bioassay directed characterization of the acute toxicity of a creosote leachate. Hazardous Waste Hazardous Mater. 7:283-291.

Norberg-King, T.J., E.J. Durhan, G.T. Ankley, and E. Robert. 1991. Application of toxicity identification evaluation procedures to the ambient waters of the Colusa Basin Drain. Environ. Tox. and Chem. In press.

Schubauer-Berigan, M.K., J.R. Dierkes, and G.T. Ankley. 1990. Toxicity identification evaluations of contaminated sediments in the Buffalo River, NY and Saginaw River, MI. National Effluent Toxicity Assessment Center Rep. No. 20-90. Environmental Research Laboratory, Duluth, MN.

Schubauer-Berigan, M.K., and G.T. Ankley. 1991. The contribution of ammonia, metals, and nonpolar organic compounds to the toxicity of sediment interstitial water from an Illinois River tributary. Environ. Toxicol. Chem. In press.

Shults, D.W., L.M. Smith, S.P. Ferraro, F.A. Roberts, and C.K. Poindexter. 1991. A comparison of methods for measuring trace organic compounds and metals in interstitial water. Water Res. In press.

Sly, P.G. 1988. Interstitial water quality of lake trout spawning habitat. J. Great Lakes Res. 14:301-315.

Stumm, W., and J.J. Morgan. 1981. Aquatic chemistry - An introduction emphasizing chemical equilibria in natural waters. John Wiley and Sons, New York. 583 pp.

Swartz, R.C., G.R. Ditsworth, D.W. Schults, and J.O. Lamberson. 1985. Sediment toxicity to a marine infaunal amphipod: Cadmium and its interaction with sewage sludge. Mar. Environ. Res. 18:133-153.

Swartz, R.C., P.F. Kemp, D.W. Schults, and J.O. Lamberson. 1988. Effects of mixtures of sediment contaminants on the marine infaunal amphipod *Rhepoxynius abronius*. Environ. Toxicol. Chem. 7:1013-1020.

Swartz, R.C., P.F. Kemp, D.W. Schults, G.R. Ditsworth, and R.J. Ozretich. 1989. Acute toxicity of sediment from Eagle Harbor, Washington, to the infaunal amphipod *Rhepoxynius abronius*. Environ. Toxicol. Chem. 8:215-222.

Swartz, R.C., D.W. Schults, T.H. DeWitt, G.R. Ditsworth, and J.O. Lamberson. 1990. Toxi-city of fluoranthene in sediment to marine amphipods: A test of the equilibrium partitioning approach to sediment quality criteria. Environ. Toxicol. Chem. 9:1071-1080.

USEPA. 1985a. Methods for measuring the acute toxicity of effluents to freshwater and marine organisms. EPA/600/485-013. U.S. Environmental Protection Agency, Cincinnati, OH.

USEPA. 1985b. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms. EPA/600/4-85-014. U.S. Environmental Protection Agency, Cincinnati, OH.

USEPA. 1985c. Ambient water quality criteria for ammonia - 1984. EPA/440/5-85-001. U.S. Environmental Protection Agency, Duluth, MN.

USEPA. 1988. Methods for aquatic toxicity identification evaluations: Phase I toxicity characterization procedures. EPA/600-3-88/034. U.S. Environmental Protection Agency, Duluth, MN.

USEPA. 1989a. Equilibrium partitioning approach to generating sediment quality criteria. EPA/440/5-89/002. U.S. Environmental Protection Agency, Washington, DC.

USEPA. 1989b. Methods for aquatic toxicity identification evaluations: Phase II toxicity identification procedures. EPA/600-3-88/035. U.S. Environmental Protection Agency, Duluth, MN.

USEPA. 1989c. Methods for aquatic toxicity identification evaluations: Phase III toxicity confirmation procedures. EPA/600-3-88/036. U.S. Environmental Protection Agency, Duluth, MN.

USEPA. 1991a. Methods for aquatic toxicity identification evaluations: Phase I toxicity characterization procedures. Second edition. EPA-600/6-91/003. Environmental Research Laboratory, Duluth, MN.

USEPA. 1991b. Methods for sediment toxicity identification evaluations. National Effluent Toxicity Assessment Center Rep. No. 08-91. Environmental Research Laboratory, Duluth, MN.

Wells, M.J.M., and J.L. Michael. 1987. Reversed-phase solid-phase extraction for aqueous environmental sample preparation in herbicide residue analysis. J. Chromatogr. Sci. 25:345-50.

Equilibrium Partitioning Approach

Christopher S. Zarba
U.S. Environmental Protection Agency
401 M Street, SW (WH-586), Washington, DC 20460
(202) 260-1326

The equilibrium partitioning (EqP) approach focuses on predicting the chemical interaction among sediments, interstitial water (i.e., the water between sediment particles), and contaminants. Based on correlations with toxicity, interstitial water concentrations of contaminants appear to be better predictors of biological effects than do bulk sediment concentrations. The EqP method for generating sediment quality criteria is based on predicted contaminant concentrations in interstitial water. Chemically contaminated sediments are expected to cause adverse biological effects if the predicted interstitial water concentration for a given contaminant exceeds the chronic water quality criterion for that contaminant.

6.1 SPECIFIC APPLICATIONS

Specific applications of EqP-based sediment quality criteria are under development. The primary use of EqP-based sediment criteria will be to identify and prevent risks associated with contaminants. Because the regulatory needs vary widely among and within U.S. EPA offices and programs, EqP-based sediment quality criteria will be used in a variety of ways.

EqP-based numerical sediment quality criteria would likely be used directly to assess risk and would be applied in a tiered approach. In tiered applications, concentrations of sediment contaminants that exceed sediment quality criteria would be considered as causing unacceptable impacts. Further testing may or may not be required, depending on site-specific and program-specific conditions. Sediment contaminants at concentrations less than the sediment criteria would not be of concern. However, sediments would not be considered safe in cases where they are suspected to contain other contaminants at concentrations above safe levels, but for which no sediment criteria exist.

Synergistic, antagonistic, or additive effects of multiple contaminants in the sediments may also be of concern. Additional testing in other tiers of the evaluation approach, such as bioassays, could be required to determine whether the sediment is safe. It is likely that such testing would incorporate site-specific considerations.

6.1.1 Current Use

Specific regulatory uses of EqP-based sediment quality criteria are under development and will be articulated in the Contaminated Sediment Management Strategy. The Science Advisory Board (SAB) has completed the review of this approach for nonionic organic contaminants. Based on the findings of this review, the method will be used for developing national sediment quality criteria. (The first five sediment quality criteria will be proposed in the *Federal Register* shortly for public comment.) At the present time, the criteria are for the protection of benthic organisms. The methodology for developing sediment criteria for metal contaminants will be presented to the SAB for review in 1993. The range of potential applications of the EqP approach is large because the approach accounts for contaminant bioavailability and can be used to evaluate most sediments.

Draft sediment criteria values have been developed for a variety of organic compounds using the EqP approach. In pilot studies at a variety of contaminated sediment sites at which site characterization and evaluation activities were undertaken, the draft criteria were used in the following ways:

- Identify extent of contamination;
- Assess the risks or potential risks associated with the sediment contamination;
- Identify responsible parties and the need for source controls; and
- Identify the environmental benefit associated with a variety of remedial options.

In addition, a number of states have used draft EqP-based sediment criteria to evaluate the potential effects of sediment contaminants found in aquatic habitats.

6.1.2 Potential Use

Potential applications of the EqP approach include a variety of ongoing activities conducted by the U.S. EPA. EqP-based sediment quality criteria could play a major role in the identification, monitoring, and cleanup of contaminated sediment sites on a national basis. This is true, in part, because EqP-based SQC establish a direct cause-and-effect relationship between a contaminant concentration and biological impacts. They could also be used to ensure that uncontaminated sites remain uncontaminated. In some cases, such sediment criteria alone will be sufficient to identify and establish cleanup levels for contaminated sediments. In other cases, it will be necessary to supplement the sediment criteria with biological sampling, testing, or other types of analysis before a decision can be made.

EqP-based sediment criteria will be particularly valuable at sites where sediment contaminant concentrations are gradually increasing. In such cases, criteria will permit an assessment of the extent to which unacceptable contaminant concentrations are being approached or have been exceeded. Comparisons of field measurements to sediment criteria will be a reliable method for providing an early warning of a potential problem. Such an early warning would provide an opportunity to take corrective action before adverse impacts occur.

Although sediment criteria developed using the EqP approach are similar in many ways to existing water quality criteria, their applications may differ substantially. In most cases, contaminants in the water column need only be controlled at the source to eliminate unacceptable adverse impacts. In contrast, contaminated sediments often have been in place for quite some time, and controlling the source of that pollution (if the source still exists) will not be sufficient to alleviate the problem. Safe removal, treatment, or disposal of contaminated sediments can also be difficult and expensive. For this reason, it is anticipated that EqP-based sediment criteria will rarely be used as mandatory cleanup levels. Rather, they will likely be used to predict or identify the degree and spatial extent of problems associated with contaminated areas, and thereby facilitate regulatory decisions.

6.2 DESCRIPTION

6.2.1 Description of Method

Concentrations of contaminants in the interstitial water correlate very closely with toxicity, whereas concentrations of contaminants bound to the sediment particles do not. The EqP method for generating sediment criteria involves predicting contaminant concentrations in the interstitial water and comparing those concentrations to quality criteria. If the predicted sediment interstitial water concentration for a given contaminant exceeds its respective chronic water quality criterion, then the sediment would be expected to cause adverse effects.

The processes that govern the partitioning of chemical contaminants among sediments, interstitial water, and biota are better understood for some kinds of chemicals than for others. Concentrations of sulfides and organic carbon have been identified as primary factors that control phase associations, and therefore bioavailability, of trace metals in sediments. However, models that can use these factors to predict research are not fully developed. Mechanisms that control the partitioning of polar organic compounds are also poorly understood. Polar organic contaminants, however, are not generally considered to be a significant problem in sediments. Partitioning of nonionic organic compounds between sediments and interstitial water is highly correlated with the organic carbon content of sediments. Also, the toxicity of nonionic organic contaminants in sediments is highly dependent on their interstitial water concentrations. Consequently, to date, the EqP approach is well developed for nonionic organic contaminants and is in the process of development for trace metals.

Interstitial water concentrations can be calculated using partition coefficients for specific nonionic organic chemicals and criteria continuous concentrations from WQC documents. The sediment quality criterion for a specific chemical is defined as the solid phase concentration that will result in an uncomplexed interstitial water concentration equal to the chronic water quality criterion for that chemical. The rationale for using water quality criteria as the effect concentrations for benthic organisms is that the sensitivity range for benthic organisms appears to be similar to the sensitivity range for water column organisms. Moreover, partition coefficients for a wide variety of contaminants are available.

The calculation procedure for nonionic organic contaminants is as follows:

$$rSQC = K_p \times cWQC$$

where:

cWQC = Criterion continuous concentration

rSQC = Sediment quality criterion (μg/kg sediment)

K_p = Partition coefficient for the chemical (L/kg sediment) between sediment and water.

Although the method for developing sediment criteria for nonionic organic contaminants has been identified, continuous refinement of the methodology is expected.

6.2.1.1 Objectives and Assumptions

Three principal assumptions underlie use of the EqP-based approach to establish sediment quality criteria:

- For sediment-dwelling organisms, the uncomplexed interstitial water concentration of a chemical correlates with observed biological effects across sediment types, and the concentration at which effects are observed is the same as that observed in a water-only exposure.

- Partitioning models permit calculation of uncomplexed interstitial water concentrations of the chemical phases of sediments controlling availability.

- Benthic organisms exhibit a range of sensitivities to chemicals that is similar to the range of sensitivities exhibited by water column organisms.

Data exist supporting each of these assumptions.

6.2.1.2 Level of Effort

6.2.1.2.1 Type of Sampling Required

Sufficient sediment chemistry sampling is required to adequately characterize the area of concern. Total organic carbon concentrations are also needed, preferably for each sampling station.

6.2.1.2.3 Types of Data Required

Analyses are needed to determine the concentrations of the contaminants of concern in the sediment (bulk sediment analysis) and the concentrations of organic carbon in the sediment.

6.2.1.2.4 Necessary Hardware and Skills

The investigator must be able to design an appropriate sampling study, conduct bulk sediment analyses, operate a pocket calculator, and understand developed values and what they protect.

6.2.1.3 Adequacy of Documentation

The method is very well documented (see Section 6.5).

6.2.2 Applicability of Method to Human Health, Aquatic Life, or Wildlife Protection

At the present time SQC do not address bioaccumulative impacts to aquatic life, wildlife, and human health. Efforts are under way to derive criteria protective of these endpoints.

6.2.3 Ability of Method to Generate Numerical Criteria for Specific Chemicals

The EqP method generates numerical criteria for a number of nonionic organic chemicals. A methodology for developing sediment criteria for metal contaminants is being developed. Draft criteria to be proposed in the *Federal Register* were developed for endrin, phenanthrene, fluoranthene, dieldrin, and acenaphthene. It is expected that three to five additional sediment criteria will be issued each subsequent year.

Methods for developing sediment criteria for metal contaminants are under development and are expected to be reviewed by the SAB in 1993.

6.3 USEFULNESS

6.3.1 Environmental Applicability

One of the principal reasons for selecting the EqP approach is that it is applicable in a wide variety of aquatic systems, which is a prerequisite for the development of national sediment quality criteria.

6.3.1.1 Suitability for Different Sediment Types

Although aspects of the EqP method are still under development, it is expected that sediment criteria for nonionic contaminants developed using this approach will be applicable to all types of sediments found in both freshwater and marine environments with organic carbon concentrations ≥0.2 percent organic carbon. Additional work is needed to clarify the best use of the EqP approach for sediments with less than 0.2 percent organic carbon.

6.3.1.2 Suitability for Different Chemicals or Classes of Chemicals

The EqP method for developing sediment criteria has been modified for different types of contaminants. Nonionic, ionic, and metal contaminants all interact with sediment particles in different ways, and partitioning models have to be modified to account for these differences. The technical approach for developing sediment criteria for nonionic organic contaminants has been well developed and is under peer review. The technical approach for developing sediment criteria for metal contaminants is under development and is expected to undergo peer review in 1993. Ionic contaminants are not believed to cause major problems in sediments, but work plans for sediment criteria development methods for these compounds have been written.

6.3.1.3 Suitability for Predicting Effects on Different Organisms

As indicated above (see Section 6.2.1), the EqP approach is based on predicted interstitial water concentrations of nonionic organic contaminants, and comparisons of these concentrations with chronic water quality criteria. Typically, water quality criteria are based on toxicity information (e.g., median lethal or median effective concentrations) for a wide number of species and are set low enough to be protective of at least 95 percent of the species tested. Consequently, exposure levels that are predicted using the EqP approach can be compared with a range of toxic effects values that are representative of the different kinds of organisms on which water quality criteria are based.

6.3.1.4 Suitability for In-Place Pollutant Control

The EqP method is suitable for in-place pollution control because it can be used to identify locations where concentrations of individual contaminants are causing adverse effects. Target cleanup levels can be identified, and the success of cleanup activities can be determined.

6.3.1.5 Suitability for Source Control

The EqP method is suitable for source control. This method predicts the concentration of a contaminant above which adverse impacts are likely. A direct measure of biological effects is not needed to identify safe levels.

6.3.1.6 Suitability for Disposal Applications

The EqP method is suitable for predicting the effects that contaminated sediments may have if moved to an aquatic site. It is not applicable to contaminated sediments that are disposed of at upland sites.

6.3.2 General Advantages and Limitations

The EqP approach offers the following advantages:

- It is consistent with existing water quality criteria;
- It establishes a cause-and-effect relationship;
- It relates risks to specific substances, and it can be used to identify probable species at risk;
- It is applicable across all types of sediments and in all types of aquatic environments, including lentic, lotic, marine, and estuarine environments;
- Only site-specific chemistry data are needed;
- Site-specific or station-specific sediment criteria can be calculated as soon as sediment chemistry data are available;
- It incorporates the large quantities of data that were used in the development of water quality criteria;
- It can be incorporated into existing regulatory mechanisms with little or no need for additional staffing or skills;
- The equilibrium partitioning theory on which it is based is well developed;
- It can be modified easily to accommodate site-specific circumstances;
- It can be used with additional development to identify risks to humans and wildlife that may occur as a result of bioaccumulation; and
- It identifies the degree of sediment contamination and permits an assessment of whether contaminant concentrations are approaching an effects level.

The EqP approach is limited in the following ways:

- Sediment criteria developed using this approach do not address possible synergistic, antagonistic, or additive effects of contaminants;
- Interim and draft sediment criteria presently exist for only 12 contaminants at this time;
- The technical approach for developing sediment criteria for metal contaminants is still under development;
- Sediment quality criteria for nonionic chemicals apply to sediments that have an organic carbon concentration ≥0.2 percent; and

- Sufficient water-only toxicity data do not exist for all contaminants of concern.

6.3.2.1 Ease of Use

The calculation of site-specific sediment criteria is relatively easy, provided that sediment chemistry data adequately characterizing the site, a partition coefficient, and water quality criteria protective of the desired organism are available.

6.3.2.2 Relative Cost

Because site-specific biological data are not needed, the costs associated with this method depend primarily on the cost of collecting site-specific chemistry data.

6.3.2.3 Tendency to Be Conservative

Sediment criteria are derived using the chronic water quality criteria as effect levels. Hence, the levels of protection afforded by sediment criteria are similar to those of water quality criteria. In general, water quality criteria are deemed to be protective of 95 percent of the organisms most of the time. Each SQC is bracketed with levels of uncertainty.

6.3.2.4 Level of Acceptance

The EqP approach and its use in deriving sediment quality criteria are the result of the efforts of many scientists who represent a variety of federal agencies, industries, environmental organizations, universities, U.S. EPA laboratories, state agencies, and other institutions. These scientists were involved in the selection of the EqP approach for generating sediment criteria and have also played a role in development of the method. Papers that discuss various aspects of this effort have been presented at scientific conferences.

6.3.2.5 Ability to Be Implemented by Laboratories with Typical Equipment and Handling Facilities

No special laboratory facilities or requirements are needed. Sediment chemistry analysis is all that is required.

6.3.2.6 Level of Effort Required to Generate Results

The necessary level of effort varies substantially from site to site and is dependent on many factors. Compared with other methods, the EqP method generates results quickly and more cost-effectively. No site-specific biological data are required.

6.3.2.7 Degree to Which Results Lend Themselves to Interpretation

All sediment evaluation procedures require some level of interpretation. However, a sediment criterion that is bracketed with an appropriate degree of uncertainty can provide pertinent information. For example, sediment chemistry data that identify concentrations below the conservative effect level for a particular contaminant could be deemed safe for that contaminant. A contaminant concentration above the upper uncertainty level could be identified immediately as contaminated, and some degree of contamination could be assigned to those sediments for the individual contaminant. Sediments whose concentration of a particular contaminant falls within the degrees of uncertainty could require more detailed interpretation and possibly additional testing.

6.3.2.8 Degree of Environmental Applicability

EqP-based sediment quality criteria can be applied directly to any contaminated sediment containing ≥0.2 percent organic carbon and nonionic chemicals for which criteria are available. Extensive data analysis and site-specific biological data are not required to use sediment criteria developed using this method. (In some cases these attributes may nonetheless be desirable.) As a result, the EqP method can be considered environmentally applicable in some cases. Because a wide variety of contaminated sediment sites exist, absolute statements regarding environmental applicability are difficult to make. However, the EqP method would be appropriate in many situations to predict bioavailability, bioaccumulation, and biological effects.

6.3.2.9 Degree of Accuracy and Precision

Each sediment criterion value developed using the EqP method will have an associated degree of uncertainty, which will vary from criterion to criterion. The principal uncertainties associated with sediment criteria developed using the EqP method are those associated with partition coefficients. Hence, each developed sediment criterion should be and is bracketed with uncertainty, thereby providing decision-makers with a greater understanding of the meaning of the developed values.

6.4 STATUS

The method for developing sediment criteria for nonionic organic contaminants has been developed and has been reviewed by the SAB on two separate occasions. Guidelines and guidance on the regulatory use of sediment criteria are under development. The method for developing sediment criteria for metal contaminants is being investigated and results are promising. The metals method is expected to be sufficiently well developed for peer review by 1993.

6.4.1 Extent of Use

Specific regulatory uses for EqP-based sediment quality criteria are being developed. A formal framework for the application of sediment criteria is not expected until EPA completes its effort to develop a contaminated sediment management strategy. The range of potential applications is very large because the need for evaluating potentially contaminated sediments arises in many contexts.

Interim sediment criteria values were developed for a variety of organic compounds. These values were used in a pilot study at a number of sites where site characterization and evaluation activities were conducted. The interim criteria were used in three ways:

- To identify the extent of contamination and responsible parties;
- To assess the risks associated with sediment contamination; and
- To identify the environmental benefits associated with a variety of remedial options.

A number of States have used interim and draft sediment criteria to evaluate the potential effects of several contaminants found in sediments in state waters. The methodologies for deriving sediment criteria have been used in a variety of situations including the evaluation of dredged material, Superfund site assessments, and the identification of appropriate cleanup levels for contaminated sediment sites.

6.4.2 Extent to Which Approach Has Been Field-Validated

Considerable effort has been made by EPA to use field sites as part of the criteria validation effort and to aid in designing regulatory programs. Table 6-1 lists ongoing and completed studies where SQC are being used to directly support sediment activities. In addition to these sites, there are other sites and situations (completed, ongoing, and planned) where the EqP is being applied to field situations. Although these efforts are not involved with criteria development efforts, they do provide valuable data on the appropriateness of the EqP.

It needs to be understood, however, that "field validation" does not describe a specific experimental protocol. The idea is to find a site that is contaminated with a single chemical and determine whether the benthic populations are degraded when the SQC is exceeded. However, there are practical difficulties. Such a field site contaminated with only one chemical must be found, and there can be no ongoing sources of the chemical since the exposure should be only from the sediment. A gradient of chemical concentration that spans the SQC concentration is necessary. The

Table 6-1. Ongoing and Completed Studies Using SQC.

Location	Chemical	Status
Huntsville, AL	DDT/DDD/DDE	Ongoing
Keweenaw Lk	Cu	Submitted for publication
Steilacoom Lk	Cu	Submitted for publication
Fox River	PCB bioaccumulation	Submitted for publication
Fox River	Metal bioaccumulation	In preparation
Foundry Cove	Cd, Ni	Published
Calumet River	Sediment partitioning	In preparation
Nationwide	Comparison of toxicity test and benthic community disruption to SQC	Ongoing
New Bedford Harbor	Bioaccumulation	Published
Narragansett Bay	Bioaccumulation	Published
Colonization Expt.	6 chemicals	Published
Colonization Expt	3 chemicals to test SQC	Ongoing
San Diego Bay	PAHs	Ongoing
Lauritzen Canal	DDT	Ongoing
Nationwide	Comparison of SQC chemicals to STORET	In the documents
Nationwide	Comparison of SQC chemicals to NOAA National Status and Trends data	In the documents

sediment type must be essentially uniform in the gradient so that only chemical concentration is changing. The benthic population must be plentiful enough so that population degradation can be observed as the SQC is exceeded. In spite of the difficulties, major field efforts are presently under way.

An intermediate level of field validation is provided by the benthic colonization experiments. The experimental design is described above. The populations that develop are determined entirely by natural recruitment. The uniformity of sediment type is guaranteed by the experimental design. The experiments last from 2 to 4 months so that the sediment can properly be called a "natural" sediment. Three benthic colonization experiments have been performed using spiked sediments. The data analysis, which is partially complete, indicates that the experiments are consistent with the SQC for the chemicals being tested.

A third type of field validation is proceeding as well. It is based on the notion that although it is not possible to prove the validity of SQC (continual accumulation of evidence in favor of its validity does not guarantee that all evidence will always be supportive), it is possible to prove that it is invalid. If sediments are collected and the state of the benthic population is evaluated relative to control sites from the same region, there are

Table 6-2. SQC Field Validation Truth Table.

	SQC Not Exceeded	SQC Exceeded
Benthic Impact	Other Chemicals	Consistent
No Benthic Impact	Consistent	Invalidates

four possibilities, which are arranged as a truth table in Table 6-2.

The correlation of the presence or lack of benthic impact with exceeding or not exceeding the SQC is consistent but not proof of causality. The observation of benthic impact where the SQC is not exceeded can be attributed to the impact of other chemicals. However, if the SQC is exceeded, with a proper accounting for the uncertainty of SQC, and no benthic impact is observed, then the SQC is invalidated. The collection of these data is an ongoing part of the SQC development effort. Analysis to date suggests that these data do not invalidate the SQC.

6.4.3 Reasons for Limited Use

The EqP method is not commonly used for the following reasons:

- The method was developed only recently, and sufficient time has not elapsed for the principles to be understood and used by others.
- Final criteria have not been issued.
- Guidance and technical support documents are in draft form and will be issued along with final criteria.

6.4.4 Outlook for Future Use and Amount of Development Needed

This method is the only procedure for derivation of sediment quality criteria that is generic across sediments, accounts for bioavailability of chemicals, and relates effects to specific chemicals. Therefore, EqP-based sediment quality criteria will be used much as water quality criteria are being used to define environmentally acceptable concentrations. Sediment quality criteria, along with sediment toxicity tests analogous to water quality criteria and whole-effluent toxicity tests, will play a major role in EPA's management of contaminated sediment.

6.5 REFERENCES

USEPA. April 1989. Briefing report to the EPA Science Advisory Board on the equilibrium partitioning approach to generating sediment quality criteria. Office of Water, Regulations and Standards, Criteria and Standards.

USEPA. February 1990. Report of the Sediment Criteria Subcommittee of the Ecological Processes and Effects Committee - Evaluation of the equilibrium partitioning approach for assessing sediment quality. A Science Advisory Board Report.

USEPA. August 1991. Analytical method for determination of acid volatile sulfide in sediment (final draft). Office of Science and Technology, Health and Ecological Criteria Division.

USEPA. August 1991. Technical basis for establishing sediment quality criteria for nonionic chemicals using equilibrium partitioning. Office of Science and Technology, Health and Ecological Criteria Division.

USEPA. November 1991. Proposed sediment quality criteria for the protection of benthic organisms: Acenapththene (draft). Office of Science and Technology, Health and Ecological Criteria Division.

USEPA. November 1991. Sediment quality criteria for the protection of benthic organisms: Dieldrin (draft). Office of Science and Technology, Health and Ecological Criteria Division.

USEPA. November 1991. Sediment quality criteria for the protection of benthic organisms: Endrin (draft). Office of Science and Technology, Health and Ecological Criteria Division.

USEPA. November 1991. Sediment quality criteria for the protection of benthic organisms: Fluoranthene (draft). Office of Science and Technology, Health and Ecological Criteria Division.

USEPA. November 1991. Sediment quality criteria for the protection of benthic organisms: Phenanthrene (draft). Office of Science and Technology, Health and Ecological Criteria Division.

Tissue Residue Approach

Philip M. Cook
U.S. Environmental Protection Agency, Environmental Research Lab-Duluth
6201 Congdon Boulevard., Duluth, MN 55804
(218) 720-5553, FTS 780-5553

Anthony R. Carlson
U.S. Environmental Protection Agency, Environmental Research Lab-Duluth
6201 Congdon Boulevard., Duluth, MN 55804
(218) 720-5523, FTS 780-5523

Henry Lee II
U.S. Environmental Protection Agency, Environmental Research Lab-Newport
Marine Science Drive, Newport, OR 97365
(503) 867-4042

In the tissue residue approach, sediment chemical concentrations that will result in acceptable residues in exposed biotic tissues are determined. Concentrations of unacceptable tissue residues may be derived from toxicity tests performed during generation of chronic water quality criteria, from bioconcentration factors derived from the literature or generated by experimentation, or by comparison with human health risk criteria associated with consumption of contaminated aquatic organisms. The tissue residue approach generates numerical criteria and is most applicable for nonpolar organic and organometallic compounds.

7.1 SPECIFIC APPLICATIONS

7.1.1 Current Use

Tissue residues of chemical contaminants in aquatic organisms, particularly fish, are frequently used as measures of water quality in both freshwater and marine systems. The tendency of organisms to bioaccumulate chemicals from water and food is one of the factors used in establishing national water quality criteria (WQC) for the protection of aquatic life (Stephan *et al.*, 1985). Nonpolar organic chemicals, which may bioaccumulate to levels toxic to organisms or render organisms unfit for human food, generally will also be found as sediment contaminants. Hydrophobic organic chemicals preferentially distribute into organic carbon in sediment and lipid in aquatic biota. The tissue residue approach has been used recently to establish the amount of reduction of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) concentration in Lake Ontario sediments necessary to attain acceptable TCDD levels in fish (Cook *et al.*, 1990). The acceptable sediment TCDD concentration is being used as a sediment criterion to determine the remedial action necessary to reduce the incremental loading of TCDD from the Hyde Park Superfund site to Lake Ontario (Carey *et al.*, 1989). Tissue residues of benthic organisms have also been used in some regulatory actions, such as the assessment of bioaccumulation potential of dredged materials (USACE, 1991).

7.1.2 Potential Use

Although tissue residues have been used more commonly to determine the potential for bioaccumulation of chemical contaminants from sediments and dredged materials, they also provide an excellent measure of "effective exposure dose": a measure of an organism's actual exposure over time to a pollutant of concern. This exposure measure may be related to the dose expected at the water quality criterion or related directly to the potential for producing chronic toxic effects. Given the ability to measure or predict tissue residues resulting from exposures in contaminated sediment systems, it is possible to establish sediment criteria based on residue-toxicity effects relationships. These relationships can provide a basis for sediment criteria that are free of uncertainties normally associated with organism exposures and sediment contaminant bioavailability. This is especially true when *in situ* measurements provide the basis for the sediment residue link to the residue-toxic effect relationship.

One example of tissue residue-toxic effects linkage is the relationship between the failure of Great Lakes lake trout (*Salvelinus namaycush*) to reproduce and bioaccumulation of TCDD and non-ortho substituted PCBs (Mac, 1988). Laboratory studies have shown significant mortality of larvae when lake trout ova contain as little as 50 ppt 2,3,7,8-TCDD (Cook *et al.*, 1990; Walker *et al.*, 1991). This residue level is found in Lake Ontario lake trout that have not successfully reproduced naturally for many years. On the basis of TCDD toxic equivalents for organochlorine components having the same mode of toxic action, residues in lake trout from Lakes Ontario and Michigan may provide a measure of the reduction in sediment contamination necessary to reduce fish tissue concentrations to a threshold presumed to allow reproduction. The same approach can be used for benthic organisms, which may have greater intersite variations in residue levels than do fish because of benthic organisms' closer association with sediments.

7.2 DESCRIPTION

7.2.1 Description of Method

The tissue residue approach involves the establishment of safe sediment concentrations for individual chemicals or classes of chemicals by determining the sediment chemical concentration that will result in acceptable tissue residues. This process involves two steps: (1) linking toxic effects to residues (dose-response relationships) and (2) linking chemical residues in specific organisms to sediment chemical contamination concentrations (exposure relationships). Methods to derive unacceptable tissue residues include at least three approaches:

- The water quality criterion-residue approach;
- The experimental approach; and
- The human health approach.

Each of these approaches is described briefly below.

Water Quality Criterion-Residue Approach—A rapid approach for determining acceptable concentrations of tissue residues involves establishing maximum permissible tissue concentrations (MPTCs) expected for organisms at the chronic water quality criterion concentration previously established for a specific pollutant. MPTCs, when not available through residue measurements obtained with toxicity tests used for water quality criteria, can be obtained by multiplying the water quality criterion by an appropriate bioconcentration factor (BCF) obtained from the literature. When a reliable empirical BCF is not available, the BCF may be predicted from an octanol-water partition coefficient or a bioconcentration kinetic model. Thus, the absence of a water quality criterion for a chemical does not eliminate this approach as long as appropriate chronic toxicity test data are available for the species of interest.

Experimental Approach—Tissue residue-toxic effects linkages can be established through a series of chronic dose-response experiments or field correlations. Although this approach has the advantage of directly determining the tissue residue-toxic effects linkages, it can be relatively time consuming and costly to implement for a large number of pollutants. The experimental approach should be used to test the assumptions of the water quality criterion-residue approach and to supplement the existing tissue residue-toxic effects database. The experimental work can be closely coupled with the experiments conducted under the bulk sediment toxicity test approach to deriving sediment quality criteria (see Chapter 3, Bulk Sediment Toxicity Test Approach).

Human Health Approach—Human health risk from consumption of freshwater fish or seafood may be used as the criterion for tissue residue acceptability. A sediment quality criterion for a specific compound can be derived by establishing an acceptable human risk level (e.g., an excess human cancer risk of $1x10^{-5}$) and then back-calculating to the sediment concentration that would result in tissue residues associated with this level of risk. The human health approach can generate sediment quality criteria lower for carcinogenic compounds (e.g., PCBs, dioxins, benzo(a)pyrene) than those criteria derived from ecological end-points.

The choice of method to determine a quantitative tissue residue-sediment contamination level relationship depends on the specific pollutants, organisms, and water systems of concern, as well as the regulatory approach (e.g., remedial action, wasteload allocation, Superfund enforcement). The linkage between organism residue and sediment chemical concentration can be made from site-specific measurements of sediment-organism partition coefficients (Kuehl *et al.*, 1987); fugacity or equilibrium partitioning model (Clark *et al.*, 1988); predictions of organism residues; or pharmacokinetic-bioenergetic model predictions of organism residues that result from uptake from food chain, water, and sediment contact (Thomann, 1989). The residue approach works best for aquatic ecosystems that are at or close to steady state with respect to the distribution of chemicals between biotic and abiotic components. Steady-state conditions are common for most sediment contaminants because of their persistence and tendency to exert long-term rather than episodic bioaccumulation and toxic effects.

7.2.1.1 Objectives and Assumptions

The objective of this approach is to generate numerical sediment quality criteria based on acceptable levels of chemical contaminants in sediment-exposed biota. This objective is parallel to that of the water quality criteria, except that organism residues provide measures of exposure to chemical contaminants rather than water concentrations of contaminants. By using tissue residues rather than interstitial water concentrations to measure dose, as in the equilibrium partitioning approach (Chapter 5), this method does not require that the organism be at thermodynamic equilibrium with respect to the sediment contamination level. The site-specific residue approach is powerful because it does not require knowledge of bioavailability relationships for each organism in the system. All interaction pathways between sediment and organisms are incorporated in the determination of organism-to-sediment contamination ratios. These can be expressed on the basis of sediment organic carbon-organism lipid for hydrophobic organic chemicals. It is assumed that reduction in sediment contaminant concentrations over time (e.g., as a result of remedial actions, wasteload allocations) will result in parallel reduction in exposure, aquatic organism residues, and, consequently, the potential for toxic effects. It is further assumed that data on residue-to-toxicity relationships can be obtained from laboratory exposures of organisms when such data are not already available and that the route of exposure responsible for residue accumulation does not influence the residue-toxicity relationships.

7.2.1.2 Level of Effort

Relatively little effort would be required to generate preliminary sediment quality criteria using MPTCs calculated from existing water quality criteria and BCFs. In the absence of appropriate water quality criteria or BCFs, the level of effort depends on the availability of tissue residue action levels and the complexity of the sediment contaminant mitigation approach to be used. Relatively little effort is required to determine the degree to which sediment contamination concentrations must be reduced for single chemicals in well-mixed systems where fish residues are uniformly unacceptable for human consumption. Much more effort is required for systems having sediment contamination "hot spots" where resident aquatic organisms are eliminated or reduced in number due to a complex mixture of sediment contaminants. Another complexity that could increase the required level of effort is the presence of sediment contaminants that are readily metabolized to chemicals of greater toxicity that are responsible for the observed adverse effects. In some cases, residue-toxic effects data would incorporate the effects of toxic metabolites.

7.2.1.2.1 Type of Sampling Required

Surface sediment samples must be analyzed for chemical contaminants of interest. Interstitial water composition does not need to be determined because the residues in biota are related to bulk sediment chemical composition. Sediment characteristics such as grain size, organic carbon content, and metal binding capacity are useful for defining sediment-to-biota relationships for different sites within an ecosystem. Biota sampling for residue analysis should include sensitive organisms when toxic effects are a concern or, in the absence of sensitive organisms, organisms whose residues will serve as biomarkers for establishing safe sediment contaminant levels.

7.2.1.2.2 Methods

The tissue residue approach, as discussed in Section 7.2, depends on determining residues in aquatic organisms that are unacceptable on the basis of toxicity to the organism or unsuitability for human or animal consumption as food. The linkage of sediment contaminant concentrations to organism residues is possible through a number of approaches including site-specific measurements, equilibrium partitioning-based predictions, and steady-state food chain models. The choice of a specific approach depends on the chemical of concern, the availability of residue-toxic effects data, the contamination history (in-place pollutant problem versus a continuing or projected sediment contamination problem), and the characteristics of the impacted ecosystem. The construction of comprehensive, systematic strategies for all potential sediment contamination assessments will be achieved through further research and development.

Toxicity identification evaluation (TIE) procedures (see Chapter 5) complement the tissue-residue approach. The TIE approach is especially useful if sediment assessment begins without knowledge of the sediment contaminants that are causing toxicity or unacceptable residues in biota. The absence of benthic species or failure of fish eggs to hatch may be attributable to acutely toxic, but non-residue-forming, chemicals (e.g., ammonia) in sediments. TIE procedures can distinguish between potential metal, nonpolar organic, polar organic, and inorganic chemical sources of toxicity in sediment pore waters or elutriates. These procedures enable a more complete assessment of the significance of residue-associated toxicity in the system.

Once potentially toxic, bioaccumulative contaminants are identified, either in sediment or in aquatic organisms associated through exposure to sediments, the toxicological significance of site-specific sediment-to-biota contaminant partition factors can be assessed. Conservative generic sediment quality criteria can be generated from residue-toxicity relationships by assuming equilibrium partitioning between the binding fractions of organisms and sediments (e.g., lipid and sediment organic carbon for nonpolar organic chemicals).

7.2.1.2.3 Types of Data Required

The tissue residue method requires identification of chemicals in the sediment that are likely to be associated with chronic environmental effects. An indirect method for identifying such chemicals and their locations is to screen aquatic organisms for residues as in the National Dioxin Study (USEPA, 1987b) or the National Study of Chemical Residues in Fish (USEPA, 1992), sponsored by EPA's Office of Water Regulations and Standards. When toxicity data are not available, either laboratory dose-response experiments or quantitative structure-activity predictions can be used to establish the toxicological significance of the tissue residues. Field surveys that indicate the absence of sensitive organisms in contaminated sediment areas are useful for establishing sediment quality criteria, especially if interspecies sensitivities to the chemicals of concern are known. Tissue residues associated with no-effect and lowest-observable-effect concentrations are needed when the sediment criterion is not based on a human health standard.

7.2.1.2.4 Necessary Hardware and Skills

Sediment and tissue analyses require commonly available chemical analytical capabilities. Some chemicals require advanced instrumental analytical techniques, such as high resolution gas chromatography/mass spectrometry.

7.2.1.3 Adequacy of Documentation

The use of tissue residues to establish sediment criteria on the basis of human health effects associated with ingestion of contaminated fish has been documented. Methods for using tissue residue-toxicity relationships to establish sediment criteria, although scientifically sound, have not been extensively documented. The various methods for predicting tissue residues in benthos and fish have been well documented.

7.2.2 Applicability of Method to Human Health, Aquatic Life, or Wildlife Protection

Tissue residue measurements are directly applicable to human risk assessment when the aquatic organism is used as human food. Because of this relationship, the tissue residue method provides a direct link between human health and sediment criteria development. Tissue residues for wildlife and aquatic organisms can be used to assess sediment toxicity when there is an established exposure linkage to the sediment. The tissue residue approach is most advantageous for sediment contaminants that adversely impact organisms such as fish that are not in direct contact with the sediment or its interstitial water. The tissue residue approach is well suited to evaluating sediment quality in systems that have aquatic food chain connections from benthos to birds experiencing eggshell thinning or genotoxic effects. The tissue residue concentration serves as a quantitative measure of sediment contaminant bioavailability, which may differ as a function of ecosystem, sediment, water, food chain, and species characteristics.

7.2.3 Ability of Method to Generate Numerical Criteria for Specific Chemicals

The tissue residue approach can be used to generate site-specific numerical criteria for nonpolar organic chemicals such as PCDDs, PCDFs, and PCBs. Tissue residues of aldrin/dieldrin (USEPA, 1980a) and endrin (USEPA, 1980b) have been used to establish water quality criteria on the basis of human health risks. The DDT and PCB water quality criteria are based on toxic effects in birds and animals as a function of fish residues (USEPA, 1980c, 1980d). Tissue residues of organometallic chemicals such as methyl mercury (USEPA, 1984) and elements such as selenium (USEPA, 1987a) have been used to establish water quality criteria and/or to predict toxic effects. There is some evidence to indicate that metal residues in sediment-dwelling aquatic organisms can reflect both metal bioavailability and potential metal toxicity. Thus, tissue residue-toxicity relationships for some elements could be used as an adjunct to the interstitial water equilibrium partitioning approach.

7.3 USEFULNESS

7.3.1 Environmental Applicability

7.3.1.1 Suitability for Different Sediment Types

There is no limitation to the suitability of this approach for different sediment types since the method is sensitive to bioavailability differences among sediments. When pelagic organisms are used to assess sediment quality, sediment variability in the water body tends to be averaged.

7.3.1.2 Suitability for Different Chemicals or Classes of Chemicals

This approach is most applicable to nonpolar organics and organometallics that bioaccumulate, are slowly metabolized, and exert chronic toxic effects or present risks to human health. This approach also could work well for chemicals that are metabolized by the organism to nontoxic forms since the parent compound residue reflects this change in toxic potential. In some cases residues of known metabolites, which are more toxic than the parent compound, can be used to establish residue-toxic effects relationships (Krahn *et al.*, 1986). The approach is not useful for assessing sediment toxicity associated with non-residue-forming toxic chemicals such as ammonia, hydrogen sulfide, and polyelectrolytes. Since there is evidence that metal residues in some sediment-dwelling organisms are indicative of both metal bioavailability and potential metal toxicity, sediment quality criteria for metals should be aided by a site-specific tissue residue approach. However, when biological species sequester metals in a nonbiologically available form, tissue residue-toxicity effects linkages may be obscured. The suitability of the method for evaluating additive, synergistic, or antagonistic effects associated with complex mixtures of sediment contaminants depends on the development of chemical mixture toxic dose-response relationships where dose is indicated by tissue residue levels.

7.3.1.3 Suitability for Predicting Effects on Different Organisms

The tissue residue approach should not be limited by species unless organism residues cannot be obtained or toxic effects cannot be determined through water quality criteria or bioassays. The key species problem is identification of sensitive species for the sediment contaminants of concern. When adequate comparative toxicity data exist, residues from tolerant organisms may be used to infer sediment criteria for sensitive organisms that are not found in association with the sediment because of toxic effects.

7.3.1.4 Suitability for In-Place Pollutant Control

Evaluation of the association of site-specific tissue residues with sediment toxic chemical concentrations provides an established method for in-place pollutant assessment for both human health and ecological risks. Comparison of tissue residues in field-collected organisms to the MPTC would be a direct estimate of ecological risk. The use of resident or caged biota for bioaccumulation potential and toxicity assessments is useful for detection of the most toxic sediments or monitoring of changes in toxicity following remedial action. By weighing the relative toxicity of bioaccumulated pollutants (e.g., by using "dioxin equivalents"), evaluation of tissue residue concentrations can help identify the pollutants most likely responsible for toxicity and their additive contribution to total sediment toxicity. This information could then be used to design the most appropriate and cost-effective mitigation response.

7.3.1.5 Suitability for Source Control

The tissue residue approach is well suited for establishing source control. Comparison of the existing or predicted tissue residue levels with MPTCs generates a quantitative estimate of the extent to which a given sediment exceeds or is below a sediment quality criterion. In conjunction with physical transport models, this information can then be used directly to determine acceptable discharge limits, wasteload allocations, or the types of remedial procedures required to achieve acceptable tissue residue levels. The Lake Ontario TCDD-Hyde Park Superfund case example described in Section 7.1.1 demonstrates the suitability of this approach for establishing source controls. The site-specific nature of this approach provides strong support for establishing controls on existing point and nonpoint sources of sediment contamination.

7.3.1.6 Suitability for Disposal Applications

When site-specific sediment-biota contaminant partition coefficients are unavailable, such as for evaluation of proposed disposal operations, the residue approach can be applied by predicting benthic tissue residues from steady-state toxicokinetic bioaccumulation models or by conducting laboratory bioaccumulation tests on the dredged material. If adverse effects on fishes, wildlife, or human health are of concern at such disposal sites, it would then be necessary to apply a trophic transfer or equilibrium partitioning model to predict tissue residues in these higher trophic levels. When the disposal site already has sediments containing the contaminants of concern, residues in existing biota may be used to predict residue levels and toxic effects that would result from additional disposal of similarly contaminated dredged material.

7.3.2 General Advantages and Limitations

7.3.2.1 Ease of Use

The application of sediment quality criteria derived from tissue residues for assessing pelagic or benthic ecological effects is fairly direct. The measured or predicted sediment concentration would simply be compared to the sediment quality criterion derived from MPTCs. The development of a tissue residue toxicity database from laboratory bioassays would allow convenient access to the required biological effects endpoints. Chemical analyses of sediment, total organic carbon, and tissue samples for assessing existing conditions require routine analytical chemistry capabilities that do not present unique problems. One potential difficulty when using tissue residues in field-collected benthos to assess in-place sediments is the difficulty in obtaining sufficient benthic biomass for chemical analysis. This problem can be avoided by conducting laboratory bioaccumulation tests on field-collected sediment or by placing caged benthic organisms in the field.

7.3.2.2 Relative Cost

Costs associated with further development of the generic tissue residue approach for sediment quality criteria include (1) development of a residue-toxicity relationship database and (2) validation of the relationships between the MPTC and chronic impacts on aquatic organisms for different chemical classes of sediment contaminants. The cost of applying the method to a particular site, however, depends on the number of sediment and biota samples to be analyzed, the availability of residue-toxicity relationship data, and the difficulty in identifying sensitive organisms. The establishment of a sediment criterion based on fish residue levels acceptable for protection of human health generally results in low analytical costs when only a few reference sediment sites are needed to characterize the system of concern.

7.3.2.3 Tendency to Be Conservative

This approach does not tend to be either conservative or liberal for prediction of ecological effects unless the system responds in a nonlinear manner to reductions in sediment contaminants. In the case of nonlinearity, the tendency would probably be toward conservatism because of the greater bioavailability of more recently introduced sediment contaminants. When human health endpoints are used to generate sediment quality criteria, the criteria may be more strict than necessary to protect resident biota.

7.3.2.4 Level of Acceptance

The tissue residue approach is accepted as a basis for regulatory decisions such as the establishment of water quality criteria for the protection of aquatic life and its uses. The direct prediction of chronic toxic effects from measured or predicted tissue residues requires validation before it can be widely endorsed. Since sediment contaminants tend to be long-term exposure problems and can bioaccumulate, residues should be acceptable for sediment criteria development. This approach should be acceptable for identifying sediments associated with a degree of exposure which exceeds that indicated as deleterious in previous experiments.

7.3.2.5 Ability to Be Implemented by Laboratories with Typical Equipment and Handling Facilities

The tissue residue approach requires analyses of only sediment and tissue residues when potentially toxic sediment contaminants are known and residue-toxicity relationship data are available. If extensive laboratory work is needed to determine chemical residue-chronic toxicity dose-response relationships for sensitive species, specialized aquatic toxicology capabilities are required. In theory, residue-toxicity-based MPTCs can be obtained for all chemicals subject to water quality criteria development.

7.3.2.6 Level of Effort Required to Generate Results

The level of effort depends on the number and nature of sediment contaminants, the complexity of the contaminant distribution pattern, and the regulatory application of the method. Some cases will require relatively few analyses of tissue and sediment residues and no toxicity testing to apply the method (e.g., to remedial action decisions, wasteload allocations).

7.3.2.7 Degree to Which Results Lend Themselves to Interpretation

Tissue residues that exceed concentrations considered safe for human exposure through seafood consumption require no interpretation when used to set residue-based sediment criteria. However, the degree of interpretation may be very large when evaluating ecotoxicological effects attributed to site-specific measurements of sediment-to-biota chemical partitioning. This interpretation problem exists for all sediment classification methods when applied on a site-specific basis. The presence of unacceptable residues in indicator organisms resident in or linked to an area of sediment contamination can be used without elaborate interpretation to determine compliance of sediments with sediment quality criteria.

7.3.2.8 Degree of Environmental Applicability

The use of site-specific tissue residues as quantitative exposure biomarkers eliminates uncertainties associated with chemical bioavailability; exposure duration, frequency, and magnitude; and toxicokinetic/bioenergetic factors. When the tissue residue approach is applied on a generic basis to generate sediment criteria for different chemicals, these uncertainties can be partially addressed through classification of sediments and exposure environments.

7.3.2.9 Degree of Accuracy and Precision

Sediment and tissue residue chemical concentrations can be determined accurately and precisely for most chemicals. Most uncertainties in sediment/organism partition coefficients are due to biological variability. Accuracy and precision can be maximized through site-specific investigations of biological factors that influence organism linkage to sediment (through food chain, water, or direct contact) and through refinement of residue-toxicity relationships.

7.4 STATUS

7.4.1 Extent of Use

Use of tissue residues to establish sediment criteria on the basis of human health effects has been documented. Tissue residues have also been used to derive water quality criteria for the protection of aquatic life and wildlife connected to the aquatic food chain. Tissue residue-toxicity data that may be used for deriving numerical sediment quality criteria for some chemicals already exist in water quality criteria documents, fish consumption advisories, and the peer-reviewed literature. Much aquatic toxicology work in progress or planned for the future could produce the necessary data if residue-based dose measurements are incorporated into research plans.

7.4.2 Extent to Which Approach Has Been Field-Validated

Sediment TCDD contamination limits have been established for Lake Ontario on the basis of fish tissue residues. This use of tissue residue to generate sediment criteria has been validated through a steady-state model (Endicott *et al.*, 1989) and a laboratory bioaccumulation study (Cook *et al.*, 1989) that demonstrated a linear relationship at steady-state between sediment contaminant concentration and bioaccumulated

TCDD in lake trout, regardless of route of uptake. Declines in DDT residues in fish and birds since its use was banned are associated with declining surficial sediment concentrations in the Great Lakes, the Southern California Bight, and elsewhere. Although other examples of studies validating the residue approach for single chemicals are available, its use for complex mixtures of chemicals in sediments to predict acceptable contaminant concentrations with ecosystem protection in mind has not been validated.

7.4.3 Reasons for Limited Use

Use of the tissue residue approach has been limited for the following reasons:

- This method is in a developmental stage and has not been formally adopted by EPA.
- Aquatic toxicology has only recently progressed to an understanding of residue-based dose-response relationships for sediment contaminants.
- Regulatory agencies, including EPA, have not yet become committed to systematic establishment and application of sediment criteria methods.
- The available and potentially available residue-based toxicity data have not been collated into a database for potential sediment criteria users.

7.4.4 Outlook for Future Use and Amount of Development Yet Needed

This method can be implemented with a minimal amount of effort in many cases, especially where a single chemical or toxicologically related family of chemicals is of concern. Guidance documents should be written and reviewed. Tissue residue criteria should be accumulated systematically for a database. The use of this method in combination with other sediment classification methods should be considered. Field validation of residue-based ecological effects predictions is essential. All sediment assessment methods should be developed with concern for identification of and application to those chemicals in the aquatic environment that are long-term sediment contaminants having chronic toxicity potential.

7.5 REFERENCES

Batterman, A.R., P.M. Cook, K.B. Lodge, D.B. Lothenbach, and B.C. Butterworth. In press. Methodology used for a laboratory determination of relative contributions of water, sediment and food chain routes of uptake for 2,3,7,8-TCDD bioaccumulation by lake trout in Lake Ontario. Chemosphere.

Carey, A.E., N.S. Shifrin, and A.C. Roche. 1989. Lake Ontario TCDD bioaccumulation study final report. Chapter 1: introduction, background, study description and chronology. Gradient Corporation, Cambridge, MA. 17 pp.

Clark, T., K. Clark, S. Pateson, D. Mockay, and R.J. Norstrom. 1988. Wildlife monitoring, modeling and fugacity. Environ. Sci. Technol. 22:120-127.

Cook, P.M., A.R. Batterman, B.C. Butterworth, K.B. Lodge, and S.W. Kohlbry. 1990. Laboratory study of TCDD bioaccumulation by lake trout from Lake Ontario sediments, food chain and water. In: Lake Ontario TCDD Bioaccumulation Study - Final Report, Chapter 6. U.S. Environmental Protection Agency, Region II, New York.

Endicott, D., W. Richardson, and D. DiToro. 1989. Lake Ontario TCDD modeling report. U.S. Environmental Protection Agency, Large Lakes Research Station, Environmental Research Laboratory Duluth, Grosse Ile, MI. 94 pp.

Krahn, M.M., L.D. Rhodes, M.S. Myers, L.K. Moore, W.D. MacLeod, and D.C. Malins. 1986. Associations between metabolites of aromatic compounds in bile and the occurrence of hepatic lesions in English sole (*Parophrys velulus*) from Puget Sound, Washington.

Arch. Environ. Contam. Toxicol. 15:61-67.

Kuehl, D.W., P.M. Cook, A.R. Batterman, D. Lothenbach, and B.C. Butterworth. 1987. Bioavailability of polychlorinated dibenzo-p-dioxins and dibenzofurans from contaminated Wisconsin River sediment to carp. Chemosphere 16:667-679.

Mac, M.J. 1988. Toxic substances and survival of Lake Michigan salmonids: field and laboratory approaches. pp. 389-401. In: Toxic Contaminants and Ecosystem Health. M.S. Evans (ed). Wiley & Sons.

Stephan, C.E., D.I. Mount, D.J. Hansen, J.H. Gentile, G.A. Chapman, and W.A. Brungs. 1985. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. PB85-227040. National Technical Information Service, Springfield, VA.

Thomann, R.V. 1989. Bioaccumulation model of organic chemical distributions in aquatic food chains. Environ. Sci. Technol. 23:699-707.

USACE. 1991. Influence of sediment potential of PCBs: Field studies at the Calumet Confined Disposal Facility. Environmental Effects of Dredging Notes - EEDP-02-16, U.S. Army Corps of Engineers. U.S. Army Engineer Waterways Experimental Station, Vicksburg, MS.

USEPA. 1980a. Ambient water quality criteria for aldrin/dieldrin. EPA 440/5-80-019. NTIS number PB81-117301. U.S. Environmental Protection Agency, Washington, DC.

USEPA. 1980b. Ambient water quality criteria for endrin. EPA 440/5-80-047. NTIS number PB81-117582. U.S. Environmental Protection Agency, Washington, DC.

USEPA. 1980c. Ambient water quality criteria for DDT. EPA 440/5-80-038. NTIS number PB81-117491. U.S. Environmental Protection Agency, Washington, DC.

USEPA. 1980d. Ambient water quality criteria for polychlorinated biphenyls. EPA 440/5-80-068. NTIS number PB81-117798. U.S. Environmental Protection Agency, Washington, DC.

USEPA. 1984. Ambient water quality criteria for mercury. EPA 440/5-84-026. NTIS number PB85-227452. U.S. Environmental Protection Agency, Washington, DC.

USEPA. 1987a. Ambient water quality criteria for selenium. EPA 440/5-87-006. NTIS number PB88-142237. U.S. Environmental Protection Agency, Washington, DC.

USEPA. 1987b. The national dioxin study. Tiers 3, 5, 6, and 7. EPA 440/4-87-003. U.S. Environmental Protection Agency, Office of Water Regulations and Standards, Washington, DC.

USEPA. 1992. National study of chemical residues in fish. 2 vols. EPA 823-R-92-008a,b. U.S. Environmental Protection Agency, Office of Science and Technology, Standards and Applied Science Division, Washington, DC.

Walker, M.K., J.S. Spitbergen, J.R. Olson, and R.E. Perterson. 1991. 2,3,7,8-Tetrachlorodibenzo-p-dioxin toxicity during early life stage development of lake trout (*Salvelinus namaycush*). Can. J. Fish. Aqua. Sci. 48:875.

