EPA 910/9-89-013
Library
PB89-217913

Puget Sound Estuary Program

Briefing Report to the EPA Science Advisory Board

THE APPARENT EFFECTS THRESHOLD APPROACH

September 1988

Prepared for
U.S. Environmental Protection Agency
Region 10 - Office of Puget Sound
Seattle, Washington

Protection Agency
Program Office
Library

BRIEFING REPORT TO THE EPA SCIENCE ADVISORY BOARD:

THE APPARENT EFFECTS THRESHOLD APPROACH

Submitted by

Office of Puget Sound
Puget Sound Estuary Program
U.S. Environmental Protection Agency, Region 10
1200 6th Avenue
Seattle, Washington 98101

Prepared by

PTI Environmental Services
3625 132nd Avenue SE
Suite 301
Bellevue, Washington 98006

under Battelle Columbus Division
EPA Contract No. 68-03-3534
PTI Contract No. C714-01

September 1988

CONTENTS

	Page
LIST OF FIGURES	iv
LIST OF TABLES	v
1. INTRODUCTION	1
EPA REGION 10 CHARGE TO THE EPA SCIENCE ADVISORY BOARD	1
REGULATORY NEEDS FOR SEDIMENT QUALITY VALUES	2
SELECTION OF A SEDIMENT QUALITY VALUE APPROACH FOR PUGET SOUND	4
EVALUATION OF RELIABILITY	6
2. THE CONCEPT OF AET	9
DESCRIPTION OF THE AET APPROACH	9
INTERPRETATION OF AET	10
Relationships Among Chemical-Specific AET	10
Dose-Response Relationships and AET	12
Influence of Environmental Factors on AET Interpretation	15
Interactive Effects and AET	17
Unmeasured Chemicals and AET	17
Matrix Effects and Bioavailability	18
3. GENERATION OF AET VALUES FOR PUGET SOUND	19
PUGET SOUND DATABASE	19
Biological Data	20
Bioassay Tests	20
Benthic Infauna Analyses	26
Chemical Data	27
Guidelines for Data Treatment	28

Bioassay Data 28
Benthic Infauna Analyses 29
Chemical Data 29

VALIDATION TEST METHODS 30

VALIDATION RESULTS AND DISCUSSION 31

Characteristics of Biologically Impacted Stations Predicted as Nonimpacted by AET 38

Amphipod Bioassay Stations 38
Benthic Infauna Stations 43

Relative Performance of Dry Weight- and TOC-Normalized AET 44

Sediment-Water Contaminant Exchange 46
The Uniformity of Organic Matter 47

4. APPLICATION OF AET IN PUGET SOUND SEDIMENT MANAGEMENT PROGRAMS 48

COMMENCEMENT BAY NEARSHORE/TIDEFLATS SUPERFUND INVESTIGATION 49

PUGET SOUND DREDGED DISPOSAL ANALYSIS 49

URBAN BAY TOXICS ACTION PROGRAM 50

PUGET SOUND WATER QUALITY MANAGEMENT PLAN 50

PROPOSED STATE SEDIMENT QUALITY STANDARDS 51

5. SUMMARY 52

6. REFERENCES 53

FIGURES

Number		Page
1	Measures of reliability (sensitivity and efficiency)	7
2	The AET approach applied to sediments tested for lead and 4-methyl phenol concentrations and toxicity response during bioassays	11
3	Hypothetical example of dose-response relationship resulting from laboratory exposure to single chemicals X and Y	13
4	Hypothetical example of toxic response resulting from exposure to environmental samples of sediment contaminated with chemicals X and Y	14
5	Hypothetical example of AET calculation for chemical X based on classification of significant and nonsignificant responses for environmental samples contaminated with both chemicals X and Y	16
6	Location of sampling sites for Apparent Effects Threshold data sets in Puget Sound	23
7	Sensitivity of dry-weight and total organic carbon-normalized AET	45

TABLES

Number		Page
1	Summary of data sets used to evaluate Puget Sound AET	21
2	Summary of selected chemical concentrations and key sediment characteristics for biological stations from reference and nonreference areas of Puget Sound	24
3	1988 Puget Sound AET for selected chemicals (normalized to dry weight)	32
4	1988 Puget Sound AET for selected chemicals (normalized to total organic carbon)	35
5	Sensitivity and efficiency results for four validation tests conducted with Puget Sound AET (normalized to dry weight- and total organic carbon)	39
6	Characteristics of stations at which significant amphipod mortality was found but no effects were predicted	41
7	Characteristics of stations at which significant benthic effects were found but no effects were predicted	42

1. INTRODUCTION

In response to growing concerns about widespread biological impacts associated with chemically contaminated sediments in Puget Sound, the Washington Department of Ecology (Ecology) and U.S. Environmental Protection Agency (EPA) Region 10, have been actively pursuing the development of numerical sediment quality standards. This effort has involved the identification and evaluation of a variety of methods that could be used as the basis for sediment management decision-making.