Freshwater Benthic Macroinvertebrate Community Structure and Function

Wayne S. Davis
U.S. Environmental Protection Agency Region V, Environmental Sciences Division
77 West Jackson (SQ-14J), Chicago, IL 60604
312/FTS 886-6233

Joyce E. Lathrop
College of DuPage, Division of Natural Sciences
22nd at Lambert Road, Glen Ellyn, IL 60137

The community, or assemblage, structure and function of benthic macroinvertebrates is used extensively to evaluate the quality of water resources and characterize causes and sources of impacts in lotic (flowing water) and lentic (standing water) freshwater ecosystems. (Marine benthic community structure is discussed in Chapter 9.) Benthic macroinvertebrates are relatively sedentary organisms that inhabit or depend on the sedimentary environment for their various life functions. Therefore, they are sensitive to both long-term and short-term changes in habitat, sediment, and water quality. This chapter discusses assessment of benthic macroinvertebrates to determine sediment quality in conjunction with an integrated approach for assessing the quality of the water resources. This integrated approach uses sediment chemistry, sediment toxicity, habitat quality, and benthic macroinvertebrate community (assemblage) structure and function to evaluate sediment quality, similar to the approaches now used to evaluate surface water quality. The structural assessment relates to the numeric taxonomic distribution of the community, and the functional assessment involves trophic level (feeding group) and morphological assessment. This chapter addresses the specific benthic community assessment methods that are available, or being developed, to complement the chemical and toxicological portions of the sediment quality assessment.

8.1 SPECIFIC APPLICATIONS

8.1.1 Current Use

Freshwater benthic macroinvertebrate communities are used in the following ways to assess the quality of the water resource (sediments, water, and habitat):

- Identification of the quality of ambient sites through a knowledge of the pollution tolerances and life history requirements of benthic macroinvertebrates;
- Establishment of criteria and standards based on community performance at multiple reference sites throughout an ecoregion or other regionalization categories;
- Comparison of the quality of reference (or least impacted) sites with test (ambient) sites;
- Comparison of the quality of ambient sites with historical data to identify temporal trends; and
- Determination of spatial gradients of contamination for source characterization.

8.1.1.1 Ecological Uses

Benthic macroinvertebrate community (assemblage) structure and function assessments have many different applications. Site-specific knowledge of surface water quality, habitat quality, sediment chemistry, and sediment toxicity provide the best context in which to interpret benthic community assessment data. The objectives of each particular study should determine the types of related data necessary. Alone, benthic macroinvertebrates can be used to screen for potential sediment contamination based on spatial gradients in community structure, but they should not be used alone to definitively determine sediment quality. Benthic macroinvertebrate data must be integrated with other available data to determine sediment quality. Benthic macroinvertebrate often provide the most important piece of information on sediment quality. Care must be exercised to collect representative samples to minimize problems with data interpretation due to natural variations. For example, collections should not be made after floods or other physical disturbances that may physically alter or remove benthic assemblages.

Benthic macroinvertebrate community structure and function have been used extensively to characterize freshwater ambient conditions and impacts from various sources. Documented changes in benthic community structure have resulted from crude oil exposure in ponds and streams (Rosenberg and Wiens, 1976; Mozley, 1978; Mozley and Butler, 1978; Cushman, 1984; Cushman and Goyert, 1984) and heavy metal contamination of lake sediments and streams (Winner *et al.*, 1975, 1980; Wentsel *et al.*, 1977; Moore *et al.*, 1979; Wiederholm, 1984a, 1984b; Waterhouse and Farrell, 1985). Benthic macroinvertebrates have been used extensively to identify organic enrichment in lentic systems (Cook and Johnson, 1974; Krieger, 1984; Rosas *et al.*, 1985) and lotic systems (Richardson, 1928; Gaufin and Tarzwell, 1952; Hynes, 1970; Hilsenhoff, 1977, 1982, 1987, 1988). Benthic community responses to pesticides (van Dyk *et al.*, 1975; Webb, 1980; Penrose and Lenat, 1982; Yasuno *et al.*, 1985), acid- and mine-stressed lotic environments (Simpson, 1983; Armitage and Blackburn, 1985), thermally stressed water bodies (Crossman *et al.*, 1984), and urban and highway runoff impacts (Smith and Kaster, 1983; Dupuis *et al.*, 1985; Denbow and Davis, 1986) have also been documented. Chironomidae (midge) larvae were even found to transport substantial amounts of PCBs from contaminated sediments to the terrestrial environment (Larsson, 1984).

8.1.1.2 Regulatory Uses

Assessment of benthic macroinvertebrate community (assemblage) structure and/or function has been used as a regulatory tool for a number of years (Davis, 1990). In 1987, USEPA hosted the First National Workshop on Biological Monitoring and Criteria (USEPA, 1988a, 1988b), which addressed the use of benthic macroinvertebrates, as well as fish, in EPA and State regulatory programs. This workshop formally initiated EPA's efforts toward development and implementation of "biological criteria" based on benthic macroinvertebrate, fish, and habitat assessments. These biological criteria, which have been predominantly based on the macroinvertebrates, are designed to determine whether a specific water body or water body segment is meeting its designated use for aquatic life (i.e., water quality standards).

EPA requires the development of biological criteria and adoption by States into their water quality standards by September 30, 1993 (USEPA, 1991a, 1990b). This requirement has been supported by a formal policy (USEPA, 1990c), program guidance (USEPA, 1992a), and technical guidance and support documents (USEPA, 1991a, 1991b, 1991c, 1991d, 1991e; 1992b, 1992c). Several States currently use benthic macroinvertebrates as a regulatory tool, either alone or in combination with other ecological parameters (Ohio EPA, 1990, USEPA, 1991c, 1991e). USEPA also supports the use of benthic macroinvertebrates as a primary environmental indicator for surface waters that EPA should use to track compliance with Clean Water Act objectives (Abe *et al.*, 1992; USEPA, 1990d, 1990e).

Under the Clean Water Act, as amended in 1987, benthic macroinvertebrates are used for the following:

- Measurement of the restoration and maintenance of biological integrity in surface waters (section 101);
- Development of water quality criteria based on biological assessment methods when numerical criteria for toxicity have not been established [section 303(c)(2)(B)];

- Production of guidance and criteria based on biological monitoring and assessment methods [section 304(a)(8)];

- Development of improved measures of the effects of pollutants on biological integrity (section 105);

- Production of guidelines for evaluating nonpoint sources (NPS) [section 304(f)];

- Listing of waters that cannot attain designated uses without additional NPS controls (section 319);

- Listing of waters unable to support balanced aquatic communities [section 304(l)];

- Assessment of lake trophic states and trends (section 314);

- Production of biennial reports on the extent to which waters support balanced aquatic communities [section 305(b)]; and,

- Determination of the effect of dredge and fill disposal on balanced wetland communities (section 404).

Benthic macroinvertebrates and biological criteria have also been used to evaluate on-site and off-site ecological impacts from hazardous waste sites. Environmental assessment of a Superfund site is done in accordance with EPA's responsibility to protect public health and the environment under the Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (CERCLA) as amended by the Superfund Amendments and Reauthorization Act of 1986 (SARA). The regulation that enables EPA to carry out its responsibilities under CERCLA/SARA is the National Contingency Plan (NCP).

The NCP calls for the identification and mitigation of environmental impacts of these sites and the selection of remedial actions that are "protective of environmental organisms and ecosystems." Federal and state laws and regulations that aid in this process are potentially "applicable or relevant and appropriate requirements" (ARARs). Compliance with these laws and regulations increasingly requires that the site's ecological effects be evaluated and measures be taken to mitigate those adverse effects.

The Clean Water Act, as amended by the 1987 Water Quality Act, is another ARAR and major federal regulation that requires the maintenance and restoration of the chemical, physical, and biological integrity of the Nation's waters. Most Superfund sites potentially affect surface waters and need to be assessed for both on-site and off-site effects. A detailed discussion of the legal and technical requirements for environmental assessments at Superfund sites can be found in EPA's *Risk Assessment Guidance for Superfund: Environmental Evaluation Manual* (USEPA, 1989a). As EPA focuses on watershed and water body impacts regardless of the programmatic sources and causes, the use of benthic macroinvertebrates for assessing the health of surface water systems will increasingly become important.

8.1.2 Potential Use

The use of benthic macroinvertebrates to assess sediment contamination will be most successful when combined with sediment chemistry and toxicity results, as in the "integrated" Sediment Quality Triad approach (see Chapter 10). Benthic macroinvertebrates will best indicate in-place pollutant control needs through a site-specific knowledge of surface water quality, habitat quality, and sediment chemistry and toxicity. Habitat quality assessments will help establish reasonable expectations for benthic community structure and function. Alone, benthic macroinvertebrates can be used to screen for potential sediment contamination and source identification by displaying spatial gradients in community structure, but they should not be used alone to definitively determine sediment quality or to develop chemical-specific guidelines. Benthic macroinvertebrate data must be integrated with other available data to determine sediment quality as well as the quality of the overall water resource.

8.2 DESCRIPTION

8.2.1 Description of Method

The benthic macroinvertebrate community structure and function assessment involves the following steps:

(1) Establishment of data quality objectives, selection of sample sites and frequency of collection in Quality Assurance Program Plan;

(2) Collection of benthic macroinvertebrates in the field (artificial or natural substrates);

(3) Sorting the organisms from debris (field or laboratory);

(4) Identification to the lowest taxon necessary (varies depending on the study objectives);

5) Multimetric or composite index quantification (e.g., taxa richness, number of individuals, indicator organism count, structural indexes and ratios, functional characteristics of taxa);

(6) Assessment of the relationship with other environmental measurements including numeric habitat quality assessment (e.g., correlations, habitat requirements) and expectations;

(7) Comparison with a local or regional "reference" site (e.g., similarity indexes, nonparametric analyses); and

(8) Complete documentation of the study methods, results, database management, and discussion of the relevance of the data.

8.2.1.1 Objectives and Assumptions

The primary objective of benthic macroinvertebrate community (assemblage) structure and function analyses is to provide data and information to assist in determining the quality of the sediment/water environment. This determination can then be used for the purposes described above in Section 8.1 (Specific Applications).

It is assumed that benthic macroinvertebrates can provide consistent and accurate assessments of sediment/water quality at a given sample location or water body. Specifically, the following assumptions are implicit in this objective:

- The benthic macroinvertebrates are relatively sedentary, especially compared to fish communities, and they depend on the sedimentary (or benthic) environment for their life functions.

- Chemical and physical perturbations of the sediments or bottom waters affect benthic macroinvertebrates since they are dependent on the benthic environment for completion of their life cycles, and they are therefore sensitive to changes in sediment and water quality.

- Benthic macroinvertebrates physically interact with the sediments to cause chemical exchange between the sediment and the overlying water, and therefore tend to reflect sediment quality as well as water quality.

- Minimum habitat quality exists below which the community structure and function will perform poorly regardless of the chemical contaminants present or not present.

- The optimal use of benthic macroinvertebrates as sediment quality indicators is as part of an integrated sediment quality assessment approach using sediment chemistry, sediment toxicity, and benthic community structure and function.

Equally important assumptions apply to actual benthic macroinvertebrate sampling strategy, collection, identification, data reduction, interpretation of results, and report preparation. It is assumed that all U.S. EPA-supported studies have an adequate

Quality Assurance Project Plan (QAPP) and that all benthic macroinvertebrate community data are reproducible and collected in a manner to minimize data interpretation problems with natural variations; the methods must be consistent within each study. Specific QA procedures that should be established early in benthic macroinvertebrate community studies include the following:

- Rationale for sample location selection;
- Sample collection methods, sorting, and storage procedures;
- Taxonomic proficiency evaluations using either U.S. EPA check-samples from Cincinnati-ERL or state-developed check-samples, in addition to voucher collections from each study area and a list of the taxonomic references used;
- Multimetric data analysis techniques used to objectively assess the data, including the structural and functional measures; and
- Nonparametric or parametric (as appropriate) statistical methods used to compare site results.

Each Regional U.S. EPA Quality Assurance Office can provide the details of QAPP requirements. Further discussion of quality assurance measures can be found in Klemm *et al.* (1990), Bode (1988), Ohio EPA (1989b), and Stribling (1991).

8.2.1.2 Level of Effort

The level of effort required to conduct freshwater benthic macroinvertebrate community studies is comparable with chemical/physical water quality measurements and bioassays and has been thoroughly discussed in Plafkin *et al.* (1989) and Ohio EPA (1990a). However, rapid benthic community assessment techniques can range from 1 to 5 hours per site if laboratory identifications are not required (Plafkin *et al.*, 1989). As expected, the greatest time expenditure is in the travel to and from the site and in the sorting and identification of the organisms.

Separating the organisms from debris and sorting the organisms into taxonomic categories can take up to 15 hours per sample, with an additional 12 hours for identification, for very enriched sites with high numbers of individuals among several taxa. In such extreme situations, subsampling may be preferred. More typically, the time spent would be about 3 hours for sorting (more time for dredge and artificial substrate samples and less time for dip-net samples), 2 hours for preparing the samples (e.g., clearing and then mounting the chironomids on microscope slides), and 6 hours for identifying the organisms to the lowest possible taxonomic level. An experienced taxonomist with appropriate keys may average only 2-4 hours per site. This typical time equates to about 11 hours per site after the samples have been collected. These estimates are only a general guide to the time it may take to perform the identifications and are meant to help assess potential or actual project costs.

8.2.1.2.1 Type of Sampling Required

The specific sampling methods to be used are dictated by the study needs. Debate will continue regarding the use of "quantitative" and "qualitative" sampling methods, but each method is acceptable contingent upon how well it will satisfy study objectives, reproducibility of the data, and consistency of collection. Typically, benthic macroinvertebrate data are quantified by the surface area of the sampler or sediment being collected. However, benthic macroinvertebrates can be quantified in other ways depending on the objectives of the study. For example, if the objective is to determine the number and types of taxa in a study area, rather than the number of individuals within each taxon, then using a dip-net in various habitats within the study area until no new taxa are encountered could be considered quantitative with relation to the number of taxa and time expended. Examples of programs using data quantified by methods other than surface area of the sampler or substrate include those described by Pollard (1981), Hilsenhoff (1982, 1987, 1988), Cummins and Wilzbach (1985), Bode and Novak (1988), Cummins (1988), Hite (1988), Lenat (1988), Maret (1988), Penrose and Overton (1988), Plafkin *et al.* (1989), and Shakelford (1988). The success of each sampling effort depends on a thorough understanding of the data quality objectives of that study and the implementation of a quality assurance program.

8.2.1.2.2 Methods

EPA (Klemm *et al.*, 1990) recently published *Macroinvertebrate Field and Laboratory Methods for Evaluating the Biological Integrity of Surface Waters*, which thoroughly addresses methodology. Most state environmental regulatory programs have a Quality Assurance Project Plan describing the field methods and standard operating procedures for collecting and evaluating benthic macroinvertebrates (Bode, 1988; Illinois EPA, 1987; Ohio EPA, 1989a, 1989b). This information should be obtained to ensure acceptance and comparability of study results with those obtained by the state agency. If this information is not available, then field methods and standard operating procedures from other existing programs should be used. Since several different collection and analysis methods are used throughout the country depending on water body type (lotic vs. lentic), habitat type, substrate type, and familiarity with specific methods, it is not practical to recommend any single sampling method. The general quality assurance requirements the use of any one particular method is that the method produce data that are reproducible, consistently used within the program, and applicable by different investigators (Klemm *et al.*, 1990).

Methods Summary—In soft freshwater sediments the most common method used to collect benthos is with a grab sampler such as a Ponar (15 x 15 cm or 23 x 23 cm) or Ekman dredge (15 x 15 cm, 23 x 23 cm, or 30 x 30 cm), each of which provides a quantitative sample based on the surface area of the sampler. The smaller of the surface area sizes are most commonly used for freshwater studies because of their relative ease of manipulation. The Ekman dredge is not as effective in areas of vegetative debris, but is much lighter than the Ponar and easier to use in softer substrates. Artificial substrates (Hester-Dendy using several 3-inch plates and spacers attached by an eyebolt; or substrate/rock-filled baskets) provide a consistent habitat for the benthos to colonize in both soft-bottomed and stony areas. Artificial substrates can be used in almost any water body and have been successfully used to standardize results despite habitat differences (APHA *et al.*, 1989; DePauw, 1986; Hester and Dendy, 1962; Ohio EPA, 1989b; Rosenberg and Resh, 1982, 1991), but the major drawback to using the artificial substrates is the 4- to 8-week period for instream colonization. This would require at least two visits for each study site—one to place the samplers and one to remove them.

A variety of methods for sampling hard-bottomed lotic systems are available. Colonization of substrates and comparisons of the artificial and natural substrate methods have been described (Beckett and Miller, 1982; Chadwick and Canton, 1983; Crossman and Cairns, 1974; Lenat, 1988; Ohio EPA, 1989b; Peckarsky, 1986; Plafkin *et al.*, 1989; Shepard, 1982). If quantification by sediment or sampler surface area is needed, a Surber-type square-foot sampler (Surber, 1937, 1970) with a #30-mesh (0.589-mm openings) can be used. The traveling kick-net (or dip-net) method, also using a #30-mesh net, can be used to quantify the sample collected by the amount of time spent sampling and the approximate surface area sampled (Pollard, 1981; Pollard and Kinney, 1979). The Surber-type and kick-net methods can each be used to provide consistent, reproducible samples, but both are limited to wadable streams. The Surber sampler's optimal effectiveness is limited to riffles, whereas kick- or dip-net sampling can be used in all available habitats. Although dip-net samplers have been effectively used to sample riffles and other relatively shallow habitats to determine taxa richness, presence of indicator organisms, relative abundances, similarity between sites, and other information, they do not provide definitive estimates of the number of individuals or biomass per surface area.

For sediment evaluations of lotic systems, a combination of artificial substrate (e.g., Hester-Dendy) and natural substrate (dip-net) sampling is recommended. This combination allows comparison of the benthos communities independent of habitat so that sediment/water quality effects can be better assessed.

Sampling Strategy—Sampling strategies have been addressed by Klemm *et al.* (1990), Millard and Lettenmaier (1986), Plafkin *et al.* (1989), Rosenberg and Resh (1991), Sheldon (1984), and USEPA 1990b, 1990c). Special monitoring strategies have been prepared for EPA's Environmental Monitoring and Assessment Program (EMAP), which employs a probabilistic sample design (USEPA, 1991f); the intensive watershed surveys of the U.S. Geological Survey (Leahy *et al.*, 1990); and forestry activities in the Pacific Northwest (USEPA, 1991g). Regardless of the study objectives for regulatory use, reference (least-impacted) sites will be required for comparison with the results from test (ambient) sites. Reference sites can be established on a site-specific or regional basis. It is preferable to use a regionalization approach because the level of confidence in the results is greater using an increased number of reference sites, which allows for a verification that the sites truly are least-impacted reference sites. Regionalization (ecoregions, watersheds) has been successfully used in a number of State programs to support biological criteria development for benthic macroinvertebrates (Gallant *et al.*, 1989, Ohio EPA, 1990, Arkansas DPCE, 1987, Hughs *et al.*, 1990, USEPA, 1991c, 1991e).

When using site-specific reference sites to detect spatial differences in sediment/water quality, or to characterize sources of pollution, the best strategy is to collect samples in similar habitats upstream and downstream of suspected pollution sources or other areas of interest for ambient monitoring such as high-quality or wild and scenic streams (USEPA, 1992b). A minimum of two upstream sites and three downstream sites of the suspected pollutant source(s) should be sampled; however, many programs are limited to only one upstream site and one or two downstream sites. If habitats vary too widely, then artificial substrates should be placed at each site, with multihabitat dip-net sampling done when the substrates are placed instream and retrieved, to complement the artificial substrate data.

To best detect temporal trends, a fixed station network should be established near the area of interest and sampled consistently at least one season each year. A reference location should also be sampled at the same times to ensure that differences found in the results can be attributed to changes in water quality near the site. It is strongly recommended that a set of reference sites be developed within each ecoregion (or by other regionalization methods) and that those reference sites be sampled seasonally to better understand site-specific seasonal variability. Sampling should be done each year during similar flow conditions and should not be conducted for at least 1 or 2 weeks after a major rainfall because of the potential for physical disturbances of the substrate resulting in potentially lower biological integrity ratings.

Seasonal distributions are always a concern for ensuring the collection of a representative sample. Therefore, *routine* sampling or monitoring is optimal during the seasons indicated in Plafkin *et al.* (1988), and long-term monitoring should strive for consistent sampling seasons. The benthic macroinvertebrate discussion group at the 1987 National Workshop on Instream Biological Monitoring and Criteria agreed that the biologically optimal time of year for sampling in lotic systems was during the latter part of the season(s) that demonstrate a stable base-flow (normal flow) and temperature regime (Davis and Simon 1988).

Sample Replication—Sample replication is a component of a good Quality Assurance Program Plan. Recommendations and discussion regarding sample replication can be found in Plafkin *et al.* (1989), Klemm *et al.* (1990), and USEPA (1992b). Statistical derivation of the number of samples required to decrease the variability of the data have been discussed by Green (1978), Merritt *et al.* (1984), Resh and Price (1984), and Klemm *et al.* (1990). These methods generally rely on a prior knowledge of the variability of the data. This prior knowledge is often not available nor practical to obtain from a programmatic view (e.g., the cost of initial sampling to estimate variability and required number of replicates may be prohibitive). Another problem with statistically determining the number of samples needed is the assumption that the data follow a specific distribution such as normal or lognormal, which is not necessarily true for biological samples. Also, the variability, as measured by the variance or standard deviation, could be different for each descriptive index analyzed (number of taxa versus number of individuals, etc.).

Field Methods—Field sampling methods have been adequately addressed by many manuals, including the new USEPA macroinvertebrate field and laboratory manual (Klemm *et al.*, 1990), the ASTM methods for sampling benthos (ASTM, 1988), Ohio EPA's Field Methods Manual (Ohio EPA, 1989b), Standard Methods (APHA *et al.*, 1989), USEPA's Rapid Bioassessment Protocols (Plafkin *et al.*, 1989), and USEPA's Superfund Field Compendium (USEPA, 1987). The following decisions will need to be made once the sample gear is chosen:

- Whether samples will be picked from debris and sorted in the field;
- Which preservative should be used;
- Whether a stain (rose bengal) will be added to the sample to facilitate separating the organisms from debris;
- Whether the samples need to be shipped and whether they require a chain-of-custody form (as in Superfund samples); and
- The type of sample containers and labeling of the containers required.

Sorting—There are many discussions elsewhere of techniques for sample sorting and preparation of slides for identification. Klemm *et al.* (1990), Merritt *et al.* (1984), Pennack (1978) and APHA *et al.* (1989) offer excellent guidance for sample sorting. Hynes (1970, 1971) stated that the earlier stages of benthos are retained by a 0.2-mm mesh size (approximately the size of a #75 standard sieve), and APHA *et al.* (1989) and Klemm *et al.* (1990) defined the benthos by a mesh size of 0.595 (standard sieve #30), which is now standard practice. However, some types of Chironomidae and other small benthos pass through the #30-mesh sieve but are be retained by the #40-mesh sieve. It is therefore recommended that samples be passed through a #30-mesh sieve and that the materials washed through be passed through a #40-mesh sieve; the material retained in both sieves should then be sorted (Ohio EPA, 1989b). Once the material is washed through the sieves the organisms should be separated from the vegetation and other debris in a white enamel pan. As the materials are separated, the organisms can be placed in different vials for the major taxa.

Taxonomy—The level to which the taxonomy should be taken is dependent on the objectives of the study. For a system reconnaissance or screening survey, it is generally not necessary to go beyond the family level (Hilsenhoff, 1988; Illinois EPA, 1987; Plafkin *et al.*, 1989; Resh, 1988). For studies attempting to identify designated use impairment and/or evaluate impacts from a specific source, the recommended minimum level of taxonomic detail should follow the list by Ohio EPA (1989b). Ohio EPA has successfully implemented numeric biocriteria based on this taxonomic detailing. This strategy is to expend the effort to differentiate those taxa which are better water resource quality indicators and for which taxonomic keys and expertise are readily available. The level of taxonomic detailing must be consistent within the program and applied for each study site. Species-level identifications for all organisms are not necessary for a successful program, and they commonly depend on the availability of local keys. General keys available for genus-level identifications include Merritt and Cummins (1984) for insects, Peckarsky *et al.* (1990) for insects and other invertebrates, Pennack (1978, 1989) for all common invertebrates, Wiederholm (1983) for midges, and Klemm (1985) for annelids (oligochaetes and leeches). Klemm *et al.* (1990) provide an excellent list of taxonomic references for other general and specific uses. Regional U.S. EPA or state biologists should be contacted to determine which of the hundreds of other taxonomic keys are available for specific taxa, both nationally and regionally.

8.2.1.2.3 Types of Data Required

The types of data analyses that are required to meet program objectives directly affect the types of data required. A list of the families of taxa present may be sufficient to meet some program objectives. Under other circumstances, species-level taxonomy and enumerations may be required. The necessary data required to conduct different types of analyses can be obtained from the following discussion of data analysis methods.

One of the most inconsistent and perplexing aspects of a freshwater benthic macroinvertebrate community assessment is the numeric representation and analysis of the data collected. Structural community measures such as richness values, diversity and biotic indexes, and enumerations have been used almost exclusively. Indicator organisms have been used to establish many of the biotic indexes but also have the potential to differentiate among types of impacts. Recently, functional community measures based on feeding groups such as shredder, collector, scraper, and predator (Cummins and Merritt, 1984) have gained wider application and acceptance due to their sensitivity in detecting system perturbation on food resources. Sediment and water quality assessments based on the benthic macroinvertebrate community should use a complementary mix of both structural and functional measures. It is strongly recommended that a multimetric technique be used (Plafkin *et al.*, 1989; Ohio EPA, 1990a) so any *single* index value or observation will not substantially influence the results. Discussions of various data analysis techniques can be found in Hawkes (1979), Cairns (1981), Klemm *et al.* (1990), Washington (1984), and Resh and Jackson (1990).

Composite Indexes—Composite or multimetric indexes combine selected structural or functional measures, or "metrics," in a cumulative scoring system, as was done with the Index of Biotic Integrity (IBI) for the fish community (Karr *et al.*, 1986). These composite, or multimetric, indexes are highly recommended and are among the most used assessment techniques for development of biological criteria for both benthic macroinvertebrates and fish.

Karr and Kerans (1992) provide an outstanding discussion of the process of developing metrics proposed for use in an invertebrate IBI. They evaluated 28 potential metrics for inclusion and have eliminated 10 from further consideration. The metrics fall into three categories: taxa richness and community composition, trophic and functional feeding group, and abundance.

Ohio EPA (1989b, 1990a) successfully developed a similar index for invertebrates using the following 10 structural metrics, adjusted for drainage area size with each ecoregion, to derive a final Invertebrate Community Index (ICI) score:

(1) Total number of taxa;

(2) Total number of mayfly taxa;

(3) Total number of caddisfly taxa;

(4) Total number of dipteran taxa;

(5) Percent mayflies;

(6) Percent caddisflies;

(7) Percent Tribe Tanytarsini midges;

(8) Percent other dipterans and non-insects;

(9) Percent tolerant organisms; and

(10) Total number of qualitative EPT taxa.

The ICI score is part of Ohio EPA's numeric biocriteria for designated use attainment, and it was developed using artificial and natural substrate data for 232 "least-impacted" reference sites. A statistical validation of the ICI using a factor analysis technique showed high correlations between the factor analysis scores and the ICI scores and little redundancy between the metrics (Davis and Lubin, 1989).

U.S. EPA (Plafkin *et al.*, 1989) developed a composite index for rapid assessments in lotic systems using the following two functional and six structural metrics:

(1) Taxa richness;

(2) Modified Hilsenhoff biotic index;

(3) Ratio of scrapers and filtering collectors (functional);

(4) Ratio of EPT and Chironomidae abundances;

(5) Percent contribution of dominant taxon;

(6) EPT index;

(7) Community similarity index; and

(8) Ratio of shredders to total number of organisms (functional).

These Rapid Bioassessment Protocols (RBPs) recommend conducting single-habitat (riffle) dip-net sampling. The scores are based on a percentage of the metric values found at a reference site, rather than comparison of the results based on "optimal" values for each metric. Modifications to the RBPs can include use of multiple reference sites. The RBPs are flexible and can be modified for different geographical locations, as evidenced by the use of different metrics in Arkansas (Shakelford, 1988) and New York (Bode and Novak, 1988). The success of the RBPs is in the use of the composite index for rapid assessments that allows for three levels of taxonomic work (i.e., order, family, or genus/species levels). Order and family taxonomy do not require laboratory taxonomy and may be done in the field. The RBPs normally use single-habitat (riffle) sampling and a 100-organism count in the field. However, they can be adapted for most program uses; for example, by employing multihabitat sampling and/or various count limitations. To be applicable to a state's program, the RBPs should undergo a rigorous validation effort within that state.

Diversity Indexes—When diversity indexes were introduced, they were used widely because of their ability to reduce the complex benthic community measurements into a single value that could be used by nonbiologist decision-makers. Diversity indexes are based on measuring the distribution of the number of individuals among the different taxa, and use methods that result in enumerations by surface area. The most common diversity index used for water quality studies is the Shannon, or Shannon-Wiener Index (Shannon and Weaver, 1949) as shown below:

$$\textit{Shannon's } H' = \sum_{i=1}^{s} \left(\frac{n_i}{n}\right) \ln\left(\frac{n_i}{n}\right)$$

where:

n_i	=	Total number of individuals in the i^{th} taxon
n	=	Total number of individuals
s	=	Total number of taxa.

(Washington (1984) provides a good explanation of how the index derived the name Shannon-Wiener Index rather than Shannon-Weaver Index.) Theoretically, higher community diversity indicates better water quality (Wilhm, 1970). However, low diversity may be caused by factors other than water quality impacts, such as extremes in weather (floods or droughts), poor habitat, or seasonal fluctuations. Although diversity indexes such as the Shannon-Wiener Index still remain in widespread use (Washington, 1984), their limitations in accurately addressing a variety of perturbations has decreased their reliability (Cooke, 1976; Hilsenhoff, 1977; Hughs, 1978; Chadwick and Canton, 1984; Washington, 1984; Mason *et al.*, 1985; Resh, 1988). Kaesler *et al.*, (1978) demonstrated that the popular Shannon's Index was actually not the preferred index for aquatic ecology studies and recommended the use of Brillouin's (1962) Index. Resh (1988) reported that diversity indexes showed varied results in detecting changes in water quality and that they are not the optimal measures of water quality. However, diversity indexes can provide additional information as to the community composition and should be reported if the data are available. Reliance on these indexes as the only, or predominant, measure on which water pollution control decisions are based is not valid. Washington (1984) provides an outstanding review of the history and uses of diversity indexes.

Biotic Indexes—Biotic indexes use pollution tolerance scores for each taxon, weighted by the number of individuals assigned to each tolerance value. If desired, relative abundance measures can be used in biotic indexes. An example of a widely used biotic index (Hilsenhoff, 1977, 1982) is as follows:

$$\textit{Biotic Index} = \sum_{i=1}^{s} \frac{n_i}{n}$$

where:

- n_i = Number of individuals in taxon i
- a_i = Tolerance value assigned to taxon i
- n = Total number of individuals in the sample.

Tolerance values can be found in Hilsenhoff (1987) or can be generated by regional-specific knowledge of the organisms' tolerances. Typical ranges of organism index values are 0-5, 0-10, or 0-11, with the higher numbers indicating greater tolerance to pollutants. Community indexes are generally limited to lotic systems impacted by organic enrichment (Woodiwiss, 1964; Chandler, 1970; Hilsenhoff, 1977; Murphy, 1978; DePauw *et al.*, 1986) or other general perturbations (Hawkes, 1979). Biotic indexes based on a specific population, rather than community, are addressed in the "Indicator Organisms" discussion below. Although the first widely applied biotic index in this country was developed by Beck (1955) for Florida streams, the Hilsenhoff Biotic Index (Hilsenhoff, 1977, 1982) has gained great popularity and has been updated to revise the scoring system from a range of 0-5 to 0-11 (Hilsenhoff, 1987) and to include a family-level biotic index (Hilsenhoff, 1988). Because the biotic indexes rely heavily on known pollution tolerances of the taxa, Washington (1984), Mason *et al.* (1985), and Hawkes (1979) preferred the biotic indexes over the diversity indexes for water quality assessments. The success of the Hilsenhoff Biotic Index prompted use of the index, or modifications of it, in several state programs (e.g., Wisconsin, Illinois, New York, North Carolina) and EPA (Plafkin *et al.*, 1989) programs. Unfortunately, tolerance values are not available for many taxa because they tend not to exhibit water quality preferences, and the assessments are generally limited to organic enrichment. Washington (1984) provides an outstanding review of the history and uses of these indexes.

Indicator Organisms—Indicator organisms have played a key role in the development of biotic indexes for both lotic and lentic systems. One of the first classifications based on indicator organisms was done in the Illinois River by Richardson (1928). Simpson and Bode (1980), Bode and Simpson (1982), and Rae (1989), among many others, used Chironomidae as indicator organisms for a variety of toxicants in stream systems. Hawkes (1979) provides an excellent review of the use of benthic macroinvertebrates for stream quality assessments, and Wiederholm (1980) does the same for lake systems. Data analyses for benthic macroinvertebrates in lentic systems have not been as progressive as those in lotic systems with regard to composite indexes and have relied extensively on enumerations, diversity indexes, richness values, and indicator organisms (Fitchko, 1986). Howmiller and Scott (1977), Krieger (1984), and Lauritsen *et al.* (1985) used oligochaete communities to establish a Great Lakes trophic index. Lafont (1984) also used oligochaetes to indicate fine sediment pollution. Brinkhurst *et al.* (1968) and Winnell and White (1985) used chironomids to develop a similar index for the Great Lakes, and Courtemanch (1987) classified Maine lakes using chironomid larvae similar to the studies of Saether (1979) and

Aagaard (1986) in European lakes. Hart and Fuller (1974) presented pollution ecology data for a number of freshwater benthic macroinvertebrates, as did U.S. EPA's pollution tolerance information series on Chironomidae (Beck, 1977), Trichoptera (caddisflies) (Harris and Lawrence, 1978), Ephemeroptera (mayflies) (Hubbard and Peters, 1978), and Plecoptera (stoneflies) (Surdick and Gaufin, 1978). Washington (1984) also reviewed population-based biotic indexes.

Richness Measures—Richness measures are based on the presence or absence of selected taxa. Commonly used measures include the total number of taxa, the number of EPT (Ephemeroptera, Plecoptera, and Trichoptera), and the number of families. The higher the richness value is, the better the quality of the system. Richness measures have been shown to have low variability and high accuracy in identifying impact (Resh, 1988) and should be applied in each study.

Enumerations—Enumerations involve obtaining a sample quantified by surface area to obtain specific abundances of each taxon. Examples include the number of total individuals, number of EPT individuals, ratios of number of individuals within a taxon to the total number of individuals (Ohio EPA, 1989a; Resh, 1988), and ratios of the number of individuals within one taxonomic group (e.g., EPT) to the number of individuals within another taxonomic group (e.g., Chironomidae) (Plafkin *et al.*, 1989; Resh, 1988). Interpretation of the enumeration ratios can be difficult without prior validation. Most possible enumerations comparing individual taxa to the total number of individuals are done for many studies, although the results may not be presented. The percent contribution of the individuals within a taxon at a sample site can be compared with the percent contribution at the reference sites to detect a change in community structure. Resh (1988) concluded that the seven common enumerations he tested had extremely high variability and unacceptably low accuracy in detecting various impacts, and he suggested that they are not as useful for detecting environmental change as richness measures or the family biotic index. Although the measures Resh (1988) used may not be optimal for widespread use, they may still provide insight into changes in the community structure. Ohio EPA (1989a) has successfully used enumerations for the percentage of mayflies, caddisflies, Tanytarsini midges, tolerant organisms, and "other" dipterans combined with non-insect individuals as a basis for their state biocriteria.

Similarity Indexes—Community similarity indexes measure the similarity between benthic communities at a reference and a study site, with high similarity indicating little change, or impact, between the two sites. The use of similarity indexes has been reviewed by Brock (1977) and Washington (1984). The simplest indexes to apply are those which use only the types of taxa found, not the abundance of the organisms within each taxon. The Jaccard Index (1908) and Van Horn's Index (1950) are examples of the simpler indexes. Van Horn's Index, used by Ohio EPA (1989b), is as follows:

$$\textit{Similarity}\ (c) = \frac{2w}{(a+b)}$$

where:

a	=	Number of taxa collected at one site
b	=	Number of taxa collected at the other site
w	=	Number of taxa common to both stations.

A value over 6.5 or 7.0 indicates good similarity. Plafkin *et al.* (1989) use the Jaccard Index in the rapid bioassessment protocols (RBPs). Other indexes such as the percent similarity (Brock 1977) and the Bray-Curtis (1957) use the abundance of organisms.

Functional Information—Community function measurements based on habitat, trophic structure, and other ecological measures were described by Kaesler *et al.* (1978) and used by Rooke and

Mackie (1982a) as the "ecological community analysis" (ECA). Rooke and Mackie (1982b) reported the ECA to provide more information on environmental quality than diversity or biotic indexes, but the ECA was very time-consuming and not practical for rapid assessments. However, Cummins and Wilzbach (1985) and Cummins (1988) describe a rapid assessment method based on sampling coarse particulate organic matter and determining the functional feeding groups described in Merritt and Cummins (1984). This method is recommended in EPA's RBPs (Plafkin *et al.*, 1989). Rabeni *et al.* (1985) also described the usefulness of a functional feeding group approach to provide a "more ecologically sound classification of water quality" during their development of a biotic index for paper mill impacts. Another useful measure of function is observations of the incidence of morphological deformities in benthic macroinvertebrates, similar to the observations made for Karr's Index of Biotic Integrity (IBI) for fish (Karr *et al.*, 1986). Deformities have been associated with exposure of metals and organic compounds to Chironomidae (Cushman, 1984; Cushman and Goyert, 1984; Wiederholm, 1984b; Warwick, 1985; Warwick *et al.*, 1987) and Trichoptera (Simpson, 1980; Petersen and Petersen, 1983). Karr and Kerans (1992) are developing an invertebrate IBI and have evaluated 10 trophic and functional feeding group metrics. This promising work is continuing.

Statistical Approaches—Various statistical approaches have been applied to determine whether the benthic community at a study site varies from that at a reference or other site. An excellent discussion of this issue appears in Klemm *et al.* (1990) and USEPA (1992b). Depending on the chosen endpoints of the study, rigorous statistical analysis may not be necessary. For instance, if the endpoint is the number of taxa or richness measures, the variability is generally quite low and accuracy quite high. In this case, the differences between two communities would need to be evaluated based on study objectives. A "statistical" difference between two communities will not always indicate whether more subtle changes in community composition are occurring or whether mitigation may be warranted before a statistical change occurs. Sometimes when that change occurs, it is too late to protect the community. USEPA (1992b) has an outstanding discussion on applying uncertainty to decision-making. The same data evaluation procedures apply to both the marine and freshwater systems. The reader is referred to the statistical discussion in Chapter 9 (marine benthic community structure).

Bivariate and multivariate analysis are often applied in benthic studies to define relationships between and among variables. Examples of these analyses include analysis of variance (ANOVA), correlations, regressions (including multiple regressions), and the two-sample t-test. A major drawback to these methods is the assumption that the data follow a statistical distribution such as a normal or lognormal distribution. This assumption is often invalid when dealing with biological populations and communities.

Alternatively, nonparametric analyses may be conducted. Such analyses are not based on assumptions about a specific distribution of the data. Examples of such tests include the chi-square test, binomial tests, rank correlations, or tests comparable to the t-test such as the Mann-Whitney test. Whichever statistical methods are employed, all data assumptions must be clearly stated and objectives known.

8.2.1.2.4 Necessary Hardware and Skills

The hardware needed for field collection includes samplers (e.g., dredges, dip-nets), sieves, benthic macroinvertebrate containers, forceps, white enamel pans, ethanol preservative, and appropriate personal gear (e.g., hip boots or chest-waders, life vest if needed, and first aid kits). For the laboratory, standard biological laboratory equipment should be available, such as microscopes (both dissecting and compound), forceps, microscope slides and cover slips, ethanol, potassium hydroxide, mounting media, and sieves. A personal computer (containing a 20-MB or larger hard drive) is important for storing and analyzing the data.

Trained benthic macroinvertebrate field biologists and taxonomists are needed for benthic community assessments. At least one should be proficient at identifications beyond the family level. That taxonomist should remain involved until the proficiency of the identifier in reaching family-level identifications is ensured. A minimum of a Master of Science degree in a related discipline is usually required for the taxonomist to have learned the necessary skills. However, adequate training is commonly available through taxonomy courses and workshops that can provide the necessary proficiency without an advanced degree. A demonstration of proficiency by accurately identifying a check sample prepared by U.S. EPA or a state agency is important. A trained benthic ecologist is necessary to compile and interpret the data. Although it would be ideal if the benthic ecologist had a rigorous statistical background, consultation with a statistician should be adequate.

8.2.1.3 Adequacy of Documentation

There is ample documentation of both field methods and analytical techniques. The *Journal of the North American Benthological Society* is a prime source of this information, as is technical exchange at professional meetings. Furthermore, there is a large volume of published and unpublished material that documents use of this method (USEPA 1992d, 1991e, 1990g, 1989f, 1988a).

8.2.2 Applicability of Method to Human Health, Aquatic Life, or Wildlife Protection

This method is directly applicable to the protection of aquatic life since it is based on direct measurements of benthic macroinvertebrates. This method is directly applicable to the protection of those aquatic organisms (e.g., fish) and wildlife that directly feed on benthic macroinvertebrates (e.g., small mammals and wading shorebirds). It is indirectly applicable to other wildlife that depend on benthos at other levels in the food chain. This method is also indirectly applicable to the protection of human health since benthic macroinvertebrates can serve as indicators of toxicant impacts that may affect humans via bioaccumulation pathways

8.2.3 Ability of Method to Generate Numerical Criteria for Specific Chemicals

This method is used in conjunction with sediment toxicity and chemistry data to characterize toxicant impacts and assist with determining the appropriate levels at which the toxicants should be controlled. By itself, however, this method would not be used to generate chemical-specific criteria.

8.3 USEFULNESS

8.3.1 Environmental Applicability

Benthic macroinvertebrates have been routinely used to assess environmental quality in a variety of geographical areas and ecoregions, as was discussed in Section 8.1.

8.3.1.1 Suitability for Different Sediment Types

Assessment of the freshwater benthic macroinvertebrate community structure is well suited for evaluating different sediment types since the benthos inhabit all substrates (Merrit and Cummins, 1984). Comparisons should be made among benthic communities of similar substrate since different types and numbers of organisms will inhabit different types of substrates.

8.3.1.2 Suitability for Different Chemicals or Classes of Chemicals

Benthic macroinvertebrate communities are routinely used to assess potential impacts caused by many different chemicals or classes of chemicals. In addition to the uses described in Section 8.1.1.1 of this chapter, many benthic organisms are used to indicate stresses from specific chemicals or classes of chemicals (Brinkhurst *et al.*, 1968; Hart and Fuller, 1974; Saether, 1979; Simpson and Bode, 1980; Wiederholm, 1980; Bode and Simpson, unpublished; Winnell and White, 1985; Aagaard, 1986; and Fitchko, 1986).

8.3.1.3 Suitability for Predicting Effects on Different Organisms

The use of benthic macroinvertebrates as indicator organisms has already been discussed. Benthic macroinvertebrates can be used to predict the effects on other aquatic organisms because if the benthic macroinvertebrate community is impacted, then the impact is likely to be, or already has been, detrimental to other organisms.

8.3.1.4 Suitability for In-Place Pollutant Control

Benthic macroinvertebrates will best indicate in-place pollutant control needs through a site-specific knowledge of surface water quality, habitat quality, and sediment chemistry and toxicity. Alone, the benthic macroinvertebrates can be used to screen for potential sources of sediment contamination based on spatial gradients in community structure, but they should not be used alone to definitively determine sediment quality or to develop chemical-specific guidelines. The benthic data must be integrated with other available data to determine sediment quality using a "weight-of-evidence" approach.

8.3.1.5 Suitability for Source Control

Benthic macroinvertebrates have been extensively used for source characterization and control in many of the state and U.S. EPA monitoring programs involving spatial surveys upstream and downstream of suspected sources (Ohio EPA, 1987; Bode and Novak, 1988; Courtemanch and Davies, 1988; Fiske, 1988; Maret, 1988; Penrose and Overton, 1988; Shakelford, 1988; USEPA, 1991c, 1988a, 1988b; Fandrei, 1989). If a detrimental change is detected in the benthic macroinvertebrate community and that change can be attributable to a source, then control measures can be implemented through the NPDES permit program. Many states aggressively pursue this action.

8.3.1.6 Suitability for Disposal Applications

The discussion presented in Section 9.3.1.6 of Chapter 9 (marine benthic macroinvertebrate community structure) is applicable to fresh water. Recently benthic community assessments have been required by U.S. EPA Region V, as stated in the *Draft Interim Guidance for the Design and Execution of Sediment Sampling Efforts Relating to Navigational Maintenance Dredging in Region V - May 1989* (USEPA, 1989d). In this guidance, benthic macroinvertebrate assessments are advised for areas that are suitable for open-lake disposal or for sediments that are difficult to characterize. All benthic community assessments will be made in concert with sediment chemistry and toxicity evaluations.

8.3.2 General Advantages and Limitations

The advantage of using the benthic macroinvertebrates community assessment approach to determining sediment quality is that it provides an economical and accurate indication of the health of the system under study, and it is based on direct observation rather than theoretically derived data. The major limitation is the difficulty in relating the findings to the presence of individual chemicals and specific concentrations of those chemicals for numeric in-place pollutant management. This method should be integrated with sediment chemistry and toxicity information.

8.3.2.1 Ease of Use

The equipment requirements for benthic surveys is minimal and inexpensive compared to those for chemical/physical analyses or even toxicity tests. The organisms are easy to obtain, but difficult to sort and identity. All materials needed for benthic assessments are easily obtained through chemical and biological supply companies and require no special mechanical setup or calibration.

8.3.2.2 Relative Cost

The cost for benthic macroinvertebrate assessments is economical compared to that for chemistry or toxicological evaluations. Ohio EPA (1990a) provided a cost of about $700 to conduct a benthic assessment at one sample site. However, this cost included overhead (e.g., rent, office equipment), all travel expenses, time spent in the field, and report preparation. Ohio EPA conducts artificial substrate (composite of five substrates) sampling in addition to natural substrate (multi-habitat) sampling at each site. Their cost of $1,099 ($824 for artificial substrates and $275 for qualitative samples) was quite economical compared to chemical/physical testing ($1,653) or bioassay testing ($3,000 to $12,000) for each site. Plafkin *et al.* (1989) discussed staff requirements for sample collection and analysis.

The most expensive items are the samplers and the microscopes to identify the organisms. However, most state programs and contractors have this equipment available for other program needs. The fieldwork can be conducted during the time it takes to collect a sediment sample. The most time-consuming aspect is the laboratory sorting and identifications, which may average 11 hours per site. However, this process compares favorably with the amount of time required to set up and run a toxicity test or to prepare and analyze chemical variables.

8.3.2.3 Tendency to Be Conservative

The benthic macroinvertebrate community assessment provides a conservative measure, since the community is responding to both temporal and spatial perturbations. There are few chances, if any, of obtaining a result indicating a high-quality community when an impact occurs. Because of influences other than sediment/water quality, it is more common to observe an impacted community when there is no sediment/water quality impact. Although the primary focus is on community-level information, changes in individual populations could also be addressed. However, the ecological significance of population changes may not be evident until the community is affected.

In a review of surface water chemistry and benthic macroinvertebrate community assessments over 800 water body segment sites in Ohio, biocriteria based on benthic macroinvertebrates were more sensitive (conservative) indicators of water quality (Ohio EPA, 1990b). In 49.5 percent of the segments, the benthic and fish assessment revealed impacts not detected by chemical water quality standards violations. In 47.4 percent of the sites, the chemical and biological assessment supported one another. Only 2.8 percent of the sites did not have a biological impact when the chemistry indicated that there would be one.

8.3.2.4 Level of Acceptance

Benthic macroinvertebrate community assessments of sediment/water quality have been used in freshwater systems since the early 1900s (Richardson, 1928). Most of the methods employed today have been widely accepted for use, although the use of function measurements is not as well documented. Perhaps the single most important demonstration of the level of acceptance of benthic assessments is the growing regulatory use and establishment of numerical biological criteria in state water quality standards.

8.3.2.5 Ability to Be Implemented by Laboratories with Typical Equipment and Handling Facilities

The only special pieces of equipment required are the samplers and sieves, which are easily obtained from biological supply warehouses. Most biological laboratories will have dissecting and compound microscopes, chemical reagents, microscope slides and cover slips, forceps, and any other materials needed. The laboratory's capability to identify benthic macroinvertebrates is less common. Taxonomy is not a widespread skill and is more likely to be found in consulting firms than in analytical laboratories.

8.3.2.6 Level of Effort Required to Generate Results

Depending on the study objectives and level of effort needed, results can be generated in written form in as little as 1 day (Plafkin *et al.* 1989) or in several months. For example, Ohio EPA processes over 500 individual benthic samples each year, identifies the organisms, and prepares reports for regulatory use in less than 1 year, with fewer than three full-time employees in their benthic macroinvertebrate unit. The critical period is the turnaround time for the taxonomy.

With artificial substrates, an additional 6-week colonization period is required; unless a rapid assessment or moderate sized study is done, a written report including interpretation of results will typically require between 6 months and 1 year.

8.3.2.7 *Degree to Which Results Lend Themselves to Interpretation*

It is never advisable to have an individual without training in benthic ecology interpret benthic data. Once the benthic ecologist provides a report with recommendations, the results can be easily implemented into a management strategy. Although several numerical indexes that appear simple to use are available, data interpretation relies on all of the information generated for a study, including chemical, physical, and toxicological measurements, as well as indicator organisms and function measures.

8.3.2.8 *Degree of Environmental Applicability*

Benthic macroinvertebrate community structure and function is used extensively to evaluate sediment and water quality and characterize impacts in lotic and lentic freshwater ecosystems.

8.3.2.9 *Degree of Accuracy and Precision*

Since benthic macroinvertebrates are measured directly, this method is highly accurate for characterizing sediment/water quality effects on aquatic life. There is little chance, if any, that a high-quality community will be indicated when an impact actually occurs (Type II error with a null hypothesis of no community change). Because of influences other than sediment/water quality, it is more common to observe an impacted community when there is no sediment/quality impact (Type I error with a null hypothesis of no community change). For environmental pollution control, a Type II error is much more serious than a Type I error, which is conservative. To reduce the possibility of a Type II error, the data (including chemistry and toxicity) must be interpreted by a trained benthic ecologist. Resh (1988) and USEPA (1992b) reviewed the levels of accuracy and precision for several of the data analysis techniques.

To ensure as much accuracy and precision in the data as possible, a detailed Quality Assurance Program Plan should be established and followed. Careful and consistent field and laboratory protocols are necessary. It is also necessary to sample during optimal conditions, which can minimize the effects of natural variations in the data. However, the natural variability, especially seasonal, is reduced when using a community-level approach rather than a population-level approach.

8.4 STATUS

Sections 8.1.1 (Current Uses) and 8.3 (Usefulness) describe the status of the discipline.

8.4.1 Extent of Use

This method is widely used in both regulatory and nonregulatory sediment and water quality programs. It has been used to assess impacts due to organic enrichment and a variety of chemical classes in both lotic and lentic systems. Benthic macroinvertebrate community assessments are the most widely used instream biological measures in state water quality programs.

8.4.2 Extent to Which Approach Has Been Field-Validated

Since it is an *in situ* study, field validation occurs when the approach can consistently and accurately assess environmental quality. Most benthic studies employ reference stations and rely on other environmental data to validate the method. The documentation provided in this paper should present adequate documentation of the method's validity.

8.4.3 Reasons for Limited Use

Benthic macroinvertebrate community assessments are very common in freshwater systems because of their relatively low cost and high information output.

8.4.4 Outlook for Future Use and Amount of Development Yet Needed

The outlook for the future use of benthic macroinvertebrate community structure and function in sediment quality assessment is very good because of the recognition that benthic macroinvertebrates provide substantial information that the chemistry and toxicity data alone cannot provide. With the Clean Water Act mandate to maintain and restore biological integrity, benthic community assessments can help determine whether sediment quality is impairing the designated uses and biotic integrity. With the increasing reliance on numerical biocriteria, additional sediment quality problems will be identified. The area where development is most needed is in combining benthic community assessments with chemical and toxicological data in an integrated approach for assessing sediment quality. In addition, the functional measures, which also hold much promise for sediment assessments, need to be validated more thoroughly.

8.5 REFERENCES

Aagaard, K. 1986. The Chironomidae fauna of north Norwegian lakes with a discussion of community classification. Hol. Ecol. 9:1-12.

Abe, J., Davis, W., Flanigan, T., Schwarz, A., and M. McCarthy. 1992. Environmental indicators for surface water quality programs - pilot study. EPA-905/R-92/001. U.S. Environmental Protection Agency, Chicago, IL.

APHA *et al.* 1989. Standard methods for the examination of water and wastewater. 17th ed. American Public Health Association, American Water Works Association, and the Water Pollution Control Federation, Washington, DC.

Arkansas DPCE. 1987. Physical, chemical and biological characteristics of least-disturbed reference streams in Arkansas' ecoregions. Arkansas Department of Pollution Control and Ecology.

Armitage, P.D., and J.H. Blackburn. 1985. Chironomidae in a pennine stream system receiving mine drainage and organic enrichment. Hydrobiologia 121:165-172.

ASTM. 1988. Annual book of ASTM standards: water and environmental technology. Vol. 11.04. American Society for Testing and Materials, Philadelphia, PA. 963 pp.

Beck, W.M., Jr. 1955. Suggested method for reporting biotic data. Sew. Ind. Wastes 27:-1193-1197.

Beck, W.M., Jr. 1977. Environmental requirements and pollution tolerance of common freshwater Chironomidae. EPA-600/4-77/024. U.S. Environmental Protection Agency, Office of Research and Development, Cincinnati, OH.

Beckett, D.C., and M.C. Miller. 1982. Macroinvertebrate colonization of multiplate samplers in the Ohio River: the effect of dams. Can. J. Fish. Aquat. Sci. 39:1622-1627.

Bode, R.W. 1988. Quality assurance workplan for biological stream monitoring in New York State. Stream Biomonitoring Unit, Bureau8 of Monitoring and Assessment, Division of Water, New York State Department of Environmental Conservation, Albany, N.Y. 58 pp.

Bode, R.W., and M.A. Novak. 1988. Proposed biological criteria for New York State streams. pp. 42-48. In: Proceedings of the First National Workshop on Biological Criteria - Lincolnwood, Illinois, December 2-4, 1987. T.P. Simon, L.L. Holst, and L.J. Shepard (eds). EPA-905/9-89/003. U.S. EPA Region 5 Instream Biocriteria and Ecological Assessment Committee, Chicago, IL. 129 pp.

Bode, R.W., and K.W. Simpson. 1982. Communities of Chironomidae in large lotic systems: impacted vs. unimpacted. Unpublished paper presented at the 30th Annual Meeting of the North American Benthological Society in Ann Arbor, MI, May 18, 1982. 15 pp.

Bray, J.R., and J.T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecol. Monogr. 27:325-349.

Brillouin, L. 1962. Science and information theory. Academic Press, New York, NY. pp. 1-347.

Brinkhurst, R.O., A.L. Hamilton, and H.B. Herrington. 1968. Components of the bottom fauna of the St. Lawrence Great Lakes. Great Lakes Inst. Report 33. University of Toronto, Toronto, Canada. 50 pp.

Brock, D.A. 1977. Comparison of community similarity indices. J. Wat. Pollut. Control Fed. 49:2488-2494.

Cairns, J., Jr. 1981. Introduction to biological monitoring. pp. 375-409. In: Water Quality Management: The Modern Analytical Techniques. H.B. Mark, Jr., and J.S. Mattson (eds.). Marcel Dekker, Inc. New York, NY.

Chadwick, J.W., and S.P. Canton. 1983. Comparison of multiplate and surber samplers in a Colorado mountain stream. J. Freshwater Ecol. 2:287-292.

Chadwick, J.W., and S.P. Canton. 1984. Inadequacy of diversity indices in discerning metal mine drainage effects on a stream invertebrate community. Wat. Air Soil Pollut. 22:217-223.

Chandler, J.R. 1970. A biological approach to water quality management. J. Wat. Pollut. Control Fed. 4:415-422.

Cook, D.G., and M.G. Johnson. 1974. Benthic macroinvertebrates of the St. Lawrence Great Lakes. J. Fish. Res. Board Can. 31:763-782.

Cooke, S.E.K. 1976. Quest for an index of community structure sensitive to water pollution. Environ. Pollut. 11:269-288.

Courtemanch, D.L. 1987. Trophic classification of Maine lakes using benthic Chironomidae fauna. Paper presented at the 7th International Symposium of North American Lake Management Society, Orlando, FL. 20 pp.

Courtemanch, D.L., and S.P. Davies. 1988. Implementation of biological standards and criteria in Maine's water classification law. pp. 4-9. In: Proceedings of the First National Workshop on Biological Criteria - Lincolnwood, Illinois, December 2-4, 1987. T.P. Simon, L.L. Holst, and L.J. Shepard (eds.). EPA-905/9-89/003. U.S. EPA Region 5 Instream Biocriteria and Ecological Assessment Committee, Chicago, IL. 129 pp.

Crossman, J.S., and J. Cairns. 1974. A comparative study between two different artificial substrate samplers and regular sampling techniques. Hydrobiologia 44:517-522.

Crossman, J.S., J.R. Wright, and R.L. Kaesler. 1984. Consolidation of baseline information, development of methodology, and investigation of thermal impacts on freshwater shellfish, insects, and other biota. EPA-600/7-84/042. Prepared by Tennessee Valley Authority for U.S. EPA Office of Research and Development, Washington, DC. 159 pp.

Cummins, K.W. 1988. Rapid bioassessment using functional analysis of running water invertebrates. pp. 49-54. In: Proceedings of the First National Workshop on Biological Criteria - Lincolnwood, Illinois, December 2-4, 1987. T.P. Simon, L.L. Holst, and L.J. Shepard (eds.). EPA-905/9-89/003. U.S. EPA Region 5 Instream Biocriteria and Ecological Assessment Committee, Chicago, IL. 129 pp.

Cummins, K.W., and M.A. Wilzbach. 1985. Field procedures for analysis of functional feeding groups of stream macroinvertebrates. Contr. 1611 to Appalachian Environmental Research Laboratory. University of Maryland, Frostburg, MD. 21 pp.

Cushman, R.M. 1984. Chironomid deformities as indicators of pollution from a synthetic, coal-derived oil. Freshwater Biology 14:179-182.

Cushman, R.M., and J.C. Goyert. 1984. Effects of a synthetic crude oil on pond benthic insects. Environ. Pollut. (Ser. A) 33:163-186.

Davis, W.S. 1990. Forward: A Historical Perspective on Regulatory Biology. In: Proceedings of the 1990 Midwest Pollution Control Biologists Meeting, Chicago, Illinois, April 10-13, 1990. pp. i-xii. W.S. Davis (ed.). USEPA Region V, Environmental Sciences Division, Chicago, IL. EPA-905/9-90-005.

Davis, W.S., and T.P. Simon. 1988. Sampling and data evaluation requirements for fish and macroinvertebrate communities. pp. 89-97. In: Proceedings of the First National Workshop on Biological Criteria - Lincolnwood, Illinois, December 2-4, 1987. T.P. Simon, L.L. Holst, and L.J. Shepard (eds.). EPA-905/9-89/003. U.S. EPA Region 5 Instream Biocriteria and Ecological Assessment Committee, Chicago, IL. 129 pp.

Davis, W.S., and A.L. Lubin. 1989. A statistical validation of Ohio EPA's invertebrate community index. Draft. Paper presented at the First Midwest Pollution Control Biologists Meeting, U.S. EPA Region V, February 14-17, 1989, Chicago, IL. 15 pp.

Denbow, T.A., and W.S. Davis. 1986. Highway runoff water quality training course student workbook. Chapter 7. In: Water Quality Impacts. U.S. Dept. Transportation, Federal Highway Administration, McLean, VA.

DePauw, N., D. Roels, and A.P. Fontoura. 1986. Use of artificial substrates for standardized sampling of macroinvertebrates in the assessment of water quality by the Belgian biotic index. Hydrobiologia 133:237-258.

Dupuis, T.V., P. Bertram, J. Meyer, M. Smith, N. Kobriger, and J. Kaster. 1985. Effects of highway turnoff on receiving waters. Volume II: Results of field monitoring program. Prepared by Rexnord for the Federal Highway Administration, McLean, VA.

Fandrei, G. 1989. Personal Communication. Minnesota Pollution Control Agency, St. Paul, MN.

Fiske, S. 1988. The use of biosurvey data in the regulation of permitted nonpoint discharges in Vermont. pp. 67-74. In: Proceedings of the First National Workshop on Biological Criteria - Lincolnwood, Illinois, December 2-4, 1987. T.P. Simon, L.L. Holst, and L.J. Shepard (eds.). EPA-905/9-89/003. U.S. EPA Region 5 Instream Biocriteria and Ecological Assessment Committee, Chicago, IL. 129 pp.

Fitchko, J. 1986. Literature review of the effects of persistent toxic substances on Great Lakes biota. Report of the Health of Aquatic Communities Task Force, International Joint Commission, Windsor, Ontario. 256 pp.

Gallant, A.L., T.R. Whittier, D.P. Larsen, J.M. Omernick, and R.M. Hughs. 1989. Regionalization as a tool for managing environmental resources. U.S. Environmental Protection Agency, Environmental Research Laboratory, Corvallis, OR, EPA/600/3-89/060. 151 pp.

Gaufin, A.R., and C.M. Tarzwell. 1952. Aquatic invertebrates as indicators of stream pollution. Pub. Health. Report 67:57.

Green, R.H. 1978 Optimal impact study design and analysis. pp. 3-28. In: Biological Data in Water Pollution Assessment: Quantitative and Statistical Analyses. K.L. Dickson, J. Cairns, Jr., and R.L. Livingston (eds.). ASTM STP 652. American Society for Testing and Materials, Philadelphia. PA.

Harris, T.L., and T.M. Lawrence. 1978. Environmental requirements and pollution tolerance of Trichoptera. EPA-600/4-78/063. U.S. Environmental Protection Agency, Office of Research and Development, Cincinnati, OH.

Hart, C.W., Jr., and S.L.H. Fuller (eds.). 1974. Pollution ecology of freshwater invertebrates. Academic Press, Inc. London. 389 pp.

Hawkes, H.A. 1979. Invertebrates as indicators of river water quality. Chapter 2, pp. 1-45. In: Biological Indices of Water Quality. A. James, and L. Evison (eds). John Wiley and Sons, New York, NY.

Hester, F.E., and J.B. Dendy. 1962. A multiplate sampler for aquatic macroinvertebrates. Trans. Amer. Fish. Soc. 91:420.

Hilsenhoff, W.L. 1977. Use of arthropods to evaluate water quality of streams. Technical Bulletin No. 100. Wisconsin Department of Natural Resources, Madison, WI. 15 pp.

Hilsenhoff, W.L. 1982. Using a biotic index to evaluate water quality in streams. Technical Bulletin No. 132. Wisconsin Department of Natural Resources, Madison, WI. 23 pp.

Hilsenhoff, W.L. 1987. An improved biotic index of organic stream pollution. Great Lakes Entomologist 20:31-39.

Hilsenhoff, W.L. 1988. Rapid field assessment of organic pollution with a family-level biotic index. J. N. Am. Benthol. Soc. 7:65-68.

Hirsch, R.M., W.M Alley, and W.G. Wiber. 1988. Concepts for a national water quality assessment program. U.S. Geological Survey circular 1021. U.S. Department of the Interior.

Hite, R.L. 1988. Overview of stream quality assessments and stream classification in Illinois. pp. 98-125. In: Proceedings of zthe First National Workshop on Biological Criteria - Lincolnwood, Illinois, December 2-4, 1987. T.P. Simon, L.L. Holst, and L.J. Shepard (eds.). EPA-905/9-89/003. U.S. EPA

Region 5 Instream Biocriteria and Ecological Assessment Committee, Chicago, IL. 129 pp.

Howmiller, R.P., and M.A. Scott. 1977. An environmental index based on relative abundance of oligochaete species. J. Wat. Pollut. Control Fed. 49:809-815.

Hubbard, M.D., and W.L. Peters. 1978. Environmental requirements and pollution tolerance of Ephemeroptera. EPA-600/4-78/061. U.S. Environmental Protection Agency, Office of Research and Development, Cincinnati, OH.

Hughs, B.D. 1978. The influence of factors other than pollution on the values of Shannon's diversity index for benthic macroinvertebrates in streams. Wat. Res. 12:359-364.

Hughs, R.M., T.R. Whittier, C. M. Rohm, and D.P. Larsen. 1990. A regional framework for establishing recovery criteria. Environmental Management, 14(5): 673-683.

Hynes, H.B.N. 1970. The ecology of running waters. University of Toronto Press. 555 pp.

Hynes, H.B.N. 1971. Benthos of flowing water. pp. 66-80. In: Secondary Productivity in Freshwaters. W.T. Edmondson, and G.C. Winberg (eds.). IBP Handbook No. 17. Blackwell Scientific Publ., Oxford, U.K.

Illinois EPA. 1987. Quality assurance and field methods manual. Section C. Macroinvertebrate monitoring. Illinois Environmental Protection Agency, Division of Water Pollution Control, Springfield, IL.

Jaccard, P. 1908. Nouvelles recherches sur la distribution florale. Bull. Soc. Vaud. Sci. Nat. XLIV(163):223-269.

Kaesler, R.L., E.E. Herricks, and J.J. Crossman. 1978. Use of indices of diversity and hierarchial diversity in stream surveys. pp. 92-112. In: Biological Data in Water Pollution Assessment: Quantitative and Statistical Analyses. K.L. Dickson, J. Cairns, Jr., and R.L. Livingston (eds.). ASTM STP 652. American Society for Testing and Materials, Philadelphia, PA.

Karr, J.R. and B.L. Kerans. 1992. Components of biological integrity: Their definition and use in development of an invertebrate IBI. In: T. Simon and W. Davis (eds.). Proceedings of the 1991 Midwest Pollution Control Biologists Meeting. pp. 1-16. EPA-905/R-92/003. U.S. EPA Region 5, Chicago, IL.

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

Klemm, D.J. 1985. A guide to the freshwater Annelida (Polychaeta, Naidid and Tubificid Oligochaeta, and Hirudinea) of North America. Kendall/Hunt Publ., Dubuque, IA. 198 pp.

Klemm, D.J., P.A. Lewis, F. Fulk, J.M. Lazorchak. 1990. Macroinvertebrate Field and laboratory methods for evaluating the biological integrity of surface waters. U.S. Environmental Protective Agency, Office of Research and Development, EPA/600/4-

Krieger, K.A. 1984. Benthic macroinvertebrates as indicators of environmental degradation in the southern nearshore zone of the central basin of Lake Erie. J. Great Lakes Res. 10:197-209.

Lafont, M. 1984. Oligochaete communities as biological descriptors of pollution in the fine sediments of rivers. Hydrobiologia 115:127-129.

Larsson, P. 1984. Transport of PCBs from aquatic to terrestrial environments by emerging chironomids. Environ. Pollut. (Ser. A) 34:283-289.

Lauritsen, D.D., S.C. Mozley, and D.S. White. 1985. Distribution of oligochaetes in Lake Michigan and comments on their use as indices of pollution. J. Great Lakes Res. 11:67-76.

Leahy, P.P., Rosenshein, J.S., and D.S. Knopman. 1990. Implementation plan for the National Water Quality Assessment Program. U.S. Geological Survey Open File Report 90-174. Reston, Virginia.

Lenat, D.R. 1988. Water quality assessment of streams using a qualitative collection method for benthic macroinvertebrates. J. N. Am. Benthol. Soc. 7:222-233.

Maret, T. 1988. A stream inventory process to classify use support and develop biological standards in Nebraska. pp. 55-66. In: Proceedings of the First National Workshop on Biological Criteria - Lincolnwood, Illinois, December 2-4, 1987. T.P. Simon, L.L. Holst, and L.J. Shepard (eds.). EPA-905/9-89/003. U.S. EPA Region 5 Instream Biocriteria and Ecological Assessment Committee, Chicago, IL. 129 pp.

Mason, W.T., P.A., Lewis, and C.I. Weber. 1985. An evaluation of benthic macroinvertebrate biomass methodology. Part 2. Field assessment and data evaluation. Environ. Monitor. Assess. 5:399-422.

Merritt, R.W., and K.W. Cummins (eds). 1984. An introduction to the aquatic insects of North America. 2d ed. Kendall/Hunt Publ., Dubuque, IA. 441 pp.

Merritt, R.W., K.W. Cummins, and V.H. Resh. 1984. Collection, sampling, and rearing methods for aquatic insects. pp. 11-26. In: R.W. Merritt, and K.W. Cummins (eds.). An Introduction to the Aquatic Insects of North America. 2d ed. Kendall/Hunt Publ., Dubuque, IA. 90/030, 256 pp.

Millard, S.P., and D.P. Lettenmaier. 1986. Optimal design of biological sampling programs using the analysis of variance. Est. Coast. Shelf Sci. 22:637-656.

Moore, J.W., V.A. Beaubien, and D.J. Sutherland. 1979. Comparative effects of sediment and water contamination on benthic invertebrates in four lakes. Bull. Environ. Contam. Toxicol. 23:840-847.

Mozley, S.C. 1978. Effects of experimental oil spills on Chironomidae in Alaska tundra ponds. Verh. Internat. Verein. Limnol. 20:1941-1945.

Mozley, S.C., and M.G. Butler. 1978. Effects of crude oil on aquatic insects of tundra ponds. Arctic 31:229-241.

Murphy, P.M. 1978. The temporal variability in biotic indices. Environ. Poll. 17:227-236.

Ohio EPA. 1987. Biological criteria for the protection of aquatic life: Volume I. The role of biological data in water quality assessment. Ohio Environmental Protection Agency, Division of Water Quality Monitoring and Assessment, Surface Water Section, Columbus, OH. 44 pp.

Ohio EPA. 1989a. Biological criteria for the protection of aquatic life. Volume II. Users manual for biological field assessment of Ohio surface waters. Update of 1987 manual. Ohio Environmental Protection Agency, Division of Water Quality Planning and Assessment, Columbus, OH.

Ohio EPA. 1989b. Biological criteria for the protection of aquatic life: Volume III. Standardized biological field sampling and laboratory methods for assessing fish and macroinvertebrate communities. Update of 1987 manual. Ohio Environmental Protection Agency, Division of Water Quality Planning and Assessment, Ecological Assessment Section, Columbus, OH.

Ohio EPA. 1990a. The cost of biological field monitoring. (Updated 1991.) Ohio Environmental Protection Agency, Division of Water Quality Planning and Management, Columbus, OH.

Ohio EPA. 1990b. The use of biocriteria in the Ohio EPA surface water monitoring and assessment program. Ohio Environmental Protection Agency, Division of Water Quality Planning and Assessment, Ecological Assessment Section, Columbus, OH.

Peckarsky, B.L. 1986. Colonization of natural substrates by stream benthos. Can. J. Fish. Aquat. Sci. 43:700-709.

Peckarsky, B.L., P.R. Fraissinet, M.A. Penton, and D.J. Conklin, Jr. 1990. Freshwater macroinvertebrates of northeastern North America. Cornell University Press, Ithaca, NY. 442 pp.

Pennack, R.W. 1978. Freshwater invertebrates of the United States. 2d ed. John Wiley and Sons, Inc., New York. 803 pp.

Pennak, R.W. 1989. Freshwater invertebrates of the United States (3rd edition) - Protozoa to Mollusca. John Wiley and Sons, Inc., New York. 628 pp.

Penrose, D.L., and D.R. Lenat. 1982. Effects of apple orchard runoff on the aquatic macrofauna of a mountain stream. Arch. Environ. Contam. Toxicol. 11:383-388.

Penrose, D.L., and J.R. Overton. 1988. Semiqualitative collection techniques for benthic macroinvertebrates: uses for water pollution assessment in North Carolina. pp. 77-88. In: Proceedings of the First National Workshop on Biological Criteria - Lincolnwood, Illinois, December 2-4, 1987. T.P. Simon, L.L. Holst, and L.J. Shepard (eds). EPA-905/9-89/003. U.S. EPA Region 5 Instream Biocriteria and Ecological Assessment Committee, Chicago, IL. 129 pp.

Petersen, L.B.M., and R.C. Petersen. 1983. Anomalies in hydropsychid capture nets from polluted streams. Freshwater Biology 13:-185-191.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughs. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. U.S. Environmental Protection Agency, Office of Water, EPA/444(440)/4-39-001, Washington, DC.

Pollard, J.E. 1981. Investigator differences associated with a kicking method for sampling macroinvertebrates. J. Freshwater Ecol. 1:215-224.

Pollard, J.E., and W.L. Kinney. 1979. Assessment of macroinvertebrate monitoring techniques in an energy development area: a test of the efficiency of three macrobenthic sampling methods in the White River. EPA-600/7-79/-163. U.S. Environmental Protection Agency, Office of Research and Development, Las Vegas, NV. 26 pp.

Rabeni, C.F., S.P. Davies, and K.E. Gibbs. 1985. Benthic invertebrate response to pollution abatement: structural changes and functional implications. Wat. Res. Bull. 21:489-497.

Rae, J.G. 1989. Chironomid midges as indicators of organic pollution in the Scioto River basin, Ohio. Ohio J. Sci. 89:5-9.

Rankin, E.T. 1989. The Qualitative Habitat Evaluation Index (QHEI): Rationale, Methods, and Application. Division of Water Quality Planning and Assessment, Ecological Assessment Section, Columbus, OH.

Resh, V.H. 1988. Variability, accuracy, and taxonomic costs of rapid assessment approaches in benthic biomonitoring. Draft. Paper presented at the 1988 North American Benthological Society Technical Information Workshop, Tuscaloosa, AL.

Resh, V.H., and D.G. Price. 1984. Sequential sampling: A cost effective approach for monitoring benthic macroinvertebrates in environmental impact assessments. Environ. Manage. 8:75-80.

Resh, V.H. and J.K. Jackson. 1991. Rapid assessment approaches to biomonitoring using benthic macroinvertebrates. In: D.M. Rosenberg and V.H Resh (eds.) Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall. New York Press

Richardson, R.E. 1928. The bottom fauna of the middle Illinois River, 1913-1925. Bull. Illinois Natural History Survey 17:387-472.

Rooke, J.B., and G.L. Mackie. 1982a. An ecological analysis of lotic environments: I. Design and testing. J. Freshwat. Ecol. 1:421-432.

Rooke, J.B., and G.L. Mackie. 1982b. An ecological analysis of lotic environments: II. Comparison to existing indices. J. Freshwat. Ecol. 1:433-442.

Rosas, I., M. Mazari, J. Saavedra, and A.P. Baez. 1985. Benthic organisms as indicators of water quality in Lake Patzcuaro, Mexico. Water Air Soil Pollut. 25:401-414.

Rosenberg, D.M., and A.P. Wiens. 1976. Community and species responses of Chironomidae (Diptera) to contamination of fresh waters by crude oil and petroleum products, with special reference to the Trail River, Northwest Territories. J. Fish. Res. Board Can. 33:1955-1963.

Rosenberg, D.M., and V.H. Resh. 1982. The use of artificial substrates in the study of freshwater benthic macroinvertebrates. pp. 175-236. In: Artificial Substrates. J. Cairns, Jr. (ed.). Ann Arbor Science Publishers, Ann Arbor, MI.

Saether, O.A. 1979. Chironomidae communities as indicators of water quality. Hol. Ecol. 2:65-74.

Shakelford, B. 1988. Rapid bioassessments of lotic macroinvertebrate communities: biocriteria development. Arkansas Department of

Pollution Control and Ecology, Little Rock, AR. 45 pp.

Shannon, C.E., and W. Weaver. 1949. The mathematical theory of communication. The University of Illinois Press, Urbana, IL. pp. 19-27, 82-83, 104-107.

Sheldon, A.L. 1984. Cost and precision in a stream sampling program. Hydrobiologia 111:147-152.

Shepard, R.B. 1982. Benthic insect colonization of introduced substrates in the Sangamon River, Illinois. Trans. Ill. Acad. Sci. 75:15-27.

Simpson, K.W. 1980. Abnormalities in the tracheal gills of aquatic insects collected from streams receiving chlorinated or crude oil wastes. Freshwater Biology 10:581-583.

Simpson, K.W. 1983. Communities of Chironomidae (Diptera) from an acid-stressed headwater stream in the Adirondack Mountains, New York. Mem. Amer. Entomol. Soc. 34:-315-327.

Simpson, K.W., and R.W. Bode. 1980. Common larvae of Chironomidae (Diptera) from New York State streams and rivers - with particular reference to the fauna of artificial substrates. New York State Department of Health, New York State Museum Bull. No. 439. Albany, NY. 105 pp.

Smith, M.E., and J.L. Kaster. 1983. Effect of rural highway runoff on stream benthic macroinvertebrates. Environ. Pollut. Ser. A. 32:157-170.

Stribling, J.B. 1991. Generic quality assurance project plan guidance for bioassessment/biomonitoring programs. Draft report prepared for U.S. EPA Environmental Monitoring and Systems Laboratory, Cincinnati, OH.

Surber, E.W. 1937. Rainbow trout and bottom fauna production in one mile of stream. Trans. Amer. Fish. Soc. 66:193-202.

Surber, E.W. 1970. Procedure in taking stream bottom samples with the stream square foot bottom sampler. Proc. Conf. Southeastern Assoc. Game Fish. Comm. 23:587-591.

Surdick, R.F., and A.R. Gaufin. 1978. Environmental requirements and pollution tolerance of Plecoptera. EPA-600/4-78/062. U.S. Environmental Protection Agency, Office of Research and Development, Cincinnati, OH.

USEPA. 1973. Biological field and laboratory methods for measuring the quality of surface waters and effluents. C.I. Weber (ed.). EPA-670/4-73/001. U.S. Environmental Protection Agency, National Environmental Research Center, Cincinnati, OH.

USEPA. 1987. A compendium of Superfund field operations methods. Section 12, Biology/Ecology. EPA-540/P-87/001. U.S. Environmental Protection Agency, Office of Emergency and Remedial Response, Washington, DC.

USEPA. 1988a. Proceedings of the First National Workshop on Biological Criteria - Lincolnwood, Illinois, December 2-4, 1987. U.S. Environmental Protection Agency, Region V, Instream Biocriteria and Ecological Assessment Committee, Chicago, IL, EPA-905/9-89/003, 129 pp.

USEPA. 1988b. Report of the National Workshop on Instream Biological Monitoring and Criteria. U.S. Environmental Protection Agency, Region V Instream Biological Criteria Committee, USEPA Office of Water, Washington, DC., 34 pp.

USEPA. 1989a. Risk assessment guidance for Superfund - Environmental evaluation manual. Interim final. U.S. Environmental Protection Agency, Office of Emergency and Remedial Response, Washington, DC. EPA-540/1-89/-001A

USEPA. 1989b. Ecological assessment of hazardous waste sites. EPA-600/3-89/013, U.S. Environmental Protection Agency, Office of Research and Development, Corvallis, Oregon.

USEPA. 1989c. Ecological risk assessment methods: A review and evaluation of past practices in the Superfund and RCRA programs. U.S. Environmental Protection Agency, Office of Policy, Planning, and Evaluation. Washington, DC. EPA-230/03-89-044.

USEPA. 1989d. Interim guidance for the design and execution of sediment sampling efforts related to a navigational maintenance dredging in Region V - May 1989. U.S. Environmental Protection Agency, Region V, Chicago, IL.

USEPA. 1989e. The nature and extent of ecological risks at Superfund sites and RCRA facilities. U.S. Environmental Protection Agency, Office of Policy, Planning, and Evaluation. Washington, DC. EPA-230/03-89-043.

USEPA. 1989f. Proceedings of the 1989 Midwest Pollution Control Biologists Meeting, Chicago, Illinois, February 14-17, 1989. W.S. Davis and T.P. Simon (eds.). U.S. Environmental Protection Agency, Region V Instream Biocriteria and Assessment Committee, Chicago, IL. EPA-905/9-89-007. 153 pp.

USEPA. 1989g. Summary of ecological risks, assessment methods and risk management decisions in Superfund and RCRA. U.S. Environmental Protection Agency, Office of Policy Analysis, Washington, DC. EPA-230-/03-89-046.

USEPA. 1990a. A guide to the Office of Water Accountability System and regional evaluations. U.S. Environmental Protection Agency, Office of Water, March 1991, Washington, DC.

USEPA. 1990b. Biological criteria: national program guidance for surface waters. U.S. Environmental Protection Agency, Office of Water, EPA-440/5-90-004, Washington, DC.

USEPA. 1990c. Policy on the use of biological assessments and criteria in the water quality program. DRAFT. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA. 1990d. Environmental monitoring and assessment program: ecological indicators. EPA/600/3-90/060. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC.

USEPA. 1990e. Feasibility report on environmental indicators for surface water programs. U.S. Environmental Protection Agency, Office of Water Regulations and Standards, and the Office of Policy, Planning and Evaluation, Washington, DC.

USEPA. 1990f. Monitoring Lake and Reservoir Restoration: Technical Supplement to The Lake and Reservoir Restoration Guidance Manual. U.S. Environmental Protection Agency, Office of Water, EPA -440/4-90/007, Washington, DC.

USEPA. 1990g. Proceedings of the 1990 Midwest Pollution Control Biologists Meeting, Chicago, Illinois, April 10-13, 1990. W.S. Davis (ed.). U.S. Environmental Protection Agency, Region V, Environmental Sciences Division, Chicago, IL. EPA-905/9-90-005. 142 pp.

USEPA. 1991a. A guide to the Office of Water Accountability System and regional evaluations. U.S. Environmental Protection Agency, Office of Water, March 1991, Washington, DC.

USEPA. 1991b. Technical support document for water quality - based toxics control. U.S. Environmental Protection Agency, Office of Water EPA/505/2-90/060, Washington, DC.

USEPA. 1991c. Biological criteria: State development and implementation efforts. U.S. Environmental Protection Agency, Office of Water, Washington, DC. EPA-440/5-91-00.

USEPA. 1991d. Biological criteria: Guide to technical literature. EPA-440/5-91-004. U.S. Environmental Protection Agency, Office of Water. Washington, DC.

USEPA. 1991e. Biological criteria: Research and regulation - proceedings of a symposium. EPA-440/5-91-005. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA. 1991f. Surface waters monitoring and research strategy - fiscal year 1991. EPA/-600/3-91/002. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC.

USEPA. 1991g. Monitoring guidelines to evaluate effects of forestry activities on streams in the Pacific Northwest and Alaska. EPA/-910/9-91-001. U.S. Environmental Protection Agency, Seattle, WA.

USEPA. 1992a. Final Draft. Procedures for initiating narrative biological criteria. July 1992. U.S. Environmental Protection Agency, Office of Water, Office of Science and Technology, Washington, DC.

USEPA. 1992b. Draft. Biological criteria: Technical guidance for survey design and statistical evaluation of biosurvey data. U.S.

Environmental Protection Agency, Office of Water, Office of Science and Technology, Washington, DC.

USEPA. 1992c. Biological criteria: Technical guidance document for streams. Draft No. 4. U.S. Environmental Protection Agency, Office of Water, Office of Science and Technology. Washington, DC.

USEPA. 1992d. Proceedings of the 1991 Midwest Pollution Control Biologists Meeting. T. Simon and W. Davis (eds.). EPA-905/R-92/-003. U.S. Environmental Protection Agency, Chicago, IL.

Van Dyk, L.P., C.G. Greeff, and J.J. Brink. 1975. Total population density of Crustacea and aquatic Insecta as an indicator of fenthion pollution of river water. Bull. Environ. Contam. Toxicol. 14:426-431.

Van Horn, W.M. 1950. The biological indices of stream quality. Proc. 5th Ind. Waste. Conf., Purdue Univ. Est. Ser. 72:215.

Warwick, W.F. 1985. Morphological abnormalities in Chironomidae (Diptera) larvae as measures of toxic stress in freshwater ecosystems: indexing antennal deformities in *Chironomus* Meigen. Can. J. Fish. Aquat. Sci. 42:1881-1914.

Warwick, W.F., J. Fitchko, P.M. McKee, D.R. Hart, and A.J. Burt. 1987. The incidence of deformity in *Chironomus* sp. from Port Hope Harbour, Lake Ontario. J. Great Lakes Res. 13:88-92.

Washington, H.G. 1984. Diversity, biotic and similarity indices: a review with special relevance to aquatic ecosystems. Water Res. 18:653-694.

Waterhouse, J.C., and M.P. Farrell. 1985. Identifying pollution related changes in chironomid communities as a function of taxonomic rank. Can. J. Fish. Aquat. Sci. 42:406-413.

Webb, D.W. 1980. The effects of toxaphene piscicide on benthic macroinvertebrates. J. Kansas Entomol. Soc. 53:731-744.

Wentsel, R., A. McIntosh, and V. Anderson. 1977. Sediment contamination and benthic macroinvertebrate distribution in a metal-impacted lake. Environ. Pollut. 14:187-193.

Wiederholm, T. 1980. Use of benthos in lake monitoring. J. Wat. Pollut. Control Fed. 52:537-547.

Wiederholm, T. (ed.). 1983. Chironomidae of the holarctic region. Keys and diagnoses.Part I. Larvae. Entomologica Scandinavica. Supplement 19:1-457.

Wiederholm, T. 1984a. Responses of aquatic insects to environmental pollution. pp. 508-557. In: The Ecology of Aquatic Insects. V.H. Resh and D.M. Rosenberg (eds.). Praeger Publishers, New York, NY. 625 pp.

Wiederholm, T. 1984b. Incidence of deformed chironomid larvae (Diptera:Chironomidae) in Swedish Lakes. Hydrobiologia 109:243-249.

Wihlm, J.L. 1970. Range of diversity in benthic macroinvertebrate populations. J. Wat. Pollut. Control Fed. 42:R221-224.

Winnell, M.H., and D.S. White. 1985. Trophic status of southeastern Lake Michigan based on the Chironomidae (Diptera). J. Great Lakes. Res. 11:540-548.

Winner, R.W., J.S. Van Dyke, N. Caris, and M.P. Farrell. 1975. Response of a macroinvertebrate fauna to a copper gradient in an experimentally-polluted Stream. Verh. Internat. Verein. Limnol. 19:2121-2127.

Winner, R.W., M.W. Boesel, and M.P. Farrell. 1980. Insect community structure as an index of heavy-metal pollution in lotic ecosystems. Can. J. Fish. Aquat. Sci. 37:647-655.

Woodiwiss, F.S. 1964. The biological system of stream classification used by the Trent River Board. Chem. Ind. 11:443-447.

Yasuno, M., Y. Sugaya, and T. Iwakuma. 1985. Effects of insecticides on the benthic community in a model stream. Environ. Pollut. (Ser. A) 38:31-43.

Marine Benthic Community Structure Assessment

Betsy Striplin, Gary Braun, and Gordon Bilyard
Tetra Tech, Inc.
11820 Northup Way, Suite 100E, Bellevue, WA 98005
(206) 822-9596

Benthic communities are communities of organisms that live in or on the sediment. In most benthic community structure assessments, primary emphasis is placed on determining the species that are present and the distribution of individuals among those species. These community attributes are emphasized largely for pragmatic reasons. Although it is relatively simple to collect, identify, and enumerate benthic organisms, it is very difficult to determine first-hand the spatial distributions of species and individuals within the benthic habitat, or the functional interactions that occur among the resident organisms or between the resident organisms and the abiotic habitat. Hence, information on benthic community composition and abundance is typically used in conjunction with information in the scientific literature to infer the distributions of species and individuals in three-dimensional space and the functional attributes of the community. Because all of the major structural and functional attributes of benthic communities are affected by sediment quality in generally predictable ways, benthic community structure assessment is a valuable tool for evaluating sediment quality and its effects on a major biological component of marine, estuarine, and freshwater ecosystems.

Benthic habitats may be broadly divided into hard-bottom habitats and soft-bottom habitats. Many types of each exist in marine, estuarine, and freshwater ecosystems. Hard-bottom habitats include rocky shorelines and bottoms of lentic and lotic systems, rocky intertidal and subtidal habitats in marine and estuarine systems, and coral reefs. Soft-bottom habitats include mud and sand habitats in marine, estuarine, and freshwater systems; marine, estuarine, and freshwater macrophyte beds; freshwater wetlands; and estuarine salt marshes. Each of these habitats requires different sample collection methods and different survey design considerations. The emphasis of this chapter is on assessments of marine benthic community structure in soft-bottom habitats as an indicator of sediment quality. Freshwater benthic invertebrate community structure is discussed in Chapter 8.

9.1 SPECIFIC APPLICATIONS

Assessment of benthic community structure is an *in situ* method that can be used alone, as part of other approaches [e.g., Sediment Quality Triad (see Chapter 10) and Apparent Effects Threshold (AET) (see Chapter 11)], or in combination with other sediment assessment techniques (e.g., sediment toxicity bioassays). It is commonly used in three ways to assess impacts to benthic communities and sediment quality:

- To compare test and reference stations, for the purpose of determining the spatial extent and magnitude of such impacts;
- To identify spatial gradients of impacts; and
- To identify temporal trends at the same locations through time.

By definition, benthic communities include all organisms living on or in the bottom substrate. For practical reasons, assessments of benthic community structure in soft sediments usually rely on the macrofauna (i.e., organisms retained on a 1.0- or 0.5-mm sieve) and to a lesser extent the meiofauna (i.e., multicellular organisms that pass through a 1.0- or 0.5-mm sieve). Reasons for the more limited use of meiofauna are twofold:

- Although they may be sampled quantitatively, their small size makes working with them difficult, and the taxonomy of many of the groups (e.g., nematodes) is not well known.
- The functional attributes of the various meiofaunal taxa are poorly known, and it is therefore difficult to interpret the importance of the presence or absence of the various taxa in relation to environmental quality. (For example, knowledge of meiofaunal taxa that respond positively or negatively to organic enrichment of the sediments is extremely limited.)

Difficulties in quantitatively sampling other size classes of benthic organisms such as the megafauna (i.e., large organisms that are typically measured in centimeters) and the microfauna (i.e., microbes) usually preclude them from consideration in assessments of benthic community structure. Furthermore, although the functional importance of sediment microbes has been studied, their structural and functional characteristics have not been used as indicators of sediment quality.

9.1.1 Current Use

Assessments of benthic community structure have been used to describe reference conditions, baseline conditions, and the effects of natural and anthropogenic disturbances. Selected examples of current uses of this approach are provided below.

Organic Enrichment—Pearson and Rosenberg (1978) performed an extensive review of benthic community succession in relation to organic enrichment of marine and estuarine sediments. Based on that review, they developed a generalized model of structural community changes (i.e., numbers of species, abundances, biomass) in relation to organic enrichment, and identified opportunistic and pollution-tolerant species that are indicative of organic enrichment. Concepts developed by Pearson and Rosenberg (1978) have subsequently been used by many investigators to assess the degree of organic enrichment that has occurred in a variety of soft-bottom habitats. For example, Dauer and Conner (1980) assessed the effects of sewage inputs on benthic polychaete populations in a Florida estuary by collecting information on the total number of individuals, total biomass, and average number of species. They compared the sewage-affected site with a reference site and examined the response of individual species to organic enrichment. In another study in Florida, Grizzle (1984) identified indicator species based on life history responses to organic enrichment and other physicochemical changes. The taxa identified as indicator species in enriched areas were generally characterized by opportunistic life history strategies. Vidakovic (1983) assessed the influence of domestic sewage on the density and distribution of meiofauna in the Northern Adriatic Sea. He concluded that raw domestic sewage did not have a negative influence on the density and distribution of meiofauna, but the nematode/copepod ratio (Parker, 1975) indicated that these stations were under stress.

Contamination Due to Toxic Metals and Metalloids—Rygg (1985a, 1986) assessed benthic community structure in Norwegian fjords where the disposal of mine tailings had resulted in metals contamination of the sediment. His studies showed an inverse relationship between concentrations of metals in the sediment and the species richness and abundance of the benthic macroinvertebrate fauna. Bryan *et al.* (1987) examined population distributions of the oyster *Ostrea edulisy*, the polychaete *Nereis diversicolor*, and the cockle *Cerastoderma edule* in relation to wastes from metals mining in the Fal Estuary. They concluded that the distribution of species is dependent on their ability to tolerate copper and zinc, and on the capabilities of a population to develop a resistance to metals and thereby maintain their original distribution range.

Contamination Due to Toxic Organic Compounds—Toxic organic compounds are frequently associated with municipal discharges, industrial effluents, and storm drains. These discharges may also result in organic enrichment and contamination by metals or metalloids. The following benthic studies provided evaluations of sediment quality in areas primarily affected by toxic organic compounds:

- Creosote contamination. Tagatz *et al.* (1983) examined the benthic communities that colonized uncontaminated sediments and sediments contaminated with three different concentrations of creosote (177, 844, and 4,420 μg/g) in field and laboratory aquaria to assess the effects of marine-grade creosote on community structure. Numbers of individuals and numbers of species in field-colonized communities were significantly lower in all three creosote-contaminated sediments than in the controls. In the laboratory-colonized communities only the two higher creosote concentrations had reduced numbers of individuals and species. Distribution of individuals within species was similar for the laboratory and field assemblages of animals.

- Oil contamination. Elmgren *et al.* (1983) determined that acute effects of the *Tsesis* oil spill were noted after 16 days on both the macrofauna and meiofauna. Initial recovery was noted 2 yr after the spill. However, the authors predicted that complete recovery would require at least 5 yr. Jackson *et al.* (1989) investigated the effects of spilled oil on the Panamanian coast and found that shallow subtidal reef corals and the infauna of seagrass beds had experienced extensive mortality. After 1.5 yr, only some of the organisms in areas exposed to the open sea had recovered. Clifton *et al.* (1984) performed field experiments in Willapa Bay, Washington, and found that oil in the sediments modified the burrowing behavior of infaunal benthos.

Dredging and Construction-Related Activities—Swartz *et al.* (1980) examined species richness and species abundances just before dredging occurred in Yaquina Bay, Oregon, and for 2 yr after dredging. Benthic community recolonization was followed from the appearance of opportunistic taxa through their replacement by less tolerant taxa. Rhoads *et al.* (1978) examined the influence of dredge-spoil disposal on benthic infaunal succession in Long Island Sound by classifying species into groups based on their appearance in a disturbed area. They suggested that the "equilibrium community is less productive than a pioneering stage" and suggested that productivity may be enhanced through managed disturbances. The abundance of polychaetes, molluscs, and crustaceans is currently used to help assess potential biological effects of dredged material disposal by the Puget Sound Dredged Disposal Analysis Program (SAIC, 1991; Striplin *et al.*, 1991).

Natural Disturbances—Most studies of natural disturbances have assessed the recovery of benthic communities after the disturbance (e.g., large storms and associated wave activity, oxygen depletion, salinity reductions, El Niño). For example, Dobbs and Vozarik (1983) sampled stations before and after Storm David and observed that the number of species decreased after the storm. They also documented changes in the rank order of the dominant taxa. Santos and Simon (1980) examined defaunation of benthic communities before, during, and after annual hypoxia in Biscayne Bay. They documented that recolonization occurs fairly rapidly after the defaunation period. Oscillations in macrobenthic populations in the shallow waters of the Peruvian coast were examined by Tarazona *et al.* (1988). Fluctuations in density, biomass, species composition, and diversity were attributed to the El Niño of 1982-1983.

Assessment of benthic community structure is also used as a component of other sediment quality assessment tools. Along with sediment chemistry and sediment toxicity bioassays, it is one of three components of the Sediment Quality Triad (see Chapter 10). It is also a component of the Apparent Effects Threshold approach (see Chapter 11).

9.1.2 Potential Use

To date, benthic community assessments performed to evaluate sediment quality have focused on the relationships between community variables (e.g., numbers of species, total abundance, biomass) and measures of sediment quality (e.g., organic content, concentrations of chemical contaminants). Only for organic enrichment have individual species been identified that are indicative of various degrees of sediment alteration [see for example Pearson and Rosenberg (1978), Word *et al.* (1977)]. Moreover, for only a very few species has the autecological relationship between organic enrichment of the sediments and an individual species been explored. [For example, Fabrikant (1984) explored the autecology of the bivalve mollusc *Parvilucina tenuisculpta* in relation to organic enrichment of the sediments in the Southern California Bight.] A tremendous potential exists, however, for identifying species that are indicative (by their persistence, enhanced abundance, reduced abundance, or absence) of sediment contaminants at various concentrations. The identification of such taxa will not be simple because of the complex ecological interactions that occur within benthic communities, and because sediments are frequently contaminated with a mixture of chemicals. A first step in this process might be to attempt to identify species or suites of species that could be used to separate the effects of sediment organic enrichment from sediment contamination by toxic substances.

Another potential use of benthic community assessments would be to predict recovery of benthic habitats following the execution of remedial actions at contaminated sites. To date, it has not been possible to use extant benthic community structure to predict recovery because the only model that relates benthic community structure to sediment quality [i.e., the Pearson and Rosenberg (1978) model] is not quantitative. Quantification of this model and the development of quantitative models for other sediment contaminants will be required before benthic community assessments can be used to predict sediment quality. A valuable byproduct of such models would be the ability to predict the capacity of the remediated area to support higher trophic level organisms that forage on benthic organisms, including commercially and recreationally harvested demersal fishes.

9.2 DESCRIPTION

9.2.1 Description of the Method

An assessment of benthic community structure typically involves a field survey that includes replicated sampling at each station; sorting and identification of the organisms to species or lowest possible taxon; analyses of the numbers of taxa, numbers of individuals, and sometimes biomass in each sample; and identification of the dominant taxa. Results of the field survey are then interpreted in conjunction with other sediment variables (e.g., sediment grain size, total organic carbon) that were collected concurrently with the benthic samples.

9.2.1.1 Objectives and Assumptions

The objective of the benthic community structure approach is to identify degraded and potentially degraded sediments by examining the communities of organisms that inhabit those sediments. This empirical approach assumes the following:

- Because benthic infauna are generally sedentary, benthic community structure reflects the chemical and physical environment at the sampling location.

- Benthic community structure may be altered in a predictable manner over time and space by chemical or physical disturbances.

- The execution of proper data collection and analysis methods can reduce natural variability of benthic infaunal data and enable the detection of trends in sediment quality.

9.2.1.2 Level of Effort

The level of effort required to assess benthic community structure is relatively high. Regardless of the analytical methods, a field survey is required to collect the organisms. The sorting and identification process is labor-intensive and usually expensive. Program objectives will determine whether the data analyses are simple or complex.

9.2.1.2.1 Type of Sampling Required

The type of sampling required to collect benthic organisms is dependent on the objectives of the sampling program and on the area under study. Usually, the objective of a benthic sampling program is to study the characteristics of and the variation in the benthic community that occupies specific sampling stations. In this case, all organisms present in the sediment at that location are sampled together: those that normally reside in the surface few centimeters of sediment and those that normally reside deeper in the sediment (e.g., 5-15 cm below the surface). In some instances, a sampling program may have a different objective. For example, sampling for the Benthic Resources Analysis Technique (BRAT) (Lunz and Kendall, 1982) involves collecting box core samples and determining the biomass (and possibly the communities) present in specific sediment strata (i.e., 0-2 cm, 2-5 cm, 5-10 cm, and 10-15 cm below the sediment surface). In that technique, the benthic data are compared with the benthic organisms consumed by bottom-dwelling fish (as determined by gut content analyses of fish captured in the same area) to determine the food value of the benthos.

Characteristics of the area under study also influence the type of sampling. In intertidal or littoral environments where sampling stations can be occupied by walking to the site, samples are usually collected using a hand-held corer. If stations are located in subtidal areas, then remote sampling from a vessel is performed using a box corer or grab sampler. Sediment grain size may influence final selection of the sampler. Some samplers (i.e., many box corers) perform poorly in sandy sediments, whereas others (i.e., van Veen grab, Smith-McIntyre grab) perform adequately in a greater range of sediment types (i.e., fine to medium sand, silt, silty clay). Methods and equipment for sampling infaunal communities are further described in several publications (Word, 1976; Swartz, 1978; Eleftheriou and Holme, 1984; Nalepa *et al.*, 1988). Blomqvist (1991) provides an extensive review of quantitative sampling methods, including a detailed bibliography of pertinent papers.

Program objectives and knowledge of benthic communities in the study area will influence selection of the sieve size through which sediment samples will be washed. It is important that the sieve mesh sizes be appropriate for the community under study (e.g., 64 μm for meiofauna, 0.5 or 1.0 mm for macrofauna). Generally, the chances of retaining most macrofauna species and individuals (and therefore increasing sampling accuracy) are improved by the use of a finer mesh (but, see Bishop and Hartley, 1986). However, sieve size is an important determinant of the cost and level of effort necessary to obtain quantitative data. Very little difference in the field processing time exists between use of a 0.5-mm and a 1.0-mm sieve when sieving sediments finer than coarse sand, but laboratory analyses are much more time-consuming when the smaller mesh is used because it retains more abiotic materials and many smaller organisms.

9.2.1.2.2 Methods

Methods for collecting data on benthic community structure are divided into three categories: program design, field methods, and laboratory methods. Each of these categories is briefly discussed below.

Program design includes the selection of station locations, level of replication, type of sampler, screen size, data analysis methods (discussed later), and quality assurance/quality control (QA/QC) procedures. The selection of station locations will directly influence the usefulness of the resulting data. Stations that will be compared to one another (including reference stations) should be situated in areas with similar hydrography, water depth, and grain size to minimize the natural variability in benthic community composition that can be attributed to these factors. However, such station placement is not always attainable because of altered grain size distributions that often result from contaminant sources.

Selection of the number of replicates is an important component of program design because the accuracy and precision with which benthic community variables are estimated depend in part on the size of the sample (including all replicates). For example, the abundance of a single taxon is generally a less accurate descriptive variable than is the abundance of the total taxa because of the greater variability typically associated with one taxon in comparison with the sum of all taxa. The total area sampled among the replicates at each station should be large enough to estimate a given variable within the limits of accuracy and precision that are acceptable to meet study objectives. A single sample may be useful for general distributional or trends analyses (Cuff and Coleman, 1979), but the inherent patchiness of benthic communities makes collection of a sufficient number of replicate samples (a minimum of 3-5, depending on study objectives and sampler area) necessary to ensure statistical reliability (see Elliott, 1977). Within a study area, adequate sample size may be determined by maximizing the number of species collected or by minimizing the error associated with the mean for the variable in question (Gonor and Kemp, 1978). Additional research on replication is presently being conducted by EPA in Newport, Oregon, under the direction of S. Ferraro (Swartz, R.C., 15 March 1989, personal communication).

Power analysis can assist in determining the appropriate number of replicates. A power analysis includes consideration of the minimum detectable difference in selected biological variables (i.e., the minimum difference in mean values of a variable at several stations that can be detected statistically, given a certain level of variability about those mean values) and the power of the statistical test to be used. The power of the test is especially important because it defines the probability of correctly detecting experimental effects (e.g., differences in biological variables among sampling stations). For a specified variance associated with a biological variable, the statistical power of a test and the minimum detectable difference among sampling areas can be expressed as a function of sample size. The allocation of sampling resources (stations, replication, and frequency) can then be determined with regard to available resources, practicality of design, and desired sensitivity of the subsequent analyses. Discussions and examples of this approach are found in Winer (1971), Saila *et al.* (1976), Cohen (1977), Moore and McLaughlin (1978); Bros and Cowell (1987), Ferraro *et al.* (1989), Kronberg (1987), Tetra Tech (1987), Self and Mauritsen (1988), and Vezina (1988).

A potential drawback to use of power analysis is that it requires *a priori* knowledge of variability in the benthic communities that will be studied. If such variability is not known and cannot be estimated, then the number of replicates will probably reflect either funding limitations or generally approved sampling methods. For example, Eleftheriou and Holme (1984) and Swartz (1978) recommend that an area of 0.5 m^2 be sampled to assess species composition in coastal and estuarine regions. Most studies of benthic community structure routinely involve five replicate 0.1-m^2 grab samples. A single 0.1-m^2 grab sample may be sufficient to obtain "useful descriptive information" for use in cluster analyses (Word, 1976). However, a single sample precludes direct estimates of within-group variance for statistical analyses. Because individuals are distributed logarithmically among the species of a benthic community (Preston, 1948; Sanders, 1968; Gray and Mirza, 1979), species collected in the second and successive replicates that were not collected in any of the previous replicates most often will be numerically "rare." Note that "rare" is not synonymous with "unimportant." Hence, a single 0.1-m^2 sample is generally not adequate to characterize benthic community structure and function. In general, five 0.1-m^2 grab samples are recommended for determining benthic community structure, unless evaluation of site-specific data (i.e., a power analysis) indicates that sufficient sensitivity can be obtained with fewer samples, or that a greater number is required due to extreme spatial heterogeneity. (Note that at least three samples are required for parametric statistical analyses.)

Another aspect of program design is selection of the appropriate degree of navigational accuracy. For baseline or distributional studies, repeatable station location may not be a high priority, and methods such as Loran C may be sufficient.

However, for monitoring programs where reoccupation of exact stations is important (e.g., disposal site monitoring), a more accurate positioning method (e.g., an electronic distance-measuring device or Mini-Ranger) may be required.

A quantitative sampling device and an appropriate mesh size must be selected to ensure that size classes of organisms appropriate for assessing sediment quality are collected. Selection of a sampler and sieve are discussed above, in Section 9.2.1.2.1.

Field and laboratory methods must be conducted according to rigorous QA/QC protocols. Field methods include collecting, sieving, and preserving the samples. Samples are typically preserved in a solution of 10 percent buffered formalin for at least 24 h. Laboratory methods include rinsing the formalin solution from the samples within 7-10 days, followed by storage in 70 percent ethanol. Samples are sorted under a dissecting microscope during which all organisms are removed from the samples and placed in vials for identification and enumeration of individual taxa. The time required to sort and identify a benthic sample varies greatly depending on the sieve size, sample area, and sediment composition. Sorting may take as little as 1 h for a 0.1-m^2 sample sieved through a 1.0-mm screen, or as much as 12 h if wood chips or other debris are present. The time needed to identify organisms in a sample depends on the number of organisms (which is a function of sieve size, habitat, or degree of contamination) and number of taxa present. The number of hours needed to identify organisms in a sample may range from 1 to over 10 h.

In addition to the collection of samples for analysis of benthic community structure, separate sediment samples should be collected at all stations for conventional sediment chemistry variables (e.g., sediment organic content, sediment grain size distribution). Because organic carbon content and sediment grain size naturally affect the composition of benthic communities, measurement of these variables will assist in determining whether benthic communities are affected by reduced sediment quality.

9.2.1.2.3 Types of Data Required

The two primary structural attributes of any benthic community are the distribution of species and individuals in three-dimensional space, and the distribution of individuals among species and higher taxa. Given an understanding of these two structural attributes, it is possible to infer functional attributes of the benthic community, including trophic relationships, primary and secondary productivity, and interactions between the resident biota and the abiotic habitat. The data required for analysis of the structural and functional attributes include the number of taxa (identifications should be to the lowest taxonomic level possible), the abundance of each taxon, biomass (depending on program objectives), and conventional sediment chemistry variables. However, collection of the appropriate data does not ensure proper evaluation of the structural and functional attributes. The selection and implementation of data analyses are equally important, and are discussed in the remainder of this section. The data analyses presented in this section address primarily structural components of benthic communities. However, functional attributes can be inferred from many of those structural attributes.

Various types of data analyses are used to describe benthic community structure, depending on the objectives of the particular program. However, several descriptive values are common to most program objectives. All organisms collected in each sample are enumerated (i.e., total abundance), and abundances of major taxonomic groups are usually summarized. Depending on the level of identification, abundances of individual taxa, numbers of taxa, and lists and abundances of pollution-tolerant and pollution-sensitive taxa in each sample may be developed. Biomass of major taxonomic groups and total biomass are sometimes reported. The composition of the numerically dominant taxa are analyzed when species level identifications are performed. In addition, descriptive indexes such as diversity [the distribution of individuals among species; see Washington (1984) for additional definitions of diversity], evenness (the evenness with which individuals are distributed among taxa), and dominance (the degree to which one or a few species dominate the community) are usually calculated.

Most programs evaluate the temporal or spatial differences in benthic community structure. Typically, comparisons of one or more indexes are made at the same station over time and compared to a baseline value, or comparisons are made between stations in a study area and stations in a reference area. If an adequate number of samples is collected (i.e., three or more), statistical tests such as t-tests or Analysis of Variance (ANOVA) (or their nonparametric analogues) are often performed to determine whether significant spatial or temporal differences exist among benthic communities.

Besides univariate (i.e., single-variable) statistical analyses, multivariate (i.e., multiple-variable) analyses are frequently performed (e.g., Boesch, 1977; Green and Vascotto, 1978; Gauch, 1982; Shin, 1982; Long and Lewis, 1987; Ibanez and Dauvin, 1988; Nemec and Brinkhurst, 1988a,b; Stephenson and Mackie, 1988). Multivariate analyses include classification methods (i.e., grouping similar stations into clusters) and ordination methods [i.e., representing sample or species relationships as faithfully as possible in a low-dimensional (two-four dimensions) space]. [See Gauch (1982) for an overview of multivariate methods.] Multivariate techniques group data and display them on a two-dimensional plot or dendrogram so that stations exhibiting similar communities are located closer to one another than to stations with dissimilar communities. The numerical and graphical results can then be compared with physical and chemical data collected concurrently to determine whether those variables correlate with trends in benthic communities. A commonly used classification technique involves first computing a matrix of similarity indexes that represent the degree of similarity in species composition between two stations. Commonly used similarity indexes include Bray-Curtis, Canberra metric, and Euclidian distance indexes. The similarity matrix is then entered into a clustering algorithm (e.g., pair-wise averaging, flexible sorting) to produce a dendrogram depicting similarities among stations. Commonly used ordination techniques include principal components analysis, detrended correspondence analysis, and discriminant function analysis. Bernstein and Smith (1986) developed an index of benthic community change along pollution gradients that is derived from results of ordination analysis. The index (called Index 5) is a measure of change from reference conditions.

Benthic community surveys generate large data matrices. These data matrices are often reduced by the elimination of certain species (Boesch, 1977) prior to performing multivariate analyses. A variety of methods exist for reducing data matrices (see Stephenson *et al.*, 1970, 1972, 1974; Day *et al.*, 1971; Clifford and Stephenson, 1975).

Both parametric statistical tests and multivariate analyses may involve data transformations. Transformations of the original data may be necessary for one or more of the following reasons:

- Benthic data sets are usually characterized by large abundances of a few species and small abundances of many species;
- The distribution of individuals among species tends to be lognormal; and
- Sampling effort may be inconsistent (Boesch, 1977).

The two basic types of transformations are strict transformations and standardizations. Strict transformations are alterations of the original values (e.g., species abundances) without reference to the range of values within the data. Commonly used transformations are square root, logarithmic, and arcsine (Sokal and Rohlf, 1981). Standardizations are alterations that depend on some property of the data under consideration. A common standardization is the conversion of values to percentages.

Benthic data are transformed to better meet the assumptions of parametric tests (e.g., normality, homogeneity of variances). In multivariate analyses, data are often transformed using logarithms [e.g., log (x+1)] because of the presence of zero scores. This transformation is also applied when population variance estimates are positively correlated with mean values (Sokal and Rohlf, 1981). Clifford and Stephenson (1975) discuss in detail the effects of transformations on commonly used resemblance measures.

Benthic community structure is usually compared with chemical and physical data that are collected concurrently. These comparisons may take the form of simple linear correlations, correlations with cluster groups, or correlations using multivariate techniques such as discriminant analyses. Multiple discriminant analysis attempts to isolate groups of similar stations so that variables responsible for the separation of groups can be identified. Results may be used to determine whether differences in community structure are due to variations in sediment grain size, variations in other physical characteristics of the environment, or changes in sediment quality due to toxic substances or organic materials.

The use of different methods and analyses may result in different interpretations of the same data. For example, use of the same data with different standardization methods in a classification analysis can yield very different results (Austin and Grieg-Smith, 1968). Generally, the more analyses that are conducted on the data, the higher the probability of interpreting the data accurately.

9.2.1.2.4 Necessary Hardware and Skills

The hardware needed to perform a benthic community assessment is fairly common and should be readily available. Equipment includes field collection gear (e.g., sampling vessel, appropriate sampler, sieves, sample storage containers, buffered fixative) and standard biological laboratory equipment (e.g., microscopes, sieves, hydrometers or pipets, and a balance). More specialized equipment includes a muffle furnace for determining total volatile solids concentrations, a taxonomic reference collection, and a taxonomic reference library. Computer equipment and appropriate software are required to make studies cost-effective. A microcomputer is sufficient for most analyses, but some complicated multivariate analyses may require the use of a minicomputer or mainframe computer.

Trained benthic taxonomists are required to ensure accurate identifications. Some computer programming and some level of data management are usually required. A trained benthic ecologist is required to synthesize and interpret the data. However, the amount of training depends on the required level of interpretation. For example, interpretation of several multivariate methods would require a higher level of training than interpretation of descriptive indexes.

9.2.1.3 Adequacy of Documentation

Many different approaches and methods are used to analyze benthic data, some of which have their origins in classical terrestrial community ecology. Because analysis of benthic community structure is a relatively old assessment tool, literally thousands of papers have been written about the method. Several books and protocols have also been developed to describe field and laboratory techniques [e.g., Holme and McIntyre (1984), Puget Sound Protocols (Tetra Tech, 1986b), U.S. EPA 301(h) protocols (Tetra Tech, 1986a)]. However, a comprehensive document that describes standardized procedures for analyzing and interpreting benthic community data is lacking.

The most commonly used interpretive approaches include measures of diversity and classification. Sometimes a general consensus exists on the best techniques to use within an approach (e.g., widespread use of Shannon-Wiener diversity index, although there is debate as to whether this is a suitable index for environmental impact analysis). Despite this consensus, studies do not necessarily follow a specified format. Program objectives tend to dictate the types of hypotheses posed and analyses used. Many relatively new and exciting approaches have been proposed for assessing benthic community structure. However, most are relatively untested and are not widely used [e.g., benthic resource analysis technique (Lunz and Kendall, 1982), abundance-biomass comparison (Warwick, 1986; Warwick *et al.*, 1987), infaunal trophic index (Word, 1978, 1980), nematode:copepod ratio (Amjad and Gray, 1983; Lambshead, 1984; Shiells and Anderson, 1985; Raffaelli, 1987), lognormal distribution (Gray and

Mirza, 1979), Index 5 (Bernstein and Smith, 1986)]. Each of these methods has shown promise in some situations, but more testing and validation are needed before any can gain universal acceptance.

Very few assessments of the information gained from analyses of data at the species level vs. the major taxa level have been undertaken. Warwick (1988) evaluated the results of ordinations run on various hierarchical levels of taxonomic data for five data sets. Three of the data sets were of macrofauna (from Loch Linne, Clyde Sea, and Bay of Morlaix); one was of nematodes from the Clyde Sea; and the last was of copepods from Oslofjord that were subjected to different levels of particulate organic material. He reported that in none of those five cases was there any substantial loss of information at the family level, and that in two cases the sample groupings related more closely to the gradient of pollution at the phylum level than at the species level. Warwick tentatively suggested that "anthropogenic effects modify community composition at a higher taxonomic level than natural environmental variables, which influence the fauna more by species replacement." Warwick's paper appears to be the only published work to support the use of higher taxonomic groups for analysis purposes. In cases where only major taxa level data have been collected (e.g., PTI and Tetra Tech, 1988), it has been difficult to determine differences in community structure between impacted areas and reference areas, and to establish causes of community alterations. Although it would be a cost-saving approach, use of higher taxonomic levels to assess benthic communities is currently not an accepted approach in the United States.

9.2.2 Applicability of Method to Human Health, Aquatic Life, or Wildlife Protection

The assessment of benthic community structure is directly applicable to the protection of aquatic life. Because benthic organisms are aquatic, assessments of benthic community structure provide a direct measure of the condition of aquatic life. Furthermore, because benthic organisms are consumed by other aquatic organisms (e.g., fish), assessing the condition of benthic communities provides information on other aquatic organisms.

Assessment of benthic community structure is both directly applicable to the protection of some wildlife (e.g., wading shorebirds that feed on the benthic infauna) and indirectly applicable to the protection of other wildlife (e.g., fish-eating wildlife). A substantial decrease in abundance of benthic organisms may result in the loss of food and a reduction in the value of certain habitat to wildlife. For example, distributions of demersal fishes have been shown to be affected by changes in the composition of benthic infaunal communities (e.g., see Kleppel *et al.*, 1980), as has the distribution of the starfish *Astropecten verilli* (Striplin, 1987).

Assessment of benthic community structure may be directly or indirectly applied to the protection of human health. When changes in community structure are caused by the presence of toxic contaminants, the bioaccumulation of those contaminants in more tolerant species may sometimes be postulated. Those contaminated benthic infauna may directly affect human health if they are ingested (e.g., shellfish contamination), or may indirectly affect human health if contaminants are transferred through the food web to humans (e.g., consumption of contaminated demersal fish).

9.2.3 Ability of Method to Generate Numerical Criteria for Specific Chemicals

Benthic community structure as a stand-alone assessment method cannot presently generate numerical criteria for specific chemicals, nor is it likely that it will without extensive research. However, it is an integral component of other methods that generate numerical criteria (e.g., Apparent Effects Threshold, Sediment Quality Triad). The great number of factors influencing benthic community structure at a given site generally precludes isolation of chemical-specific effects.

9.3 USEFULNESS

Assessment of benthic community structure has become a valued tool for determining sediment quality. It is recognized as the only *in situ* measure that provides information on changes in ecological relationships among species that inhabit potentially contaminated sediment. Its usefulness will continue both as an assessment method on its own and as a component of other sediment quality assessment tools.

9.3.1 Environmental Applicability

This method is applicable in a variety of environments. As a tool for assessing sediment quality, it has been used to assess the effects of known or suspected contaminants (e.g., industrial or municipal discharges, oil spills). The results of such studies reveal the geographic extent of the problem area and the type and severity of contamination.

9.3.1.1 Suitability for Different Sediment Types

Benthic community structure is well suited for assessing spatial and temporal effects of chemical contamination and/or organic enrichment in a variety of sediment types. However, to the extent possible, benthic communities occupying different types of sediment should not be compared. Considerable research has shown that the structure of benthic communities in coarse sediments differs from that in fine sediments (see Rhoads and Young, 1970; Rhoads and Boyer, 1982). Briefly, species recruiting into soft, silty sediments must be able to tolerate the deposition of fine particulate material. These environments tend to be inhabited by subsurface deposit-feeding organisms, whereas sandy environments tend to be inhabited by both surface suspension-feeding species and subsurface-dwelling species. Therefore, the experimental design of a benthic survey must reflect that the functional attributes of benthic communities in silty and sandy environments fundamentally differ.

When reference stations are used as the basis for determining differences in community structure between nonimpacted and potentially impacted stations, the reference and test stations should exhibit, to the extent possible, similar sediment characteristics (as well as similar water depths because benthic communities naturally vary by depth). However, it is not always possible for the reference and test stations to have sediment that has a similar composition; for example, dredged material at a dump site may have different characteristics than native sediment surrounding the dump site. If the experimental design is based on sampling the same stations through time to assess temporal change, then presumably sediment grain size would remain constant. If the objective is to sample along a potential gradient of chemical contamination or organic enrichment, then all stations should have similar grain sizes and water depths. However, this is not always possible because the source of contamination may alter the natural grain size distribution of the sediments.

Benthic community structure is also a suitable technique for assessing the presence of anaerobic sediments caused by poor flushing or excessive organic loading. The success of this approach will once again hinge on comparing benthic community structure between stations with similar grain sizes and water depths.

9.3.1.2 Suitability for Different Chemicals or Classes of Chemicals

Analysis of benthic community structure is frequently used to determine effects of chemicals present in the sediment. However, it is not used as a method to quantify the relative concentrations of individual chemicals or classes of chemicals present in sediment. Although individual species may react to certain chemicals, these reactions are not quantifiable at the community level. The Apparent Effects Threshold approach (Chapter 10) incorporates changes in abundance of major taxa for specific chemicals.

Benthic communities respond predictably to general categories of contamination. For example, metals contamination of sediments results in decreased species diversity (Rygg, 1985a, 1985b, 1986). Organic enrichment, which leads to an increase in the food supply, generally results in increased diversity and abundance at slight to moderate levels of enrichment (Pearson and Rosenberg, 1978; Rygg, 1986). However, beyond some level of organic enrichment, diversity and abundance decrease with continued organic loading (Pearson and Rosenberg, 1978). In an area receiving both organic enrichment and toxic contaminants, it may be difficult to distinguish the effects of these forms of pollution from each other. Additional research is greatly needed to help separate the effects of multiple sources of contaminants.

9.3.1.3 Suitability for Predicting Effects on Different Organisms

Changes in benthic communities that result from the presence of organic enrichment or chemical pollutants may be useful indicators of the potential effects of that pollution on predators of the infauna (see Kleppel, 1982; Striplin, 1987). However, using benthic community structure to predict specific effects on potential predators (such as benthic-feeding fish or shorebirds) may be difficult. Information on trophic relationships, competition, and predation is often not available. The capability to predict the effects of altered prey communities on predators may improve with research on these topics. Factors such as food quality, distribution of food, interactions among species, and distribution of prey will all be important components of this research.

9.3.1.4 Suitability for In-Place Pollutant Control

Benthic community structure has not been used to set sediment quality goals or criteria for polluted marine sediments. Benthic communities naturally express sufficient spatial and temporal variability to eliminate them from consideration as a goal or criterion-setting variable. However, benthic communities are an integral part of other approaches to assess sediment quality (see Chapters 10, and 11, and 12) in which benthic community structure is the only *in situ* biological measure.

9.3.1.5 Suitability for Source Control

Benthic community assessments can provide valuable information for certain aspects of source control. Benthic communities can assist the identification of outfalls that discharge toxic chemicals or high organic loads. Depending on the nature of the material being discharged, benthic communities may be diverse and abundant if the material is organically enriched or may be depauperate if the material has high levels of toxic contaminants. Because benthic communities are not currently useful for identifying specific chemicals or classes of chemicals present in the sediment, they are of limited value for identifying specific sources of contaminants.

Following the control of sources, benthic community structure may be used to monitor long-term recovery of the receiving environment (Tetra Tech, 1988). It is not recommended as an indicator of the immediate effects of controlling sources because the sediment will remain contaminated until the sediment is actively remediated, or until bioturbation and natural deposition of uncontaminated particulates dilute the contaminated sediment. Furthermore, this assessment technique would be useful only in areas where other uncontrolled sources would not obscure sediment recovery due to the controlled source. Where source control has occurred, or is planned on a regional level, establishment of one or more stations for the analysis of long-term trends in benthic community structure is recommended as a method of monitoring regional sediment recovery. The concentration and type of the contaminants and the hydrodynamics of the study area will govern the length of time over which recovery will occur (Perez, K., 1 May 1989, personal communication).

9.3.1.6 Suitability for Disposal Applications

Regulations concerning biological testing of sediment that is dredged under sections 401 and

404 of the Clean Water Act do not include assessments of benthic community structure. Benthic communities inhabit only the upper layers of sediment that will be dredged. Because sediment quality near the sediment surface may not reflect sediment quality throughout the depth of sediment to be dredged, benthic communities are unable to provide information that is suitable for assessing the entire volume of sediment that will be dredged. Chemical analyses, laboratory bioassays, and bioaccumulation studies can, however, be used to assess sediment quality throughout the dredging depth. Section 102 of the Marine Protection Research and Sanctuary Act does call for monitoring of benthic community structure in areas where dredged material is disposed.

The International Joint Commission (IJC) recommends use of benthic communities to determine whether areas of concern exist in sediments that require dredging (IJC, 1988a, 1988b). However, they do not discuss whether benthic community structure would be used to determine the suitability of dredged material for open-water disposal.

Analysis of benthic community structure is appropriate for postdisposal monitoring of confined and unconfined disposal sites and for monitoring recovery of areas that were dredged. As part of the Puget Sound Dredged Disposal Analysis (PSDDA) postdisposal monitoring program, benthic community structure is used to monitor the potential transport of disposed material away from the disposal site (SAIC, 1991; Striplin *et al.*, 1991). The purpose of this aspect of the monitoring program is to determine whether benthic communities are altered near the disposal site and, if so, whether the changes are due to offsite migration of the disposed material. Benthic community structure was also incorporated into the proposed monitoring program for confined aquatic disposal sites to confirm recolonization of the clean sediment cap and to monitor cap integrity at the Commencement Bay Nearshore/Tideflats Superfund site in Tacoma, Washington (Tetra Tech, 1988). As described earlier, Swartz *et al.* (1980) documented recovery in Yaquina Bay, Oregon, following dredging. Rhoads *et al.* (1978) suggested that periodic disturbance such as dredging and disposal may enhance benthic productivity.

9.3.2 General Advantages and Limitations

General advantages of using benthic community structure to determine sediment quality include its inherent capability to provide an ecological basis for evaluation of sediment quality. It is an empirical rather than a theoretical approach. However, as with most assessment techniques involving field studies, the evaluation of benthic communities is costly and time-consuming. The information gained is often not suitable for specific management decisions because of the lack of numerical management criteria and the method's inability to identify specific chemicals responsible for an impact. However, the technique has been incorporated into other predictive techniques (see Chapters 10, 11, and 12) that provide information more easily used by resource managers.

9.3.2.1 Ease of Use

Assessments of benthic community structure require field collections, extensive laboratory work, and data analysis and interpretation by trained benthic ecologists. It is difficult to argue that the method is easy to use, especially in comparison to other methods that rely on established criteria. However, the use of benthic community structure as a sediment quality assessment tool is widely accepted, and trained benthic ecologists are available throughout the country. By using highly experienced individuals to conduct the field, laboratory, and data analysis work, potential problems (such as generating "noisy" data that obscure real trends, or arriving at different interpretations using the same data) should not occur.

9.3.2.2 Relative Cost

The relative cost of conducting an assessment of benthic communities is less than the cost to develop and implement other sediment quality assessment techniques such as the Apparent Effects Threshold and equilibrium partitioning approaches. However, once sediment quality values have been generated, the relative cost of conducting a benthic survey is greater than the cost of analyzing sediment for contaminant concentrations and comparing those data to the values to determine sediment quality. Sediment toxicity bioassays are generally less costly than analysis of replicated benthic samples. Because the Triad approach requires synoptic analyses of sediment chemistry, sediment toxicity, and benthic communities, it is more costly to implement than simply an analysis of benthic communities. It also provides broader information from which to determine sediment quality.

The objectives of benthic community assessment programs strongly influence cost by dictating the number of stations and number of replicates per station. The cost per replicate is relatively high (i.e., \$400-\$1,000), but varies greatly depending on the size of the area sampled, the screen size, the level of the taxonomic identifications, and the environment sampled.

9.3.2.3 Tendency to Be Conservative

Benthic community structure is a moderately conservative measure of sediment quality. Because benthic community structure reflects the collective response of all species, responses of individual species that are susceptible to degradation in sediment quality may not be obvious at the community level because of the lack of response in other species that are more tolerant of environmental degradation. Changes to numerous species or dominant species must occur before changes at the community level are evident. If assessments of sediment quality were made using individual species instead of communities, they could be either conservative by relying on sensitive species or not conservative by relying on tolerant species.

9.3.2.4 Level of Acceptance

Benthic community assessments have been used as a sediment quality assessment tool for several decades in North America, Europe, and Australia, as well as in South Africa, China, and Japan. The method has gained widespread acceptance because of its inherent capability to assess sediment quality at the community level, thereby documenting ecological response to sediment perturbations.

Many methods may be used to analyze benthic community data, as discussed above. Some of these methods have gained far wider acceptance than have other, sometimes newer, approaches. The most widely accepted types of analyses include measures of abundance, numbers of taxa, diversity, similarity, community classification, and the abundance of sensitive and tolerant species. Other analytical methods include the lognormal distribution (Gray and Mizra, 1979), the use of major taxa instead of species-level data (Warwick, 1988), and the Infaunal Trophic Index (Word, 1978, 1980). Each of these may be appropriate for certain types of perturbations, but have yet to gain widespread acceptance.

9.3.2.5 Ability to Be Implemented by Laboratories with Typical Equipment and Handling Facilities

Many laboratories either have the essential equipment for conducting benthic community surveys or can readily obtain this equipment. However, locating qualified taxonomists to oversee the sorting and to identify the organisms may be difficult. Taxonomists require several years of training and experience before they are considered experts in their respective taxonomic fields. They also require access to a reference museum of verified organisms to assist in their identifications. A thorough taxonomic library containing original descriptions of species is also an integral component of taxonomic laboratories.

9.3.2.6 Level of Effort Required to Generate Results

The level of effort required to conduct a benthic community survey is dependent on the objectives of the program, which may affect the number of stations, number of replicates per station, taxonomic level of the identifications, and data analysis procedures. Regardless of those objectives, a field effort is required; the samples must be sorted, identified, and enumerated; and the resulting data must be analyzed. This process typically requires several months, but it is not unusual for it to require a full year for a very large sampling effort, or for a program in which the samples require large sorting or identification times. For example, the sorting time for samples collected from deep water silt and clay may be 1-2 h, whereas that for samples from shallow sandy sites might be 4-6 h because shallow sandy areas typically contain more abiotic material. If wood chips are present in the sample, then the sorting time can easily exceed 12 h, depending on the volume of wood chips.

9.3.2.7 Degree to Which Results Lend Themselves to Interpretation

The interpretation of benthic community data requires an expert who is familiar with the natural history of the fauna and the statistical techniques that are routinely used to analyze the data. Interpretation of the many data points generated by this approach may require many weeks before meaningful trends are recognized. The inherent variability of benthic communities has so far prevented the development of specific benthic criteria for use in assessing pollutant-related trends in sediment quality.

9.3.2.8 Degree of Environmental Applicability

The assessment of benthic community structure is a direct measure of the environmental effects of pollutants and, as such, is highly applicable as a method to assess sediment quality. Its applicability lies in its ability to provide information on the effects of pollutants on ecological processes within the sedimentary environment.

9.3.2.9 Degree of Accuracy and Precision

Provided that sufficient funding is available to collect and process the necessary numbers of replicate samples, analysis of benthic community structure is accurate (defined as how well the data represent true field conditions) and precise (defined as the consistency and reliability of the samples). The resulting data are obtained directly from the populations under study. Other sediment quality assessment methods described in this compendium are not direct measures of field conditions and therefore are less likely to be as accurate and precise.

Many factors in the design of a benthic community survey directly influence the degree of accuracy and precision of the resulting data. These factors include station placement, number of replicates, appropriateness of reference areas, sampler, sieve mesh size, sampling interval, quality of taxonomy, and the type and quality of the data analysis. The best way to ensure high degrees of accuracy and precision is to conduct a pilot study in the area of interest prior to designing a major field survey. The pilot survey will provide information on variability within benthic communities, which then directly affects the required number of replicates and station placement. The analysis of data from a pilot study may also help generate different hypotheses that may alter the sampling and analysis plans to better define the communities.

9.4 STATUS

Many methods to assess sediment quality rely on benthic community structure as a measure of potential ecological effects of pollutants. Benthic community structure has been incorporated into programs with vastly different objectives because the resident biota are sensitive indicators of many kinds of environmental perturbations. Aspects of the status of benthic community structure as a sediment quality assessment tool are discussed in this section.

9.4.1 Extent of Use

Assessment of benthic community structure has been a valued tool in marine, estuarine, and freshwater environments for several decades. Many of the early programs examined benthic communities from an academic viewpoint. Since the 1970s, benthic community structure has been used as a measure of sediment quality. Since then this method has been used to determine the effects of municipal effluents, industrial discharges, eutrophication, organic enrichment, oil spills, and mine tailings disposal (see Section 9.1.1). It has also been used to determine the suitability of sediments for dredged material disposal, to monitor dredged material disposal sites, and to monitor recovery of impacted areas following the cessation of contaminant loading.

9.4.2 Extent to Which Approach Has Been Field-Validated

Because benthic community structure is an *in situ* sediment quality assessment tool, it does not require additional field validation.

9.4.3 Reasons for Limited Use

Although conducting studies of benthic community structure is a common practice, the cost and amount of time required to generate usable results may prevent the method from being implemented by all who could benefit from its use. In fact, the method has been deleted from some programs due solely to cost (Bilyard, 1987). In some situations, costs and time have been reduced by taking the identifications only to the major taxonomic level. This reduction of taxonomic detail frequently reduces the usefulness of the information (Warwick, 1988), which exacerbates a perception by some resource managers that the data are too variable to be useful. Detecting trends within benthic data is not a simple process. However, the proper design and implementation of a field survey will radically increase the probability of producing valuable data and results.

9.4.4 Outlook for Future Use and Amount of Development Yet Needed

The outlook for the future use of benthic community structure as a sediment quality assessment tool is particularly bright because of the continuing development of new data analysis methods by researchers in North America and Europe. The objective of these methods is generally to reduce cost or variability within the data by relating aspects of the distributions of organisms or organism biomass to specific kinds of environmental perturbations. Gray and Mirza (1979) determined that the lognormal distribution of individuals was altered in a predictable manner in the presence of slight organic pollution. A more recent method for detecting pollution effects on marine benthic communities is the species abundance/biomass comparison (ABC) method developed by Warwick (1986). This method proposes that the relationship between the number of individuals among species and the distribution of biomass among species changes in a predictable manner in the presence of organic pollution. Beukema (1988) evaluated the ABC method in an intertidal habitat in the Dutch Wadden Sea and determined that the method "cannot be applied to tidal flat communities without reference to long-term and spatial series of control samples." Yet another benthic community assessment method that remains under development is the Infaunal Trophic Index proposed by Word (1978, 1980). That method is based on changes in the feeding ecology of benthic infauna in relation to organic enrichment. The Benthic Resource Assessment Technique, developed by Lunz and Kendall (1982), quantifies the effects of changes in benthic communities on fish resources. Although the BRAT technique is not a direct assessment of benthic community structure, it provides important information on the relationships among benthic communities and higher level predators, and describes how those relationships may change in the presence of pollutants.

A radically different approach to interpreting long-term changes in benthic community structure involves use of a sediment profile camera. Rhoads and Germano (1986) developed the REMOTS® (remote ecological mapping of the seafloor) system. They use a vessel-deployed sediment-profile camera to photograph vertical sections of the sediment. Although REMOTS®

cannot determine the species composition of the benthic community, it can document relationships between organisms and sediment. Rhoads and Germano (1986) characterized the successional stages of benthic communities and suggested that mapping these stages will permit the detection of changes in benthic communities. When this information is collected as part of a preliminary survey, it can be used to assist in the design of a cost-efficient benthic community survey for obtaining geochemical and biological information.

Additional research is needed on some fundamental aspects of benthic community assessment. These include the development of guidelines for the identification of reference sites or reference values and additional studies into the usefulness of identifying infauna to various taxonomic levels. U.S. EPA is presently examining some aspects of these questions through the Clean Water Act section 301(h) program, including examination of the degree of variability in benthic communities in contaminated and reference areas, development of a quantitative definition of "balanced indigenous populations," and assessment of the effects of overlapping contaminant sources on benthic infaunal communities.

The sediment profile camera has been used for a variety of other purposes including assessing the relationships between sediment quality and eutrophication (Day *et al.*, 1987; Revelas *et al.*, 1987; Rhoads, D.C., 1 May 1989, personal communication), monitoring the perimeter of dredged material disposal sites (Rhoads, D.C., 1 May 1989, personal communication; Diaz, R.J., 1 May 1989, personal communication), and evaluating the overwintering habitat of blue crabs in Chesapeake Bay (Schaffner and Diaz, 1988). With further research, the sediment profile camera may be used for other applications concerning aspects of benthic community structure and sediment quality.

9.5 REFERENCES

Amjad, S., and J.S. Gray. 1983. Use of the nematode/copepod ratio as an index of organic pollution. Mar. Poll. Bull. 14:178-181.

Austin, M.P., and P. Grieg-Smith. 1968. The application of quantitative methods to vegetation survey. II. Some methodological problems of data from rain forest. J. Ecol. 56:827-844.

Beukema, J.J. 1988. An evaluation of the ABC method (abundance-biomass comparison) as applied to macrozoobenthic communities living on tidal flats in the Dutch Wadden Sea. Mar. Biol. 99:425-433.

Bernstein, B.B., and R.W. Smith. 1986. Community Approaches to Monitoring. IEEE Oceans '86 Conference Proceedings, Washington, DC, September 23-25, 1986. pp. 934-939.

Bilyard, G.R. 1987. The value of benthic infauna in marine pollution monitoring studies. Mar. Poll. Bull. 18:581-585.

Bishop, J.D.D., and J.P. Hartley. 1986. A comparison of the fauna retained on 0.5 mm and 1.0 mm meshes from benthic samples taken in the Beatrice Oilfield, Moray Firth, Scotland. Proc. Royal Soc. Edinburgh. 91B:247-262.

Blomqvist, S. 1991. Quantitative sampling of soft-bottom sediments: problems and solutions. Mar. Ecol. Prog. Ser. 72:295-304.

Boesch, D.F. 1977. Application of numerical classification in ecological investigations of water pollution. EPA 600/3-77-033. U.S. Environmental Protection Agency, Corvallis, OR. 115 pp.

Bros, W.E., and B.C. Cowell. 1987. A technique for optimizing sample size (replication). J. Exp. Mar. Biol. Ecol. 114:63-71.

Bryan, G.W., P.E. Gibbs, L.G. Hummerstone, G.R. Burt. 1987. Copper, zinc, and organotin as long-term factors governing the distribution of organisms in the Fal Estuary in Southwest England. Estuaries 10:208-219.

Clifford, H.T., and W. Stephenson. 1975. An introduction to numerical classification. Academic Press, San Francisco, CA. 229 pp.

Clifton, H.E., K.A. Kvenvolden, and J.P. Rapp. 1984. Spilled oil and infaunal activity-modification of burrowing behavior and redistribution of oil. Mar. Environ. Res. 11:111-136.

Cohen, J. 1977. Statistical power analysis for the behavioral sciences. Academic Press, New York, NY.

Cuff, W., and N. Coleman. 1979. Optional survey design: lessons from a stratified random sample of macrobenthos. J. Fish. Res. Bd. Can. 36:351-361.

Dauer, D.M., and W.G. Conner. 1980. Effects of moderate sewage input on benthic polychaete populations. Est. Mar. Sci. 10:335-346.

Day, B., L.C. Schaffner, R.J. Diaz, and J. Ryther, Jr. 1987. Long Island Sound sediment quality survey and analyses. Prepared for National Oceanic and Atmospheric Administration, Rockville, MD. Evans-Hammilton, Inc., Seattle, WA. 113 pp. + appendices.

Day, J.H., J.G. Field, and M.P. Montgomery. 1971. The use of numerical methods to determine the distribution of the benthic fauna across the continental shelf of North Carolina. J. Anim. Ecol. 40:93-125.

Diaz, R.J. 1 May 1989. Personal communication (phone by Ms. Betsy Day, Tetra Tech, Inc., Bellevue, WA, regarding uses of the sediment profile camera system). Virginia Institute of Marine Science, Gloucester Point, VA.

Dobbs, F.L., and J.M. Vozarik. 1983. Immediate effects of a storm on coastal infauna. Mar. Ecol. Prog. Ser. 11:273-279.

Eleftheriou, A., and N.A. Holme. 1984. Macrofauna techniques. pp. 140-216. In: Methods for the Study of Marine Benthos. N.A. Holme and A.D. McIntyre (eds.). Blackwell Scientific Publications, Oxford, U.K.

Elliott, J.M. 1977. Some methods for the statistical analysis of samples of benthic invertebrates. 2d ed. Freshwater Biological Association. Titus Wilson & Son Ltd., Kendal, U.K. 156 pp.

Elmgren, R., S. Hansson, U. Larsson, B. Sundelin, and P.D. Boehm. 1983. The "Tsesis" oil spill: acute and long-term impact on the benthos. Mar. Biol. 73:51-65.

Fabrikant, R. 1984. The effect of sewage effluent on the population density and size of the clam *Parvilucina tenuisculpta*. Mar. Poll. Bull. 15:249-253.

Ferraro, S.P., F.A. Cole, W.A. DeBen, and R.C. Swartz. 1989. Power-cost efficiency of eight macrobenthic sampling schemes in Puget Sound, Washington, USA. Can. J. Fish. Aquat. Sci., 46:2157-2165.

Gauch, H.G. 1982. Multivariate analysis in community ecology. Cambridge University Press, New York, NY. 298 pp.

Gonor, J.J., and P.F. Kemp. 1978. Procedures for quantitative ecological assessments in intertidal environments. EPA 600/3-78-078. U.S. Environmental Protection Agency, Corvallis, OR. 104 pp.

Gray, J.S., and F.B. Mirza. 1979. A possible method for the detection of pollution-induced disturbance on marine benthic communities. Mar. Poll. Bull. 10:142-146.

Green, R.H., and G.L. Vascotto. 1978. A method for analysis of environmental factors controlling patterns of species composition in aquatic communities. Water Res. 12:583-590.

Grizzle, R.E. 1984. Pollution indicator species of macrobenthos in a coastal lagoon. Mar. Ecol. Prog. Ser. 18:191-200.

Holme, N.A., and A.D. McIntyre (eds.). 1984. Methods for the study of marine benthos. Blackwell Scientific Publications, Oxford, U.K. 387 pp.

Ibanez, F., and J. Dauvin. 1988. Long-term changes (1977-1987) in a muddy fine sand *Abra alba-Melinna palmatay* community from the western English Channel: multivariate time-series analysis. Mar. Ecol. Prog. Ser. 49:65-81.

International Joint Commission. 1988a. Procedures for the assessment of contaminated sediment problems in the Great Lakes. IJC, Windsor, Ontario, Canada. 140 pp.

International Joint Commission. 1988b. Options for the remediation of contaminated sediments in the Great Lakes. IJC, Windsor, Ontario, Canada. 78 pp.

Jackson, J.B.C., J.D. Cubit, B.D. Keller, V. Batista, K. Burns, H.M. Caffey, R.L. Caldwell, S.D. Garrity, C.D. Getter, C. Gonzales, H.M. Guzman, K.W. Kaufman, A.H. Knap, S.C. Levings, M.J. Marshall, R. Steger, R.C. Thompson, and E. Weil. 1989. Ecological effects of a major oil spill on Panamanian coastal marine communities. Sci. 243:37-44.

Kleppel, G.S., J.Q. Word, and J. Roney. 1980. Demersal fish feeding in Santa Monica Bay and off Palos Verdes. pp. 309-318. In: Coastal Water Research Project Biennial Report 1979-1980. Southern California Coastal Water Research Project, El Segundo, CA.

Kronberg, I. 1987. Accuracy of species and abundance minimal areas determined by similarity area curves. Mar. Biol. 96:555-561.

Lambshead, P.J.D. 1984. The nematode/ copepod ratio, some anomalous results from the Firth of Clyde. Mar. Poll. Bull. 15:256-259.

Long, B., and J.B. Lewis. 1987. Distribution and community structure of the benthic fauna of the north shore of the Gulf of St. Lawrence described by numerical methods of classification and ordination. Mar. Biol. 95:93-101.

Lunz, J.D., and D.R. Kendall. 1982. Benthic resource analysis technique, a method for quantifying the effects of benthic community changes on fish resources. pp. 1021-1027. In: Conference Proceedings on Marine Pollution, Oceans 1982. National Oceanic and Atmospheric Administration, Office of Marine Pollution Assessment, Rockville, MD.

Mirza, F.B., and J.S. Gray. 1981. The fauna of benthic sediments from the organically enriched Oslofjord, Norway. J. Exp. Mar. Biol. Ecol. 54:181-207.

Moore, S.F., and D.B. McLaughlin. 1978. Design of field experiments to determine the ecological effects of petroleum in intertidal ecosystems. Water Res. 12:1091-1099.

Nalepa, T.F., M.A. Quigley, and R.W. Ziegler. 1988. Sampling efficiency of the ponar grab in two different benthic environments. J. Great Lakes Research 14:89-93.

Nemec, A.F.L., and R.O. Brinkhurst. 1988a. Using the bootstrap to assess statistical significance in the cluster analysis of species abundance data. Can. J. Fish. Aquatic Sci. 45:965-970.

Nemec, A.F.L., and R.O. Brinkhurst. 1988b. The Fowlkes-Mallows statistic and the comparison of two independently determined dendrograms. Can. J. Fish. Aquatic Sci. 45:971-975.

Parker, H.R. 1975. The study of benthic communities. A model and review. Elsevier Oceanography Series 9. Elsevier, Amsterdam.

Pearson, T.H., and R. Rosenberg. 1978. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. Oceanogr. Mar. Biol. Annu. Rev. 16:229-311.

Perez, K. 1 May 1989. Personal communication (phone by Ms. Betsy Day, Tetra Tech, Inc., Bellevue, WA, regarding mesocosm experiments to determine rates of benthic recovery). U.S. Environmental Protection Agency, Environmental Research Laboratory, Narragansett, RI.

Preston, F.W. 1948. The commonness, and rarity, of species. Ecology 29:254-283.

PTI and Tetra Tech. 1988. Elliott Bay Action Program: Analysis of toxic problem areas. Draft Report. Prepared for the U.S. Environmental Protection Agency, Region X, Office of Puget Sound. Tetra Tech, Inc., Bellevue, WA.

Raffaelli, D. 1987. The behavior of the nematode/copepod ratio in organic pollution studies. Mar. Environ. Res. 23:135-152.

Revelas, E.C., D.C. Rhoads, and J.D. Germano. 1987. San Francisco Bay sediment quality survey and analyses. Prepared for National Oceanic and Atmospheric Administration, Rockville, MD. Science Applications International Corporation, Newport, RI. 127 pp. + appendices.

Rhoads, D.C. 1 May 1989. Personal communication (phone by Ms. Betsy Day, Tetra Tech, Inc., Bellevue, WA, regarding uses of the REMOTS® sediment profile camera system). Science Applications International Corporation, Woods Hole, MA.

Rhoads, D.C., and L.F. Boyer. 1982. The effects of marine benthos on physical properties of sediments: a successional perspective. pp. 3-52. In: Animal-Sediment Relations. P.L. McCall and M.J.S. Trevesz (eds.). Plenum Press.

Rhoads, D.C., and J.D. Germano. 1986. Interpreting long-term changes in benthic community structure: A new protocol. Hydrobiologia.

Rhoads, D.C., and D.K. Young. 1970. The influence of deposit-feeding organisms on sediment stability and community trophic structure. J. Mar. Res. 28:150-178.

Rhoads, D.C., P.L. McCall, and J.Y. Yingst. 1978. Disturbance and production on the estuarine seafloor. Amer. Sci. 66:577-586.

Rygg, B. 1985a. Distribution of species along pollution-induced diversity gradients in benthic communities in Norwegian fjords. Mar. Poll. Bull. 12:469-474.

Rygg, B. 1985b. Effect of sediment copper on benthic infauna. Mar. Ècol. Prog. Ser. 25:83-89.

Rygg, B. 1986. Heavy-metal pollution and log-normal distribution of individuals among species in benthic communities. Mar. Poll. Bull. 17:31-36.

SAIC. 1991. PSDDA 1990 monitoring: Post-disposal surveys of Elliot Bay and Port Gardner. Final Report. Prepared for Washington Department of Natural Resources. Prepared by Science Applications International Corporation, Bothell, WA.

Saila, S.B., R.A. Pikanowski, and D.S. Vaughan. 1976. Optimum allocation strategies for sampling benthos in the New York Bight. Est. Coast. Mar. Sci. 4:119-128.

Sanders, H.L. 1968. Marine benthic diversity: a comparative study. Amer. Nat. 102:243-282.

Santos, S.L., and J.L. Simon. 1980. Response of soft-bottom benthos to annual catastrophic disturbance in a south Florida estuary. Mar. Ecol. Prog. Ser. 3:347-355.

Schaffner, L.C., and R.J. Diaz. 1988. Distribution and abundance of overwintering blue crab *Callinectes sapidus* in the lower Chesapeake Bay. Estuaries 11:68-72.

Self, S.G., and R.H. Mauritsen. 1988. Power/sample size calculations for generalized linear models. Biometrics 44:79-86.

Shiells, G.M., and K.J. Anderson. 1985. Pollution monitoring using the nematode/copepod ratio, a practical application. Mar. Poll. Bull. 16:62-68.

Shin, P.K.S. 1982. Multiple discriminant analysis of macrobenthic infaunal assemblages. J. Exp. Mar. Biol. Ecol. 59:39-50.

Sokal, R.R., and F.J. Rohlf. 1981. Biometry. 2d ed. W.H. Freeman and Company, San Francisco, CA, 859 pp.

Stephenson, M., and G.L. Mackie. 1988. Multivariate analysis of correlations between environmental parameters and cadmium concentrations in *Hyallella azteca* (Crustacea: Amphipoda) from central Ontario lakes. Can. J. Fish. Aquatic Sci. 45:1705-1710.

Stephenson, W., W.T. Williams, and G.W. Lance. 1970. The macrobenthos of Moreton Bay. Ecol. Managr. 40:459-494.

Stephenson, W., W.T. Williams, and S.D. Cook. 1972. Computer analyses of Petersen's original data on bottom communities. Ecol. Monogr. 42:387-415.

Stephenson, W., W.T. Williams, and S.D. Cook. 1974. The benthic fauna of soft bottoms, Southern Moreton Bay. Mem. Qd. Mus. 17:73-123.

Striplin, B.D., D.R. Kendall, and J.D. Lunz. 1991. Environmental conditions at two PSDDA open-water disposal sites: do they match the predictions? Proceedings, Puget Sound Research '91. p. 281-288.

Striplin, P.L. 1987. Resource utilization by *Astropecten verrilli* along gradients of organic enrichment. M. Sc. Thesis. California State University at Long Beach, Long Beach, CA. 108 pp. + appendices.

Swartz, R.C. 1978. Techniques for sampling and analyzing the marine macrobenthos. EPA 600/3-78-030. U.S. Environmental Protection Agency, Corvallis, OR. 27 pp.

Swartz, R.C., W.A. DeBen, F.A. Cole, and L.C. Bentsen. 1980. Recovery of the macrobenthos at a dredge site in Yaquina Bay, Oregon. pp. 391-408. In: Contaminants and Sediments, Vol. 2. R. Baker (ed.). Ann Arbor Science, Ann Arbor, MI.

Swartz, R.C. 15 March 1989. Personal communication (phone by Ms. Betsy Day, Tetra Tech, Inc., Bellevue, WA regarding status of replication study using samples collected during the Everett Harbor Action Program survey). U.S. Environmental Protection Agency, Newport, OR.

Tagatz, M.E., G.R. Plaia, C.H. Deans, and E.M. Lores. 1983. Toxicity of creosote-contaminated sediment to field-and laboratory-colonized estuarine benthic communities. Environ. Tox. Chem. 2:441-450.

Tarazona, J., H. Salzwedel, and W. Arntz. 1988. Oscillations of macrobenthos in shallow waters of the Peruvian central coast induced by El Niño 1982-83. J. Mar. Res. 46:593-611.

Tetra Tech. 1986a. Quality assurance/quality control (QA/QC) for 301(h) monitoring programs: guidance on field and laboratory methods. Prepared for the U.S. Environmental Protection Agency, Office of Marine and Estuarine Protection, Marine Operations Division, Washington, DC. Tetra Tech, Inc., Bellevue, WA.

Tetra Tech. 1986b. Recommended protocols for measuring selected environmental variables in Puget Sound. Prepared for the Puget Sound Estuary Program, U.S. Environmental Protection Agency, Region X, Seattle, WA. Tetra Tech, Inc., Bellevue, WA.

Tetra Tech. 1987. Technical support document for ODES statistical power analysis. Prepared for Marine Operations Division, Office of Marine and Estuarine Division, Office of Marine and Estuarine Protection, U.S. Environmental Protection Agency. Tetra Tech, Inc., Bellevue, WA. 34 pp. + appendices.

Tetra Tech. 1988. Commencement Bay nearshore/tideflats feasibility study. Prepared for Washington Department of Ecology and U.S. Environmental Protection Agency. Tetra Tech, Inc., Bellevue, WA.

Tilley, S., D. Jamison, J. Thornton, B. Parker, and J. Malek. 1988. Management plans technical appendix. Prepared for Puget Sound Dredged Disposal Analysis. U.S. Army Corps of Engineers, Seattle, WA.

Vezina, A.F. 1988. Sampling variance and the design of quantitative surveys of the marine benthos. Mar. Biol. 97:151-155.

Vidakovic, J. 1983. The influence of raw domestic sewage on density and distribution of meiofauna. Mar. Poll. Bull. 14:84-88.

Warwick, R.M. 1986. A new method for detecting pollution effects on marine macrobenthic communities. Mar. Biol. 92:557-562.

Warwick, R.M. 1988. The level of taxonomic discrimination required to detect pollution effects on marine benthic communities. Mar. Poll. Bull. 19:259-268.

Warwick, R.M., T.H. Pearson, and Ruswahyuni. 1987. Detection of pollution effects on marine macrobenthos: further evaluation of the species abundance/biomass method. Mar. Biol. 95:-193-200.

Washington, H.G. 1984. Diversity, biotic, and similarity indices. A review with special relevance to aquatic ecosystems. Water Res. 18:653-694.

Winer, B.J. 1971. Statistical principles in experimental design. McGraw-Hill Book Company, New York, NY.

Word, J.Q. 1976. Biological comparison of grab sampling devices. pp. 189-194. In: Coastal Water Research Project Annual Report. Southern California Coastal Water Research Project, El Segundo, CA.

Word, J.Q. 1978. The infaunal trophic index. pp. 19-39. In: Coastal Water Research Project Annual Report for 1978. Southern California Coastal Water Research Project, El Segundo, CA.

Word, J.Q. 1980. Classification of benthic invertebrates into infaunal trophic index feeding groups. pp. 103-121. In: Coastal Water Research Project. Biennial Report of the years 1979-1980. W. Bascom (ed.). Southern California Coastal Water Research Project, Long Beach, CA.

Word, J.Q., B.L. Myers, and A.J. Mearns. 1977. Animals that are indicators of marine pollution. pp. 199-206. In: Coastal Water Research Project Annual Report. Southern California Coastal Water Research Project, El Segundo, CA.

Sediment Quality Triad Approach

Peter M. Chapman
E.V.S. Consultants Ltd.
195 Pemberton Avenue, North Vancouver, BC, Canada V7P 2R4
Phone (604) 986-4331, FAX (604) 662-8548

The Sediment Quality Triad (Triad) approach is an effects-based approach to describe sediment quality. It typically incorporates measures of sediment chemistry, sediment toxicity, and benthic infauna communities, although other variables can be used. This combination method is both descriptive and numeric. It is most commonly used to describe sediment qualitatively, but has also been used to generate chemical-specific sediment quality criteria (Chapman, 1986, 1989; Long, 1989). One application of the Triad approach, the Apparent Effects Threshold (AET), is described in detail in the following chapter (Chapter 11).

10.1 SPECIFIC APPLICATIONS

10.1.1 Current Use

The Triad approach can be used to determine the extent of pollution-induced degradation of sediments in a non-numerical, multiple-chemical mode (e.g. Chapman *et al.*, 1986, 1987a, 1991a; Chapman and Power, 1990; Chapman, 1990). It can also be used to determine numerical sediment quality criteria directly (e.g. Chapman, 1986, 1989) and, through manipulations, to determine AET values (see Chapter 11).

The AET is only one possible method of evaluating triad data and is directed solely at determining numeric sediment quality values (Chapman *et al.*, 1991b, 1991c). The triad approach has been used in marine coastal waters on the west coast of North America (e.g., Puget Sound, San Francisco Bay, and Vancouver Harbor, Canada), in the Gulf of Mexico, in freshwater environments including the Great Lakes, and in the North Sea (Long and Chapman, 1985; Chapman, in press; Chapman *et al.*, 1986, 1987a, in press; Chapman and Power, 1990; Cross *et al.*, 1991, in review). Current uses of the Triad approach are summarized in Table 10-1 and discussed in Section 10.3.1, Environmental Applicability.

10.1.2 Potential Use

The Sediment Quality Triad approach can also be used to meet the following objectives:

- To identify problem areas of sediment contamination where pollution-induced degradation is occurring;
- To prioritize and rank degraded areas and their environmental significance; and
- To predict where such degradation will occur based on levels of contamination and toxicity.

The Triad approach can be used in any number of situations and is not restricted to aquatic sediments. For example, it can be used in water column work with phytoplankton and in terrestrial hazardous waste dump studies with other organisms of concern. Other uses are described in Section 10.3.1. A complete description of the Triad in the context of integrated assessments is provided in Chapman *et al.*, 1991b.

10.2 DESCRIPTION

10.2.1 Description of Method

The Triad approach consists of three components (Figure 10-1):

- Sediment chemistry—to measure chemical contamination;

Table 10-1. Current Uses of the Sediment Quality Triad Approach.

Use	Comment	General Locations Where Implemented*
Prioritize areas for remedial actions	Most common usage to date	PS, GM, SF, VH, FW
Determine size of areas	Assuming increasing importance	PS
Verify quality of reference areas	Assuming increasing importance	PS
Determine contaminant concentrations always associated with effects	Common usage; can result in numerical sediment quality criteria and setting of standards	PS, NS
Describe ecological relationships between sediment properties and biota at risk	Along with setting standards and criteria, provides for proactive approach to environmental protection	PS, VH, FW, NS

*PS = Puget Sound, various locations (Long and Chapman, 1985).
GM = Gulf of Mexico, oil platform (Chapman *et al.*, 1991a; Chapman and Power, 1990).
SF = San Francisco Bay, various locations (Chapman *et al.*, 1986, 1987a).
VH = Vancouver Harbor, Canada, various locations (Chapman *et al.*, 1989; Cross *et al.*, 1991; Cross *et al.*, in review).
FW = Various freshwater environments (Malueg *et al.*, 1984; Chapman unpublished data; Rogers, North Texas State, unpublished data; Wiederholm *et al.*, 1987).
NS = North Sea (Chapman, in press; Chapman *et al.*, in press).

- Sediment bioassays—to measure toxicity;
- *In situ* biological variables — to measure *in situ* alteration (e.g., a change in benthic community structure).

The three components provide complementary data. No single component of the Triad approach can be used to predict the measurements of the other components. For instance, sediment chemistry provides information on contamination but not on biological effects. Sediment bioassays provide direct evidence of sediment toxicity. However, the laboratory conditions under which bioassays are conducted may not accurately reflect field conditions of exposure to toxic chemicals. *In situ* alteration of resident biota measured by infauna community analyses provides direct evidence of contaminant-related effects in the environment, but only if confounding effects not related to pollution (e.g., competition, predation, recruitment cycles, sediment type, salinity, temperature, recent dredging) can be excluded. In particular, because the toxicity of a chemical substance in sediments may vary with its concentration and with the conditions within a specific sediment, the importance of any particular concentration of a chemical or suite of chemicals in sediments cannot be determined solely from chemical measurements. Sediment conditions include grain size, organic content, pH, Eh, chemical form, and presence of other chemicals.

The three components of the Triad approach integrate chemical and biological response data. They also provide the strongest evidence for identifying pollution-induced degradation. For instance, if there are high levels of sediment contamination, toxicity, and biological alteration, the burden of evidence indicates degradation. Conversely, low levels of sediment contamination, toxicity, and biological alteration indicate non-degraded conditions. Conclusions that can be drawn from intermediate responses are listed in Table 10-2.

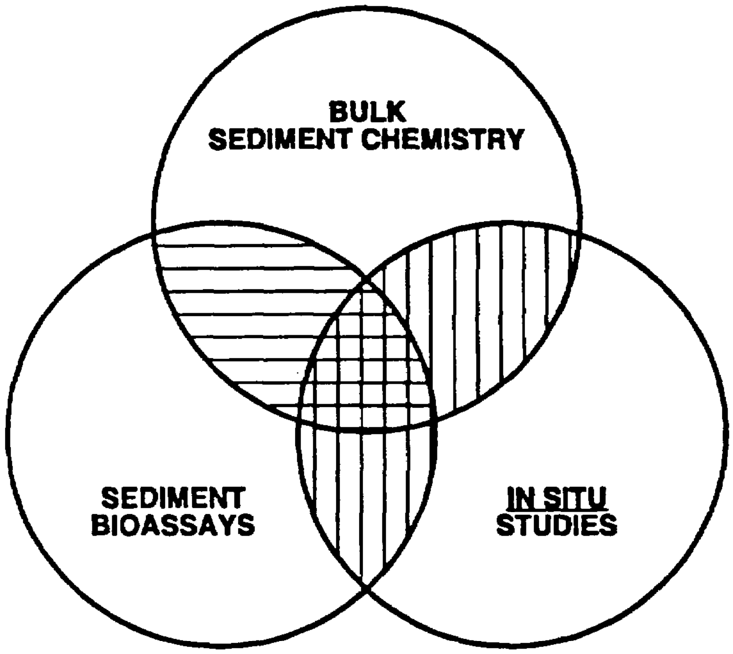


Figure 10-1. Conceptual Model of the Sediment Quality Triad.
The Triad combines data from chemistry, toxicity bioassays, and in situ *studies. Chemistry and bioassay estimates are based on laboratory measurements with field-collected sediments.* In situ *studies generally include, but are not limited to, measures of benthic community structure. Areas where the three facets of the Triad show the greatest overlap (in terms of either positive or negative results) provide the strongest data for determining sediment quality criteria.*

10.2.1.1 Objectives and Assumptions

The objectives of the Triad approach are to independently measure sediment contamination, sediment toxicity, and biological alteration, and then use the burden of evidence to assess sediment quality based on all three sets of measurements.

The following assumptions apply:

- The approach allows for (1) the interactions between contaminants in complex sediment mixtures (e.g., additivity, antagonism, synergism); (2) the actions of unidentified toxic chemicals; and (3) the effect of environmental factors that influence biological responses (including toxicant concentrations).
- Selected chemical contaminant concentrations are appropriate indicators of overall chemical contamination.
- Bioassay test results and values of selected benthic community structure variables are appropriate indicators of biological effects.

These components are presently often treated in an additive manner, with each having equal

Table 10-2. Possible Conclusions Provided by Using the Sediment Quality Triad Approach.

Possible Outcome	Contamination	Toxicity	Alteration	Possible Conclusions
1.	+	+	+	Strong evidence for pollution- induced degradation
2.	-	-	-	Strong evidence for absence of pollution-induced degradation
3.	+	-	-	Contaminants are not bioavailable
4.	-	+	-	Unmeasured chemicals or conditions exist that have the potential to cause degradation
5.	-	-	+	Alteration is probably not due to toxic chemical contamination
6.	+	+	-	Toxic chemicals are stressing the system
7.	-	+	+	Unmeasured toxic chemicals are causing degradation
8.	+	-	+	Chemicals are not bioavailable or alteration is not due to toxic chemicals

*+ = Measured difference between test and control or reference conditions.
- = No measurable difference between test and control or reference conditions.

weight because there is insufficient information available to assign weightings.

10.2.1.2 Level of Effort

Ideally, the Triad approach would be based on the use of synoptic data. Sediments for analysis of toxicity should come from the same composited homogenate, as originally detailed by Chapman (1988), ideally from field rather than solely laboratory test replicates. Benthic infauna samples should be collected at the same sampling locations. Chemistry and bioassay sediments are collected (usually by remote grab), transferred to a solvent-rinsed glass or stainless steel bowl, and thoroughly homogenized by stirring with a glass or stainless steel spatula until textural and color homogeneity are achieved. The homogenized sediments are then placed in appropriate sampling containers. In general, chemistry and bioassay samples should include field rather than laboratory replication. Benthic infaunal samples are collected at the same location. In the absence of initial sampling to determine the optimum level of replication at a site, five field replicate benthic samples are recommended per station (see Chapter 8, Methods). Coincident rather than synoptic sampling is possible (e.g., Long and Chapman, 1985); however, spatial heterogeneity in sediment contamination and toxicity make such data difficult to interpret (Swartz *et al.*, 1982).

Adequate quality QA/QC measures must be followed in all aspects of the study, from field sampling through laboratory analyses and data entry. Detailed QA/QC procedures are available through international (e.g., Keith *et al.*, 1983) and regional publications (e.g., Tetra Tech, 1986a).

The first component of the Triad involves identification and quantification of inorganic and organic contaminants present in the sediments. Chemical analytes measured are generally restricted by equipment, technology, and the availability of funds and facilities. Local concerns and existing data also affect target analytes measured. Cost, if a factor, must be balanced against the need for an analytical database sufficiently large

Table 10-3. Example Analytes and Detection Limits for Use in the Chemistry Component of Sediment Quality Triad Approach.

Analyte	Detection Limit	Analyte	Detection Limit
Conventionals (mg/kg, dry)			
Grain size	n/a	Biphenyl	5
TOC[a]	n/a	Perylene	5
Sulfides	0.5	Coprostanol	10
Acid volatile sulfide (AVS)[b]	n/a	Ammonia	0.5
Inorganics (mg/kg, dry)		op'-DDD	0.15
Arsenic	0.05	op'-DDE	0.25
Iron	2.5	op'-DDT	0.15
Chromium	1.0	pp'-DDD	0.15
Copper	0.5	pp'-DDE	0.10
Cadmium	0.05	pp'-DDT	0.10
Lead	0.05	Dieldrin	0.10
Mercury	0.01	Heptachlor	0.10
Nickel	1.0	Hexachlorobenzene	0.10
Silver	0.05	Lindane	0.15
Selenium	0.05	Mirex	0.10
Zinc	0.5	PCBs[d]	2.5
		PCP[e]	1.0
		TCP[f]	1.0
Organics (μg/kg, dry)			
LPAH[c]	5		
Benzo(a)pyrene	10		
Benzo(e)pyrene	10		
Benzo(a)anthracene	10		
Chrysene	10		
Dibenzoanthracene	16		
Fluoranthene	5		
Pyrene	5		

The detection limits are the instrumental estimates. Actual detection limits may be higher because of matrix effects.

[a] TOC = total organic carbon.
[b] AVS = AVS methodology is described by the U.S. EPA (1991); modifications are expected. Contact Christopher Zarba at (202) 475-7326 to obtain latest protocols.
[c] LPAH = low-molecular-weight polycylic aromatic hydrocarbons (includes acenaphthene, anthracene, naphthalene and methylated naphthalenes, fluorene, phenanthrene, and methylated phenanthrenes).
[d] PCBs = polychlorinated biphenyls.
[e] PCP = pentachlorophenol.
[f] TCP = tetrachlorophenol.

to allow determination of the presence (or absence) of known toxicants of concern.

An example of some of the types and classes of compounds required to provide a reasonable characterization of chemical contamination is shown in Table 10-3.

Total organic carbon and grain size are measured to provide a basis for normalizing the data to different types of sediments. Acid volatile sulfides (AVS) provide information for determining metals availability from sediments. Coprostanol, an indicator of human waste, can be

Table 10-4. Possible Static Sediment Bioassays.

Bioassay	Duration	Endpoint	Amount of Sediment Required (L)
Marine Waters			
Rhepoxynius abronius[a] (adult amphipod)	10 days	Survival, avoidance	1.5
Bivalve Larvae development	48 hours	Survival, development	0.5
Neanthes sp. (juvenile polychaetes)	20 days	Survival, growth	2.0
Fresh Waters			
Hyalella azteca (adult amphipod)	10 days	Survival, avoidance	1.5
Daphnia magna (water flea)	10 days	Survival, reproduction	0.5
Chironomus tentans (juvenile insect)	25 days	Survival, growth	1.5
Estuarine Waters			
Eohaustorius estuarius (adult amphipod)	10 days	Survival, avoidance	1.5

[a] Note: Other options include but are not necessarily restricted to *Ampelisca abdita, Corophium volutator, Grandidierella japonica, Foxiphalus xiximeus.*

measured to differentiate sewage inputs from industrial inputs.

The second Triad component involves identification and quantification of toxicity based on laboratory tests using field-collected sediments. Ideally, one would test the toxicity of the sediments to all ecologically and commercially important fauna living in or associated with the sediments. For logistical reasons, a small number of bioassays is conducted to cover as wide a range as possible of organism type, life cycle, exposure route, and feeding type. The number of tests undertaken is affected by the same constraints as those mentioned for sediment chemistry analyses.

Possible static sediment bioassays that provide a reasonable characterization of the degree of toxicity are shown in Table 10-4. Obvious omissions from this list include full life-cycle chronic tests, and genotoxic or cytotoxic response tests. Such tests merit consideration for inclusion when proven accepted methods become available (e.g., Long and Buchman, 1989).

The final Triad component involves the evaluation of *in situ* biological alteration. Generally, this component is provided by benthic infauna community data because benthic organisms are relatively sessile and location-specific. Histopathology of bottom fish has also been used for this Triad component (Chapman, 1986), but for areawide rather than site-specific studies, because these fish are relatively mobile. Several variables in combination are effective in characterizing benthic community structure for the Triad approach: numbers of taxa, numerical dominance, total abundance, and percentage composition of major taxonomic groups. In the marine environment, this last category includes any or all of polychaetes, amphipods, molluscs, and echinoderms. In the freshwater environment, oligochaetes, chironomids, and other major insect groups would fit into the last category.

Sediment chemistry, toxicity, and benthic infauna data are combined in the Triad approach to assess the degree of degradation of each station and of each site (see Figure 10-1). All data are compared on a quantitative basis and can be normalized to reference site values by converting them to ratio-to-reference (RTR) values as described by Chapman *et al.* (1986, 1987a) and Chapman (1990). The reference site chosen (either *a priori* or *a posteriori*) is generally the least contaminated site of those sampled, and ideally its sediment and other characteristics (e.g., water depth) would be similar to those of the other sites. To determine RTR values, the values of specific variables (e.g., normalized concentration of a particular metal, percent mortality in a particular bioassay, number of taxa) are divided by the corresponding reference values. This process normalizes the data so that they can be compared even when, for instance, there are large differences in the units of measurement. The reference site may be a single station (whose RTR value is 1.0 by definition) or an area containing several stations for which data are averaged.

The RTR criterion is based but does not depend on the assumption that the reference site concentrations are indicative of reference or background conditions. The degree to which chemical concentrations are elevated above the mean reference concentrations at a selected site is used as the criterion for selecting chemicals most likely to be anthropogenically enriched and of concern. An index of contamination can be calculated for each station by separately determining RTR values for groups of similar chemicals (e.g., metals, PAH, chlorinated organics) and then, assuming additivity, combining these values as a single mean chemistry RTR value. Similarly, an index of toxicity can be calculated by combining bioassay RTR values as a single mean value. Finally, an index of biological alteration can be calculated in the same manner as is toxicity, using benthic community structure data. The indexes of contamination can be used to rank stations. These summary ranks can also be compared with the ranks generated using the sediment bioassay and infaunal data.

The composite RTR values for each Triad component can also provide useful visual indexes. These values can be plotted on scales with a common origin and placed at 120 degrees from each other such that each of the three values becomes the vertex of a triangle. The relative degree of degradation is derived by calculating and comparing the areas of the triangles for each station or site. Examples of such triaxial plots are shown in Figure 10-2, for the eight possible situations shown in Table 10-2. These plots also provide a visual guide to the characteristics of background or reference stations. Because reference data usually involve a site containing more than one reference station, RTR comparisons should also be made against individual reference stations. Alden (1992) provides a method for determining confidence limits for such triaxial plots. Non-RTR methods of Triad data analysis are outlined in Section 10.2.1.2.3, Types of Data Required.

10.2.1.2.1 Type of Sampling Required

As described, synoptic sampling is preferred for all three Triad components. Any reasonable sampling procedure can be used if it provides suitable sediment samples for quantifying sediment contamination, toxicity, and biological alteration. To date, studies have used remote samplers such as a 0.1-m^2 Van Veen grab operated from a vessel.

10.2.1.2.2 Methods

Typical variables included in the chemical analyses and sediment bioassays are listed in Tables 10-3 and 10-4, respectively. Details for benthic infauna analyses are provided in Chapter 8. Although unit costs vary, costs are generally on the order of $1,500 for three separate replicated (n=5) sediment bioassays, $1,500 for unreplicated

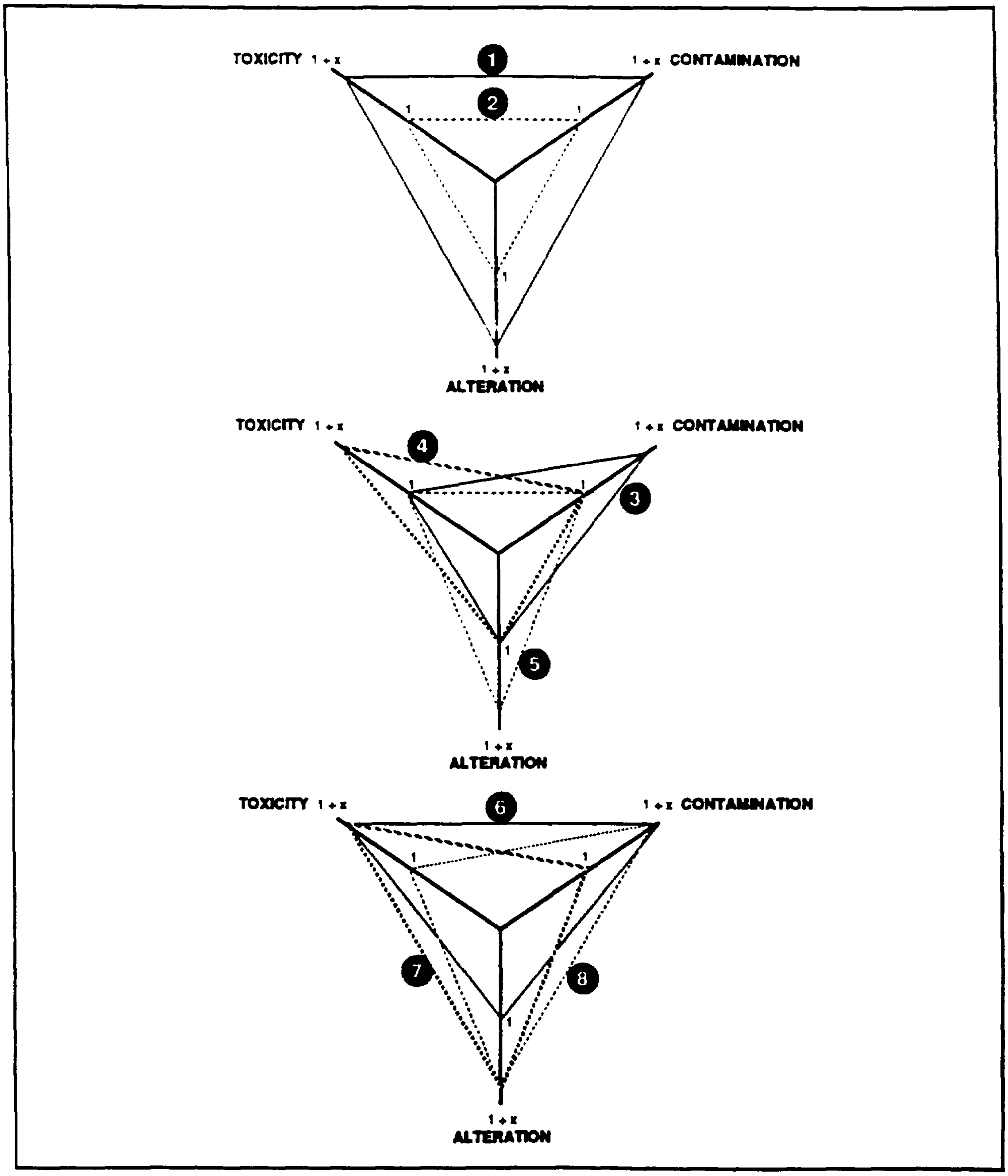


Figure 10-2. Sediment Quality Trial Triaxial Plots for the Eight Possible Situations Shown in Table 10-2. *The Sediment Quality Triad determined, in the example situation, for each of the eight possible outcomes described in Table 10-2. Toxicity, contamination, and alteration are shown normalized to Ratio-to-References values as described by Chapman et al. (1986, 1987a), 1.0 = reference conditions. Note that the exact symmetry in these examples would not be routinely expected in actual studies.*

chemical analyses, and $2,500 for replicated (n=5) benthos.

10.2.1.2.3 Types of Data Required

Standard measurements of chemistry, toxicity, and biological alteration are required. These measurements can then be combined, as described above. Detailed data calculations and analyses are as follows:

Data Calculations - Benthic Data

- Calculate/determine endpoints
  - taxa richness
  - total abundance
  - numerical dominance
  - species diversity
  - mean abundances of all species of major taxa (e.g. polychaetes, amphipods, chironomids, oligochaetes)

- Cluster Analysis
  - e.g., using mean numbers of individuals per taxa present at each station tested.

Data Calculations - Chemistry

- Bulk concentration normalized to dry weight

- Organic carbon normalized concentration of organic compounds

- Normalize to percent fines, sand, silt, and clay

- AVS normalized concentration of metals (DiToro *et al.*, 1990; DeWitt *et al.*, 1990)

- Summarize means, standard deviations, ranges for each parameter at each site.

Data Calculations - Bioassay

- Between station differences in mean response, ANOVA, multiple comparison tests.

- Paired comparison with control response.

- Comparison of mean response with lower prediction limit (LPL) (DeWitt *et al.*, 1988); this comparison addresses possible grain-size effects on amphipods.

Non-RTR Methods of Triad Data Analysis

The traditional reduction technique of calculating RTRs (by translating resultant measures to proportions of comparable values obtained for the reference site) has the following problems (Cross *et al.*, 1991; Cross *et al.*, in review):

- Substantial loss of information during the conversion of multivariate data into single proportional indexes;

- Loss of any spatial relational information;

- Inability to statically assess significance of spâtial impacts; and

- Requirement of an appropriate reference station.

In addition, Triad results could be strongly influenced by the presence of unmeasured toxic contaminants that may or may not covary with measured chemicals (Chapman, 1990). The RTR approach is useful in specific situations and with defined limitations; however, the following options are useful for reducing or removing the problems identified.

Ranking—In addition to RTRs, rankings can also be assigned to biological, chemical, and toxicological data for statistical comparisons of the data. Using the chemistry data as an example, the sample with the lowest level of a chemical is scored as 1 and the highest is scored with a number that is equal to the number of time periods or samples that are to be ranked. Tied data should be scored by calculating an average of the tied ranks. Each site will have a rank for each biological, chemical, and toxicological parameter. An overall mean rank for each site can be calculated using each of the parameters. This effectively determines how each site compares to each of the other sites.

Average ranks for biological, chemical, and toxicological data can also be calculated and can be compared using Kendall's coefficient of concordance (Zar, 1984). High concordance will indicate that biological, chemical, and toxicological parameters are changing in the same direction (improving or degrading). Low concordance will indicate that biological, chemical, toxicological data are changing independently of each other.

Multivariate Analysis—Multivariate analysis comprises data matrix preparation, analysis independent of the Triad components, analysis concurrent with the Triad components, and Mantel's test. Each of these is briefly described here.

Data Matrix Preparation

For each Triad component, data are standardized to common units where possible and incorporated into separate matrices for analysis and interpretation.

Benthos: Data are abundance of each taxon per grab sample; transformed to log (x+1).

Chemistry: Values less than the detection limit are omitted to maintain the integrity of the matrix. Remaining data are log-transformed.

Bioassay: Because of the number of independent bioassays and differing endpoints (e.g., mortality, avoidance, reburial, etc.), these data cannot be standardized to common internal units. Various transformations (arsine square root, log, etc.) may be used as required.

Independent Analysis of the Triad Components

Each matrix is analyzed separately to determine environmental impact as provided by each independent approach. Community classification analysis may be performed for each data matrix using cluster analysis. "Boot-strapping" techniques developed by Nemec and Brinkhurst (1988a, 1988b) can be used to test whether clusters of samples differ significantly from each other.

Concurrent Analysis of the Triad Components

The ecological ordination technique, principal components analysis (PCA), can be used to examine relationships between benthos community structure, toxicology, and the physical-chemical attributes of the bottom sediments, (Cross *et al.*, 1991, in review). PCA is used to reduce the multidimensionality of the benthos data, creating two variables (principal component or PC) from the original matrix of many variables (taxon abundances). These PCs can then be correlated with PCs derived from physical-chemical data or bioassay results, or with individual physical or chemical parameters. High correlations among PCs from the three Triad components indicate agreement or concordance of impact assessments.

Correlations of PCs from benthic data (or bioassay data) with individual chemical parameters can be used to assess or develop sediment quality criteria. The impacts associated with existing criteria can be expressed as a PC score for benthic data, calculated from a regression of these scores on chemical concentrations. Sediment quality criteria could also be developed by predicting the chemical concentration associated with a significant impact on the benthic community, provided that "significant impact" could be unequivocally associated with a particular PC score or range of scores.

Mantel's Test

Another method that can be used to determine whether different components of the Triad are related is Mantel's test (Mantel, 1967; Legendre and Fortin, 1989). Mantel's test uses a randomization procedure that calculates the probability that two distance matrices are more similar than would be expected by chance alone. Multivariate (or univariate) distance between each of the sites (observations) can be calculated using data from each component of the Triad. For example, to develop a distance matrix based on toxicity test results, each of the toxicology variables would be used to develop the distance. Similar matrices would be calculated for benthos and chemistry data.

The randomization procedure ensures that the relationships between two distance matrices are real and not spurious. The distance between two stations (A and B) is always partially related to the distance between these two and other stations (e.g., A and C, B and C). Mantel's test avoids the possibility of spurious correlations by calculating correlations between the two matrices based on random samples, and comparing the actual correlation with the distribution based on the random samples.

10.2.1.2.4 Necessary Hardware and Skills

Appropriate sampling equipment and trained field and laboratory personnel are required for chemical analyses, toxicity testing, and benthic infaunal analyses. Although the equipment required can be both costly and sophisticated, it is commonly necessary for sediment contamination investigations. The necessary equipment, facilities, and expertise are generally available through a wide variety of government, university, commercial, and private facilities.

10.2.1.3 Adequacy of Documentation

Documentation for use of this method is provided by Long and Chapman (1985), Chapman (1986, 1989, 1990), and Chapman *et al.* (1986, 1987a, 1991a, 1991b). Other investigators have also successfully applied this method (cf. Chapman *et al.*, 1991c).

10.2.2 Applicability of Method to Human Health, Aquatic Life, or Wildlife Protection

This approach is directly applicable to the protection of aquatic life. To date, only benthic invertebrates and fish have been used to assess *in situ* biological effects and sediment toxicity. Protection of aquatic life may indirectly protect wildlife (e.g., wading birds feeding on benthos) and humans (e.g., via consumption of aquatic life). The approach can be directly applicable to human health and wildlife protection if the Triad components are redirected towards issues such as bacterial contamination and toxic contaminant bioaccumulation. For instance, Triad could be used in three ways to address bacterial problems: (1) measure bacterial contamination in water or sediment, (2) measure bacterial diseases or concentrations in tissues, and (3) perform laboratory tests to quantify relationships between sediment/water concentrations and effects. Toxic contaminant bioaccumulation could be addressed by these uses of the Triad approach: (1) measure toxic contaminant concentrations in water or sediment, (2) measure bioconcentration/biomagnification in tissues, and (3) perform laboratory tests to determine effects related to bioconcentration and biomagnification.

10.2.3 Ability of Method to Generate Numerical Criteria for Specific Chemicals

The Triad approach has been used to generate criteria for three contaminants: lead, PAH, and PCBs (Chapman, 1986). These criteria were developed in Puget Sound by examining large data sets to identify contaminant areas and concentrations that were associated with no or minimal biological effects. The criteria fall within a factor of 2 to 10 of values generated for these contaminants by the screening-level concentration (see Chapter 11, Section 11.1.1.), the AET approach (see Chapter 11), and laboratory toxicity methods (Chapman *et al.*, 1987b). As detailed by Chapman (1989), the AET application of the Triad concept provides criteria for benthic infauna and each bioassay conducted, whereas the latter combines all bioassay and *in situ* biological effects data to provide a single value, interpretation, or analysis. However, there has been little work since Chapman (1986) on development of the

Triad approach for the production of numerical sediment quality criteria separate from AET.

10.3 USEFULNESS

10.3.1 Environmental Applicability

Although the Triad approach is both labor-intensive and expensive, its strengths render it extremely cost-effective for the level of information provided. First, it provides empirical evidence of sediment quality (based on observation, not theory). Second, it allows ecological interpretation of physical, chemical, and biological properties (i.e., interpretation of how these relate to the real environment). Third, it uses a preponderance-of-evidence approach rather than relying on single measurements (i.e., all the data are considered). Because of the comprehensive nature of Triad studies, additional follow-up studies are usually not necessary. Finally, the data generated by the Triad approach can be used to generate effects-based classification indexes.

The Triad approach enables investigators to estimate the size of degraded and nondegraded areas. It also provides a test of the quality of reference areas (i.e., do contamination or biological effects occur?). Standards in the form of sediment quality criteria (Chapman, 1986, 1989; PTI, 1988a, 1988b) can be set from the contaminant concentrations that are always associated with effects, using the AET application of the Triad. The Triad approach also provides the information necessary to describe the ecological relationships between sediment properties and biota at risk from sediment contamination.

The Triad approach has been used in dredging studies to support dredged material disposal siting and disposal decisions (Chapman, unpublished). In multiplying the relative degree of degradation at a site by the volume of sediment to be dredged, investigators can compare different sites, provided that the same reference area is used. This comparison helps investigators determine whether dredging will affect useful habitat or result in material unacceptable for ocean disposal. Similarly, potential disposal sites can be compared with each other and with the material to be dredged, and then compared to acceptability criteria for various uses and options. This application of the Triad approach replaces similar but less useful comparisons based solely on the total mass of chemical contaminants to be dredged.

In areas where benthic communities have been eliminated or drastically changed because of a natural event (e.g., storms, oxygen depletion) or physical anthropogenic impact (e.g., recent dredging, boat scour), the other two Triad components (sediment chemistry and toxicity) provide information when conventional univariate approaches would prove deficient. Such cases emphasize the need to use knowledge of an area in making any type of environmental assessment, including the Sediment Quality Triad.

The Triad approach can be used to discern and ultimately to monitor regional trends in sediment quality. Such information is necessary to delineate areas that are excessively contaminated with toxic chemicals affecting the biota and, therefore, most in need of remedial action. Pilot studies of this nature have been conducted in Puget Sound and San Francisco Bay (Long and Chapman, 1985; Chapman, 1986; Chapman *et al.*, 1986, 1987a) and in Europe (e.g., Chapman, in press; Chapman *et al.*, in press).

3.1.1 Suitability for Different Sediment Types

The Triad approach can be used with all sediment types, including sands, muds, aerobic sediments, and anaerobic sediments. It includes sediment characterization with physical parameters [e.g., grain size, acid volatile sulfides (AVS), and total organic carbon (TOC)] that may be important in interpreting the Triad compounds. For example, caution must be used in interpreting the results of toxicity tests in sediments that remain anaerobic in the laboratory despite aeration. Specifically, organisms will die from lack of oxygen, making it difficult to distinguish that mortality from toxicity due to high concentrations of contaminants.

10.3.1.2 Suitability for Different Chemicals or Classes of Chemicals

The Triad approach can be used with all chemicals or classes of chemicals, provided that bioassay organisms and tests are appropriate for all chemicals. For this reason, a battery of bioassay tests is recommended. Caution must be used when testing sediment extracts that may be specific to certain chemical classes. Interpretation of the results must be restricted to only those chemicals.

10.3.1.3 Suitability for Predicting Effects on Different Organisms

Application of the Triad approach can be limited by the organisms in the environment if the *in situ* effects are determined primarily by the same species used in the bioassay tests. In other words, all biological effects data are based on a single species. In such cases, independence of the infaunal community analyses and bioassay test results cannot be assumed. Hence, more than one bioassay test is recommended. Ideally, the tests would include a wide variety of organisms, life stages, feeding types, and exposure routes.

10.3.1.4 Suitability for In-Place Pollutant Control

The Triad approach provides a comprehensive approach to in-place pollutant control because it allows for assessment of all potential interactions between chemical mixtures and the environment. This method is comprehensive because it includes the measurements of multiple chemicals as well as the potential toxic effects of both measured and unmeasured chemicals.

10.3.1.5 Suitability for Source Control

The Triad approach is as suitable for source control as it is for in-place pollutant control. It can be an environmental complement to toxicity reduction evaluation (TRE) programs that involve chemical and toxicity investigations of sediments, and effluents and other discharges.

10.3.1.6 Suitability for Disposal Applications

The Triad approach has been used for disposal applications, including Navy Homeporting work in San Francisco Bay. In that study, the Triad approach clearly separated potential dredge sites from one another in terms of the relative level of pollution. Although the Triad was not used in the final decision because of other considerations, decision-makers were able to use information provided by the Triad to compare the suitability of dredging and disposal options.

10.3.2 General Advantages and Limitations

The following are the major advantages of the Triad approach:

- Combines three separate components to provide a preponderance-of-evidence approach;
- Does not require *a priori* assumptions concerning the specific mechanisms of interaction between organisms and toxic contaminants;
- Can be used to develop sediment quality values (including criteria) for any measured contaminant or a combination of contaminants, including both acute and chronic effects;
- Provides empirical evidence of sediment quality;
- Can be used for any sediment type;
- Allows ecological interpretation of both physical-chemical and biological properties; and
- Does not usually required follow-up when a complete study is conducted.

The following are the major limitations to the Triad approach:

- Statistical criteria have not been fully developed for use with the Triad approach (but see Section 10.2.1.2.3, Types of Data Required);
- Rigorous criteria for calculating single indexes from each of the sediment chemistry, bioassay, and *in situ* biological effects data sets have not been developed (but may not be required);
- A large database is required;
- If the approach is used to determine single-chemical criteria, results could be strongly influenced by the presence of unmeasured toxic contaminants that may or may not covary with measured chemicals;
- Methods for sediment bioassay testing need to be standardized;
- Sample collection, analysis, and interpretation are labor-intensive and expensive; and
- The choice of a reference site is often made without adequate information on how degraded the site may be.

10.3.2.1 Ease of Use

The Triad approach is relatively easy to use and understand. The concept is straightforward. A high level of chemical and biological expertise is required to obtain the data for the three separate Triad components. However, many laboratories or groups of laboratories possess the required expertise.

10.3.2.2 Relative Cost

Relative cost can be evaluated in either dollars or environmental damage. The Triad approach may not prevent environmental damage, but it can be used to identify contaminated areas for future remediation. In terms of dollars, the Triad approach requires substantial resources to be implemented properly, although step-wise, tiered use of Triad components is possible. Measured against the potential environmental damage due to toxic contamination and the costs of remediation, the Triad approach can be extremely cost-effective.

10.3.2.3 Tendency to Be Conservative

The Triad approach provides objective data with which to determine and sometimes to predict environmental damage. Its predictive ability allows for, but does not require, conservatism on the part of the decision-makers.

10.3.2.4 Level of Acceptance

The Triad approach is gaining a high level of acceptance in various parts of North America and in Europe (Forstner *et al.*, 1987; Chapman, in press). In addition, Canada has conducted Triad studies in Vancouver to determine the suitability of this approach for implementation of the new Canadian Environmental Protection Act (Cross *et al.*, 1991; Cross *et al.*, in review).

10.3.2.5 Ability to Be Implemented by Laboratories with Typical Equipment and Handling Facilities

All aspects of the Triad approach (i.e., benthic infaunal studies, sediment chemistry analyses, sediment toxicity bioassays) can be conducted by any competent, specialist laboratory that is reasonably well equipped. The major requirements are adequate QA/QC procedures for chemical measurements; appropriate detection limits; and, for biological analyses, taxonomic experts and a taxonomic reference library or museum.

10.3.2.6 Level of Effort Required to Generate Results

Different levels of effort will generate different levels of results. For instance, results can be generated by simply measuring one or two chemicals, determining the number of infauna present, and conducting a single sediment toxicity bioassay. However, the applicability of these results may be severely limited. Consequently, multiple chemicals including inorganic and organic compounds should be measured, and *in situ* biological alteration and sediment toxicity should be measured multiple times. Although it is possible to use previously collected nonsynoptic data to derive results in a "paper" study (Long and Chapman, 1985), fieldwork and synoptic sampling generate the most useful results.

10.3.2.7 Degree to Which Results Lend Themselves to Interpretation

Beyond the general conclusions noted in Table 10-2, expert judgment is required to implement and interpret the Triad approach. In particular, the definition of "minimal" and "severe" biological effects is required to establish chemical-specific criteria. The Triad approach reflects the complexity of the issues that must be addressed to assess environmental quality.

10.3.2.8 Degree of Environmental Applicability

As discussed, the Triad approach has an extremely high degree of environmental applicability (see Section 10.3.1).

10.3.2.9 Degree of Accuracy and Precision

The accuracy and precision of the Triad approach have not been quantitatively determined. It is expected to have a high degree of accuracy and precision, although these parameters will vary with those of the constituent components.

10.4 STATUS

10.4.1 Extent of Use

Development of the formalized Triad concept has occurred relatively recently (Long and Chapman, 1985; Chapman, 1986, 1990; Chapman *et al.*, 1986, 1987a, 1988, 1991a). The Triad approach has been used directly to establish sediment quality criteria (Chapman, 1986) and, through data manipulations, to determine AET values for sediment quality criteria (Tetra Tech, 1986a; PTI, 1988a, 1988b).

The Triad has been used to identify spatial and temporal trends of pollution-induced degradation. Indexes developed using the Triad approach can be numeric (as described in Chapter 11 for the AET application of the Triad concept) or primarily descriptive (see Figure 2, Chapman *et al.*, 1987a). In either case, the Triad approach provides an objective identification of sites where contamination is causing discernible harm (cf. Power *et al.*, 1991).

10.4.2 Extent to Which Approach Has Been Field-Validated

Because the Triad approach measures *in situ* biological alteration in the field, field validation is an integral part of each complete Triad investigation.

10.4.3 Reasons for Limited Use

As previously described, the Triad approach is being used in the United States, Canada, and Europe for marine, estuarine, and freshwater areas. It is not being used in small projects because of the cost and expertise required for full implementation.

10.4.4 Outlook for Future Use and Amount of Development Yet Needed

The following areas of the Triad approach require development:

- Determining the appropriateness of the various endpoints of different bioassays, selected chemical contaminants, selected measures of benthic community structure, and other potential measures of *in situ* biological alteration;
- Determining the appropriateness of an additive treatment of the data (e.g., summing bioassay responses to provide a single index for toxicity);

- Further development of statistical criteria;
- Development of rigorous criteria for determining, where and if appropriate, composite indexes for each of the three Triad components; and
- Continued standardization of methods for sediment toxicity bioassays.

Even without development of these areas, the Triad approach provides valuable information. The argument has been made (Chapman *et al.*, 1986, 1987a) that the Triad approach provides objective information on which to judge the extent of pollution-induced degradation. For this reason the Triad approach will likely be used much more widely in future.

10.5 REFERENCES

Alden, R. W. II. 1992. Uncertainty and sediment quality assessments: I. Confidence limits for the Triad. Environ. Toxicol. Chem. 11:637-644.

Chapman, P.M. 1986. Sediment quality criteria from the Sediment Quality Triad - an example. Environ. Toxicol. Chem. 5: 957-964.

Chapman, P.M., R.N. Dexter, S.F. Cross, and D.G. Mitchell. 1986. A field trial of the Sediment Quality Triad in San Francisco Bay. NOAA Technical Memorandum NOS OMA 25. National Oceanic and Atmospheric Administration, San Francisco, CA. 127 pp.

Chapman, P.M., R.N. Dexter, and E.R. Long. 1987a. Synoptic measures of sediment contamination, toxicity and infaunal community structure (the Sediment Quality Triad) in San Francisco Bay. Mar. Ecol. Prog. Ser. 37:75-96.

Chapman, P.M., R.C. Barrick, J.M. Neff, and R.C. Swartz. 1987b. Four independent approaches to developing sediment quality criteria yield similar values for model contaminants. Environ. Toxicol. Chem. 6:723-725.

Chapman, P.M. 1988. Marine sediment toxicity tests. pp. 391-402. In: Chemical and Biological Characterization of Sludges, Sediments, Dredge Spoils, and Drilling Muds. J.J. Lichtenberg, F.A. Winter, C.I. Weber, and L. Fradkin (eds.). ASTM STP 976. American Society for Testing and Materials, Philadelphia, PA.

Chapman, P.M. 1989. Current approaches to developing sediment quality criteria. Environ. Toxicol. Chem. 8: 589-599.

Chapman, P.M., C.A. McPherson, and K.R. Munkittrick. 1989. An assessment of the ocean dumping tiered testing approach using the Sediment Quality Triad. Unpublished report prepared for Environmental Protection Canada. E.V.S. Consultants, North Vancouver, BC., Canada.

Chapman, P.M., and E.A. Power. 1990. Sediment toxicity evaluation. American Petroleum Institute Publication No. 4501. 209 pp.

Chapman, P.M. 1990. The Sediment Quality Triad approach to determining pollution-induced degradation. Sci. Total Environ. 97/8:815-825.

Chapman, P. M. In press. Pollution status of North Sea sediments—An international scientific study. Mar. Ecol. Prog. Ser.

Chapman, P.M., R.N. Dexter, H.A. Andersen, and B.A. Power. 1991a. Evaluation of effects associated with an oil platform, using the Sediment Quality Triad. Environ. Toxicol. Chem. 10:407-424.

Chapman, P. M., E. A. Power, and G. A. Burton, Jr. 1991b. pp. 313-340. Chapter 14: Integrative assessments in aquatic ecosystems. In: Contaminated Sediment Toxicity Assessment. G. A. Burton Jr. (ed.). Lewis Publishers, Chelsea, Michigan.

Chapman, P.M., E.R. Long, R. C. Swartz, T.H. DeWitt, and R. Pastorok. 1991c. Sediment toxicity tests, sediment chemistry and benthic ecology do provide new insights into the significance and management of contaminated sediments - a reply to Robert Spies. Environ. Toxicol. Chem. 10:1-4.

Chapman, P.M., R.C. Swartz, B. Roddie, H. Phelps, P. van den Hurk and R. Butler. In press. An international comparison of sediment toxicity tests in the North Sea. Mar. Ecol. Prog. Ser.

Cross, S.F., J.M. Boyd, P.M. Chapman, and R.O. Brinkhurst. 1991. A multivariate approach for defining spatial impacts using the Sediment Quality Triad. p. 886. In: Proceedings of the 17th Annual Aquatic Toxicity Workshop, P.M. Chapman, F. S. Bishay, E. A. Power, K. Hall, L. Hardking. D. McLeavy, M. Nassichuk and W. Knapp (eds.). Can. Tech. Rept. Fish. Aquat. Sci. 1774.

Cross, S. F., J. M. Boyd, P. M. Chapman, and R. O. Brinkhurst. (In review). A multivariate approach to assessing the spatial extent of benthic impacts established using the Sediment Quality Triad. Environ. Toxicol. Chem.

DeWitt, T. H., G. R. Distworth, and R. C. Swartz. 1988. Effects of natural sediment features on survival of the Phoxocephalid amphipod, *Rhepoxynius abronius*. Mar. Environ. Res. 24:99-124.

DeWitt, D. M., J. D. Mahony, D. J. Hansen, K. J. Scott, M. B. Hicks, S. M. Mayr, and M. S. Redmond, 1990. Toxicity of cadmium in sediments: the role of acid volatile sulfide. Environ. Toxicol. Chem. 9:1487-1502.

DiToro, D.M., J.D. Mahony, D.J. Hansen, K.J. Scott, M.B. Hicks, S.M. Mayr, and M.S. Redmond. 1990. Toxicity of cadmium in sediments: The role of acid volatile sulfide. Environ. Toxicol. Chem. 9: 1487-1502.

Forstner, V.U., F. Ackermann, J. Alberti, W. Calmano, F.H. Frimmel, K.N. Kornatzki, R. Leschber, H. Rossknecht, U. Schleichert, and L. Tent. 1987. Qualitatskriterien fur Gewassensedimente - Allgemeine Problematik und internationaler stand der Diskussion. Wasser-Abwasser-Forsch 20:54-59.

Keith, L.H., W. Crummett, J. Deegan, Jr., R.A. Libby, J.K. Taylor, and G. Wentler. 1983. Principles of environmental analysis. Anal. Chem. 55:2210-2218.

Legendre, P and M.J. Fortin. 1989. Spatial pattern and ecological analysis. Vegetatio 80:107-138.

Long, E. R. 1989. The use of the Sediment Quality Triad in classification of sediment contamination. pp. 78-93. In: Marine Board, National Research Council Symposium/Workshop on contaminated marine sediments.

Long, E.R., and M.F. Buchman. 1989. An evaluation of candidate measures of biological effects for the National Status and Trends Program. NOAA Technical Memorandum 105 pp. NOS OMA 45: National Oceanic and Atmospheric Administration, Rockmille, MD.

Long, E.R., and P.M. Chapman. 1985. A sediment quality triad: measures of sediment contamination, toxicity and infaunal community composition in Puget Sound. Mar. Poll. Bull. 16:405-415.

Malueg, K.W., G.S. Schuytema, D.F. Krawczyk, and J.H. Gakstatter. 1984. Laboratory sediment toxicity tests, sediment chemistry and distributions of benthic macroinvertebrates in sediments from the Keweenaw Waterway, Michigan. Environ. Toxicol. Chem. 3:233-242.

Mantel, N. 1967. The detection of disease clustering and generalized regression approach. Cancer Res. 27:200-209.

Nemec, A.F.L., and R.O. Brinkhurst. 1988a. Using the bootstrap to assess statistical significance in the cluster analysis of species abundance data. Can. J. Fish. Aquat. Sci. 45:965-970.

Nemec, A.F.L., and R.O. Brinkhurst. 1988b. The Fowlkes-Mallows statistic and the comparison of two independently determined dendrograms. Can J. Fish. Aquat. Sci. 45:971-975.

Power, E. A., K. R. Munkittrick, and P. M. Chapman. 1991. An ecological impact assessment framework for decision making related to sediment quality. pp. 48-64. In: Aquatic Toxicity and Risk Assessment: Fourteenth Volume. M. A. Mayers and M. G. Barron (eds.). ASTM STP 1124. American Society for Testing and Material, Philadelphia, PA.

PTI Environmental Services, Inc. 1988a. Sediment quality values refinement: Tasks 3 and 5 -1988 update and evaluation of Puget Sound AET. Unpublished report prepared for Tetra Tech, Inc. for the Puget Sound Estuary Program, EPA Contract No. 68-02-43441. PTI Environmental Services, Inc., Bellevue, WA.

PTI Environmental Services, Inc. 1988b. Briefing report to the EPA Science Advisory Board: the Apparent Effects Threshold approach. Unpublished report prepared for Battelle Columbus Division, EPA Contract No. 68-03-3534. PTI Environmental Services, Inc., Bellevue, WA.

Swartz, R.C., W.A. DeBen, K.A. Sercu, and J.O. Lamberson. 1982. Sediment toxicity and the distribution of amphipods in Commencement Bay, Washington, USA. Mar. Poll. Bull. 13:359-364.

Tetra Tech. 1986a. Recommended protocols for measuring selected environmental variables in Puget Sound. Prepared for the Puget Sound Estuary Program, U.S. Environmental Protection Agency, Region X, Seattle, Washington. Tetra Tech, Inc., Bellevue, WA.

Tetra Tech. 1986b. Development of sediment quality values for Puget Sound. Prepared for Resource Planning Associates and U.S. Army Corps of Engineers, Seattle District, for the Puget Sound Dredged Disposal Analysis Program. Tetra Tech, Inc., Bellevue, WA.

U.S. EPA. 1991. Analytical method of determination of acid volatile sulfide in sediment. U.S. Environmental Protection Agency, Criteria and Standards, Washington, DC.

Wiederholm, T., A-M. Wiederholm, and G. Milbrink. 1987. Bulk sediment bioassays with five species of fresh-water oligochaetes. Water, Air and Soil Pollut. 36: 131-154.

Zar, J. H. 1984. Biostatistical Analysis, 2d ed. Prentice-Hall, Englewood Cliffs, NJ.

Apparent Effects Threshold Approach

John Malek
Office of Puget Sound, U.S. Environmental Protection Agency Region X
1200 Sixth Avenue, Seattle, WA 98101
(206) 553-1286

In the Apparent Effects Threshold (AET) approach, empirical data are used to identify concentrations of specific chemicals above which specific biological effects would always be expected. Following the development of AET values for a particular geographic area, they can be used to predict whether statistically significant biological effects are expected at a station with known concentrations of toxic chemicals.

11.1 SPECIFIC APPLICATIONS

11.1.1 Current Use

At present, the AET approach is being used by several programs as guidelines for the protection of aquatic life in Puget Sound. These guidelines are the culmination of cooperative planning and scientific investigations that were initiated by several federal and state agencies in the early and mid-1980s.

Three programs and applications of the AET approach are highlighted below. Notably, all these programs involve an element of direct biological testing in conjunction with the use of AET values, in recognition of the fact that no approach to chemical sediment quality values is 100 percent reliable in predicting adverse biological effects. An underlying strategy in many of these programs was to develop two sets of sediment quality values based primarily on AET values:

- One set of values identifies low chemical concentrations below which biological effects are improbable.
- A second set of values identifies higher chemical concentrations above which multiple biological effects are expected.

The programs incorporate direct biological testing in concentration ranges between these two extremes to serve as a "safety net" (i.e., to account for the uncertainty of chemical predictions) for potential adverse effects or anomalous situations at "moderate" chemical concentrations.

Commencement Bay Nearshore/Tideflats Superfund Investigation

Commencement Bay is a heavily industrialized harbor in Tacoma, WA. Recent surveys have indicated over 281 industrial activities in the nearshore/tideflats area. Comprehensive shoreline surveys have identified more than 400 point and nonpoint source discharges in the study area, consisting primarily of seeps, storm drains, and open channels. A remedial investigation (RI) under Superfund, started in 1983, revealed 25 major sources contributing to sediment contamination, including major chemical manufacturing, pulp mills, shipbuilding and repair, and smelter operations. Adverse biological effects were found in sediments adjacent to these sources.

The AET approach was developed during the course of the RI to assess sediment quality using chemical and biological effects data [i.e., depressions in the number of individual benthic taxa, presence of tumors and other abnormalities in bottom fish, and several laboratory toxicity tests (amphipod mortality, oyster larvae abnormality, bacterial bioluminescence)]. AET values were also used in the subsequent feasibility study (FS) to identify cleanup goals and define volumes of contaminated sediment for remediation. The AET values used in the FS were generated from a reduced set of biological effects indicators, which comprised depressions in total benthic abundance, amphipod mortality, oyster larvae abnormality, and bacterial luminescence.

Puget Sound Dredged Disposal Analysis Program

In 1985, the Puget Sound Dredged Disposal Analysis (PSDDA) program was initiated to develop environmentally safe and publicly acceptable options for unconfined, open-water disposal of dredged material. PSDDA is a cooperative program conducted under the direction of the U.S. Army Corps of Engineers (Corps) Seattle District, U.S. EPA Region X, the Washington Department of Ecology (Ecology), and the Washington Department of Natural Resources (WDNR). AET values were used to develop chemical-specific guidelines to determine whether biological testing on contaminated dredged material is needed. Results of the biological testing help determine suitable disposal alternatives.

Above a specified chemical concentration (i.e., the screening-level concentration or SLC) biological testing is required to determine the suitability of dredged material for unconfined, open-water disposal. Based primarily on AET values for multiple biological indicators, a higher "maximum level concentration" was also identified. Above this latter concentration, failure of biological tests is considered to be predictable. However, an optional series of biological tests can be conducted under PSDDA to demonstrate the suitability of such contaminated material for unconfined, open-water disposal (Phillips et al., 1988).

Urban Bay Toxics Action Program

The Urban Bay Toxics Action Program is a multiphase program to control pollution of urban bays in Puget Sound. The program includes steps to identify areas where contaminated sediments are associated with adverse biological effects, specify potential pollution sources, develop an action plan for source control, and form an action team for plan implementation. Initiated in 1984 by Ecology and U.S. EPA Region X's Office of Puget Sound, the program is a major component of the Puget Sound Estuary Program (PSEP). Substantial participation has also been provided by the Puget Sound Water Quality Authority (Authority) and other state agencies and local governments. Major funding and overall guidance for the program is provided by U.S. EPA Office of Wetlands, Oceans and Watersheds.

In the PSEP urban bay program, AET values are used in conjunction with site-specific biological tests during the assessment of sediment contamination to define and rank problem areas. Source control actions are well under way, but sediment remediation has not yet begun at any of the sites (PTI, 1988).

11.1.2 Potential Use

The AET approach to determining sediment quality can also be used as follows:

- To determine the spatial extent and relative priority of areas of contaminated sediment;
- To identify potential problem chemicals in impacted sediments and, as a result, to focus cleanup activities on potential sources of problem contaminants;
- To define and prioritize laboratory studies for determining cause-effect relationships; and
- With appropriate safety factors or other modifications, to screen sediments in regulatory programs that involve extensive biological testing.

Proposed regulations for sediment contamination are currently under review in Puget Sound. These regulations may include use of AET values to develop statewide sediment quality standards. Ecology is currently developing a suite of sediment management standards, as mandated by the Puget Sound Water Quality Authority (1988) in its 1989 Management Plan. The proposed standards are based in part on AET values. Development of these standards (Becker *et al.*, 1989) relies heavily on the past and ongoing efforts described in Section 11.1.1 and involves active participation by Ecology, U.S. EPA, the Authority, WDNR, the Corps (Seattle District), and various public interest groups. The draft regulation currently under development affects only sediments in Puget Sound. As additional data become available from other locations, the adopted regulation will eventually be broadened and modified to include the entire state.

11.2 DESCRIPTION

11.2.1 Description of Method

AET values are derived using a straightforward algorithm that relates biological and chemical data from field-collected samples. For a given data set, the AET for a given chemical is the sediment concentration above which a particular adverse biological effect (e.g., depressions in the total abundance of indigenous benthic infauna) is always statistically significant ($P \leq 0.05$) relative to appropriate reference conditions. The calculation of an AET for each chemical and biological indicator is conducted as follows:

(1) Collect "matched" chemical and biological effects data—Conduct chemical and biological effects testing on subsamples of the same field sample. (To avoid unaccountable losses of benthic organisms, benthic infaunal and chemical analyses are conducted on separate samples collected concurrently at the same location.)

(2) Identify "impacted" and "nonimpacted" stations—Statistically test the significance of adverse biological effects relative to suitable reference conditions for each sediment sample. Suitable reference conditions are established by sediments exhibiting very low or undetectable concentrations of any toxic chemicals, an absence of other adverse effects, and physical characteristics that are directly comparable with those of the test sediments.

(3) Identify AET using only "nonimpacted" stations—For each chemical, the AET can be identified for a given biological indicator as the highest *detected* concentration among sediment samples that do not exhibit statistically significant effects. (If the chemical is undetected in all nonimpacted samples, then no AET can be established for that chemical and biological indicator.)

(4) Check for preliminary AET—Verify that statistically significant biological effects are observed at a chemical concentration higher than the AET; otherwise, the AET should be regarded only as a preliminary minimum estimate.

(5) Repeat Steps (1)-(4) for each biological indicator.

The AET approach for a group of field-collected sediment samples is shown in Figure 11-1. The samples were collected at various locations and were analyzed for (1) toxicity in a laboratory bioassay and (2) the concentrations of a suite of chemicals, including lead and 4-methylphenol. Based on the results of bioassays conducted on the sediments from each station, two subpopulations of all sediments are represented by bars in the figure:

- Sediments that did not exhibit statistically significant ($P > 0.05$) toxicity relative to reference conditions ("nonimpacted" stations) and
- Sediments that exhibited statistically significant ($P \leq 0.05$) toxicity in bioassays relative to reference conditions ("impacted" stations).

Over the observed range of concentrations for these sediment samples (horizontal axis in Figure 11-1), the sediments fall into two groups for each chemical:

- At low to moderate concentrations, significant sediment toxicity occurred in some samples, but not in others.
- At concentrations above an apparent threshold value, significant sediment toxicity occurred in all samples.

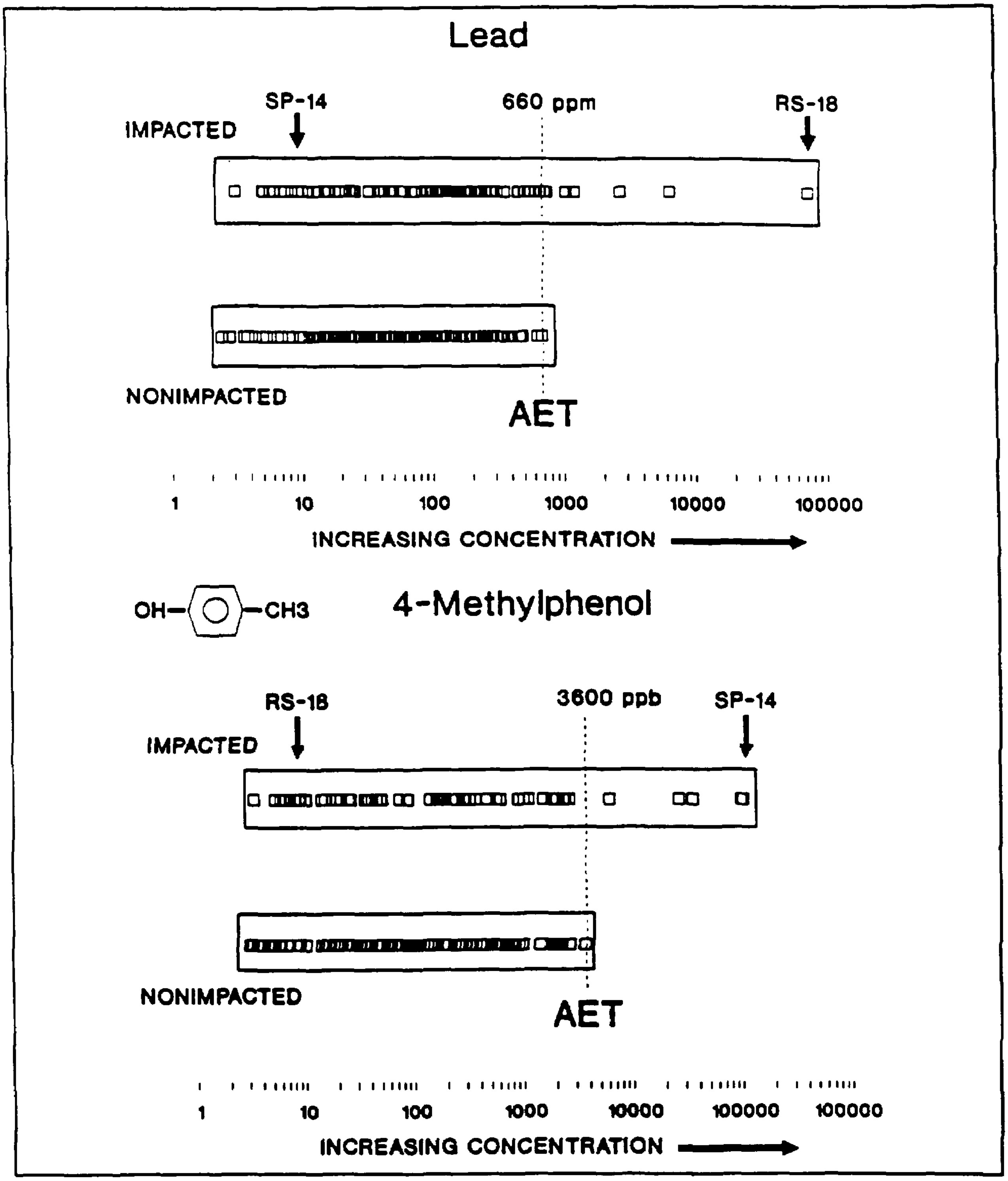


Figure 11-1. The AET approach for a group of field-collected sediment samples.
The AET approach applied to sediments tested for lead and 4-methylphenol concentrations and toxicity response during bioassays.

The AET value is defined for each chemical as the highest concentration of that chemical in the sediments that did *not* exhibit sediment toxicity. Above this AET value, significant sediment toxicity was *always* observed in the data set examined. Data are treated in this manner to reduce the weight given to samples in which factors other than the contaminant examined (e.g., other contaminants, environmental variables) may be responsible for the biological effect.

For each chemical, additional AET values could be defined for other biological indicators that were tested (e.g., other bioassay responses or depressions in the abundances of certain indigenous benthic infauna).

11.2.1.1 Objectives and Assumptions

The objective of the AET approach is to identify concentrations of contaminants that are associated exclusively with sediments exhibiting statistically significant biological effects relative to reference sediments. AET value generation is a conceptually simple process and incorporates the complexity of biological-chemical interrelationships in the environment without relying on *a priori* assumptions about the mechanisms of these interrelationships. Although the AET approach does not require specific assumptions about mechanisms of the uptake and toxic action of chemicals, it does rely on more general assumptions regarding the interpretation of matched biological and chemical data for field-collected samples, as described below:

- For a given chemical, concentrations can be as high as the AET value and not be associated with statistically significant biological effects (for the indicator on which the AET was based).

- When biological impacts are observed at concentrations below an AET value for a given chemical, it is assumed that the impacts may be related to another chemical, chemical interactive effects, or other environmental factors (e.g., sediment anoxia).

- The AET concept is consistent with a relationship between increasing concentrations of toxic chemicals and increasing biological effects (as observed in laboratory exposure studies).

The assumptions in interpreting environmental data are demonstrated below with actual field data. Using Figure 11-1 as an example, sediment from Station SP-14 exhibited severe toxicity, potentially related to a greatly elevated concentrations of 4-methylphenol (7,400 times reference levels). The same sediment from Station SP-14 contained a relatively low concentration of lead that was well below the AET for lead (Figure 11-1). Despite the toxic effects associated with the sample, sediments from many other stations with higher lead concentrations than Station SP-14 exhibited no statistically significant biological effects. These results were interpreted to suggest that the effects at Station SP-14 were potentially associated with 4-methylphenol (or a substance with a similar environmental distribution) but were less likely to be associated with lead. A converse argument can be made for lead and 4-methylphenol in sediments from Station RS-18.

Applied in this manner, the AET approach helps to identify measured chemicals that are potentially associated with observed effects at each biologically impacted site and eliminates from consideration chemicals that are far less likely to be associated with effects (i.e., the latter chemicals have been observed at higher concentrations at other sites without associated biological effects). Based on the results for lead and 4-methylphenol, bioassay toxicity at five of the impacted sites shown in the figure may be associated with elevated concentrations of 4-methylphenol, and toxicity at eight other sites may be associated with elevated concentrations of lead (or similarly distributed contaminants).

As illustrated by these results, the occurrence of biologically impacted stations at concentrations below the AET of a single chemical does not imply that AET values in general are not protective against biological effects, only that single chemicals may not account for all stations with biological effects. By developing AETs for multiple chemicals, a high percentage of all stations with biological effects are accounted for with the AET approach (see Section 11.3.2.9 and USEPA, 1988).

AETs can be expected to be more predictive when developed from a large, diverse database with wide ranges of chemical concentrations and a wide diversity of measured chemicals. Data sets that have large concentration gaps between stations and/or do not cover a wide range of concentrations must be scrutinized carefully (e.g., to discern whether chemical concentrations in the data set exceed reference concentrations) to determine whether AET generation is appropriate.

11.2.1.2 Level of Effort

11.2.1.2.1 Type of Sampling Required

Collection of field data for initial generation of AETs is a labor-intensive and capital-intensive process. The exact level of sampling effort required depends on the amount and variety of data collected (e.g., the number of samples collected, the diversity of biological indicators that are tested, and the range of chemicals measured). One means of minimizing these costs is to compile existing data that meet appropriate quality assurance criteria. There are no definitive requirements for the size and variety of the database, although a study of the predictive abilities of the AET approach with Puget Sound data (Barrick et al., 1988) resulted in the following recommendations for data collection:

- Collect or compile chemical and biological effects data from 50 stations or more (and from suitable reference areas).
- Bias the positioning of stations to ensure sampling of various contaminant sources (e.g., urban environments with a range of contaminant sources and, preferably, with broad geographic distribution) over a range of contaminant concentrations (preferably over at least 1-2 orders of magnitude).
- Conduct chemical tests for a wide range of chemical classes (e.g., metals, nonionic organic compounds, ionizable organic compounds). To generate AETs on an organic carbon-normalized basis, total organic carbon (TOC) measurements are required in all sediments.
- Ensure that detection limits of <100 ppb (lower if possible) are attained for organic compounds. High detection limits (i.e., insensitive analyses) can obscure the occurrence of chemicals at low to moderate concentrations; as noted previously, only detected data are used in AET calculations. Metals are naturally occurring substances, and most metals concentrations typically exceed routine detection limits.

The only strict requirement for field sampling of data for AET generation is the collection of "matched" chemical and biological data (as described at the beginning of Section 11.2.1). Matched data sets should be used to reduce the possibility that uneven (spatially variable) sediment contamination could result in associating biological and chemical data that are based on dissimilar sediment samples. Because the toxic responses of stationary organisms (e.g., bioassay organisms confined to a test sediment, or infaunal organisms largely confined to a small area) are assumed to be affected by direct association with contaminants in the surrounding environment, it is considered essential that chemical and biological data be collected from nearly identical subsamples from a given station.

11.2.1.2.2 Methods

Methodological details for the generation of AET values are described at the beginning of Section 11.2.1.

11.2.1.2.3 Types of Data Required

Two fundamental kinds of data analysis are required for AET generation:

- Statistical analysis of the significance of biological effects relative to reference conditions (i.e., classification of stations as impacted or nonimpacted for each biological indicator) and

- Generation of an AET value for each chemical and biological indicator (essentially a process of ranking stations based on chemical concentration).

Additional kinds of data analysis needed for AET generation are quality assurance/quality control (QA/QC) review of biological and chemical data, and evaluation of the appropriateness of reference area stations. These topics have been described elsewhere (e.g., Beller et al., 1986; Barrick et al., 1988).

The AET method does not intrinsically require a specific method of statistical analysis for determination of significance of biological effects relative to reference conditions. Existing Puget Sound AETs have relied largely on pairwise t-tests; details of statistical analyses performed for the generation of Puget Sound AET have been described elsewhere (USEPA, 1988; Barrick *et al.*, 1988; Beller *et al.*, 1986). For example, the following steps were used to determine the statistical significance of amphipod mortality bioassay results (Swartz *et al.*, 1985) in field-collected sediments:

- All replicates from all stations in the reference area used for each study were pooled, and a mean bioassay response and standard deviation were calculated.

- Results from each potentially impacted site were then compared statistically with the reference conditions using pairwise analysis.

- The F_{max} test (Sokal and Rohlf, 1969) was used to test for homogeneity of variances between each pair of mean values.

- If variances were homogenous, then a t-test was used to compare the two means.

- If variances were not homogenous, then an approximate t-test (Sokal and Rohlf, 1969) was used to compare the two means.

- Statistical significance was tested with a pairwise error rate of 0.05 to ensure consistency among studies of differing sample sizes.

Data analyses that have been applied to other biological indicators are described elsewhere (Beller *et al.*, 1986; Barrick *et al.*, 1988). Notably, comparisons to reference conditions were somewhat more complicated for benthic infaunal abundances than for sediment bioassays. For benthic infaunal comparisons, reference data for each potentially impacted site were categorized so that comparisons were made with samples collected during the same season, at a similar depth, and whenever possible, in sediments with similar particle size characteristics (i.e., percentage of particles <64 µm) as those of the potentially impacted site. In this manner, statistical comparisons were normalized to account for the influence of three of the major natural variables known to influence the abundance and distribution of benthic macroinvertebrates. All benthic data were also log-transformed so that data distributions conformed to the assumptions of the parametric statistical tests that were applied. Additional data treatment methods presented elsewhere (Barrick *et al.*, 1988) are not discussed further herein, because they are not considered intrinsic to the AET approach, but rather are options to address potentially unusual matrices or biological conditions.

11.2.1.2.4 Necessary Hardware and Skills

The primary skills required for AET generation are related to the development of the biological/chemical database. Expertise in environmental chemistry is required to evaluate chemical data quality, and the need for normalization of chemical data and related factors. Biological and statistical expertise are required for the determination of statistical significance. For benthic data in particular, evaluation of appropriate reference conditions and knowledge of benthic taxonomy and ecology are necessary.

Computers are recommended for the efficient generation of AET values. A menu-driven database (SEDQUAL) has been developed for U.S.

EPA Region X that is capable of a number of data manipulation tasks, including the following: (1) storing chemical and biological data, (2) calculating AET values, (3) comparing a specified set of AET to stored sediment chemistry data to identify stations at which adverse biological effects are or are not predicted, and (4) based on such comparisons, calculating the rate of correct prediction of biological impacts. The SEDQUAL system, which requires an IBM-AT compatible computer with a hard disk, has been documented in detail in a user's manual (Nielsen, 1988). The SEDQUAL database currently includes stored data from Puget Sound (over 1,000 samples, not all of which have biological and chemical data).

11.2.1.3 Adequacy of Documentation

Various aspects of the AET approach have been extensively documented in reports prepared for U.S. EPA and other regulatory agencies, as listed below and in the reference list:

- Generation of Puget Sound AET values and evaluation of their predictive ability (Beller *et al.*, 1986; Barrick *et al.*, 1988);
- Data used to generate Puget Sound AET values (appendices of Beller *et al.*, 1986 and field surveys cited in Beller *et al.*, 1986 and Barrick *et al.*, 1988);
- Briefing report to the U.S. EPA Science Advisory Board (USEPA, 1988); and
- Policy implications of effects-based marine sediment criteria (PTI, 1987).

11.2.2 Applicability of Method to Human Health, Aquatic Life, or Wildlife Protection

The AET approach has been designed for use in evaluating potential adverse impacts to aquatic life associated with chemical contamination of sediments. By empirically determining the association between chemical contamination and adverse biological effects, predictions can be made regarding the levels of contamination that are always associated with adverse effects (i.e., the AET values). These critical levels of contamination can then be used to develop guidelines for protecting aquatic life (e.g., sediment quality values). AETs can be developed for any kind of aquatic organism for which biological responses to chemical toxicity can be measured. The protectiveness of the AET can therefore be ensured by evaluating organisms and biological responses with different degrees of sensitivity to chemical toxicity. For example, evaluations of metabolic changes (i.e., usually a very sensitive biological response) in a pollution-sensitive species would likely result in AET values that are lower and more protective than evaluations of mortality (i.e., generally a less sensitive response) in a more pollution-tolerant species. The protectiveness of AETs can also be ensured through the application of "safety factors." For example, to be protective of chronic biological responses, a factor based on an acute-chronic ratio could be applied to AETs developed on the basis of acute biological responses.

11.2.3 Ability of Method to Generate Numerical Criteria for Specific Chemicals

The AET approach is not intrinsically limited in application to specific chemicals or chemical groups. In general, the approach can be used for chemicals for which data are available. However, when using a specific data set to generate AETs, it is preferable that AET generation be limited to chemicals with wide concentration ranges (e.g., ranging from reference concentrations to concentrations near direct sources) and/or with appropriate detection frequencies (e.g., greater than 10 detections). A partial list of chemicals for which AETs have been developed is presented in Table 11-1.

11.3 USEFULNESS

11.3.1 Environmental Applicability

11.3.1.1 Suitability for Different Sediment Types

The AET approach can be applied to any sediment type in saltwater or freshwater environments for which biological tests can be conducted.

Table 11-1. Selected Chemicals for Which AETs Have Been Developed in Puget Sound.

METALS		
Antimony	Copper	Nickel
Arsenic	Lead	Silver
Cadmium	Mercury	Zinc
Chromium		
ORGANIC COMPOUNDS		
Low-Molecular-Weight PAHs	**High-Molecular-Weight PAHs**	**Chlorinated Benzenes**
Naphthalene	Fluoranthene	1,3-Dichlorobenzene
Acenaphthylene	Pyrene	1,4-Dichlorobenzene
Acenaphthene	Benz(a)anthracene	1,2-Dichlorobenzene
Fluorene	Chrysene	1,2,4-Trichlorobenzene
Phenanthrene	Benzofluoranthenes	Hexachlorobenzene (HCB)
Anthracene	Benzo(a)pyrene	
2-Methylnaphthalene	Indeno(1,2,3-c,d)pyrene	
	Dibenzo(a,h)anthracene	
	Benzo(g,h,i)perylene	
Phthalates	**Total PCBs**	**Phenols**
Dimethyl phthalate		Phenol
Diethyl phthalate		2-Methylphenol
Di-n-butyl phthalate		4-Methylphenol
Butyl benzyl phthalate		2,4-Dimethylphenol
Bis(2-ethylhexyl)phthalate		Pentachlorophenol
Di-n-octyl phthalate		
Pesticides	**Miscellaneous Extractables**	**Volatile Organics**
p,p'-DDE	Benzyl alcohol	Tetrachloroethene
p,p'-DDD	Benzoic acid	Ethylbenzene
p,p'-DDT	Dibenzofuran	Total xylenes
	Hexachlorobutadiene	
	N-Nitrosodiphenylamine	

By normalizing chemical concentrations to appropriate sediment variables (e.g., percent organic carbon), differences between different sediment types can be minimized in the generation of AETs. In practice, identification of unique or atypical sediment matrices is important in determining the general applicability of AET values generated from a specific set of data.

Differences in physical characteristics (e.g., grain size, habitat exposure) are one major factor that may account for stations not meeting predictions based on existing AET values. In Puget Sound studies, for example, fine-grained sediments dominated stations that had significant amphipod mortality that had not been predicted, and coarse-grained sediments dominated stations that had significant depressions in benthic infauna that had not been predicted by benthic AETs (Barrick *et al.*, 1988).

11.3.1.2 Suitability for Different Chemicals or Classes of Chemicals

There are no constraints on the types of chemicals for which AETs can be developed. An AET can be developed for any measured chemical (organic or inorganic) that spans a wide concentration range in the data set used to generate AETs. The availability of a wide diversity of chemical data increases the probability that toxic agents (or chemicals that covary in the environment with toxic agents) can be included in interpreting observed biological impacts.

To date, AETs have been developed for over 60 chemicals frequently detected in the environment, including 16 polycyclic aromatic hydrocarbons (PAHs); several alkylated PAHs and related nitrogen-, sulfur-, and oxygen-containing heterocycles; polychlorinated biphenyls (PCBs) (reported as total PCBs); 5 chlorinated benzenes; 6 phthalate esters; 3 chlorinated hydrocarbon pesticides; phenol and 4 alkyl-substituted and chlorinated phenols; 10 metals and metalloids; 3 volatile organic compounds; and 5 miscellaneous extractable substances. Data for other miscellaneous chemicals that were less frequently detected or analyzed for in the Puget Sound area were also evaluated for their potential use in developing AETs (e.g., resin acids and chlorinated phenols in selected sediments from areas influenced by pulp and paper mill activity).

AETs have been developed for chemical concentrations normalized to sediment dry weight and sediment organic carbon content (expressed as percent of dry weight sediment). Using a 188-sample data set from Puget Sound, AETs were also developed for data normalized to fine-grained particle content (expressed as the percent of silt and clay, or <63-μm particulate material, in dry weight of sediment). These latter AET values did not appear to offer advantages in predictive reliability over the more commonly used dry weight and TOC normalizations (Beller *et al.*, 1986).

11.3.1.3 Suitability for Predicting Effects on Different Organisms

The AET approach can be used to predict effects on any life stage of any marine or aquatic organism for which a biological response to chemical toxicity can be determined. Because the approach is empirical, relying on direct measurement of the chemical concentrations associated with samples exhibiting adverse effects, the results are directly applicable to predicting effects on the organisms used to generate the AET. The results can also be used to predict effects on nontarget organisms by ensuring that the organisms used to generate an AET are either representative of the nontarget organisms or are more sensitive to chemical toxicity than those organisms. For example, AETs generated for a species of sensitive amphipod might be considered as protective of the chemical concentrations associated with adverse effects in other species of equally or less sensitive amphipods. At the same time, these AET might be considered protective of most other benthic macroinvertebrate taxa because they are based on a member of a benthic taxon (i.e., Amphipoda) that is considered to be sensitive to chemical toxicity (Bellan-Santini, 1980). By contrast, AETs generated for a pollution-tolerant species such as the polychaete *Capitella capitata* (cf. Pearson and Rosenberg, 1978) might be considered representative for other pollution-tolerant species, but not protective for most other kinds of benthic macroinvertebrates.

11.3.1.4 Suitability for In-Place Pollutant Control

In remedial action programs, assessment tools such as the AET approach can be used to address the following specific regulatory needs:

- Provide a preponderance of evidence for narrowing a list of problem chemicals measured at a site;
- Provide a predictive tool for cases in which site-specific biological testing results are not available;
- Enable designation of problem areas within the site;
- Provide a consistent basis on which to evaluate sediment contamination and to separate acceptable from unacceptable conditions;
- Provide an environmental basis for triggering sediment remedial action; and

- Provide a reference point for establishing a cleanup goal.

Because AET values are derived from sediments with multiple contaminants, they incorporate the influence of interactive effects in environmental samples. The ability to incorporate the influences of chemical mixtures, either by design or default, is an advantage for the assessment of in-place pollutants.

11.3.1.5 Suitability for Source Control

The AET approach is well suited for identifying problem areas. Because specific cause-effect relationships are not proven for specific chemicals and biological effects, remedial actions should not be designed exclusively for a specific chemical. (This caution applies to all approaches because of the complex mixture of contaminants in environmental samples.) The link between problem areas and potential sources of contamination is established by analysis of concentration gradients of contaminants in these problem areas and the presence and composition of contaminants in sediments and source materials. The AET approach provides a means of narrowing the list of measured chemicals that should be considered for source control and provides supportive evidence for eliminating chemicals from consideration that appear to be present at a concentration too low to be associated with adverse biological effects. Reduction of the overall contaminant load to a problem area such that all measured chemicals are below their respective AETs is predicted to result in mitigation of the adverse biological effects. It is possible that such source controls may be effective because of the concomitant removal of an unmeasured contaminant.

11.3.1.6 Suitability for Disposal Applications

The evaluation of potential biological impacts associated with the disposal of dredged material is an important component in the designation of disposal sites and review of disposal permits for dredged material. AET values provide a preponderance of evidence in determining a "reason to believe" that sediment contamination could result in adverse biological effects. Hence, the AET approach is a useful tool for assessing the need for biological testing during the evaluation of disposal alternatives. It is assumed that AET values generated for in-place sediments provide a useful prediction of whether adverse biological effects will occur in dredged material after disposal at aquatic sites.

11.3.2 General Advantages and Limitations

11.3.2.1 Ease of Use

In this section, "use" is treated as both generation and application. The ease of generating AET values depends on the status of the data to be used for AET generation (i.e., whether field data have been collected and whether statistical significance has been determined for biological indicators). It is recommended that a search for existing data be conducted as part of determining the need for collecting new samples. The existing database of matched biological and chemical data from Puget Sound comprises over 300 samples. Collection of new field data (e.g., for application outside of Puget Sound) would require a considerable expenditure of effort, as would the statistical analysis of a large number of samples. However, if data are available and statistical analyses have been performed, the generation of AET values is very easy with the SEDQUAL database (described in Section 11.2.1.2.4). The menu-driven system allows for a considerable amount of flexibility in choosing stations and biological indicators to be included in AET generation. Application of AET (i.e., comparison of AET values to chemical concentrations in field samples) is also very easy when using SEDQUAL, provided that the field data have been computerized. Application of AET values to chemical data presented in existing literature is also straightforward.

11.3.2.2 Relative Cost

The cost of developing AET values can span a wide range, depending on the stage of database development and the numbers and kinds of chemicals and biological indicators used. The least costly means of developing the values is to use existing chemical and biological information, thus minimizing the expenses associated with field sampling and laboratory analyses. (Selective sampling to confirm whether existing AET values are applicable would still be useful.) The historical database could be based on the pooled results from various studies conducted in a region, providing that each study passed QA/QC performance criteria and satisfied the prerequisites of the AET approach (e.g., matched chemical and biological measurements and the ability to discriminate adverse biological effects).

If the historical database is judged inadequate to generate AETs for a region, then the costs of field measurements of chemical concentrations in sediments and associated biological effects must be incurred to develop the database. These costs can vary substantially, depending on the chemicals and biological indicators evaluated. Costs would be minimized if evaluations were based on a limited range of chemicals and a single, inexpensive biological test. It is recommended that the approach be based on a relatively wide range of chemicals, and if possible, several kinds of biological indicators.

The existing database for the Puget Sound region is based on a wide range of chemicals (i.e., U.S. EPA priority pollutants and other selected chemicals) and four kinds of biological indicators. The costs for developing AETs varied considerably among the four indicators. For example, laboratory costs for the least expensive indicator (Microtox bioassay) were approximately $200 per station, whereas costs for the most expensive indicator (abundances of benthic macroinvertebrates) were as high as $1,800 per station. Therefore, within the existing database, the range of costs for biological testing spanned almost 1 order of magnitude.

Once AET values have been generated, use of these values to predict the occurrence of biological effects is relatively inexpensive. Chemical data may be compared to AET values by using the SEDQUAL database or through manual data manipulations.

11.3.2.3 Tendency to Be Conservative

The empirical, field-based nature of the AET approach precludes definitive *a priori* predictions of its tendency to be either over- or underprotective of the environment. The occurrence of biologically impacted stations at concentrations below the AET of a given chemical (see Figure 11-1) may appear to be underprotective. However, the occurrence of impacted stations at concentrations below the AET of a single chemical does not imply that AETs in general are not protective against biological effects, only that single chemicals may not account for all stations with biological effects. If AETs are developed for multiple chemicals, the approach can account for a high percentage of stations with adverse biological effects.

To date, AETs have been developed for acute sediment bioassays of mortality in adult amphipods, developmental abnormality in larval bivalves, and metabolic alterations in bacteria. All of these organism/endpoint combinations are considered to be sensitive to chemical toxicity. AETs have also been generated for *in situ* reductions in the abundances of benthic macroinvertebrates. Because these reductions incorporate chronic (i.e., long-term) exposure to contaminants, they can also be considered as sensitive measures of the effects of chemical toxicity. However, a more protective approach would be to use the lowest of the four kinds of AET for each chemical as the concentration on which predictions are made. Alternatively, the protectiveness of any kind of AET could be modified by developing sediment quality values based on "safety factors" applied to existing AETs.

11.3.2.4 Level of Acceptance

The AET approach has been accepted by several federal and state agencies in the Puget Sound region as one tool in providing guidelines for regulatory decisions. U.S. EPA has used AET values to develop sediment quality values with which to evaluate the potential toxicity of contaminated sediments in urban bays. PSDDA has used AET values as a tool to develop chemical guidelines for determining whether biological testing is necessary for dredged sediments proposed for unconfined, open-water disposal. Ecology has used AET to develop sediment management standards. These standards were promulgated by the State of Washington and approved by EPA Region X in 1991. The standards are being used by a number of water quality programs (e.g., source control, remediation).

Several major characteristics influence the acceptability of the AET approach. The most attractive characteristic of the approach is probably the reliance on empirical information based on field-collected sediments or indigenous organisms, and exposure of laboratory test organisms to environmental samples. A second attractive feature of the approach is the setting of an AET at the chemical concentration in the data set above which adverse biological effects are always observed. This characteristic provides consistency that, with a representative database used to generate AETs, enhances the preponderance of evidence of adverse effects in the environment. The AET values can be updated as new information is collected. The AET approach can also be applied to an existing database in new regions, providing certain prerequisites are met by the database (e.g., synoptic measurement of chemical and biological data, and QA/QC guidelines).

A limitation of the AET approach is that field-based approaches do not directly assess cause-effect relationships. Because sediments in the environment are often contaminated with a complex mixture of chemicals, it is difficult when using field-collected sediment for any approach to relate observed biological effects to a single chemical. The approach also requires selection of appropriate normalized chemical data to address the bioavailability of contaminants to organisms. Organic carbon normalization may be most appropriate for nonpolar organic contaminants based on theoretical considerations. In addition, nonprotective AETs could be generated if unusual matrices (e.g., slag) that anomalously restrict bioavailability are included in the database used to generate the AETs, or if biological test results are incorrectly classified. Recommended data treatment guidelines for chemical and biological data are discussed by Barrick et al. (1988). The AET approach was reviewed by the U.S. EPA Science Advisory Board (SAB, 1989), which noted the method had "major strengths in its ability to determine biological effects and assess interactive chemical effects."

11.3.2.5 Ability to Be Implemented by Laboratories with Typical Equipment and Handling Facilities

If applicable data do not already exist, the development of AET values requires a relatively extensive amount of field sampling and laboratory analysis. The chemical analyses required for development of AET represent standard analytical procedures. A laboratory with appropriately trained staff should be able to conduct the necessary benthic community analyses and sediment bioassays. Specific methods for performing the chemical and biological tests that were used to develop Puget Sound AET are detailed in the Puget Sound Protocols (Tetra Tech, 1986). These efforts can be minimized by using historical data whenever possible. Once AETs are developed, their routine implementation is relatively easy. In addition, they can be easily updated as additional data become available.

11.3.2.6 Level of Effort Required to Generate Results

As noted in Section 11.3.2.1, the SEDQUAL database facilitates AET generation and application. After field data have been collected, the most time-consuming task is data entry and verification. Entry of chemical and biological data for 50 samples requires roughly 16 person-hours (assuming 75 chemicals have been measured and biological effects are being coded simply as "impacted" or "nonimpacted"). Generating a set of AET values for a given biological indicator, 75 chemicals, and 50 stations takes approximately 0.75-1 h of computer time on SEDQUAL (and about 5 min of labor to set up the analysis). To compare a set of AET (for 75 chemicals) to a 50-sample set of field data takes approximately 0.5-0.75 h of computer time on SEDQUAL (and roughly 5 min of labor to set up the analysis). SEDQUAL is capable of comparing any kind of chemical sediment criteria to field data, but requires that the numerical criteria be entered in the database.

11.3.2.7 Degree to Which Results Lend Themselves to Interpretation

The manner in which the AET approach can be used to interpret matched biological and chemical data from field-collected sediments is described in Section 11.2.1. As noted previously, the use of AET can help investigators eliminate chemicals from further consideration (as the cause of an observed effect); however, the approach cannot identify specific cause-effect relationships. Because the AET approach is empirical, it is not well suited to identifying specific toxic agents or elucidating mechanisms of biological uptake and metabolism. However, certain general relationships could be examined on an *a posteriori* basis with the AET approach (e.g., testing the relative importance of different ways of normalizing chemical concentration data in predicting adverse biological effects).

A number of environmental factors may complicate the interpretation of the data. Although the AET concept is simple, the generation of AET values based on environmental data incorporates many complex biological-chemical interrelationships. For example, the AET approach incorporates the net effects of the following factors that may be important in field-collected sediments:

- Interactive effects of chemicals (e.g., synergism, antagonism, and additivity);
- Unmeasured chemicals and other unmeasured, potentially adverse variables; and
- Matrix effects and bioavailability (i.e., phase associations between contaminants and sediments that affect bioavailability of the contaminants, such as the incorporation of PAH in soot particles).

The AET approach cannot quantify the individual contributions of interactive effects, unmeasured chemicals, or matrix effects in environmental samples, but AET values may be influenced by these factors. AET values are expected to be reliable predictors of adverse effects that could result from the influence of these environmental factors if the samples used to generate AETs are representative of samples for which AET predictions are made. Alternatively, isolated occurrences of such environmental factors in a data set used to generate AETs may limit the predictive reliability of those AET values. If confounding environmental factors render the AET approach unreliable, then this should be evident from validation tests in which biological effects are predicted in actual environmental samples.

A more detailed discussion of the interpretation of AETs and the confounding effects of environmental factors is presented in U.S. EPA (1988).

11.3.2.8 Degree of Environmental Applicability

The AET approach has a high degree of environmental applicability based on its reliance on chemical and biological measurements made directly on environmental samples. Such information provides tangible evidence that various chemical concentrations either are or are not associated with adverse biological effects in typically complex environmental settings.

The environmental applicability of the AET approach has been quantified for the four kinds of AET developed for Puget Sound by evaluating the reliability with which each kind of AET predicted the presence or absence of adverse biological effects in field samples collected from Puget Sound (USEPA, 1988). The overall reliability of the four tests ranged from 85 to 96 percent, indicating that all four kinds of AETs were relatively accurate at predicting the presence or absence of effects for samples from the existing database. This high level of reliability suggests that AETs have a relatively high degree of environmental applicability in Puget Sound, and it has been a primary factor in the use of the AET approach by agencies in the Puget Sound region. AET values generated for Puget Sound have also been used as examples of effects-based sediment criteria to provide an initial estimate of the magnitude of potential problem areas in coastal regions of the United States for the U.S. EPA Office of Policy Analysis (PTI, 1987).

11.3.2.9 Degree of Accuracy and Precision

In this section, accuracy is considered to be the ability of AET to predict biological effects and precision represents the expected variability (uncertainty range) for a given AET value for a given data set.

In previous evaluations of the AET approach and other sediment quality values using field-collected data, the accuracy of the approach was defined by two qualities:

- Sensitivity in detecting environmental problems (i.e., are *all* biologically impacted sediments identified by the predictions of the chemical sediment criteria?)

- Efficiency in screening environmental problems (i.e., are *only* biologically impacted sediments identified by the predictions of the chemical sediment criteria?).

Sensitivity is defined as the proportion of all stations exhibiting adverse biological effects that are correctly predicted using sediment criteria. Efficiency is defined as the proportion of all stations predicted to have adverse biological effects that actually are impacted. Ideally, a sediment criteria approach should be efficient as well as sensitive. For example, a sediment criteria approach that sets values for a wide range of chemicals near their analytical detection limits will likely be conservative (i.e., sensitive) but inefficient. That is, it will predict a large percentage of sediments with biological effects. It will also predict impacts at many stations where there are no biological effects, but chemical concentrations are slightly elevated. The concepts of sensitivity and efficiency are illustrated in Figure 11-2.

The overall reliability of any sediment criteria approach addresses both sensitivity and efficiency. This measure is defined as the proportion of all stations for which correct predictions were made for either the presence or absence of adverse biological effects:

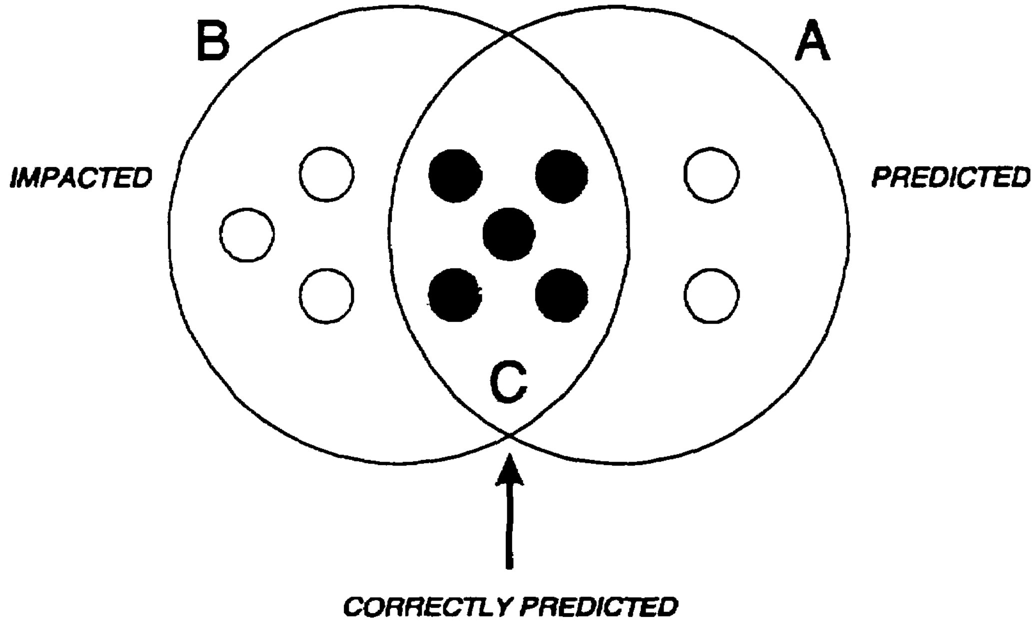
$$\textit{Overall reliability} = \frac{\textit{All stations correctly predicted as impacted} + \textit{All stations correctly predicted as nonimpacted}}{\textit{Total number of stations evaluated}}$$

High reliability results from correct prediction of a large percentage of the impacted stations (i.e., high sensitivity, few false negatives) and correct prediction of a large percentage of the nonimpacted stations (i.e., high efficiency, few false positives). An assessment of AET reliability was recently conducted using a large database comprising samples from 13 Puget Sound embayments (Barrick *et al.*, 1988). These evaluations suggest that the AET approach is relatively sensitive for the biological indicators tested and also relatively efficient. For example, 68-83 percent sensitivity and 55-75 percent efficiency were observed when AETs generated from a 188-sample data set were evaluated with an independent 146-sample data set. The ranges of sensitivity and efficiency cited above represent the ability of benthic infaunal AET values to predict statistically significant depressions in the abundances of benthic infauna in field-collected samples and the ability of amphipod mortality bioassay AET values to predict statistically significant mortality in bioassays conducted on field-collected sediment.

Precision of the AET approach has not been as intensively investigated as accuracy. AET values are the result of parametric statistical procedures (i.e., determination of the significance of biological effects relative to reference conditions) and nonparametric methods (e.g., ranking of stations by concentration), and thus are not amenable to the routine definition of confidence intervals. However, the degree of AET precision is considered to depend on the following factors:

- The concentration range between the AET (determined by a nonimpacted station) and the next highest concentration that is associated with a statistically significant effect;

- Classification error associated with the statistical significance of biological indi-



SENSITIVITY = C/B x 100 = 5/8 x 100 = 63%
EFFICIENCY = C/A x 100 = 5/7 x 100 = 71%

FOR A GIVEN BIOLOGICAL INDICATOR:

A ALL STATIONS PREDICTED TO BE IMPACTED
B ALL STATIONS KNOWN TO BE IMPACTED
C ALL STATIONS CORRECTLY PREDICTED TO BE IMPACTED

Figure 11-2. Measures of reliability (sensitivity and efficiency).

cator results (i.e., whether a station is properly classified as impacted or nonimpacted, as related to Type I and Type II statistical error);

- The weight of evidence or number of observations supporting a given AET value; and
- The analytical error associated with quantification of chemical results.

Detailed discussion of these factors is provided in Beller *et al.* (1986).

One approach used in Puget Sound to estimate the uncertainty range around the AET value was to define the lower limit as the concentration at the nonimpacted station immediately below the AET and to define the upper limit as the concentration at the impacted station immediately above the AET. These limits are based largely on probabilities of statistical classification error. For data sets with large concentration gaps between stations, such uncertainty ranges will be wider and precision will be poorer than for data sets with more continuous distributions. The number of stations used to establish an AET would be expected to have a marked effect on AET uncertainty because small data sets would tend to have less continuous distributions of chemical concentrations than large data sets. Based on analyses conducted with Puget Sound data, the magnitude of the AET uncertainty for 10 chemicals or chemical groups that are commonly detected is typically less than one-third to one-half of the value of the AET itself (considering both amphipod mortality bioassay and benthic infaunal AET data). Based on quality assurance information for these data, analytical error is probably a minor component of overall precision, particularly for metals.

11.4 STATUS

11.4.1 Extent of Use

The AET approach is used by several agencies and sediment management programs in the Pacific Northwest to provide guideline values for regulatory decisions. The State of Washington has developed sediment management standards primarily using the AET approach but also including equilibrium partitioning values. These standards were promulgated by the State and approved by EPA, Region X, in 1991 and are currently being implemented in a variety of programs. The standards are the culmination of cooperative planning and scientific investigations by several federal and state agencies throughout the 1980's, including:

- Superfund investigations at Commencement Bay and Eagle Harbor;
- Puget Sound Dredged Disposal Analysis (PSDDA);
- Urban Bay Toxics Action Program; and
- Puget Sound Water Quality Authority Management Plan.

A key result of these efforts has been the recognition by regulators of two separate levels of sediment contamination and has led to the development of two sets of sediment quality values. This separation in management use of sediment values arose from the sensitivity and efficiency concepts of reliability previously discussed. This management decision was made because it was determined that none of the available approaches for developing sediment quality values would result in 100 percent sensitive and 100 percent efficient values. Different strategies have been used by different programs for use of AET-generated values. In general, the lowest AET (termed LAET) for any of the biological tests is used to establish the lower level where there is little concern of sediment contamination (e.g., the goal for remedial actions). The AET approach has developed higher chemical levels (termed HAET), above which adverse effects are predicted for all the biological tests. In most regulatory programs, direct biological testing is allowed to resolve the differences in predictions of these two sets of sediment quality values (i.e., prediction of adverse biological effect by highly sensitive sediment quality values, which at lower chemical concentrations are not predicted by highly efficient sediment quality values). To date, such sediment quality values developed were for and used in marine and estuarine environments. The State of Washington and EPA, Region X, are gathering chemical and biological data to potentially develop companion values for freshwater sediments.

Other efforts are under way outside Puget Sound and the Pacific Northwest to develop sediment quality values using the AET approach. These include California and the Great Lakes region in the United States, and the countries of Canada, New Zealand, and Australia internationally.

11.4.2 Extent to Which Approach Has Been Field-Validated

As described in U.S. EPA (1988), the reliability of AETs generated from Puget Sound data was evaluated with tests of sensitivity and efficiency (defined in Section 11.3.2.9). Tests of the sensitivity and efficiency of the AET approach were carried out in several steps, as described below:

- The chemical database was subdivided into groups of stations that were tested for the same biological effects indicators. Specifically, all chemistry stations with associated amphipod bioassay data were grouped together (287 stations), all chemistry stations with associated benthic infaunal data were grouped together (201 stations), all chemistry stations with associated oyster larvae bioassay data were grouped together (56 stations), and all chemistry stations with associated Microtox bioassay data were grouped together (50 stations). Stations with more than one biological indicator were included in each appropriate group.

- The stations in each group were classified as impacted or nonimpacted based on the appropriate statistical criteria (i.e., F_{max} and t-tests at alpha = 0.05).

- Several tests of reliability were conducted at this point:

  - Test 1: AET values (dry weight) were generated with the entire Puget Sound database available in 1988, and sensitivity and efficiency tests were performed against the same database for each biological indicator.

  - Test 2: The test described above was repeated in two parts: (a) using TOC-normalized AET values for nonionic organic compounds and dry weight-normalized AET values for all other compounds (i.e., ionizable organic compounds, metals, and metalloids), and (b) using TOC-normalized data for all chemicals. Test 2 allowed for *a posteriori* evaluation of the relative success of dry weight and TOC normalization for nonionic organic chemicals.

  - Test 3: Because the efficiency of the AET based on the entire Puget Sound database is 100 percent by constraint (as in Tests 1 and 2), predictive efficiency was estimated by the following procedure. For each biological indicator, a single station was sequentially deleted from the total database, AETs were recalculated for the remaining data set, and biological effects were predicted for the single deleted station. The predictive efficiency was the cumulative result for the sequential deletions of single stations. For example, the 287-sample database for amphipod bioassay results can be used to provide a 286-sample independent database for predicting (in sequence) effects on all 287 samples.

  - Test 4: In this test, independent data sets were used to generate and test AETs to confirm the sensitivity and efficiency measurements in Tests 1 and 3. AETs (dry weight) generated with 188 stations from diverse geographic regions in Puget Sound were tested with a completely independent set of 146 Puget Sound stations.

In addition, the influence of geographic location and other factors on AET predictive ability were examined (Barrick *et al.*, 1988). Further testing of Puget Sound AET values using matched biological/chemical data from other geographic areas is desirable before recommending direct application of the Puget Sound values in other geographic regions.

11.4.3 Reasons for Limited Use

The AET approach is being increasingly used outside of Puget Sound and the Pacific Northwest to evaluate and compare different classes of sediments and to develop bay-, site-, or region-specific sediment quality values for a variety of regulatory uses. Because the approach is based on empirical data, direct application of values from

Puget Sound or another area to a specific bay, site, or region usually encounters some conflicting or confounding data. Because regional reference areas are used to determine the significance of adverse biological effects in the AET approach, the AET developed for one region may be overprotective or underprotective of the resources in the other area. Additionally, the mix of chemicals in one region's sediments may not be the same in another region. The use of the AET *approach* and use of specific AET *values* should not be confused.

Development of site-specific AETs for other geographic areas may require additional sampling. Because many past studies were not multidisciplinary, measurements were often made only for chemistry or biology rather than for both kinds of information. In such cases, there will be a limited amount of appropriate historical data that can be used to develop AETs. The integration or comparison of AET data sets among different regions can also be restricted because appropriate biological indicators for generating AETs may vary among regions.

11.4.4 Outlook for Future Use and Amount of Development Yet Needed

The following two approaches to AET development could be particularly beneficial in expanding the use of this approach:

- Use of laboratory cause-effect (spiking) studies to evaluate AET predictions on a chemical-specific basis and

- Use of a large set of matched biological/chemical data from different geographic areas to test the predictive ability of AET and to test the "precision" of AET values based on data sets from different areas.

The AET approach was presented to (USEPA, 1988) and reviewed by the U.S. EPA Science Advisory Board (SAB, 1989). The SAB noted major strengths and limitations of the method and provided recommendations that would improve the validity of the AET values. The method was considered to contain sufficient merit for use in developing location-specific sediment quality values. Because of the specificity of the method, i.e., the empirical applications at specific localities, under specific environmental conditions, the approach seemed less useful for development of general, broadly applicable (i.e., national) sediment quality criteria.

11.5 REFERENCES

Barrick, R.C., S. Becker, L. Brown, H. Beller, and R. Pastorok. 1988. Sediment quality values refinement: 1988 update and evaluation of Puget Sound AET. Volume I. Final Report. Prepared for Tetra Tech, Inc. and U.S. Environmental Protection Agency Region X, Office of Puget Sound. PTI Environmental Services, Bellevue, WA. 74 pp. + appendices.

Becker, D.S., R.P. Pastorok, R.C. Barrick, P.N. Booth, and L.A. Jacobs. 1989. Contaminated sediments criteria report. Prepared for the Washington Department of Ecology, Sediment Management Unit. PTI Environmental Services, Bellevue, WA. 99 pp. + appendices.

Bellan-Santini, D. 1980. Relationship between populations of amphipods and pollution. Mar. Poll. Bull. 11:224-227.

Beller, H.R., R.C. Barrick, and D.S. Becker. 1986. Development of sediment quality values for Puget Sound. Prepared for Resource Planning Associates, U.S. Army Corps of Engineers, Seattle District, and Puget Sound Dredged Disposal Analysis Program. Tetra Tech, Inc., Bellevue WA. 128 pp. + appendices.

Nielsen, D. 1988. SEDQUAL users manual. Prepared for Tetra Tech, Inc. and U.S. Environmental Protection Agency Region X, Office of Puget Sound. PTI Environmental Services, Bellevue, WA.

Pearson, T.H., and R. Rosenberg. 1978. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. Oceanogr. Mar. Biol. Annu. Rev. 16: 229-311.

Phillips, K., P. Jamison, J. Malek, B. Ross, C. Krueger, J. Thornton, and J. Krull. 1988. Evaluation procedures technical appendix-Phase 1 (Central Puget Sound). Prepared for Puget Sound Dredged Disposal Analysis by the Evaluation Procedures Work Group. U.S. Army Corps of Engineers, Seattle, WA.

Puget Sound Water Quality Authority. 1988. 1989 Puget Sound Water Quality Management Plan. Puget Sound Water Quality Authority, WA. 276 pp.

PTI. 1987. Policy implications of effects-based marine sediment criteria. Prepared for American Management Systems and U.S. Environmental Protection Agency, Office of Policy Analysis. PTI Environmental Services, Bellevue, WA.

PTI. 1988. Elliott Bay Action Program: 1988 action plan. Prepared for Tetra Tech, Inc. and U.S. Environmental Protection Agency. PTI Environmental Services, Bellevue, WA. 43 pp. + appendices.

Sokal, R.R., and F.J. Rohlf. 1969. Biometry. W.H. Freeman and Company, San Francisco, CA. 859 pp.

State of Washington, Department of Ecology. 1991. Chapter 173-204, Washington Administrative Code, Sediment Management Standards. Olympia, WA.

Swartz, R.C., W.A. DeBen, J.K. Phillips, J.D. Lamberson, and F.A. Cole. 1985. Phoxocephalid amphipod bioassay for marine sediment toxicity. pp. 284-307. In: Aquatic Toxicology and Hazard Assessment: Proceedings of the Seventh Annual Symposium. R.D. Cardwell, R. Purdy, and R.C. Bahner (eds.). ASTM STP 854. American Society for Testing and Materials, Philadelphia, PA.

Tetra Tech. 1986. Recommended protocols for measuring selected environmental variables in Puget Sound. Final report. Prepared for U.S. Environmental Protection Agency, Region X, Office of Puget Sound, Seattle, WA. Tetra Tech, Inc., Bellevue, WA.

USEPA, 1989. Science Advisory Board. Report of the Sediment Criteria Subcommittee, Evaluation of the Apparent Effects Threshold (AET) Approach for Assessing Sediment Quality. SAB-EETFC-89-027. Office of the Administrator, Science Advisory Board, Washington, DC.

USEPA. 1988. Briefing report to the EPA Science Advisory Board. Prepared for Battelle and U.S. Environmental Protection Agency, Region X, Office of Puget Sound. PTI Environmental Services, Bellevue, WA. 57 pp.

A Summary of the Sediment Assessment Strategy Recommended by the International Joint Commission

Philippe Ross
The Citadel, Department of Biology
Charleston, SC 29409
(803) 792-7875

The International Joint Commission (IJC) Sediment Subcommittee has published a document entitled *Procedures for the Assessment of Contaminated Sediment Problems in the Great Lakes* (IJC, 1988a). An overview of the IJC strategy for assessing contaminated sediments is provided in this chapter. However, because it would be inappropriate to reproduce all, or substantially all, of the document in this chapter, the interested reader is referred to the IJC (1988a) document itself for an explanation of details that are not provided herein.

12.1 SPECIFIC APPLICATIONS

12.1.1 Current Use

The IJC (1988a) document is intended as guidance for the assessment of contaminated sediments in the Great Lakes. Its first application is in a work plan for sediment investigations at Great Lakes areas of concern (AOCs, as identified by the IJC). Section 118(c)(3) of the Water Quality Act of 1987 calls for U.S. EPA's Great Lakes National Program Office to survey at least five AOCs as part of a 5-yr study and demonstration program called ARCS (Assessment and Remediation of Contaminated Sediments). The strategy recommended by IJC (1988a) will be applied through a series of activities involving physical mapping and characterization, sampling, chemical analyses, toxicity testing, and *in situ* community analysis. The assessment began in 1989 and was completed in 1991. The ARCS program also seeks to improve upon the IJC (1988a) approach by comparing various test methods and by evaluating cost-effective reconnaissance and screening methods.

12.1.2 Potential Use

Other AOCs will eventually be evaluated in the process of developing remedial action plans. It is possible that other Great Lakes harbors, rivers, and estuaries will be added to the list of AOCs, in which case remedial action plans would have to be developed there. In addition, the guidance document could potentially be used to assess suspected sediment contamination outside the Great Lakes basin.

12.2 DESCRIPTION

12.2.1 Description of Method

12.2.1.1 Objectives and Assumptions

In response to the need for a common approach to the assessment of contaminated sediments, the IJC's Sediment Subcommittee has developed a strategy based on protocols that emphasize biological monitoring. The approach is intended for use in comprehensive assessments of areas (e.g., bays, harbors, rivers, other depositional zones) where sediment contamination and the need for remedial action are suspected. While the suggested strategy attempts to minimize the cost and expertise, the assessments are relatively large undertakings appropriate to situations where large-scale remedial actions might be contemplated. In such cases, the cost of conducting accurate assessments would be justified if the subsequent remedial options could cost far more than the assessments. It was not the primary intent of the subcommittee to provide guidance for small-scale decision-making activities, such as sample-by-sample disposal of dredged material from navigation channels. Nevertheless, some of the component methods described could be useful and cost-effective in this regard. The first major assumption, therefore, is that the scope of the study in question is sufficient to warrant a large-scale integrated investigation.

Another fundamental assumption is that the ultimate concern of a problem assessment focuses on whether sediment contaminants are exerting biological stress or are being bioaccumulated. Accepting this assumption, it follows that adequate assessments of sediment quality should involve components of chemistry, toxicity, and infaunal community structure (Chapman and Long 1983), a concept frequently referred to as the Sediment Quality Triad approach (see Chapter 9). The proposed strategy has the following objectives:

- To provide accurate assessments of specific problems by using a modified "triad" approach, which integrates chemical, physical, and biological information;

- To perform tasks in a sequence so that the results from each technique can be used to reduce subsequent sampling requirements and costs;

- To provide adequate proof of linkage between the contamination and the observed biological impact;

- To quantify problem severity, thereby enabling intercomparisons between and within areas of investigation (thus allowing development of a priority list for remedial actions and the objective selection of appropriate remedial options);

- To consider the effects on different species and different trophic levels, since biological impairment may occur in the water column and the sediments if resuspension occurs and since there is no such thing as the universal "most-sensitive species" (Cairns, 1986).

The IJC approach is an integrated strategy that provides the necessary data to identify sediment-associated contamination as the problem source, specify effects, rank problem severity, and assist in the selection of remedial options. While the assessment portion of the document identifies a set of the best currently available assessment tools (see Section 12.2.1.2.2), it is assumed that decisions will be made based on the circumstances unique to each AOC. There is no substitute for experience (expert judgment), and it is also assumed that appropriate expertise will be assembled before the assessment study plan is formulated.

12.2.1.2 Level of Effort

12.2.1.2.1 Type of Sampling Required

The IJC (1988a) approach involves two stages. Stage I, the initial assessment, is used for areas where an inadequate or outdated database exists. In the IJC document, Stage I is not subdivided, while Stage II is broken into Phases I, II, III, and IV. Stage I uses only *in situ* assessment techniques and criteria: a limited physical description of the area (e.g., basin size and shape, bathymetry) and the sediments, bulk chemical analyses, resident benthic community organization (e.g., family-level identifications), fish contaminant body burdens (one important species, selected by expert judgment), and external abnormalities on collected specimens. Any one of the following criteria provides sufficient justification for proceeding to Stage II:

- Concentrations of metals above background levels in sediments;

- Concentrations of hazardous persistent organic compounds above best available detection levels in sediments;

- Concentrations of hazardous persistent organic compounds above detection levels in fish or benthos;

- The absence of a healthy benthic community (e.g., absence of clean water organisms such as amphipods or mayflies, presence of a community dominated by oligochaetes, the complete absence of invertebrates); and

- Presence of external abnormalities in fish.

These conditions must be supported by evidence that the observed situation is not due to a major sediment perturbation, such as dredging or substrate modification.

Available data may preclude the need for a Stage I assessment. The cost and effort that Stage I entails should be avoided if there is already strong evidence of a contamination problem.

When a probable sediment contamination problem is identified, either through the initial assessment or from the examination of existing data, then Stage II, the detailed assessment, should be undertaken. The detailed assessment consists of four phases, which together define the sediment problem in the most cost-effective manner. The phases are not inflexible protocols, but rather logical groupings of work units. The expert investigator should be responsible for the final study design.

In Phase I of Stage II, extensive information on the physical composition of the sediments is collected. These data are used to define areas or zones of homogeneity within a study area. Knowledge of these zones allows sampling requirements for Phase II to be estimated.

In Phase II of Stage II, the benthic community structure is examined to the lowest possible taxonomic level (e.g., species or variety), along with the surficial sediment chemistry (e.g., pH, total organic carbon, redox potential, metals, extractable organic compounds). Phase II results can be combined with Phase I data to reduce the sampling effort in the next phase.

In Phase III of Stage II, a battery of laboratory bioassays (e.g., Microtox, algal, daphnid, benthic invertebrate, fish, Ames test) are performed on a smaller number of sediment samples than those in the Phase II sample set. Since fresh sediment must be collected for this phase, precision position-finding equipment is required to relocate previously sampled sites. Phase III costs can be reduced by performing acute lethality bioassays on a sediment sample before proceeding to tests that measure chronic or sublethal effects. Also in Phase III, sediment cores are collected, dated, and sectioned for stratified chemical analyses and bioassays. Finally, adult fish are examined histopathologically for internal (e.g., liver) tumors. In relatively confined geographical areas, Phases II and III may be combined because further sampling may be more costly than conducting additional bioassays and relocating Phase II sets for Phase III sampling may be difficult. In this case, Phase II sampling will include extra material for Phase III.

In the fourth and final phase of Stage II sediment dynamics (e.g., accumulation, resuspension, movement) and factors affecting them are quantified. All of the foregoing information is necessary for the selection of appropriate remedial options. For example, depositional history, as revealed by sampling sediment cores, and sediment dynamics are critical pieces of information in the selection and cost evaluation of remedial options.

Criteria that clearly indicate when some form of remedial action must be considered (based on the results of Stage II) are essential. Because of the absence of definitive sediment action criteria at time of writing, the criteria proposed by the IJC (1988a) are highly conservative, following the language of the 1978 Great Lakes Water Quality Agreement as revised in 1987 (especially Annexes 1 and 12), in order to promote maximum protection and effective restoration of the Great Lakes ecosystem. The IJC (1988a) urges that these criteria be reviewed regularly to ensure that they continue to fulfill their intended purpose.

12.2.1.2.2 Methods

During Stage I, the minimum amount of information necessary to assess potential problem sediments is collected. A variety of physical, chemical, and biological measurements are recommended, as outlined below:

- A geographical description of the area and its bathymetry is required.

- Sediment grain size - Size analysis techniques based on settling velocity (American Society for Testing and Materials, 1964; Duncan and LaHaie, 1979) are recommended. The sand fraction is removed by a 62-μm sieve and analyzed separately from the fine-grained material.

- Sediment water content - The water content can be determined during sample preparation for grain size and other analyses by comparison of sample weights before and after either freeze-drying or oven-drying (Adams *et al.*, 1980).

- Redox potential (Eh) and pH should be measured [specific methods are not recommended by IJC (1988a)].

- Organic carbon - It is recommended that total sediment organic carbon be measured as described by Plumb (1981).

- Phosphorus - Two measurements are suggested: total phosphorus, as extracted from sediment by sodium carbonate fusion or by perchloric acid digestion, and bioavailable phosphorus, as estimated by NaOH extractable phosphorus (Williams *et al.*, 1980).

- Ten metals (lead, nickel, copper, zinc, cadmium, chromium, iron, manganese, mercury, and arsenic) are recommended for routine analysis at Great Lakes AOCs. Additional metal analyses are left to the judgment of the investigator. An extraction procedure using a mix of hydrochloric and nitric acids (1:1) is suggested (Plumb, 1981).

- Persistent organic compounds - The reader is referred to the U.S. EPA (1984) protocols for broad scans and analyses of individual compounds. When the strategy was written, no standardized chemical protocols for estimating bioavailability of trace organic compounds were identified.

- External abnormalities in fish - The presence of one or more external abnormalities is often indicative of anthropogenically induced stress or damage. In the case of the brown bullhead, *Ictalurus nebulosus*, phenomena such as stubbed barbels, skin discoloration (melanoma), and skin tumors are highly correlated with liver cancer incidence (Smith *et al.*, 1988). It is recommended that locally occurring catfish (particularly *I. nebulosus*) be examined for tumors, melanoma, blindness, and barbel abnormalities during a Stage I assessment.

- Contaminant body burdens - The benthic infauna are in continuous contact with the sediments, providing a direct measure of the specific relationship between localized sediment contaminant concentrations and bioavailability. Carp are also regularly in contact with and ingest large quantities of sediments. They represent a larger spatial and temporal integration of contaminants than do the benthic infauna. Collection of adult common carp (*Cyprinus carpio*) for tissue residue analysis is recommended. Three to five fish per replicate should be composited. The number of replicates is determined using variability estimates from monitoring programs (Schmitt *et al.*, 1983) and a chosen level of precision, to calculate an idealized sample size (p. 247, Sokal and Rohlf, 1969). It is also recommended that the most abundant benthic invertebrate species (often oligochaete worms in contaminated sediments) be sampled in early summer, prior to thermal stratification. Standard U.S. EPA methods are suggested for tissue residue analysis. The problem of obtaining enough biomass for analysis (at least 1 g) is recognized.

- Benthic community structure - In a Stage I assessment, a preliminary analysis of community structure impairment is recommended. A qualitative study with minimal replication and identification only to the family level is suggested. Because it is important that rare taxa be sampled, simple techniques that employ inexpensive equipment but take large samples are recommended. This approach should suffice to identify the existence of a stressed community for the purposes of Stage I criteria (see Section 12.2.1.2.1 above).

Phase II of the detailed assessment consists of more focused analyses to supplement or complement information obtained in Stage I. Phase I of the detailed assessment focuses on physical mapping of the environment. The most important aspect of the physical assessment of a suspected contaminated sediment deposit is its three-dimensional mapping. A rectangular grid pattern is recommended for the initial mapping operation. Concurrent with bottom sampling at grid intersections, echo-sounder and side-scan sonar surveys should be performed to improve spatial resolution of sediment zones and bottom features. Detailed surveys should include piston coring for stratigraphic resolution. The grid sampling results should be examined using cluster analysis (or similar techniques), which are easy to interpret and functional with a small number of variables. Basic information required in this phase includes geographic location, areal extent, thickness and total sediment volume, average depths of overlying water, and the grain size properties of the deposit. Phase I results are used to select sampling sites for later phases.

Phase II of the detailed assessment focuses on surficial sediment chemistry and benthic community structure. Based on the previous mapping of homogeneous zones (Phase I), effort in Phase II can be expended in depositional areas and in those areas with fine-grained sediments. Surficial chemistry sampling should be coincident with the sampling for detailed benthic community structure analysis. Total organic carbon, redox potential, pH, metals, and persistent organics should be measured. Investigators are referred to Plumb (1981), Williams *et al.* (1980), and U.S. EPA (1984) for collection and analysis methods.

Since the main objective of Stage II community structure assessment is to examine subtle distinctions in stress response, more detailed taxonomic data are required in this phase than were required in Stage I. In the study design and sample collection steps, investigators are urged to follow the 10 principles of sampling set forth by Green (1979). Further guidance is given in Elliott (1977) for critical factors such as site selection, sample numbers, sampling design, and data analyses. To help investigators assess community impact, IJC (1988a) provides a partial list of literature descriptions of normal nearshore communities in habitats that most closely approximate Great Lakes AOCs. A detailed discussion of statistical methods is also included.

Phase III of the detailed (Stage II) assessment consists of obtaining additional information concerning sediment toxicity (i.e., bioassays and fish histopathology) and stratigraphic characterization of sediment cores. A suite of bioassays is proposed for toxicological evaluation of sediments:

- Microtox - an acute, liquid-phase (elutriate or pore-water) test with luminescent bacteria (Bulich, 1984);
- Algal photosynthesis - an acute, liquid-phase test using natural communities [algal fractionation bioassay (Munawar and Munawar, 1987)] or the laboratory species *Selenastrum capricornutum* (Ross *et al.*, 1988);
- Zooplankton life-cycle tests (*Daphnia magna* liquid and solid phases) monitoring growth and reproduction (Nebeker *et al.*, 1984; LeBlanc and Surprenant, 1985);
- Chronic, solid-phase tests using the benthic invertebrates *Chironomus tentans* (Nebeker *et al.*, 1984), *Hyalella azteca* (Nebeker *et al.*, 1984), or *Hexagenia limbata* (Malueg *et al.*, 1983);
- A solid-phase fish bioaccumulation test with the fathead minnow *Pimephales promelas* (Mac *et al.*, 1984)

- The liquid-phase (extract) Ames *Salmonella*/microsome assay, a bacterial mutagenicity test (Tennant *et al.*, 1987).

In addition to bioassays, histopathological examinations of indigenous adult fish (especially *Ictalurus nebulosus*), focusing on preneoplastic and neoplastic liver lesions (Couch and Harshbarger, 1985), are recommended.

Also included in Phase III (Stage II) work are chemical analyses and dating of sediment cores. Isotopic (^{14}C, ^{210}Pb, ^{55}Fe, ^{137}Cs) and biostratigraphic [i.e., ragweed (*Ambrosia*) pollen] methods are both recommended for dating sediment cores. This dating is necessary to establish the three-dimensional configuration of the contaminated sediment mass and to assign a date to the sediment depositional unit.

In Phase IV (Stage II) of the detailed assessment, studies on sediment dynamics are necessary to determine the following:

- Potential water column impacts through resuspension;
- Movement of contaminated sediment out of the AOC;
- The quality and rate of new sediment accumulation; and
- Vertical and horizontal redistribution of sediments and their contaminant burdens within an AOC.

This information is essential for the development and evaluation of a remediation plan. In the absence of practical predictive models, suspended sediment characterization (Poulton, 1987), shear strength measurements (Terzaghi and Peck, 1967), and resuspension studies (Tsai and Lick, 1986) are recommended.

12.2.1.2.3 Types of Data Required

The Stage I initial assessment should be based on aberrant macrozoobenthic community structure (ascertained from family-level taxonomic identification); metals concentrations above background levels in the surficial sediments (ascertained from dating); hazardous persistent organic compound concentrations above detection levels in carp, benthos, or surficial sediments; metals concentrations in carp or benthos, established on a case-by-case basis; and presence in fishes of external abnormalities known to have contaminant-related etiologies.

The Stage II detailed assessment should be based on a phased sampling of the physical, chemical, and biological aspects of the sediments. The biological impacts should be assessed with both field (benthic invertebrate community structure and incidence of fish liver tumors) and laboratory (battery of selected bioassays) methods. The phased sampling approach will allow subsequent testing requirements to be reduced. When Phases I and II of Stage II have revealed homogeneous zones of sediment type and similar community structure, the number of Phase III samples can be appropriately scaled down. Impairment due to sediment contamination and the probable need for remediation are established when the biomonitoring results from the detailed assessment demonstrate significant departures from controls.

Each section of IJC (1988a) contains a detailed discussion of the statistical procedures required, with references and examples. The preferred method of interpretation is left to the expert investigator in many cases.

12.2.1.2.4 Necessary Hardware and Skills

The initial assessment, and to an even greater degree the detailed assessment, requires a large array of field and laboratory equipment. Although none of the items recommended are unusual or inordinately sophisticated, one laboratory or field unit is unlikely to have all the required apparatus. Specific suggestions for hardware and skills are provided by IJC (1988a). Because this approach is intended for major sediment assessment efforts, several groups would probably have to be mobilized to contribute to the effort.

12.2.1.3 Adequacy of Documentation

Each component method described in IJC (1988a) is fully referenced in the text and accompanied by a separate bibliography. Some methods are more developed than others, and areas where additional validation or calibration is needed are clearly identified in the text.

12.2.2 Applicability of Method to Human Health, Aquatic Life, or Wildlife Protection

The IJC strategy includes direct measures of effects on benthic infauna and fishes and is thus directly applicable to aquatic biota. Existing sediment assessment methods (e.g., Apparent Effects Threshold, Sediment Quality Triad) could be used to evaluate the results of the Stage II detailed assessment and to determine whether chemically contaminated sediments have affected aquatic biota in the vicinity of AOCs. Although the IJC (1988a) strategy was not designed to assess the effects of toxic chemicals on wildlife or humans, the tissue residue data and the sediment chemistry data may be useful in preliminary evaluations of contaminant exposure to these populations. Wildlife exposure could occur through consumption of chemically contaminated prey. Human exposure could occur through consumption of chemically contaminated fish or through dermal absorption by direct contact with chemically contaminated sediments or water.

12.2.3 Ability of Method to Generate Numerical Criteria for Specific Chemicals

The document was designed to provide guidance to assessment programs. Nevertheless, since chemical, toxicological, and infaunal data are collected in the Stage II assessment, it is possible that these data could be used to develop chemical-specific criteria. For example, data from the Stage II assessment could be used to develop empirical sediment quality values (e.g., AET values) that are protective of aquatic biota in locations other than the AOC under consideration.

12.3 USEFULNESS

12.3.1 Environmental Applicability

12.3.1.1 Suitability for Different Sediment Types

The approach recommended in IJC (1988a) is suitable for any sediment type. Indeed, one of its major objectives is to characterize and provide a three-dimensional map of the contaminated sediment mass, including physical, chemical, and biological variables. The investigator is given the flexibility to choose the appropriate sampling methods for the sediment type or types in the AOC under study.

12.3.1.2 Suitability for Different Chemicals or Classes of Chemicals

The document is intended for situations where contamination is suspected, but where the toxic chemicals may or may not be identified. The methods recommended by IJC (1988a) are effective for most contaminants found in Great Lakes sediments. The broad-based nature of the approach contains sufficient flexibility to deal with anomalous situations.

12.3.1.3 Suitability for Predicting Effects on Different Organisms

The proposed strategy includes both laboratory testing and analysis of indigenous communities (i.e., fish, macrozoobenthos). In this way, laboratory results (i.e., chemistry, toxicity) that can be compared to standard conditions and literature values may be placed in the context of empirically derived effects data from the site under investigation.

12.3.1.4 Suitability for In-Place Pollutant Control

The guidance document was developed specifically for the assessment of in-place pollutant problems. It is designed to fit into the framework of evaluating and choosing remedial options by providing an adequate database on which to base such decisions. A companion document (IJC, 1988b) provides guidance in the selection of courses of remediation.

12.3.1.5 Suitability for Source Control

The detailed assessment provides an adequate framework for identifying hot spots, and for establishing significant differences from background conditions. In some cases, the resultant maps may provide further evidence of contaminant sources and migration patterns, using spatial autocorrelation techniques. Presumably, such evidence could facilitate regulation of identified sources. However, source control is not a primary objective of the IJC (1988a) strategy.

12.3.1.6 Suitability for Disposal Applications

Although the document was not intended for the use in decision-making related to the disposal of material from navigational dredging, the data generated from an initial assessment could be used to make initial disposal decisions. Other practices for the assessment of dredged material may be more cost-effective, however.

12.3.2 General Advantages and Limitations

12.3.2.1 Ease of Use

The proposed strategy is designed to be applicable to the AOC under investigation. It is intended to flexible, relying on the judgment and experience of those who apply it. A detailed assessment would be practical only in cases where a major remedial effort is contemplated.

12.3.2.2 Relative Cost

The Stage I and II assessments are costly compared to other less comprehensive methods of assessing sediment quality. However, when compared to the potential remedial costs, the assessment costs are relatively small. The sequential approach is designed to reduce sampling, analysis, and expense where possible. In many cases, the Stage I assessment need not be done. If it is clear that a sediment contamination problem exists, then the investigators may proceed directly to Stage II assessment. Alternatively, if the Stage I assessment produces no results of concern, then Stage II need not be undertaken. The cost of a detailed assessment, although relatively high, is controlled somewhat by the sequential approach to data collection. No firm cost figures are currently available, but assessments planned for priority AOCs under Section 118(c)(3) of the Water Quality Act of 1987 are projected to cost in the range of $500,000. These costs are expected to vary from site to site.

12.3.2.3 Tendency to Be Conservative

The strategy is designed to be highly protective of the environment. It combines chemical analysis, toxicity testing, and examination of indigenous communities to ensure that no significant effects are overlooked. Because the application of criteria is left to the expert judgment of the investigator, the degree of conservatism in decision-making will be variable.

12.3.2.4 Level of Acceptance

The guidance document (IJC, 1988a) does not describe a new method, but rather a combination of several types of methods, each widely accepted in its own sphere. The strategy as a whole is being used for the first time in 1989.

12.3.2.5 Ability to Be Implemented by Laboratories with Typical Equipment and Handling Facilities

None of the methods is particularly unusual or difficult, but the detailed assessment requires a breadth of expertise and resources that an individual organization may not possess. The strategy will need to be implemented by drawing on a variety of expertise in a given geographical area.

12.3.2.6 Level of Effort Required to Generate Results

The total level of effort for a detailed assessment will be relatively high in most cases. This strategy is most suitable for major evaluation projects.

12.3.2.7
Degree to which Results Lend Themselves to Interpretation

The actual statistical analysis and interpretation to generate effects conclusions are relatively complex and should be done only by trained investigators. Specific statistical protocols are not recommended. However, the reader is given an array of choices, with comments on their respective strengths and weaknesses. The ultimate decision is left to the investigator. The inclusion of chemical, toxicological, and infaunal information in the database allows the investigator to compare different types of indicators before making decisions.

12.3.2.8 Degree of Environmental Applicability

One of the strengths of a strategy that includes *in situ* community analysis is that effects data have a high degree of environmental relevance. Site-relevant species can even be substituted in the bioassay battery if necessary, and the body burden and community structure data are always site-specific.

12.3.2.9 Degree of Accuracy and Precision

The strategy proposed by the IJC (1988a) is not a single method, but rather guidance for a study design containing many options and decision points. Overall precision or accuracy values would be impossible to calculate. Nevertheless, the criteria for selecting recommended protocols included a consideration of attainable precision. In many sections, the investigator is directed to choose the required level of precision for a given measurement during the study design process. The "accuracy" of an integrated strategy is difficult to assess, but the methods recommended by the IJC (1988a) were chosen for their relevance to the Great Lakes ecosystem.

12.4 STATUS

12.4.1 Extent of Use

IJC's (1988a) document was published in December 1988 and distributed in early 1989. The strategy is intended for the Great Lakes, and was used for the first time in 1989. Most of the individual methods recommended are widely used and accepted.

12.4.2 Extent to Which the Approach Has Been Field-Validated

The first extensive field validation of the approach was conducted in 1989-1991 as part of the ARCS program under section 118(c)(3) of the Water Quality Act of 1987. The ARCS Sediment assessment reports are expected to be released in 1993.

12.4.3 Reasons for Limited Use

Most component protocols are in wide use. Because the IJC (1988a) document describes a major effort with an integrated approach, the ARCS program is the only project where an undertaking using this approach has been initiated.

12.4.4 Outlook for Future Use and Development

With the backing of both signatories to the Great Lakes Water Quality Agreement, the document seems destined for widespread use in the Great Lakes basin. As methods progress, each section of the document will be updated.

12.5 REFERENCES

Adams, D.D., D.A. Darby, and R.J. Young. 1980. Selected analytical techniques for characterizing the metal chemistry and geology of fine-grained sediments and interstitial water. In: Contaminants and Sediments. R.A. Baker (ed.) Ann Arbor Sci. Pub., Inc. Ann Arbor, MI.

American Society for Testing and Materials. 1964. Procedures for testing soils. ASTM, Philadelphia, PA. 535 pp.

Bulich, A.A. 1984. Microtox - a bacterial toxicity test with general environmental applications. pp. 55-64. In: Toxicity Screening Procedures Using Bacterial Systems. D. Lin and B.S. Dutka (eds.). Marcel Dekker, New York, NY.

Cairns, J., Jr. 1986. The myth of the most sensitive species. BioScience 36:670-672.

Chapman, P.M., and E.R. Long. 1983. The use of bioassays as part of a comprehensive approach to marine pollution assessment. Mar. Pollut. Bull. 14:81-84.

Couch, J.A., and J.C. Harshbarger. 1985. Effects of carcinogenic agents on aquatic animals: an environmental and experimental overview. Env. Carcinogenesis Rev. 3:63-105.

Duncan, G.A., and G.G. LaHaie. 1979. Size analysis procedures used in the sedimentology laboratory, NWRI. Env. Can. NWRI contribution. 23 pp.

Elliott, J.M. 1977. Some methods for the statistical analysis of samples of benthic invertebrates. Scientific Publication No. 25. Freshwater Biological Association. 160 pp.

Green, R.H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley and Sons, New York, NY. 257 pp.

IJC. 1988a. Procedures for the assessment of contaminated sediment problems in the Great Lakes. International Joint Commission, Windsor, Ontario, Canada. 140 pp.

IJC. 1988b. Options for the remediation of contaminated sediments in the Great Lakes. International Joint Commission, Windsor, Ontario, Canada. 78 pp.

LeBlanc, G.A., and D.J. Surprenant. 1985. A method for assessing the toxicity of contaminated freshwater sediments. pp. 269-283. In: Aquatic Toxicology and Hazard Assessment, Seventh Symposium. R.D. Cardwell, R. Purdy, and R.C. Bahner (eds.), ASTM STP 854. American Society for Testing and Materials, Philadelphia, PA.

Mac, M.J., C.C. Edsall, R.J. Hesselberg, and R.E. Sayers, Jr. 1984. Flow-through bioassay for measuring bioaccumulation of toxic substances from sediment. EPA DW-930095-01-0. U.S. Environmental Protection Agency, Chicago, IL. 26 pp.

Malueg, K.W., G.S. Schuytema, J.H. Gakstatter, and D.F. Krawczyk. 1983. Effect of *Hexagenia* on *Daphnia* response in sediment toxicity tests. Env. Toxicol. Chem. 2:73-82.

Munawar, M., and I.F. Munawar. 1987. Phytoplankton bioassays for evaluating toxicity of *in situ* sediment contaminants. Hydrobiologia 149:87-105.

Nebeker, A.V., M.A. Cairns, J.H. Gakstatter, K.W. Malueg, and G.S. Schuytema. 1984. Biological methods for determining toxicity of contaminated freshwater sediments to invertebrates. Env. Toxicol. Chem. 3:617-630.

Plumb, R.H., Jr. 1981. Procedures for handling and chemical analysis of sediment and water samples. Technical Report EPA/CE-81-1. U.S. Environmental Protection Agency/U.S. Army Corps of Engineers Technical Committee on Criteria for Dredged and Fill Material, U.S. Army Waterways Experiment Station, Vicksburg, MS. 471 pp.

Poulton, D.J. 1987. Trace contaminant status of Hamilton Harbor. J. Great Lakes Res. 13:193-201.

Ross, P.E., V. Jarry, and H. Sloterdijk. 1988. A rapid bioassay using the green alga *Selenastrum capricornutum* to screen for toxicity in St. Lawrence River sediments. American Society for Testing and Materials. STP 988:68-73.

Schmitt, C.J., M.A. Ribick, J.L. Ludke, and T.W. May. 1983. National pesticide monitoring program: organochlorine residues in freshwater fish, 1976-79. Fish and Wildlife Service Res. Publ. No. 152. U.S. Dept. of Interior, Washington, DC.

Smith, S.B., M.J. Mac, A.E. MacCubbin, and J.C. Harshbarger. 1988. External abnormalities and incidence of tumors in fish collected from three Great Lakes Areas of Concern. Paper presented at the 31st Conference on Great Lakes Research, McMaster University, Hamilton, Ontario. May 17-20, 1988.

Sokal, R.R., and F.J. Rohlf. 1969. Biometry. W.H. Freeman and Co., San Francisco, CA.

Tennant, R.W., B.H. Margolin, D.D. Shelby, E. Zeiger, J.K. Haseman, J. Spalding, W. Caspary,

M. Resnick, S. Stasiewicz, B. Anderson, and R. Minor. 1987. Prediction of chemical carcinogenicity in rodents from *in situ* genetic toxicity assays. Science 236:933-941.

Terzaghi, K., and R.B. Peck. 1967. Soil mechanics in engineering practice. John Wiley and Sons, New York. 729 pp.

Tsai, C.H., and W. Lick. 1986. A portable device for measuring sediment resuspension. J. Great Lakes Res. 12:314-321.

USEPA. 1984. Guidelines establishing test procedures for the analysis of pollutants under the Clean Water Act; final rule and interim final rule and proposed rule. U.S. Environmental Protection Agency. Washington, DC. Federal Register Vol. 49, No. 209, Part VIII. pp. 1-210.

Williams, J.D.H., H. Shear, and R.L. Thomas. 1980. Availability to *Scenedesmus quadricauda* of different forms of phosphorus in sedimentary materials in the Great Lakes. Limnol. Oceanogr. 25:1-11.

Summary of Sediment-Testing Approach Used for Ocean Disposal

David P. Redford
U.S. Environmental Protection Agency
499 South Capitol Street, SW (WH-556F), Washington, DC 20003
(202)260-9179

The *Evaluation of Dredged Material Proposed for Ocean Disposal—Testing Manual* (USEPA/USACE, 1991) commonly referred to as the "Green Book," was published in February 1991 by the U.S. Environmental Protection Agency (USEPA) and the U.S. Army Corps of Engineers (USACE). The Green Book contains national guidance for evaluating the suitability of dredged material for ocean disposal; it replaces the guidance of the original manual (USEPA/USACE, 1977) that was published by USEPA and the USACE in 1977. The manual stresses the use of bioassay and bioaccumulation testing as evaluative tools, and it contains technical guidance on the use of such tests. The following is a summary of the 1991 manual and the approach used by USEPA and the USACE to determine the suitability of dredged material for ocean disposal. The manual will be revised at a future date, based on the findings of an EPA Science Advisory Board (SAB) review (SAB, 1992), and changes will be made to the Ocean Dumping Regulations (referenced below).

13.1 APPLICATION

The 1991 USEPA/USACE Green Book provides updated guidance for dredging applicants, scientists, and regulators to evaluate dredged-material compliance with the 1977 U.S. Ocean Dumping Regulations [Title 40, Code of Federal Regulations (CFR), Parts 220-228]. The manual is applicable to all activities involving the transportation of dredged material for the purpose of dumping it in ocean waters outside the baseline from which the territorial sea is measured. The guidance in this manual is applicable to dredging operations conducted under permits as well as to federal projects conducted by the USACE. The procedures in this manual do not apply to activities excluded by 40 CFR 220.1.

It is important to note that the regulations are legally binding and that the guidance provided in this manual is responsive to the specific requirements of these regulations, but *the manual does not carry the force of law.* The document simply provides guidance on evaluating the potential environmental impact of dredged-material ocean disposal.

The manual is organized into tiers for efficient evaluation of the suitability of dredged material for ocean disposal. Within the tiers, specific physical, chemical, and biological tests are recommended. To meet specific regional needs, USEPA Region and USACE District offices are to develop local agreements and manuals to implement the national guidance in the 1991 Green Book (such as using local species in biological tests and screening for particular contaminants in chemical analyses).

13.1.1 Current Use

The 1991 Green Book replaces the 1977 Green Book. USEPA Region and USACE District offices are developing local agreements and regional testing manuals that implement the 1991 Green Book guidance and establish permit procedures for dredging and dredged-material disposal.

Projects that have been issued under USACE permits prior to the completion of the new local agreement/manual for the area covered by the project may continue to be evaluated according to the 1977 guidance manual and the existing local guidance. New dredged-material disposal projects, projects that have not had sampling and analysis plans approved prior to finalization of the local agreement/manual, should be evaluated under the updated guidance in the 1991 Green Book. Ongoing projects that have been approved based on 1977 Green Book guidance should be reevaluated according to 1991 Green Book guidance and the new local agreement/manual within 3 years of permit approval.

13.1.2 Potential Use

The Green Book guidance, and revisions thereof, will be applied to dredged-material evaluations for the foreseeable future.

The manual will be revised at a future date based on (1) the findings of an EPA SAB review (SAB, 1992), (2) technical advances in assessing sediment contamination and marine environmental impact, and (3) changes to the Ocean Dumping Regulations.

13.2 DESCRIPTION

Analysis of sediment to determine its suitability for ocean disposal is conducted according to the procedures in the 1991 Green Book. The 1991 Green Book recommends procedures that satisfy section 103 of the Marine Protection, Research, and Sanctuaries Act of 1972 (MPRSA), Public Law 92-532. The MPRSA was enacted to regulate ocean dumping of all materials that might adversely affect human health, the marine environment, or other legitimate uses of the oceans. In addition, the MPRSA implements the Convention on the Prevention of Marine Pollution by Dumping of Wastes and Other Matter (London Dumping Convention), of which the United States is a signatory. MPRSA section 103 specifies that all proposed operations involving the transportation and dumping of dredged material into ocean waters must be evaluated to determine the potential environmental impact of such activities. These environmental evaluations must be in agreement with the criteria published in 40 CFR Parts 220-228 and 33 CFR Parts 320-330 and 335-338.

Technical guidance on specific methods for testing dredged material is presented in the 1991 Green Book. If the results of the appropriate tests show that the proposed dredged material meets the chemical- and biological-effects criteria, and meets other requirements in the regulations, disposal of the material at a designated ocean dredged-material disposal site (ODMDS) is supported. If the test results show that the material does not meet the criteria set forth in the regulations, significant impact on the ocean environment is predicted. Significant adverse impact may include adverse consequences to the marine ecosystem and negative human-health effects from uses of the marine environment.

The manual does not present guidance for the disposal of dredged material that fails to meet the regulatory criteria. Such disposal involves management decisions and case-specific engineering work (e.g., control of dump releases, disposal-site capping, submarine burial, and predisposal treatment) that are beyond the scope of the document.

13.2.1 Description of Method

Integral to the 1991 Green Book is a tiered-testing procedure to characterize dredged material and predict its impact on the water-column and benthic environment at ODMDSs. The procedure was developed by USEPA and USACE personnel and testing-laboratory researchers, and is consistent with the requirements of the Ocean Dumping Regulations, state-of-the-art dredged-material evaluation techniques, and the realities of the testing and permitting process for new and existing projects. Knowledge of local conditions is both recommended and necessary to adapt the national guidance in the manual to specific dredged-material projects. USEPA Regions and USACE Districts are presently developing local agreements/manuals to apply the national guidance of the manual to specific dredging and disposal areas.

The tiered-testing procedure in the Green Book comprises four tiers, with decision points at each tier (Figure 13-1). Each successive tier provides increasing investigative intensity to generate the information for permitting decisions on ocean disposal.

The tiered-testing procedure is constructed to determine whether the dredged material meets the limiting permissible concentration (LPC), as defined in section 227.27 of the Ocean Dumping Regulations. The LPC for the liquid-phase concentration of dredged material in the water column is the concentration that, after allowance for initial mixing, does not exceed applicable marine water-quality criteria (WQC) or a toxicity threshold of 0.01 of the acutely toxic concentration. The LPC of the suspended particulate and solid phases is the concentration that will not cause unreasonable toxicity or bioaccumulation.

The overall tiered-testing procedure is relatively flexible. The dredged-material evaluator can enter and exit the testing procedures at any tier. However, to begin the evaluation in Tier II, III, or IV, the existing data must satisfy the requirements of the earlier tier(s). Additionally, Tier II testing for water-quality criteria (WQC) compliance is mandatory if the water-column evaluation cannot be completed within Tier I. To exit any tier before reaching a decision on LPC compliance, the dredged-material evaluator must select an option other than open-ocean disposal.

In most cases, determinations of LPC compliance can be made in Tier I, II, or III. In extraordinary cases, where LPC compliance cannot be determined by Tier III, the dredged material must be evaluated under Tier IV. Tier IV tests are case-specific investigations of potential impact of the dredged material at the ODMDS. Significant investment in the research and development of analytical methods is usually necessary to conduct Tier IV evaluations, and the applicant might select an alternative to open-ocean disposal instead of proceeding with Tier IV testing. Similarly, an applicant can try to save time and money by proceeding directly to Tier II, III, or IV if it is believed that analysis in the earlier tiers will not lead to a definitive evaluation. The only absolute requirement is that the dredged material must comply with the regulations if it is to be dumped at an ODMDS. The tiered-testing procedure facilitates this determination.

In summary, the 1991 Green Book

- Includes state-of-the-art methods to determine the potential impact of marine-sediment disposal;
- Ensures adherence to the Ocean Dumping Regulations (40 CFR Parts 220-228);
- Incorporates existing (and valuable) regional expertise and guidance into the evaluation process; and
- Provides for National consistency in evaluating dredged material for ocean disposal.

13.2.1.1 Objectives and Assumptions

The objective of the tiered-testing procedure is to determine whether the water-column and benthic LPC is met for the proposed dredged material, as defined in the Ocean Dumping Regulations. Three decision options are possible as the dredged-material evaluator proceeds through the tiers.

(1) The LPC is met; the ocean disposal option is supported; further evaluation is unnecessary.

(2) The LPC evaluation is inconclusive; the ocean disposal option is not supported; proceed to the next tier.

(3) The LPC is not met; the ocean disposal option is not supported; further evaluation is unnecessary.

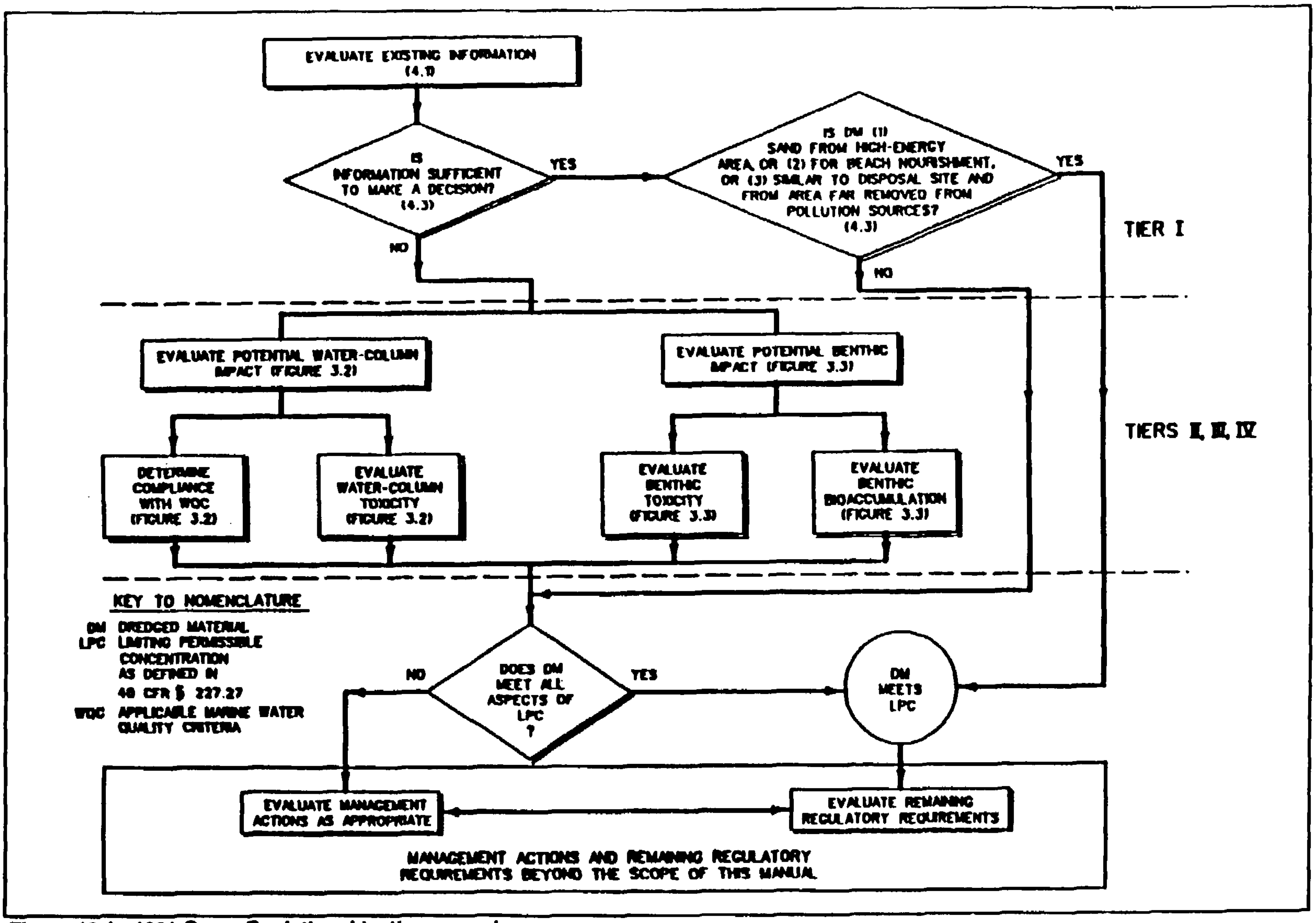


Figure 13-1. 1991 Green Book tiered-testing procedure.

Both the water-column and benthic LPC considerations must be satisfactorily resolved for the open-ocean disposal option to be supported. An inconclusive evaluation in Tiers I-III requires the dredging applicant to conduct additional testing in subsequent tiers, or to decide not to ocean-dump. However, a determination of LPC noncompliance does not necessarily exclude all possibilities for ocean disposal. Management actions might be feasible to make the dredged material meet the LPC. Management actions for dredged material that exceeds water-column or benthic LPC are *not* included in the Green Book because of the wide range of available options and the project-specific nature of such work.

It is assumed that the users of the 1991 Green Book are generally familiar with the need for and methods of dredged-material testing. The manual is not a standalone document. The guidance in the manual requires the evaluator to consult the regulations frequently (40 CFR Parts 220-228 is included in the Green Book as Appendix A) and to have a general understanding of material contained in the numerous citations and references. The guidance in the manual concentrates on data collection and decision points, and it only summarizes recommended field and laboratory procedures that can be used to obtain data. The user must refer to the original sources for most of the physical, chemical, and biological testing procedures.

13.2.1.2 *Level of Effort*

Tier I: Initial Assessments—Tier I is used to identify contaminants of concern and determine dredged-material LPC compliance through analysis of existing physical, chemical, and biological information. For many dredging projects, there is a wealth of readily available information on the proposed dredged material and on the characteristics of the disposal site. This is especially true of areas that have historically undergone maintenance dredging or have been the subject of other studies, such as fishery assessments. The available information for a given area might not be sufficient to reach a final LPC evaluation, but often there are accessible high-quality data that can supplement the results of tests in subsequent tiers and facilitate reaching an early decision with lowered expenditure of time and resources.

Whatever the source of information for Tier I evaluations, the quality of the data must be evaluated and weighed accordingly. The references in Chapter 13 of the manual, Quality-Assurance Considerations, should be consulted for guidance for evaluating the quality of data obtained from different information sources.

If the information set compiled in Tier I is complete and comparable to information that would appropriately satisfy the LPC in Tier II, III, or IV, a decision on regulatory compliance be completed without proceeding into the next tiers. For compliance determination to be completed within Tier I, the weight of evidence of the collected information must convincingly show that the dredged-material disposal either will or will not meet the LPC.

Included in Tier I is an assessment of the three exclusionary criteria in 40 CFR 227.13(b): (1) the dredged material is predominantly sand, gravel, or rock from a high-energy area; (2) the material is suitable for beach nourishment; or (3) the material is similar to the disposal site and from an area far removed from pollution sources. If one or more of the above exclusionary criteria can be satisfied, the LPC is met for the dredged material and no further evaluation is required. If none of the exclusionary criteria is met and the collected information is insufficient to reach a definitive LPC determination, the evaluation process moves to Tier II.

Tier II: Physical/Chemical Evaluations—Tier II consists of physical and chemical data evaluation. To determine marine WQC compliance, a numerical mixing model is used; to evaluate benthic-impact potential for nonpolar organic compounds, a theoretical bioaccumulation potential (TBP) calculation is used. The conceptual purpose of the tier is to provide reliable, rapid screening of impact potential without the need for further testing. This purpose is fulfilled for water-column evaluations, but at present there is no USEPA-approved single screening procedure for deposited sediment. When technically sound sediment-quality criteria (SQC) are developed and approved for dredged-material evaluation, they will be incorporated at this level.

Tier II: Water-Column Physical/Chemical Evaluations—The Tier II water-column evaluation for WQC compliance is a two-step process that includes the application of a numerical mixing model. In Step 1, the model is used as a screen; all of the contaminants in the dredged material are assumed to be released into the water column during the disposal process. If the model predicts that the concentration of contaminants of concern released into the water column is less than the applicable WQC and if no synergistic effects among the contaminants are suspected, the dredged material meets the water-column LPC and no further water-column evaluations are necessary.

If LPC compliance cannot be shown in Step 1, Step 2 is conducted. In Step 2, chemical data from an elutriate test of the dredged material are run in the model. Compared to the assumption of total contaminant release in the Step 1 screen, the elutriate data applied in Step 2 are a more precise representation of the concentration of contaminants that would actually be released into the water column during ocean disposal of dredged material.

If the model predicts in Step 2 that *any* WQC are exceeded, the water-column LPC is not met (open-ocean disposal not supported). If there are WQC for all of the contaminants of concern, if no WQC are exceeded by the Step 2 model, and if no contaminant synergistic effects are suspected, the water-column LPC is met and no further water-column evaluations are necessary (open-ocean disposal supported). If there are contaminants of concern without WQC or if synergistic effects are suspected, water-column toxicity and water-column LPC compliance must be evaluated in Tier III.

Numerical Models for Initial Mixing—Numerical models are used to evaluate dredged-material dilution during the initial-mixing phase of ocean disposal, as defined in the regulations. The 1991 Green Book recommends using the USACE Automated Dredging and Disposal Alternatives Management System (ADDAMS) models to evaluate initial mixing of dredged material at ODMDSs. ADDAMS models can be run on a personal computer with a minimum of hardware. The models account for the physical processes of dredged-material disposal at open-water disposal sites by calculating the water-column concentrations of dissolved contaminants and suspended sediments and the initial deposition of material on the bottom. Three separate ADDAMS models address different methods of disposal:

- DIFID — Disposal from an instantaneous dump
- DIFCD — Disposal from a continuous discharge
- DIFHD — Disposal from a hopper dredge

To evaluate initial mixing following ocean disposal, the appropriate model is run *for the contaminant requiring the greatest amount of dilution* to meet the LPC. The models simulate movement of the disposed material as it falls through the water column, as it is transported and diffused by the ambient current, and as it spreads over the bottom. The models have some limitations; for example, the DIFID model will not work for very shallow disposal sites where the discharge time from the barge exceeds the descent period to the bottom. However, the models can simulate a wide range of disposal options. USEPA and the USACE are in the process of field-verifying these models.

Appendix B of the 1991 Green Book is a summary of the ADDAMS models; the computer diskettes that accompany the manual contain the models themselves. ADDAMS modeling personnel at the USACE Waterways Experiment Station (WES), Vicksburg, Mississippi, are available to supply model updates, answer questions, and assist with the selection and running of the individual models.

Table 13-1. 1991 Green Book Species for Water-Column and Benthic Evaluations.

Water-Column Species	Benthic Species
■ Crustaceans	■ Crustaceans
Mysids	Infaunal Amphipods
Mysidopsis sp.[a]	*Rhepoxynius* sp.[a]
Neomysis sp.[a]	*Ampelisca* sp.[a]
Holmesimysis sp.[a]	*Eohaustorius* sp.[a]
Shrimp	*Grandiderella japonica*
Palaemonetes sp.	*Corophium insidiosum*
Penaeus sp.	Mysids
Pandalus sp.	*Mysidopsis* sp.
Crab	*Neomysis* sp.
Callinectes sapidus	*Holmesimysis* sp.
Cancer sp.	Shrimp
	Penaeus sp.
■ Fish	*Palaemonetes* sp.
Menidia sp.[a]	*Crangon* sp.
Cymatogaster aggregata[a]	*Pandalus* sp.
Cyprinodon variegatus	*Sicyonia ingentis*
Lagodon rhomboides	Crab
Leiostomus xanthurus	*Callinectes sapidus*
Citharicthys stigmaeus	*Cancer* sp.
Leuresthes tenuis	
Coryphaena hippurus	■ Fish
	Clevelandia ios
■ Zooplankton	*Atherinops affinis*
Copepods	
Acartia sp.[a]	■ Burrowing Polychaetes
Mussel larvae	*Neanthes* sp.[a]
Mytilus edulis[a]	*Nereis* sp.[a]
Oyster larvae	*Nephthys* sp.
Crassostrea virginica[a]	*Glycera* sp.
Ostrea sp.[a]	*Arenicola* sp.
Sea-urchin larvae	*Abarenicola* sp.
Strongylocentrotus	
purpura tus	■ Molluscs
Lytechinus pictus	*Yoldia limatula*
	Macoma sp.
	Nucula sp.
	Protothaca staminea
	Tapes japonica
	Mercenaria mercenaria

[a]Recommended test species.

The model output can present water-column contaminant concentrations in milligrams per liter. These concentrations are compared to the appropriate LPCs to determine compliance.

Tier II: Benthic Physical/Chemical Evaluations—As previously noted, only benthic effects attributed to nonpolar organic chemicals in the deposited sediment can be addressed in Tier II at the present time. Nonpolar organic chemicals include all organic compounds that do not dissociate or form ions. These include chlorinated hydrocarbon pesticides, other halogenated hydrocarbons, polychlorinated biphenyls (PCBs), most polynuclear aromatic hydrocarbons (PAHs), dioxins, and furans. It does not include polar organic compounds, organometals, and metals. If all of the contaminants of concern in the dredged material are *nonpolar* organic compounds, the theoretical bioaccumulation potential (TBP) can be calculated for the dredged material and the reference sediment[1] to determine benthic LPC compliance. The TBP calculation is an environmentally conservative screen, based on calculating the concentration of the nonpolar organic chemical in the sediment, the total organic-carbon concentration, and the percent lipid content of an organism of interest. If the TBP of the dredged material is not statistically greater than that of the reference material, the LPC for the nonpolar organic contaminants is met. (Acute-toxicity evaluations must be performed under Tier III unless sufficient toxicity information was obtained under Tier I.)

If any of the contaminants of concern are polar organic compounds or have suspected toxic components or if the dredged-material TBP exceeds the reference-material TBP described above, the bioaccumulation evaluation for benthic impact by the dredged material must take place in Tier III or IV. The benefit of additional tests in Tier II to screen for benthic impact is recognized by USEPA and the USACE, and new tests are under development and evaluation. When the scientific and regulatory community verifies one or more of these tests, they will be incorporated into Tier II in a future Green Book revision. Meanwhile, evaluation of benthic impact that cannot be made in Tier I must be completed in Tier III or IV.

Tier III: Biological Evaluations—Tier III tests include (1) determination of water-column toxicity and (2) assessment of contaminant toxicity and bioaccumulation from the material to be dredged. The evaluations in this tier are based on the output from Tiers I and II and comprise standardized bioassays with the organisms listed in Table 13-1.

Tier III: Water-Column Biological Evaluations—Tier III water-column tests are acute tests that evaluate the toxicity of the dissolved and suspended portions of the dredged material that remains in the water column after initial mixing. The bioassays are run if the Tier II evaluations are inconclusive, e.g., if there are not applicable WQC for all contaminants of concern or there is reason to suspect synergistic effects among the contaminants. (See Tier II.) The tests involve exposing fish, crustaceans, and zooplankton to a dilution series containing both dissolved- and suspended-sediment components of the dredged material. A typical test monitors organism mortality over a 96-h period.

The results of the bioassays are used to calculate the LC_{50} concentration of the dredged material in the water column. The LPC for this evaluation is 1 percent of the LC_{50} outside the ODMDS during the initial 4-h mixing period and anywhere in the marine environment 4 h after disposal. Following the determination of the LPC for the proposed dredged material, the data are used to run the numerical model (see model discussion above) and determine LPC compliance.

Tier III: Benthic Biological Evaluations—Benthic evaluations in Tier III consist of toxicity and bioaccumulation tests. To conduct these tests, the 1991 Green Book provides laboratory guidance on sediment preparation; treatment, reference-, and control-sediment tests; replicates; organism

[1]A reference sediment is defined as a sediment, substantially free of contaminants, that is as similar as practicable to the grain size of the dredged material and the sediment at the disposal site, and that reflects the conditions that would exist in the vicinity of the disposal site had no dredged-material disposal ever taken place, but had all other influences on sediment condition taken place.

handling; test-chamber conditions; QA considerations; and data analysis. The organisms used in the tests are surrogates for disposal-site species and are used to estimate dredged-material effects. The toxicity tests quantify mortality. If the mortality of the test species in the dredged-material bioassays is greater than the allowable percentage over the mortality in the reference-sediment bioassays, the LPC is not met. If, however, the dredged-material tests below the allowable percentage, or the increased mortality is statistically insignificant, the LPC is met.

The bioaccumulation tests evaluate the potential of benthic organisms to accumulate contaminants from the dredged material in their tissues. At the conclusion of the tests, the tissues of the organisms are analyzed for the contaminants of concern that are identified in Tier I.

Section 227.27 of the Ocean Dumping Regulations requires that benthic bioassays be conducted on dredged material with filter-feeding, deposit-feeding, and burrowing species. Infaunal amphipods, such as *Ampelisca* sp. and *Rhepoxynius* sp., are sensitive bioindicators and strongly recommended in the Green Book as the preferred species for toxicity tests. Infaunal amphipods filter-feed, deposit-feed, and, to some extent, burrow in the sediment, thereby fulfilling the three organism categories in the regulations. For bioaccumulation evaluations, the manual recommends using a burrowing polychaete (e.g., *Neanthes* sp. or *Nereis* sp.) and a deposit-feeding bivalve mollusc (e.g., *Macoma* sp. or *Yoldia limatula*). In summary, the manual recommends that at least two species be tested for acute toxicity and at least two other species for bioaccumulation evaluation. Each set of test species should cover the three species types stipulated in the regulations. The ecological and economic relevance of the organisms and the practical aspects of using the species in the laboratory, such as tolerance to grain-size ranges and seasonal availability, also must be considered when selecting the test species.

The Tier III bioaccumulation evaluation compares the contaminant level in the tissues of the organisms to two criteria: (1) the United States Food and Drug Administration (FDA) Action Levels for Poisonous or Deleterious Substances in Fish and Shellfish for Human Consumption and (2) the contaminant levels in organisms that are exposed to the reference sediment. Regardless of the statistical comparison to the reference-material test organisms, if the level in the tissues of dredged-material organisms statistically exceeds the FDA levels in any category, the LPC is not met. If the dredged-material results are lower than the FDA action levels and not statistically greater than the reference material level, the LPC for bioaccumulation is satisfied. However, if bioaccumulation exceeds that found in the reference-material tests, the test results must be evaluated against case-specific criteria. USEPA and the USACE develop the evaluative criteria case by case from local technical information that addresses the bioaccumulation aspects of the benthic criteria of section 227.13(c)(3) of the regulations.

At present, tests for chronic sublethal exposure to benthic contaminants are being developed. When the tests are approved by USEPA, they will be incorporated in Tier III in future updates to the Green Book.

Tier IV: Advanced Biological Evaluations—Tier IV consists of bioassay and bioaccumulation tests to evaluate the long-term benthic and water-column impact of dredged material. Tests at this level are selected to address specific issues for a specific dredging operation that could not be fully evaluated in the earlier tiers. Since these tests are case-specific and since they require significant time and money to complete, evaluative criteria must be agreed on in advance by USEPA and by the USACE to determine compliance with the LPC.

Conducting Tier IV benthic testing is possible with current methods, but the 1991 Green Book emphasizes that this tier is not intended for routine application. Tier IV benthic tests consume significant resources of the dredging applicant and of the regulatory authority, and a final noncompliance determination is still possible. Therefore, the applicant must weigh the options and decide whether to perform Tier IV testing or to consider an alternative that does not involve ocean dumping, such as upland disposal. If the applicant elects to proceed with Tier IV testing, the role of the regulatory authority is to design tests that lead to a definitive LPC evaluation for the project.

Under Tier IV evaluations, bioaccumulation testing measures the steady-state body burden of contaminants of concern in tissues of organisms subjected to long-term laboratory exposures or in tissues of appropriately sampled field organisms. The contaminant concentration in the tissues of dredged-material test organisms is compared against the appropriate FDA action levels and against bioaccumulation data obtained from organisms that are exposed to reference-material sediment. If contaminant bioaccumulation in the dredged-material organisms is less than the FDA levels but greater than the levels in the reference-material organisms, organisms are collected from the vicinity of the disposal site and analyzed for the contaminants of concern. If the contaminant bioaccumulation of the dredged-material organisms is lower than the steady-state body burden of the field-collected organisms, the LPC for bioaccumulation is met. If field-collected organisms have contaminant levels lower than those of the dredged-material organisms, case-specific criteria are developed to make a final LPC compliance determination for bioaccumulation.

13.2.1.2.1 Type of Sampling Required

Section 8.0 of the 1991 Green Book, Collection and Preservation of Samples, provides general information on sampling plans and sample handling, preservation, and storage.

To adequately and efficiently conduct a dredged-material evaluation, a comprehensive sampling plan should be in place before sampling begins. Sufficient amounts of sediment and water should be collected to conduct the necessary evaluations. Careful consideration of maximum allowable and recommended holding times for sediments, as well as the exigencies of resampling, should be given careful consideration. Additionally, sample size should be small enough to be conveniently handled and transported, but large enough to meet the requirements for all planned analyses. The overall confidence of the final LPC determination is based on the following three factors.

- Collecting representative samples;
- Using appropriate sampling techniques; and
- Protecting or preserving the samples until they are tested.

Table 13-2 shows the general sampling requirements to conduct dredged-material testing. Actual sampling requirements are project-specific and are determined during the development of the project plan, based on the guidance that is provided in the 1991 Green Book and in local agreements/manuals.

13.2.1.2.2 Methods

As described in Section 13.2.1.2.1 above, only existing information is evaluated in Tier I. This requires the careful compilation and analysis of such information. If the information cannot show that the proposed dredged material meets one of the exclusionary criteria, or if the information is insufficient to reach an LPC determination, physical, chemical, and biological information on the dredged material and the ODMDS must be collected in Tiers II and/or III.

Proper sample collection, handling, and preservation are critical to the accurate evaluation of Tier II and III test results. Sampling methods are usually developed by individual testing laboratories and documented in standard operating procedure (SOP) documents. Consistent use of SOPs in the field and laboratory ensure that sampling and analytical errors are minimized.

Methods necessary to conduct toxicity and bioaccumulation evaluations may include the following:

- Sieving;
- Combustion;
- Gravimetry;

Table 13-2. Sample-Collection Requirements

Tests	Water Samples			Sediment Samples		
	Disposal Site	Dredging Site	Control[a]	Dredging Site	Reference Site	Control[a]
Tier II Water Column						
Screen	□			□		
Elutriate	□	□		□		
Tier II Benthic				□	□	
Tier III Water Column	□[b]	□	□	□		
Tier III Benthic				□	□	□
Tier IV Water Column	□	□	□	□		
Tier IV Benthic				□	□	□

[a]May or may not have to be field-collected.
[b]Dilution water; disposal-site water, artificial water, or clean seawater.

- Gas chromatography (GC);
- Electron-capture detection (ECD);
- Mass spectrometry (MS);
- Graphite furnace atomic absorption spectroscopy (GFAAS);
- Atomic absorption spectroscopy (AAS);
- Inductively coupled plasma (ICP) technique;
- 96-h elutriate toxicity bioassays;
- 10-day whole-sediment toxicity bioassays;
- 10-day whole-sediment bioaccumulation tests (for trace-metals analysis only); and
- 28-day whole-sediment bioaccumulation tests.

Project-specific methods necessary to conduct Tier IV water-column and benthic evaluations may include laboratory and/or field evaluations of long-term toxicity or bioaccumulation effects of the dredged material, such as the following:

- Population-survival assessments;
- Community-change assessments; and
- Reproduction assessments.

13.2.1.2.3 Types of Data Required

As discussed in Sections 13.2.1.2.1-13.2.1.2.4 above, data required to conduct the LPC evaluations may include the following:

- Physical sediment data;
- Organic- and inorganic-chemistry sediment data;
- Organic- and inorganic-chemistry sediment-elutriate data;
- Physical-oceanography data;
- Bioassay data;
- Bioaccumulation data; and
- Field species data.

13.2.1.2.4 Necessary Hardware and Skills

The hardware and skills necessary to conduct 1991 Green Book evaluations are relatively specialized. Many federal, state, and contract laboratories have capabilities to conduct most or all of the necessary evaluations. However, to conserve time and resources, field sampling, laboratory work, data management, and analysis of the results are often conducted by separate organizations according to aptitude, cost, and scheduling parameters.

The general categories of capabilities necessary to reach a Tier III dredged-material LPC compliance determination are the following:

- Regulation and literature research;
- Field sampling at the dredging site, disposal site, and reference site;
- Physical analysis of sediment samples;
- Trace-metal (chemical) analysis of water and sediment samples;
- Organic-compound (chemical) analysis of water and sediment samples;
- Numerical modeling for initial-mixing analysis;
- Toxicity bioassay testing of elutriate samples;
- Toxicity bioassay testing of whole-sediment samples;
- Bioaccumulation testing;
- Chemical analysis of tissue samples;
- Statistical analysis of test results;
- Quality-assurance implementation (throughout evaluation); and
- Compliance determination.

13.2.1.3 Documentation

Throughout the 1991 Green Book, references are provided for the recommended sampling and testing methods, data analyses, QA procedures, and additional testing guidance. For convenience to manual users, a copy of the U.S. Ocean Dumping Regulations (40 CFR Parts 220-228) is included in the 1991 Green Book as Appendix A.

Information on documentation and recordkeeping is interspersed throughout the testing guidance. Records ensure that all aspects of the field and laboratory work are documented so that the resulting data may be properly interpreted. Dredged-material test data may be rejected if their history cannot be confidently traced.

13.2.2 Applicability of Method to Human Health, Marine Life, or Wildlife Protection

The effects-based guidance provided in the 1991 Green Book is directly applicable to the protection of human health, marine life, and wildlife because it is based on determining LPC compliance. If the testing shows that either the LPC for the water-column or benthic environment will be exceeded, ocean disposal for the proposed dredged material is not supported. In 40 CFR 227.27(a), the LPC is defined as the concentration of the liquid phase of the dredged material that will not exceed either the established WQC or 1 percent of the acutely toxic concentration following the initial-mixing phase (initial mixing is defined in 40 CFR 229.29). In 40 CFR 227.27(b), the LPCs for the suspended particulate and solid phases are defined as those concentrations "*. . . that will not cause unreasonable acute or chronic toxicity or other sublethal adverse effects based on bioassay results using appropriately sensitive marine organisms . . . or will not cause accumulation of toxic materials in the human food chain.*"

The tiered-testing procedure in the manual establishes a conservative, yet workable, decision-making process for environmentally protective dredged-material management. Dredged material that poses no risk of adverse impact is readily supported for ocean disposal early in the procedure (i.e., Tier I or II). Dredged material that has unknown impact potential is evaluated to the level required to make a definitive LPC compliance determination. Only dredged material that is shown to meet both the water-column and benthic LPC through state-of-the-art analytical techniques is supported for open-ocean disposal.

13.2.3 Ability of the Testing to Generate Numerical Criteria for Specific Chemicals

The physical, chemical, and biological data generated by the Tier II, III, and IV tests can be used to field-validate SQC that are presently under development. The state-of-the-art sampling and analytical techniques contained in the 1991 Green Book guidance will provide for increases in method reproducibility, confidence of the test data, and utility to SQC research and development projects.

13.3 USEFULNESS

13.3.1 Environmental Applicability

The guidance in the 1991 Green Book is suitable for dredged material regulated under MPRSA because it is based on biological-effects testing, which takes into account synergistic, antagonistic, and additive effects of all contaminants in the material. This approach includes both water-column and benthic impact, and assesses both toxicity and bioaccumulation. Adaptations of the guidance are also being applied to nearshore and Great Lakes dredge disposal projects, and the tiered testing framework may serve as a model for sediment assessments under other regulatory and nonregulatory programs.

13.3.1.1 Suitability for Different Sediment Types

Except for extremely coarse- or angular-grain sediments, the tiered-testing approach is suitable for all sediment types. The test organisms recommended in the manual are suitable for most medium- and fine-grain dredged material. If the dredged material being tested is composed of very coarse sediments, or the dredged material has other physical properties that are potentially incompatible with recommended test species, alternative organisms may be used if they meet 40 CFR 227.27(c) and are ecologically relevant to the disposal site. Alternative test organisms may also be necessary to avoid grain-shape insensitivities when using sediment-ingesting organisms. Noncontaminant-related mortality has been linked on at least one occasion to internal organism damage that was caused by highly angular sediment of moderate grain size (Oakland Harbor sediment; Word *et al.*, 1990). Sample handling and chemical extraction of very coarse-sediment dredged material can also cause analytical problems.

In general, few analytical problems are caused by sediment type. Grain-size problems occur rarely because (1) most large-grain-size sediment contains few contaminants and meets the LPC in either Tier I or II, and (2) the tiered-testing procedure is relatively flexible and allows for alternative evaluation methods.

13.3.1.2 Suitability for Different Chemicals or Classes of Chemical Contaminants

Since the guidance in the 1991 Green Book uses effects-based tests, it does not rely on the explicit identification of contaminants for decision-making. However, the guidance is suitable for detecting and quantifying a wide range of organic and inorganic chemicals. In Tier I of the testing procedure, target analytes are determined for the proposed dredged material. If contamination is suspected, but specific contaminants cannot be isolated in the Tier I evaluation, the manual recommends that the dredged material be scanned for a broad spectrum of contaminants. A list of 131 potential target analytes is provided in Table 9-1 of the 1991 Green Book, Priority Pollutant and 301(l) Pesticides Listed According to Structural Compound Class.

Extensive guidance for laboratory analysis of organic and inorganic compounds is provided in Section 9 of the manual, Physical Analysis of Sediment and Chemical Analysis of Sediment, Water, and Tissue Samples. Target analytes for the water and tissue analyses are the same as those for whole-sediment analyses. Guidance is also provided in Section 9 of the manual for minimizing salt interferences with the chemical analyses.

13.3.1.3 Suitability for Predicting Effects on Different Organisms

All four tiers of the tiered-testing procedure consider effects on marine organisms that are representative of organisms that are indigenous to ODMDSs and have known impact tolerances. In Tier I, information on the proposed dredged material's effect on laboratory and indigenous species is analyzed. In Tier II, the theoretical bioaccumulation potential (TBP) for nonpolar inorganic contaminants in the dredged material is calculated and compared against that of the reference site. In Tier III, water-column toxicity, benthic toxicity, and benthic bioaccumulation are determined for ecologically relevant laboratory organisms. In Tier IV, case-specific bioassays and bioaccumulation studies are conducted on laboratory and/or field organisms.

13.3.1.4 Suitability for In-Place Pollutant Control

The 1991 Green Book was developed to determine water-column and benthic LPC compliance for proposed dredged material, not for in-place management of contaminated sediments. However, the physical, chemical, and biological tests that are recommended in the tiered-testing procedure are readily adaptable to nondredging management of sediments.

The sediment data that are generated with the guidance in the manual must be of sufficiently high quality to develop LPC determinations for the dredged material. If these data show that the dredged material does not meet the LPC for ocean disposal, the same data are readily adaptable to other sediment-management uses, including in-place pollutant management.

13.3.1.5 Suitability for Source Control

The purpose of the detailed sampling and testing guidance in the 1991 Green Book is to fully characterize the dredged material that is proposed for ocean disposal. Although it is not the intended purpose, this characterization may be useful for controlling sources of contaminants that are entering the sediments.

If portions of a proposed project exceed the LPC, it benefits the applicant to isolate the compliant and noncompliant areas to economize management of the dredged material. For example, material that meets the LPC might be disposed of at an ODMDS and material that does not meet the LPC might disposed of upland. During the process of site characterization, contaminant gradients and source locations might be identified (such as occurred in New Bedford Harbor, Massachusetts) and remedial or enforcement actions can be directed as appropriate.

13.3.1.6 Suitability for Disposal Applications

As discussed in Section 13.1 above, the guidance in the 1991 Green Book is used to conduct LPC evaluations, which are in turn used to support ocean-disposal management decisions. The manual is not intended to provide guidance on other disposal options available to dredged-material managers. Some ocean and nonocean disposal options may require additional or alternative analyses of the dredged material to reach decision points. Numerous other guidance manuals on dredged-material management are available from USEPA and the USACE.

13.3.2 General Advantages and Limitations

13.3.2.1 Ease of Use

As discussed in Section 13.2.1 above, the tiered-testing procedure is relatively flexible. The dredged-material evaluator can enter and exit the testing procedures at any tier. However, to begin the evaluation in Tier II, III, or IV, the data must satisfy the requirements of the earlier tier(s). The overall ease of use of the testing procedure depends on the evaluator's familiarity with the following:

- Federal regulations pertinent to dredged-material testing and disposal;
- Sources of existing dredged-material (sediment-quality) information;
- Sampling design;
- Numerical modeling;
- Physical, chemical, and biological testing;
- Statistical analysis; and
- Quality assurance.

13.3.2.2 Relative Cost

Tiers I, II, III, and IV are ordered by increasing complexity and cost. Tier I is relatively inexpensive and consists solely of assembly and analysis of existing information. Tier IV can be very expensive, consisting of case-specific toxicity and bioaccumulation analysis, including extensive field and laboratory studies. However, significant time and resources can be saved if the earlier tiers are completed to the maximum extent possible before proceeding to the later tier(s). For example, an in-depth analysis of "grey literature" (university reports, etc.) might show the possible existence of "hot spots" within a project. The sampling plan could then be designed to appropriately sample these areas of concern during a single sampling event, thereby saving the time and expense required to conduct additional sampling at a later time. Similarly, money and time will be saved if LPC compliance for nonpolar organic contaminants can be shown in the Tier II TBP calculation rather than in the Tier III laboratory testing and analysis.

As all dredging projects contain case-specific components, it is difficult to estimate the overall cost of a typical dredged-material analysis. USEPA and the USACE predict that the updated methods in the manual would not cause a significant increase in evaluation expenses and actually might lead to lower testing costs because LPC determinations might be achieved earlier in the testing process, thereby making full-scale bioassay and bioaccumulation laboratory tests unnecessary. Also, as the recommended analytical methods become refined, market pressures will force costs lower.

13.3.2.3 Tendency to Be Conservative

As discussed in Section 13.2.2 above, the tiered-testing procedure is very protective of human health and the marine environment. It is a sequential and comprehensive analysis of the proposed dredged material's biological effects, as shown by previous studies, modeling, and laboratory testing. However, the tiered-testing procedure is an "expert system"; that is, the product of the procedure (LPC compliance determination) is only as good as the information that is integrated into it.

To reach a defensible and ecologically sound LPC evaluation, high-quality information is required. There is risk of an inaccurate compliance determination if incomplete or inaccurate information is used, or if good information is misapplied. The regulations and numerous references in the manual should be consulted, and well-trained and experienced evaluators should be involved throughout the decision-making process.

13.3.2.4 Level of Acceptance

The 1991 Green Book is the official USEPA/USACE guidance manual for determining the suitability of dredged material for ocean disposal. During the development of the updated manual, comments from USEPA and USACE Headquarters, USEPA Regions, USACE Districts, other federal agencies, port authorities, special-interest groups, and the general public were solicited, received, and addressed as appropriate. In 1990, USEPA and the USACE conducted a public meeting on the document[1] and held six regional training sessions[2] on the updated methods. The final manual is the product of extensive USEPA/USACE dredged material program experience, current state-of-the-art testing methods, and review by a wide array of individuals and agencies.

13.3.2.5 Ability to Be Implemented by Laboratories with Typical Equipment and Handling Facilities

Many evaluations recommended in the 1991 Green Book, particularly for organic and chemical analysis, require standard laboratory equipment and handling facilities. However, some laboratories have difficulty attaining accurate and precise test results for low contaminant concentrations. Agency and contract laboratories that presently do not have the capabilities to conduct precise analyses will have to make significant investments in equipment, personnel, and training. It is expected that contract laboratories will choose to specialize in only a few methods to be efficient and competitive in the dredged-material testing market. Quality assurance (QA) program development, although not equipment-intensive, is also a necessary and significant investment for testing laboratories. QA programs are necessary to ensure that sample and data integrity are of sufficient quality and defensible.

13.3.2.6 Level of Effort Required to Generate Results

The overall level of effort necessary to conduct dredged-material analysis is comparable to that required by the preceding guidance (1977 Green Book). The level of effort is relatively low in Tier I and relatively high in Tiers III and IV.

13.3.2.7 Degree to Which Results Lend Themselves to Interpretation

The analysis of raw data that are generated during the tiered-testing procedure is relatively complex, especially for bioassay and bioaccumulation test data. Interpretation of results is specifically described and decision points and values are clearly defined in the 1991 Green Book. Section 13 of the manual, Statistical Methods, presents guidance for handling the following:

- Unequal numbers of experimental animals assigned to each treatment container or loss of animals during the experiment;
- Unequal numbers of replications of the treatments (i.e., containers or aquaria);

[1]Washington, DC.

[2]Narragansett, RI; Gulf Breeze, FL; Vicksburg, MS; Newport, OR; San Francisco, CA; and Washington, DC.

- Measurements scheduled for selected time intervals but actually performed at other times;
- Different conditions of salinity, pH, dissolved oxygen, temperature, etc., among exposure chambers; and
- Differences in placement conditions of the testing containers or in the animals assigned to different treatments.

USEPA and the USACE are presently developing software and additional guidance to facilitate data interpretation for dredged-material evaluations.

13.3.2.8 Degree of Environmental Applicability

The USEPA/USACE (1991) effects-based approach used to evaluate marine sediments has wide environmental and regulatory applicability. The approach uses test organisms that

- Are sensitive to impact;
- Are reasonable representatives of indigenous ODMDS species;
- Fulfill the species categories required by 40 CFR 227.27(c,d);
- Have extensive test databases; and
- Are hardy enough to withstand laboratory procedures.

Alternative test species that meet the guidance in the 1991 Green Book may be used to avoid testing problems such as grain-size tolerance and seasonal availability. Complete elucidation and quantification of all chemical components in the sediment are useful, but not required, for regulatory decision-making. The overall approach is environmentally conservative and relatively economical.

One feature of the 1991 Green Book guidance posing environmental limitations is the numerical modeling that is used in Tier I and II water-column evaluations. The ADDAMS models are not suitable for calculating water-column impacts at disposal sites that are extremely shallow (i.e., where the discharge period from the disposal vessel is longer than the descent time to the bottom). Additionally, there is some uncertainty about the applicability of the models for extremely deep (>200 m) ODMDSs.

13.3.2.9 Degree of Accuracy and Precision

The 1991 Green Book guidance strongly emphasizes the importance of a comprehensive QA program to achieve sufficient data quality during the tiered evaluation process. QA issues are addressed in subsections throughout the data-generation sections of the manual, and Section 13, Quality-Assurance Consideration, gives guidance on the structure and components of QA programs and data-quality assessment.

The general guidance for QA program development includes information on field and laboratory sample handling, personnel training, and documentation. For chemical analyses, the guidance recommends appropriate use of method blanks, procedural blanks, matrix spike/matrix-spike duplicates (MSSD), and standard reference materials (SRM) to determine accuracy and precision of the data. For biological testing, the importance of control-sediment tests, reference-site tests, and reference-toxicant testing is discussed.

13.4 STATUS

13.4.1 Extent of Use

The 1991 Green Book guidance will be applied to all evaluations for dredged material that is proposed for disposal outside the baseline of the territorial sea (non-state waters). Until completion of ongoing work on a national testing manual for disposal shoreward of the baseline of the territorial sea (Clean Water Act section 404 waters), portions of the Green Book guidance are also expected to be applied to nearshore and internal-water dredged-material disposal projects in the United States.

13.4.2 Extent to Which the Approach Has Been Field-Validated

Large portions of the tiered-testing procedure for dredged material have been field-validated since the publication of the original guidance in 1977 by ongoing state and federal dredging programs. Several large-scale, long-term USEPA/USACE projects in the New England and West Coast regions have applied and improved on the methods in the 1977 manual. The guidance in the 1991 Green Book contains methods proven for marine sediment analyses, developed for national testing consistency, and organized into tiers for efficient compliance determination. The tiered approach for environmental monitoring of aquatic ecosystems is strongly recommended by the National Research Council (NRC, 1990).

13.4.3 Reasons for Limited Use

Only extreme time and resource constraints (national emergencies, etc.) would limit the use of the guidance in the manual. Most of the recommended procedures are already widely applied.

13.4.4 Outlook for Future Use and Development

USEPA and the USACE will continue to support and apply the guidance in the manual both nationally and regionally. Ongoing public and private research and development of evaluation methods will continue to expand federal and state dredging-program experience.

The manual will be revised at a future date based on (1) the findings of an EPA SAB review (SAB, 1992); (2) technical advances in assessing sediment contamination and marine environmental impact; and (3) changes to the Ocean Dumping Regulations.

13.5 REFERENCES

NRC. 1990. Managing troubled waters: The role of marine environmental monitoring. National Research Council. National Academy Press, Washington, DC. 125 pp.

SAB. 1992. An SAB report: Review of a testing manual for evaluation of dredged material proposed for ocean disposal. Prepared by the Sediment Criteria Subcommittee of the Ecological Processes and Effects Committee, USEPA Science Advisory Board, Washington, DC. EPA-SAB-EPEC-92-014.

USEPA/USACE. 1977. Environmental Protection Agency/United States Army Corps of Engineers Technical Committee on Criteria for Dredged and Filled Material. Ecological evaluation of proposed discharge of dredged material into ocean waters; Implementation manual for section 103 of Public Law 92-532 (Marine Protection, Research, and Sanctuaries Act of 1972). July 1977 (second printing April 1978). Environmental Effects Laboratory, United States Army Engineer Waterways Experiment Station, Vicksburg, MS. 24 pp + appendices.

USEPA/USACE. 1991. Environmental Protection Agency/United States Army Corps of Engineers. Ecological evaluation of proposed discharge of dredged material into ocean waters. January 1990. United States Environmental Protection Agency, Office of Marine and Estuarine Protection, Washington, DC 20460. USEPA-503-8-90/002. 219 pp + appendices.

Word, J.Q., J.A. Ward, J.A. Strand, N.P. Kohn, and A.L. Squires. 1990. Ecological evaluation of proposed discharge of dredged material from Oakland Harbor into ocean waters (Phase II of 42-Foot Project). Prepared for United States Army Corps of Engineers. U.S. Department of Energy Contract No. DE-AC06-76RLO 1830. September 1990.

National Status and Trends Program Approach

Edward R. Long
Coastal Monitoring and Bioeffects Assessment Division
National Oceanic and Atmospheric Administration
7600 Sand Pt. Way, NE, Seattle, WA 98115
(206) 526-6338

Donald D. MacDonald
MacDonald Environmental Sciences, Ltd.
2376 Yellow Point Road, R.R. #3, Ladysmith, BC, Canada V0R 2E0

Sediment quality criteria based on multiple methods have been recommended for broad applications in the United States (USEPA/SAB, 1989; Adams *et al.*, in press). The approach used by the National Status and Trends Program (NSTP) of the National Oceanic and Atmospheric Administration (NOAA) to develop informal, effects-based guidelines involves the identification of the ranges in chemical concentrations associated with biological effects based on a weight of evidence from many studies. In this approach, the data for many chemicals are assembled from modeling, laboratory, and field studies to determine the ranges in chemical concentrations that are rarely, sometimes, and usually associated with toxicity. The data from many of the studies of the individual approaches described elsewhere in this document are compiled and examined to develop no-effects, possible-effects, and probable-effects ranges (Figure 14-1).

14.1 SPECIFIC APPLICATIONS

14.1.1 Current Use

The NSTP Approach was used initially to develop informal guidelines for use by the National Status and Trends (NS&T) Program (Long and Morgan, 1990; Long, 1992). NOAA analyzes sediments from numerous locations nationwide as a part of its monitoring program. The guidelines were developed as tools for identifying locations in which there is a potential for toxicity to living resources for which NOAA is the federal steward. Areas in which chemical concentrations often exceeded the guidelines were identified as high priorities for investigations of toxicity with biological tests.

Environment Canada evaluated many candidate approaches to the development of sediment quality guidelines and elected to develop its national guidelines using the NSTP Approach (MacDonald and Smith, 1991; MacDonald *et al.*, 1991). The Florida Department of Environmental Regulation elected to use the NSTP Approach to develop state sediment quality guidelines as a part of its sediment management strategy (MacDonald, 1992). The California Water Resources Control Board will use the NOAA guidelines in its initial evaluations of ambient chemical data. Following that step, data from field studies, laboratory bioassays, and equilibrium partitioning models will be used to develop sediment quality objectives (Lorenzato *et al.*, 1991). Finally, the International Council for Exploration of the Sea Study Group on the Biological Significance of Contaminants in Marine Sediments has elected to adopt the NSTP Approach in the development of guidelines for participating nations (Dr. Herb Windom, Working Group on Marine Sediments, ICES, personal communication).

Guidelines developed with the NSTP Approach were used by NOAA to identify chemicals that occurred in concentrations that were sufficiently high to warrant concern and to identify sampling sites and areas in which there was a potential for toxicity (Long and Morgan, 1990;

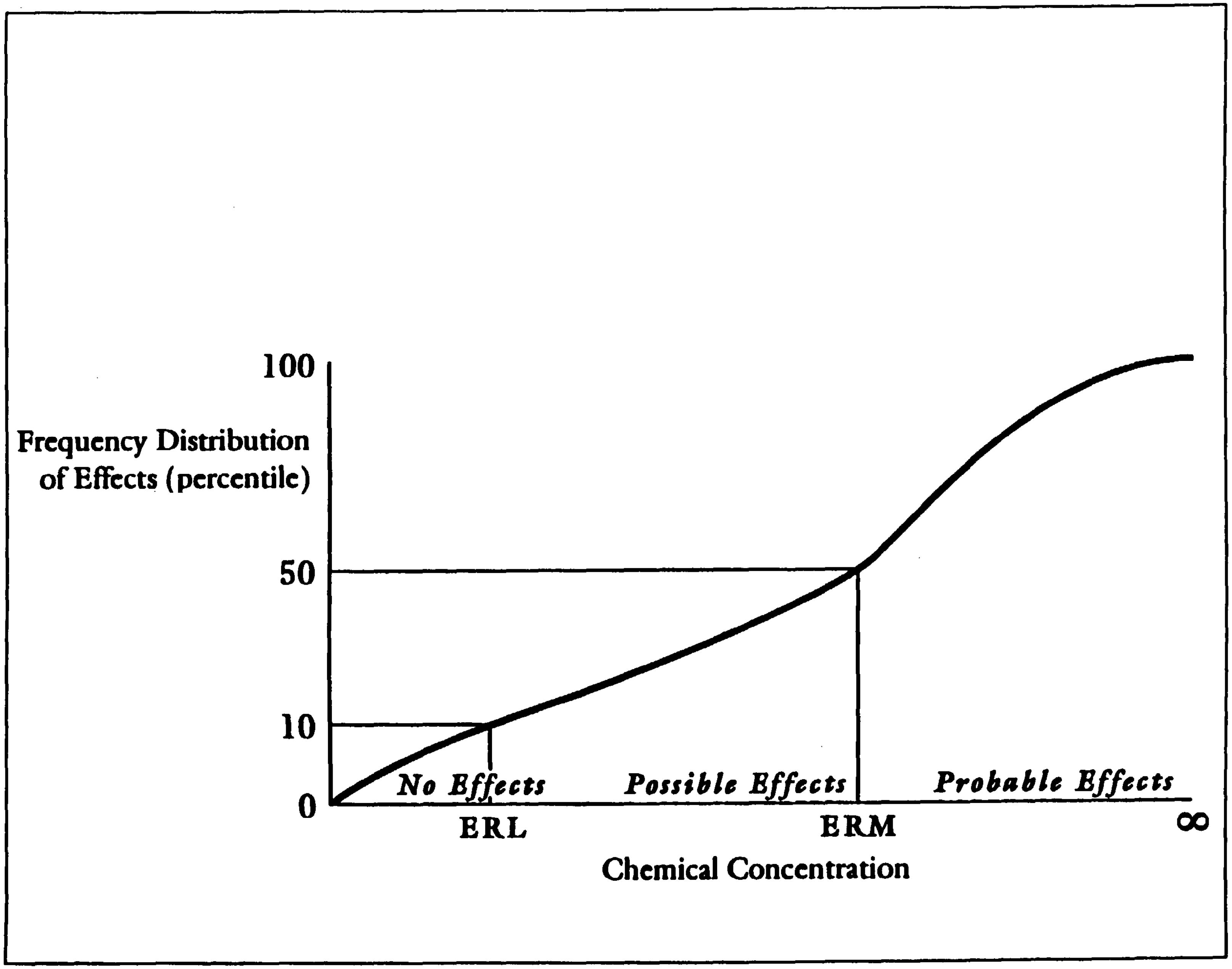


Figure 14-1. Conceptual outline of the relationship between the NSTP Approach guidelines and the no-effects, possible effects, and probable-effects ranges in chemical concentrations.

Long *et al.*, 1991; Long and Markel, 1992). It was presumed that the potential for toxicity was relatively high in areas where numerous chemicals exceeded the upper bounds of the guidelines. Likewise, it was assumedthat the potential for toxicity was relatively low in areas where none of the chemical concentrations exceeded the lower bounds of the guidelines. In those regions with the highest potential for toxicity, NOAA has implemented regional surveys of toxicity, using a battery of biological analyses and tests.

Also, NOAA has used the guidelines in assessments and prioritization of hazardous waste sites (Dr. Alyce Fritz, NOAA Hazardous Materials Response and Assessment Division, personal communication). Other agencies and consultants have used the guidelines as a means of placing ambient chemical data into perspective with respect to the potential for toxicity (for example, Squibb *et al.*, 1991 for New York/New Jersey Harbor; Mannheim and Hathaway, 1991 for Boston Harbor; Soule *et al.*, 1991 for Marina Del Rey). The Florida Department of Environmental Regulation has used the guidelines as informal tools for interpreting ambient chemical data and for identifying regional priorities for sediment quality management (MacDonald, 1992).

14.1.2 Potential Use

Potential uses of the guidelines are as follows:

- Identification of potentially toxic chemicals in ambient sediments;
- Ranking and prioritization of areas and sampling sites for further investigation;
- Assessment of potential ecological hazards of contaminated sediments;
- Design of spiked-sediment bioassay experiments;
- Description of the kinds of toxic effects previously associated with specific concentrations of chemicals;
- Quantification of the relative likelihood of toxicity over specific ranges in chemical concentrations; and
- Identification of the need for sediment management initiatives.

14.2 DESCRIPTION

14.2.1 Description of Method

The NSTP Approach involves a simple evaluation of available data to identify three ranges in concentrations for each chemical:

- No-Effects Range: The range in concentrations over which toxic effects are rarely or never observed;
- Possible-Effects Range: The range in concentrations over which toxic effects are occasionally observed; and
- Probable-Effects Range: The range in concentrations over which toxic effects are frequently or always observed.

These ranges are identified by evaluating information from numerous studies in which matching biological and chemical data were developed. The specific steps in the method are:

(1) Compile matching chemical and biological data from laboratory spiked-sediment bioassays, equilibrium-partitioning models, and field studies and determine the chemical concentrations associated with no observed effects and those associated with adverse effects.

(2) Enter the data into a database, including the type of biological test performed, the adverse effect(s) measured, the chemical concentrations associated with observations of either effects or no effects, the type of study method and approach, and the degree of concordance between the

measure of effects and the concentration of the chemical.

(3) For those analytes for which sufficient data exist, prepare data tables sorted according to ascending chemical concentrations.

(4) Arithmetically determine the no-effects range, possible-effects range, and probable-effects range for each chemical.

The steps taken to select and screen candidate data sets are described in Section 14.2.1.2.3. The approach is intended to encourage periodic updates as new data become available.

Two slightly different methods have been used to determine the three chemical ranges. First, two percentiles in the chemical concentrations associated with toxicity were derived by Long and Morgan (1990): the lower 10th percentile and the 50th percentile (median). The lower 10th percentile was identified as the Effects Range-Low (ERL), and the median was identified as the Effects Range-Median (ERM). In their evaluation of the ascending data tables, Long and Morgan (1990) used only the chemical concentrations that had been associated with toxicity (i.e., the "effects" data). The conceptual basis for this approach and the three ranges are illustrated in Figure 14-2.

Later, MacDonald (1992) identified the three ranges with a method that used both the concentrations associated with biological effects (the "effects" data) and those associated with no observed effects (the "no-effects" data). In this method, a threshold effects level (TEL) was calculated first as the square root of the product of the lower 15th-percentile concentration associated with observations of biological effects (the ERL) and the 50th-percentile concentration of the no-observed-effects data (the NER-M). A safety factor of 0.5 was applied to the TEL to define a No-Observable-Effects Level (NOEL). Next, a Probable-Effects Level (PEL) was calculated as the square root of the product of the 50th-percentile concentration of the effects data (the ERM) and the 85th-percentile concentration of the no effects data (the NER-M).

Neither of these methods is preferred or advocated over the other. The significant feature of this approach is the use of a weight of evidence developed in the ascending tables, not in the specific method of using the data tables. In addition to the two methods described here, many others could be applied to the ascending data tables to derive guidelines. The method used by MacDonald (1992) considered both the "effects" and "no-effects" data, whereas that of Long and Morgan (1990) used only the "effects" data. Different percentiles in the ascending data were used in the two methods. Despite these differences in the methods, the agreement between the NOELs and ERLs and between the PELs and the ERMs was very good, usually within a factor of 2.

In both documents, the lower of the two guidelines for each chemical was assumed to represent the concentration below which toxic effects rarely occurred. The range in concentrations between the two values was that in which effects occasionally occurred. Toxic effects usually or frequently occurred at concentrations above the upper guideline value.

As an example, Figure 14-2 compares the frequency distribution of toxic effects and no-effects data associated with concentrations of napththalene to the ERL and ERM concentrations for naphthalene. Long and Morgan (1990) reported the ERL as 340 ppb dry wt. and the ERM as 2100 ppb dry wt. for naphthalene, based on an ascending data table of 49 data points. These guidelines defined three ranges of chemical concentrations: the no-effects range (0-340 ppb); the possible-effects range (340-2100 ppb); and the probable-effects range (>2100 ppb). Only 10.5 percent of the chemical concentrations below the ERL were associated with toxic effects, suggesting that toxicity is unlikely below the ERL concentrations. In contrast, 81 percent of the chemical concentrations between the ERL and ERM values were associated with the toxic effects and 93 percent of the data points were associated with toxicity at concentrations above the ERM value.

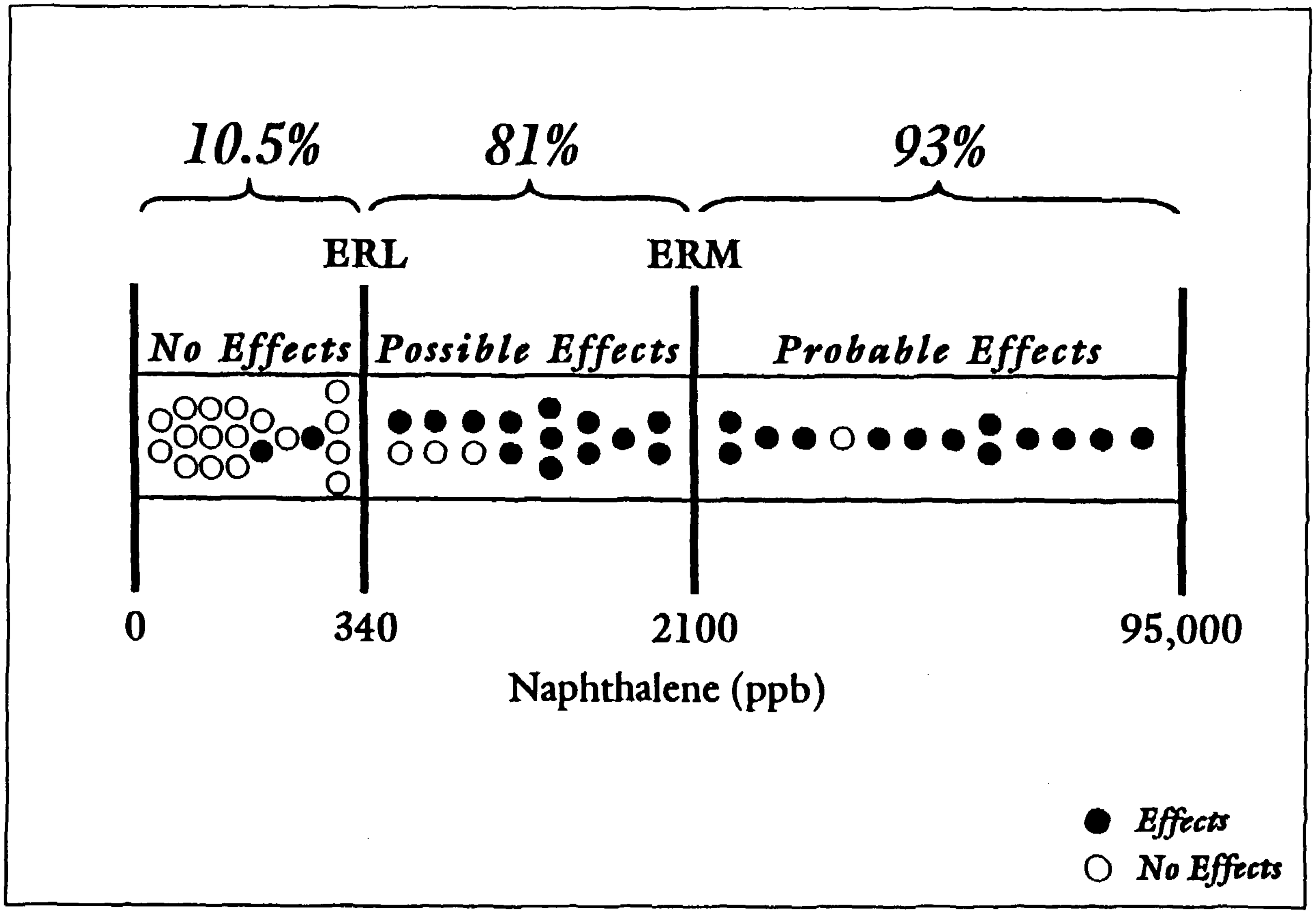


Figure 14-2. Frequency distribution of naphthalene concentration associated with toxic effects below the ERL value, between the ERL and ERM values, and above the ERM value (from Long and Morgan, 1990).

14.2.1.1 Objectives and Assumptions

The objective of the NSTP Approach is to provide informal, effects-based guidelines that are based on a weight of evidence and reported as ranges in concentrations. The guidelines are based on chemical concentrations associated with measures of biological effects, thereby providing toxicological and/or biological relevance to the guidelines. They are based on data from multiple studies and research methods, thus providing a weight of evidence. In recognition of the variability in the kinds of data that are available, they are presented as ranges, instead of absolute values, thereby providing a flexible interpretive tool with broad applicability. They are presented along with all of the supporting evidence in ascending tables, providing the user an interpretive framework for comparison with ambient data.

In this approach it is assumed that the data from all individual studies are equal in weight and credibility, although they may have involved very different methods and test endpoints. It is assumed that the methods used by the individual investigators were reasonably accurate. Most important, it is assumed that as the concentrations increase, the potential for toxicity also increases, thereby providing a conceptual basis for identifying the ranges in concentrations frequently associated with no toxic effects and those frequently associated with toxic effects. The guidelines can be formulated to account for site-specific factors that control bioavailability (see Section 14.3.1.1).

14.2.1.2 Level of Effort

14.2.1.2.1 Type of Sampling Required

The NSTP Approach relies on the use of a database compiled from a wide variety of sediment quality assessments. The database currently contains over 800 entries generated by the three major approaches to the establishment of effects-based guidelines: equilibrium-partitioning models; laboratory spiked-sediment bioassays; and various assessments of matching, field-collected, sediment chemistry, and biological effects data. The NSTP Approach was specifically designed to use existing data, therefore eliminating or minimizing the need for additional sampling. However, evaluation of the regional applicability of the guidelines could, in some cases, require further site-specific investigations, the magnitude of which could vary considerably.

14.2.1.2.2 Methods

The methods for deriving numerical sediment quality guidelines using the NSTP Approach are summarized in Section 14.2.1. Also, these methods are described by Long and Morgan (1990) and MacDonald (1992).

14.2.1.2.3 Types of Data Required

The NSTP Approach was intended to integrate a diverse assortment of information into a single database to support the derivation of numerical guidelines. Consequently, data from numerous modeling, laboratory, and field studies were collated into one database. Ideally, the database used to establish guidelines should include entries from all three of these types of approaches. Suitable data were available from a wide variety of sources. While collection and analysis of these data sets were labor-intensive, subsequent, incremental updates of the database should be relatively simple and inexpensive.

The data compiled from numerous studies were entered into the Biological Effects Database for Sediments (BEDS) by MacDonald (1992). All of the compiled data were fully evaluated prior to incorporation into the BEDS to ensure internal consistency in the database. The screening procedures used to support the development of the BEDS were designed to ensure that only relevant and high-quality data were used to derive the guidelines. No subjective biases were employed in screening the data; as many sources of data were included as possible. Candidate data from each study were evaluated to determine the acceptability of the experimental design, the test protocols, the analytical methods, and the statistical procedures that were used. Only data in which there were matched measures of sediment chemistry and biological effects were included. The database included only those data in which either statistically significant biological results were obtained or in which major differences in the biological results between samples were reported.

The BEDS currently includes over 800 data entries, mainly data from studies performed throughout North America. It was developed jointly by NOAA, Florida Department of Environmental Regulation, Environment Canada, and MacDonald Environmental Services Ltd.

In the evaluation of candidate data from field studies, only those data were used in which at least a 10-fold difference in the concentrations of at least one chemical among the samples was reported. Once this criterion was met, the data from many of the field studies were evaluated to determine the mean chemical concentrations in toxic samples (i.e., significantly different from controls) and those in nontoxic samples or in samples with relatively depauperate benthic communities (i.e., those with low abundance or species richness) versus those with more robust communities. Further, those mean concentrations in biologically affected samples that exceeded by twofold or more the mean concentrations in the background, reference, or nonaffected samples were assigned an asterisk in the ascending tables. The asterisks symbolized that a biological effect was noted and that there was a strong association between the chemical gradient and the biological gradient. Concentrations associated with nontoxic reference conditions were noted as "no effects." Those in which there was no concordance between the measures of effects and chemical concentrations were noted as "no gradient" or "no concordance." The concentrations derived in the modeling and spiked-sediment bioassays were always assigned asterisks. The concentrations with asterisks were used as "effects" data by both Long and Morgan (1990) and MacDonald (1992).

14.2.1.2.4 Necessary Hardware and Skills

The primary skills required to derive guidelines are associated with the development of the database. Expertise is required to evaluate the suitability of the biological and chemistry data, using the screening criteria. This process requires experience in the evaluation of sediment data and the methods that were used to develop the data.

The database has been developed on a personal computer and is readily transferable to other systems, but requires knowledge of the use of a computer. The database provides a means of storing and accessing all of the information that relates chemical concentrations to adverse biological effects. This information can be manipulated in this environment or exported into other formats.

14.2.1.3 Adequacy of Documentation

The NSTP Approach was documented by Long and Morgan (1990), in which the approach was peer-reviewed both within and outside NOAA. A second printing of the document was issued in 1992, following further review. A synopsis of the approach was described in a scientific journal (Long, 1992). The approach has been described orally in numerous technical and scientific forums. MacDonald and Smith (1991) and MacDonald *et al.* (1991) described the application of the approach in the development of guidelines for Canada. MacDonald (1992) described the use of the approach in a statewide sediment management strategy for Florida.

14.2.2 Applicability of Method to Human Health, Aquatic Life, or Wildlife Protection

The guidelines are intended to provide an estimate of the potential for adverse biological effects of sediment-associated contaminants on benthic organisms, based on a weight of evidence from analyses performed with multiple species and/or biological communities. They accommodate and rely on the data from tests of acute and chronic toxicity and from analyses of benthic community structure. The guidelines are based on data from many different areas and oceanographic regimes, thereby broadening their applicability. Currently, the data entered into the BEDS are from only marine and estuarine areas.

The guidelines provide a means of numerically estimating the percent frequency of biological effects over the three ranges of concentrations. The ascending tables accompanying the guidelines also provide a supplementary basis for interpreting new ambient chemical data. Also, these tables provide a visual and statistical means of estimating the relative degree of certainty in the guidelines.

The guidelines are not intended to be used for the protection of human life or wildlife. Rather, they are intended to be used in estimating the potential for adverse effects among benthic communities.

14.2.3 Ability of Method to Generate Numerical Criteria for Specific Chemicals

Long and Morgan (1990) reported numerical guidelines for 41 chemicals, including 12 trace metals, 18 polynuclear aromatic hydrocarbons (PAHs), and 11 synthetic organic compounds. MacDonald (1992) developed guidelines for 9 trace metals, total PCBs, 13 PAHs, 3 classes of PAHs, and 2 pesticides.

Conceptually, guidelines derived using this approach could be developed for any toxic chemical, provided sufficient data exist and provided the toxicity of the chemical is dose-responsive. Long and Morgan (1990) assigned a high degree of confidence to guidelines for chemicals for which data existed from many different approaches, different regions, and in which there was a good agreement in the data from different studies. MacDonald (1992) calculated guidelines only for those chemicals for which there was a minimum of 40 data points, after determining the minimum amount of data necessary to calculate reliable and consistent values. These minimum data requirements were established by iteratively calculating guidelines using data sets of increasing size (e.g., 4 to 60 data points) and determining when the estimate of the guidelines stabilized.

14.3 USEFULNESS

14.3.1 Environmental Applicability

14.3.1.1 Suitability for Different Sediment Types

The NSTP Approach can be applied equally to any sediment type that occurs in freshwater, estuarine, and marine environments. Since the database that supports the guidelines contains information from a wide variety of sediment types, the resultant guidelines are considered to be widely applicable. An increasing amount of information suggests that the bioavailability, and, therefore, toxicity, of many contaminants is controlled by such factors as TOC, AVS, and grain size. The BEDS currently accommodates the data for these variables, and, consequently, the guidelines could be normalized to the appropriate factors that control bioavailability. However, insufficient information currently exists to derive guidelines that are expressed in these terms. It is anticipated that future revisions of the guidelines will be expressed in these terms, thereby increasing their applicability.

Partly to increase the suitability of the guidelines to different sediment types, they are expressed as ranges in concentrations, not absolutes. These ranges provide a basis for evaluating chemical concentrations in the different types of sediments represented in the BEDS. In addition, the ascending data tables used to generate the guidelines can be examined to calculate frequency distributions of effects and no effects within each range of concentrations. These frequency distributions can be used as estimates of the probability of toxic effects.

14.3.1.2 Suitability for Different Chemicals or Classes of Chemicals

The approach can be applied to a wide variety of chemicals for which analytical methods are available. Thus far, numerical guidelines have been developed by Long and Morgan (1990) and by MacDonald (1992) for 43 and 28 chemicals or classes of chemicals, respectively. Data are included in the BEDS for over 200 chemicals or classes of chemicals. Guidelines could be developed for all of these substances when sufficient information becomes available.

14.3.1.3 Suitability for Predicting Effects on Different Organisms

Since the database compiled from many different studies is based on tests or analyses performed with many different species, the guidelines are widely applicable to benthic organisms.

In addition, the species studied in each investigation is(are) listed in the database; therefore, species-specific applicability can be evaluated by the users. Furthermore, the ERL values often are based on data from relatively sensitive species or life stages, and, therefore, can be used as guidelines suitable for the protection of sensitive species.

14.3.1.4 Suitability for In-Place Pollutant Control

Numerical sediment guidelines developed using the NSTP Approach can be used in a variety of ways as a tool in pollutant control. Specifically, these assessment tools respond to regulatory requirements by:

- Providing a basis for evaluating existing sediment chemistry data and ranking areas of concern and chemicals of concern in terms of their potential for causing toxicity and

- Identifying the need for further investigations, such as biological testing, to support regulatory decisions.

As is the case with all of the other approaches that rely on data collected in the field, the guidelines derived using the NSTP Approach integrate information obtained from studies of complex mixtures of contaminants and thereby consider their interactive effects. Consideration of the effects of contaminant mixtures is an advantage in the assessment of in-place pollutants in real-world conditions. However, this approach also relies on and gives equal weight to the data from equilibrium-partitioning models and laboratory spiked-sediment bioassays performed with single chemicals (see Section 14.2.1.1).

14.3.1.5 Suitability for Source Control

A reasonable amount of confidence in sediment quality guidelines is needed to justify using them in source control actions. Since the guidelines are developed with a weight of evidence compiled from many different studies, they provide a credible and defensible basis for evaluating contaminants in real-world conditions. The guidelines provide an efficient basis for identifying priority chemicals and priority areas that would benefit from source controls. In addition, the ascending tables provide a basis for estimating the probability of observing adverse effects at sites of interest, reducing the probability of effects through source controls, and evaluating the improvements in sediment quality following the implementation of source control measures.

14.3.1.6 Suitability for Dredged Material Disposal Applications

Neither the numerical guidelines nor the frameworks that have been developed for their application are intended to replace accepted testing protocols for dredged material disposal evaluations. Nonetheless, these guidelines can provide relevant tools for estimating the potential for adverse biological effects of contaminants associated with solid-phase sediments.

14.3.2 General Advantages and Disadvantages

14.3.2.1 Ease of Use

The approach has the advantage of relying on existing data. Therefore, guidelines can be developed relatively quickly and easily.

The original efforts by Long and Morgan (1990) and MacDonald (1992) to assemble the databases used to develop the guidelines were labor-intensive. Numerous reports and data sets were located, and a huge amount of data was entered into spreadsheets. However, these data now exist in a centralized, computerized database, the BEDS. Subsequent derivations of guidelines based on iterative expansions of the BEDS database should be relatively quick, easy, and inexpensive.

The guidelines are easily used and interpreted. Chemical data can be readily compared with the guidelines and with the ascending tables. The frequency of occurrence of toxicity over the no-effects, possible-effects, and probable-effects ranges can be calculated and compared with the chemical data.

Sediments in which numerous chemicals occur at concentrations that fall within the probable-effects ranges have a higher probability of being toxic than those in which most of the chemical concentrations are within the no-effects range. This type of simple interpretation makes the guidelines very easy to use.

14.3.2.2 Relative Cost

The original effort of Long and Morgan (1990) involved roughly one year of labor. The confirmation and expansion of the database by MacDonald (1992) involved more than another year of labor. The costs of subsequent iterations of the guidelines based on further expansions of the database would vary with the amount of data entered and the number of chemicals. The calculations of the guideline values themselves are very simple and quick. Also, the guidelines can be used very quickly and easily.

If the necessary data are not available for entry into a database, then the costs to generate them could be relatively high. If initiated *de novo*, modeling, bioassay, and field studies necessary to generate sufficient data could vary considerably in costs and time, depending on the amount of data needed.

14.2.3 Tendency to Be Conservative

The predictive capabilities of the guidelines have not been independently quantified. The protectiveness of the guidelines could be increased by considering data only from chronic sublethal endpoints or by applying a numerical safety factor, such as was applied in the Florida guidelines (MacDonald, 1992). Also, the guidelines would become more conservative if data were included only from areas in which toxicants were highly bioavailable.

14.3.2.4 Level of Acceptance

The NSTP Approach has been published by NOAA, following an in-house and outside peer review. It has been published in a peer-reviewed scientific journal. The approach has been used by Environment Canada and Florida Department of Environmental Regulation in the development of their respective guidelines. It has been adopted by a committee of the International Council for Exploration of the Sea for use by member nations. The State of California has adopted a similar approach to the development of sediment quality objectives (Lorenzato *et al.*, 1991).

The numerical guidelines developed by use of the approach have been used by NOAA to compare and rank the potential for toxicity at monitoring sites nationwide, within San Francisco Bay, and within Tampa Bay. Approximately 1500 copies of the report by Long and Morgan (1990) have been distributed. Users of the report have compared ambient concentrations with the guidelines in assessments of hazardous waste sites, analyses of prospective dredge material, evaluations of survey and monitoring data, and estimates of ecological risk (for example, Mannheim and Hathaway, 1991; Soule *et al.*, 1991; Squibb *et al.*, 1991). NOAA routinely uses the guidelines in its estimates of ecological risk at National Priority List hazardous waste sites. The guidelines have been used as a basis for interpretation of chemical data in court cases.

14.3.2.5 Ability to Be Implemented by Laboratories with Typical Equipment and Handling Facilities

The spreadsheets and database needed to generate the guidelines can be prepared with a personal computer and need not be very complicated. Entry of data into the database and the generation of the ascending tables are very simple. The calculations of the guidelines can be performed manually, on a desk-top calculator or a personal computer. The database can be supplemented with new data as they become available. Implementation of the approach can become more laborious and complicated if the necessary data must be generated *de novo*.

14.3.2.6 Level of Effort Required to Generate Results

As outlined in Section 14.3.2.2, the level of effort required in the development of the original set of guidelines was relatively high. Subsequent iterations of the guidelines for other purposes, other chemicals, or for the same chemicals following additions to the database would be relatively easy. Entry of new data points from spiked-sediment bioassays, equilibrium-partitioning models, or apparent effects thresholds into the database would require only a few minutes. Manipulation of raw matching data from biological and chemical analyses performed in a field study would require from a few hours to several days, depending on the size of the data set, followed by entry of the data points into the database.

14.3.2.7 Degree to Which Results Lend Themselves to Interpretation

The guidelines and the ascending data tables on which they are based can be used in a number of ways. First, the data from analyses of ambient samples can be compared visually with the two numerical guidelines to determine whether the ambient concentrations exceed either of the guidelines. Second, the ambient concentrations can be compared with the data in the ascending tables to determine the kinds of toxic effects that have been observed in previous studies at the concentrations of concern. Finally, the frequencies of toxicity in the no-effects, possible-effects, and probable-effects ranges can be used to predict the probability of toxicity associated with any contaminant concentration.

The guidelines developed thus far with this approach do not account for the effects of factors that control bioavailability of the toxicants. This is not a weakness of the approach; rather, it is a weakness of the available data. Nevertheless, this weakness may hinder interpretation of ambient data with the guidelines. The BEDS database includes a provision for entering data from analyses of acid volatile sulfides and total organic carbon (and other potential normalizers) and, therefore, would lend itself to recalculation of guidelines normalized to these factors once the necessary data become available.

An important strength of this approach is that it provides the user some flexibility in the use and interpretation of the guidelines. All of the data are provided in ascending order for the user to see and evaluate. The degree of certainty in the data can be assessed and judged by the user. Ranges in concentrations are provided, instead of rigid, single absolute values.

One of the most attractive features of this approach is the estimation of the probability of biological effects, based on the frequency distributions of effects for each chemical. For example, the data in the BEDS database indicate that only 5.8 percent of the chemical concentrations within the no-effects range for cadmium (0 to 1 mg/kg) determined by MacDonald (1992) were associated with adverse biological effects (Figure 14-3). These data suggest that there is a low probability of observing adverse effects within this range. Within the probable effects range for cadmium (>7.5 mg/kg), roughly 68 percent of the database entries were associated with adverse effects. These data suggest that there is a relatively high probability of observing adverse effects within this range. Positive concordance between frequency of effects and chemical concentrations should inspire confidence in the guideline values.

Evaluation of the guidelines for mercury reveals that a lower level of confidence should be placed on the guidelines for this element. The data in the BEDS database indicate that within the no-effects range (0 to 0.1 mg/kg), roughly 7 percent of the entries were associated with adverse effects (Figure 14-4). However, frequency distributions of effects are similar within the possible-effects range (0.1 to 1.4 mg/kg) and the probable-effects range (>1.4 mg/kg), namely 30.1 percent and 33.3 percent, respectively. Therefore, it is more difficult to adequately determine the unacceptable levels of mercury in sediments than with, say, cadmium.

14.3.2.8 Degree of Environmental Applicability

The guidelines are highly applicable to the interpretation of environmental data. They are generated with data from environmentally realistic field studies, as well as theoretical modeling studies and controlled laboratory experiments. They are generated with data from many different regions in which the mixtures and concentrations

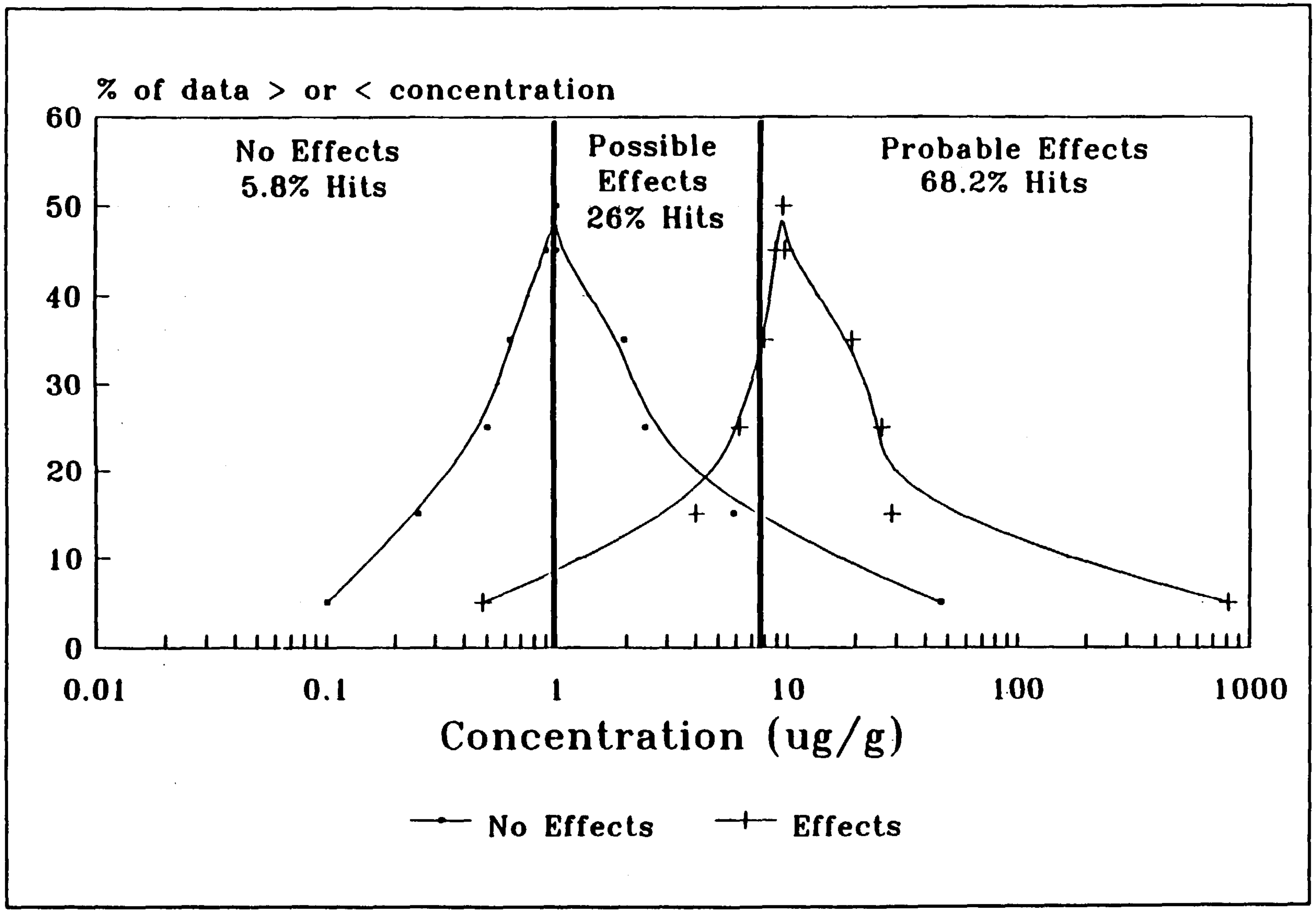


Figure 14-3. Summary of the available data in BEDS on the effect of cadmium (from MacDonald, 1992).
Percent frequency of effects data and percent frequency of no-effects data are plotted against cadmium concentrations.

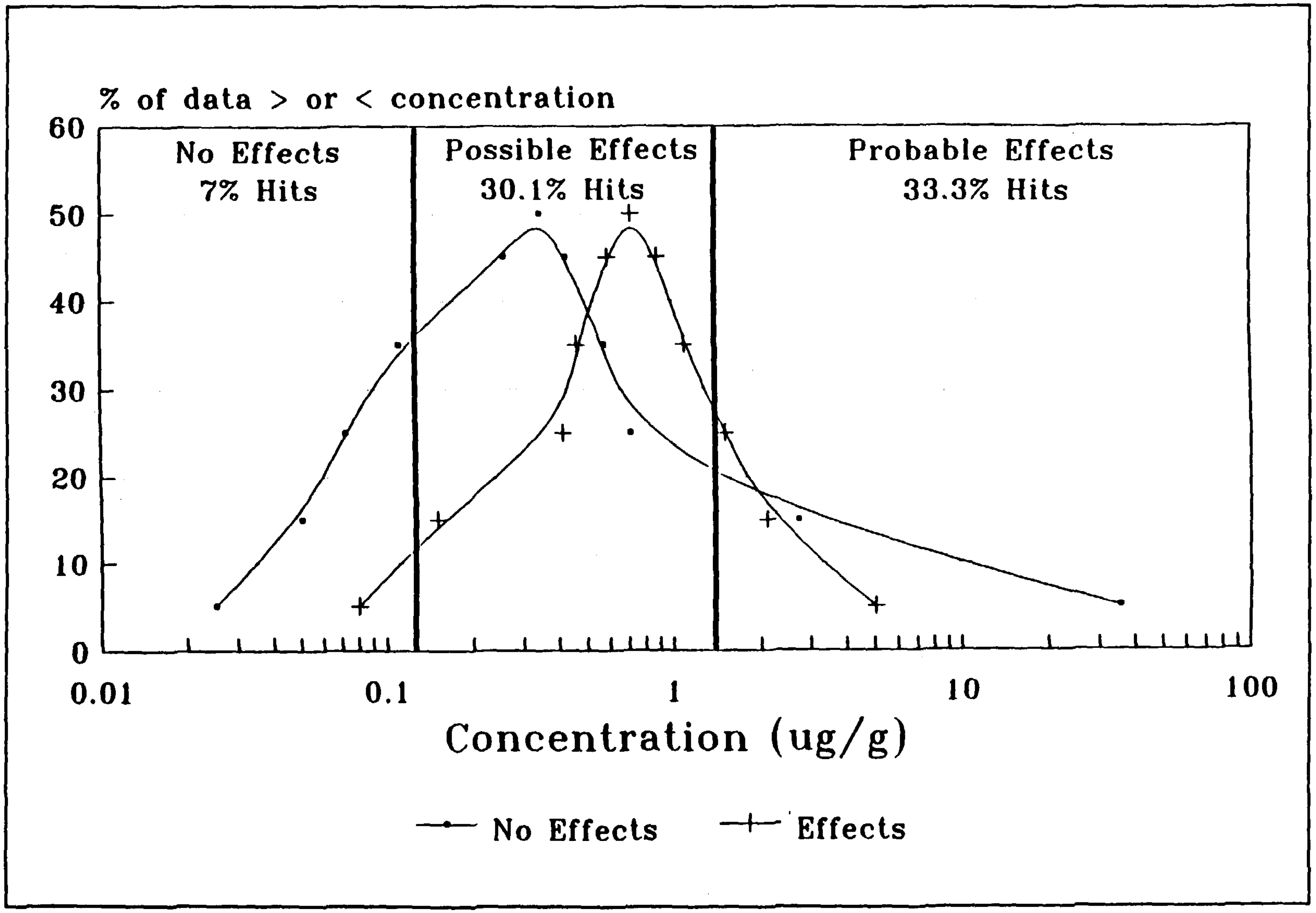


Figure 14-4. Summary of the available data in BEDS on the effects of mercury (from MacDonald, 1992).
Percent frequency of effects data and percent frequency of no-effects data are plotted against mercury concentrations.

of chemicals differ and in which sedimentological properties differ. They are generated with tests using different species with different sensitivities to toxicants. They are universally applicable in North America since they were generated with data from many regions in the United States and Canada. Confidence in the utility of the guidelines is inspired by the weight of evidence from these multiple studies.

14.3.2.9 Degree of Accuracy and Precision

By iteratively adding and removing different data sets from the ascending tables, MacDonald (1992) determined that a minimum of 40 data sets were needed to develop consistent and reliable guidelines. Clearly, some variability in the guidelines is to be expected as data are added or deleted, but, once the minimum amount of data is compiled, this variability appears to be minimal.

MacDonald (1992) generally doubled or tripled the amount of data in the ascending tables compiled by Long and Morgan (1990) mainly with new data from field studies and laboratory spiked-sediment bioassays. Also, MacDonald (1992) considered only estuarine and marine data, thereby deleting the freshwater data included in Long and Morgan (1990). The effects on the guideline concentrations of eliminating some data and adding a substantial amount of new data are illustrated in Tables 14-1 and 14-2. The ERL and ERM values, based on the Long and Morgan (1990) data tables and the larger MacDonald (1992) tables, are compared by using the methods of Long and Morgan (1990) applied to both data sets.

For 13 aromatic hydrocarbons, the average of the ratios between the two sets of guidelines was 1.5 (1.9 for the ERLs and 1.2 for the ERMs). For eight trace metals, the average of the ratios between the two sets of guidelines was 1.7. The trace metals ERL values changed more than the ERM values (average ratios of 1.9 for the ERLs and 1.5 for the ERMs).

Overall, 7 of the 23 ERL values did not change and the ratios between the two sets of ERL values ranged from 1.0 to 9.4. Also, 7 of the 23 ERM values did not change. Of the 46 values, 14 remained unchanged, 17 increased, and 15 decreased. The overall mean factors of change were less than twofold for both trace metals and PAHs. These observations suggest that the guidelines are not terribly sensitive to the addition of new data once a minimum amount has been compiled. Also, they suggest that the guidelines originally developed by Long and Morgan (1990) generally are substantiated by additional data compiled by MacDonald (1992).

The accuracy of the guidelines in predicting toxicity has not yet been quantified. However, in the Hudson-Raritan estuary, the concentrations of many chemicals quantified in previous studies (Squibb *et al.*, 1992) frequently exceeded the ERM guidelines in the Arthur Kill and rarely exceeded them in the lower Hudson River. In a recent survey funded by NOAA, sediments from the Arthur Kill were extremely toxic to amphipods and other species, whereas the sediments from the lower Hudson River were not toxic.

14.4 STATUS

14.4.1 Extent of Use

The NSTP Approach is being used by NOAA's National Status and Trends Program, by Environment Canada, and by the Florida Department of Environmental Regulation. A variation on the approach is being pursued by the California Water Resources Control Board. Other states and regional districts have inquired about the possible use of the approach.

14.4.2 Extent to Which Approach Has Been Field-Validated

Validations of the guidelines have not yet been quantified. As described in Section 14.3.2.9, the original set of guidelines generally were substantiated by the addition of considerable amounts of new data, largely from field studies performed in many regions. The concordance between predictions of toxicity with the guidelines and actual observations of toxicity has been very

Table 14-1. Ratios Between the Guideline Values for Polynuclear Aromatic Hydrocarbons Determined with Data from Long and Morgan (1990) and Those Determined with Data from MacDonald (1992).
Total number of data points available are listed (with those used to determine guidelines in parentheses).

Chemical Analyte	MacDonald (1992)	Long and Morgan (1990)	Ratio Between Two Sets of Values	Values Increased (+) Decreased (-)
Polynuclear aromatic hydrocarbons (ppb d.w.)				
Acenaphthene	n=69(30)	n=35(15)	2.0(2.0)	
ERL	16	150	9.4	-
ERM	500	650	1.3	-
Anthracene	n=88(46)	n=39(26)	2.3(1.8)	
ERL	85.3	85	1.0	•
ERM	1100	960	1.1	+
Fluorene	n=95(48)	n=44(28)	2.2(1.7)	
ERL	19	35	1.8	-
ERM	540	640	1.2	-
2-methylnathphalene	n=49(28)	n=31(15)	1.6(1.9)	
ERL	70	65	1.1	+
ERM	670	670	1.0	•
naphthalene	n=97(44)	n=50(28)	1.9(1.6)	
ERL	160	340	2.1	-
ERM	2100	2100	1.0	•
phenanthrene	n=101(51)	n=49(34)	2.1(1.5)	
ERL	240	225	1.1	+
ERM	1500	1380	1.1	+
benzo(a)anthracene	n=81(43)	n=34(30)	1.9(1.4)	
ERL	261	230	1.1	+
ERM	1600	1600	1.0	•
benzo(a)pyrene	n=89(44)	n=43(27)	2.1(1.6)	
ERL	430	400	1.1	+
ERM	1600	2500	1.6	-
chrysene	n=89(45)	n=41(27)	2.2(1.7)	
ERL	384	400	1.0	•
ERM	2800	2800	1.0	•
dibenzo(a,h)anthracene	n=76(31)	n=23(18)	2.3(1.7)	
ERL	63.4	60	1.1	+
ERM	260	260	1.0	•
fluoranthene	n=117(71)	n=51(33)	2.3(1.8)	
ERL	600	600	1.0	•
ERM	5100	3600	1.4	+
pyrene	n=93(50)	n=43(28)	2.2(1.8)	
ERL	665	350	1.9	+
ERM	2600	2200	1.2	+
total PAH	n=78(34)	n=63(34)	1.2(1.0)	
ERL	4022	4000	1.0	•
ERM	44,790	35,000	1.3	+
Mean change in PAH ERLs		1.90		
Mean change in PAH ERMs		1.17		
Overall mean change in PAH values		1.53		

Table 14-2. Ratios Between the Guideline Values for Total PCBs and Trace Metals Determined with Data from Long and Morgan (1990) and Those Determined with Data from MacDonald (1992). *Total number of data points available are listed (with those used to determine guidelines in parentheses).*

Chemical Analyte	MacDonald (1992)	Long and Morgan (1990)	Ratio Between Two Sets of Values	Values Increased (+) Decreased (-)
Polychlorinated Biphenyl (ppb d.w.)				
total PCB	n=126(50)	n=77(33)	1.6(1.5)	
ERL	22.7	50	2.2	-
ERM	180	400	2.2	-
Trace Metals (ppm d.w.)				
arsenic	n=143(27)	n=48(16)	3.0(1.7)	
ERL	8.2	33.0	4.0	-
ERM	70.0	85.0	1.2	-
cadmium	n=261(84)	n=106(36)	2.5(2.3)	
ERL	1.2	5.0	4.2	-
ERM	9.6	9.6	1.0	•
copper	n=221(76)	n=91(51)	2.4(1.5)	
ERL	34.0	70.0	2.0	-
ERM	270	390	1.4	-
chromium	n=197(37)	n=76(21)	2.6(1.8)	
ERL	81	80	1.0	•
ERM	370	145	2.6	+
lead	n=210(73)	n=83(47)	2.5(2.6)	
ERL	46.7	35.0	1.3	+
ERM	223	110	2.0	+
mercury	n=169(42)	n=76(30)	2.2(1.4)	
ERL	0.15	0.15	1.0	•
ERM	0.71	1.3	1.5	-
nickel	n=169(19)	n=56(18)	3.0(1.1)	
ERL	20.9	30	1.4	-
ERM	51.6	50	1.0	•
silver	n=96(25)	n=47(13)	2.0(1.9)	
ERL	1.0	1.0	1.0	•
ERM	3.7	2.2	1.7	+
zinc	n=214(74)	n=79(46)	2.7(1.6)	
ERL	150	120	1.25	+
ERM	410	270	1.5	+
Mean change in PAH ERLs			1.9	
Mean change in PAH ERMs			1.9	
Overall mean change in metals values		1.74		

good thus far, but the degree of concordance has not been quantified. Additional opportunities to field-validate the guidelines will be available in future studies in Tampa Bay, the Hudson-Raritan estuary, and southern California.

14.4.3 Reasons for Limited Use

The NSTP Approach initially was used by NOAA to develop informal guidelines for internal agency use. Therefore, knowledge of and access to the guidelines was limited. As interest in the guidelines increased, they were released in a government document with a limited distribution. Therefore, the main reason for the limited use of the approach has been the limited awareness of its existence. Furthermore, the equilibrium-partitioning approach to national criteria and the most successful regional approach to criteria (apparent effects thresholds in Washington) have received considerable attention. Moreover, the guidelines thus far have not considered the potential for bioavailability or bioaccumulation because of a lack of data.

14.4.4 Outlook for Future Use and Amount of Development Yet Needed

There is significant potential for the expanded use of the NSTP Approach. Canada, Florida, and California currently are using the approach to develop their respective guidelines. Since the approach relies on existing data, other region-specific guidelines could be developed easily, using the data available from specific regions. The approach can be used to validate criteria developed with other single-method approaches. The database can be accessed for specific regions or for fresh, estuarine, or marine waters.

Several types of data are needed to further develop the approach. First, additional data are needed from studies in which TOC, grain size, and acid volatile sulfides were measured. Second, additional data are needed from spiked-sediment bioassays to establish cause-effect relationships. Third, additional data are needed from field studies in which very strong chemical gradients were observed. These studies should include measures of the toxicity and chemical contamination of bulk sediments and pore water. They would benefit from toxicity identification evaluations to identify the causative agents responsible for the observed biological effects (Ankley, 1989). A number of large field surveys are under way and being planned by NOAA and will lead to additional data to be included in the database. Once these additional data are available, they could be entered into the database and used to develop updated or new guidelines.

14.5 REFERENCES

Adams, W.J., R. A. Kimerle, and J. W. Barnett, Jr. In press. Sediment quality and aquatic life assessment. Envir. Sci. and Technol.

Ankley, G. 1989. Sediment toxicity assessment through evaluation of the toxicity of interstitial water. Environmental Research Laboratory-Duluth. U.S. Environmental Protection Agency, Duluth, MN. 27 pp.

Long, E. R. 1992. Ranges in chemical concentrations in sediments associated with adverse biological effects. Mar. Pollu. Bull. 24 (1): 38-45.

Long, E.R., D. MacDonald, and C. Cairncross. 1992. Status and trends in toxicants and the potential for their biological effects in Tampa Bay, Florida. NOAA Tech. Memo. NOS OMA 58. National Oceanic and Atmospheric Administration, Seattle, WA. 77 pp.

Long, E.R., and R. Markel. 1992. An evaluation of the extent and magnitude of biological effects associated with chemical contaminants in San Francisco Bay, California. NOAA Tech. Memo. NOS OMA 64. National Oceanic and Atmospheric Administration, Seattle, WA. 86 pp.

Long, E.R., and L.G. Morgan. 1990. The potential for biological effects of sediment-sorbed contaminants tested in the National Status and Trends Program. NOAA Tech. Memo. NOS OMA 62. National Oceanic and Atmospheric Administration, Seattle, WA. 175 pp.

Lorenzato, S.G., A. J. Gunther, and J. M. O'Connor. 1991. Summary of a workshop concerning sediment quality assessment and development of sediment quality objectives. California State Water Resources Control Board, Sacramento, CA. 32 pp.

MacDonald, D.D. 1992. Development of an integrated approach to the assessment of sediment quality in Florida. Prepared for Florida Department of Environmental Regulation. MacDonald Environmental Services, Ltd. Ladysmith, British Columbia. 114 pp.

MacDonald, D.D., and S.L. Smith. 1991. A discussion paper on the derivation and use of Canadian sediment quality guidelines for the protection of freshwater and marine aquatic life. Prepared for Canadian Council of Ministers of the Environment. Environment Canada. Ottawa.

MacDonald, D.D., S.L. Smith, M.P. Wong, and P. Mudroch. 1991. The development of Canadian marine environmental quality guidelines. Report prepared for the Interdepartment Working Group on Marine Environmental Quality Guidelines and the Canadian Council of Ministers of the Environment. Environment Canada. Ottawa, Canada. 50 pp.

Mannheim, F.T., and J.C. Hathaway. 1991. Polluted sediments in Boston Harbor-Massachusetts Bay: Progress report on the Boston Harbor data management file. U.S. Dept. of the Interior, Geological Survey Open File Report 91-331. USGS, Woods Hole, MA. 18 pp.

Soule, D.F., M. Oguri, and B.H. Jones. 1991. Marine Studies of San Pedro Bay, California, Part 20F. The marine environment of Marina Del Rey. October 1989 to September 1990. Submitted to Department of Beaches and Harbors, County of Los Angeles. University of Southern California, Los Angeles, CA. 206 pp.

Squibb, K. S., J. M. O'Connor, and T.J. Kneip. 1991. New York/New Jersey Harbor Estuary Program. Module 3.1: Toxics characterization report. Prepared for U.S. Environmental Protection Agency, Region 2. NYU Medical Center, Tuxedo, NY. 65 pp.

USEPA/SAB. 1989. Evaluation of the apparent effects threshold (AET) approach to assessing sediment quality. U.S. Environmental Protection Agency Science Advisory Board. Report of the Sediment Criteria Subcommittee. U.S. EPA SAB-EETFC-89-027. 16 pp.

EPA
United States
Environmental Protection Agency
Office Of Water ((WH-556)
401 M Street SW.
Washington DC 20460

Official Business
Penalty for Private Use $300

★ 1992 ★
THE YEAR OF
CLEAN WATER
Celebration + Commitment