The Apparent Effects Threshold (AET) approach is a tool for deriving empirical sediment quality values for a range of biological indicators used to assess contaminated sediments. The purpose of this briefing report is to describe the AET approach, how the method has been used to develop sediment quality values, and how these sediment quality values are being used to make regulatory decisions in a variety of programs in Puget Sound. The overall goal of these programs is to develop technically defensible and publicly acceptable tools for sediment management.

EPA REGION 10 CHARGE TO THE EPA SCIENCE ADVISORY BOARD

Based on the predictive capabilities of the AET method, and its applicability to existing or planned regulatory activities, Ecology has concluded that the AET approach could provide a technically defensible and publicly acceptable basis for the establishment of state sediment quality standards. Subject to the provisions of Section 303 of the Clean Water Act, the proposed standards will be submitted to EPA Region 10 for review and comment. It is currently anticipated that state sediment quality standards will be submitted to EPA in draft form by January 1989.

Although the AET approach has undergone regional peer review, Region 10 and Ecology believe that an additional evaluation by an independent review board is warranted. A Science Advisory Board (SAB) review is appropriate because the development of sediment criteria and standards represents a relatively new and still evolving science. In addition, the AET approach is already being considered by other state and federal agencies for use in sediment management. The results of an SAB review will be important in determining the extent to which the development of AET values should be encouraged for other geographic locations.

Based upon these considerations, Region 10 has charged the SAB to undertake a review of the AET approach. As part of this review, it was requested that the SAB evaluate the method from two distinct perspectives:

- First, the AET approach should be evaluated as a technical concept, without consideration of the specific regulatory applications of the method in Puget Sound or the extent of the database used to generate Puget Sound AET.

- Second, the AET approach should be reviewed specifically as applied in Washington to generate sediment quality values for Puget Sound.

During review of the AET approach, both conceptually and as applied in Puget Sound, it is requested that the SAB specifically address the following questions.

1) Is the use of a statistically-based, empirical approach (i.e., field and laboratory observations) to establishing quantitative relationships between sediment contaminants and biological effects a technically acceptable means of developing sediment quality values?

2) Are the guidelines for data treatment used in generating Puget Sound AET technically appropriate?

3) Does the AET approach enable generation of sediment quality values for a wide range of the chemical contaminants that could be considered problematic in marine sediments?

4) Does the AET approach adequately incorporate consideration of a range of biological organisms and ecological endpoints?

5) Does the AET approach adequately incorporate direct measures of sediment toxicity?

6) Are the biological indicators used in Puget Sound for the development of AET values appropriate for the assessment of sediment contamination?

7) Does the AET approach adequately incorporate, either by design or default, the influences of complex mixtures of contaminants typically found in environmental samples?

8) Is the AET approach likely to generate sediment quality values that are applicable to field conditions?

9) Is the approach used to evaluate the predictive reliability of AET values, as applied to field data, scientifically defensible?

10) Is the AET approach technically defensible and appropriate for use in regulatory decision-making?

In addition, it was requested that the SAB provide recommendations to the region concerning the following technical issues:

1) Appropriateness of normalization factors (e.g., organic carbon) for the development of sediment quality standards.

2) The need for, and role of, site-specific biological testing as a means of verifying AET predictions in the context of regulatory programs.

REGULATORY NEEDS FOR SEDIMENT QUALITY VALUES

Sediment in many areas of Puget Sound is contaminated with potentially toxic substances such as petroleum-derived compounds, chlorinated organic compounds, and

metals. Sources of these contaminants include runoff from urban streets, industrial discharges, and municipal sewage treatment plants. Shallow sills at the north and south ends of Puget Sound inhibit the exchange of water and promote recirculation of contaminants. It is estimated that less than 20 percent of the contaminants discharged to Puget Sound are transported seaward to the Strait of Juan de Fuca; most of the persistent contaminants are deposited in urban embayments within Puget Sound. Sediment contamination in Puget Sound has been associated with impacts to benthic infauna, and development of tumors and other abnormalities in bottom-dwelling fish (Long and Chapman 1985; Barrick et al. 1985, 1986; Malins et al. 1982, 1984; Swartz et al. 1982). In addition, fish, crabs, and bivalves in contaminated areas have been observed to accumulate pollutants in their muscle tissue and organs (Dexter et al. 1982; Gahler et al. 1982; Barrick et al. 1985; Yake and Norton 1986; Ginn and Barrick 1988; Beller et al. 1988a; Pastorok et al. 1988). In several of these areas (e.g., Elliott Bay, Commencement Bay, and Eagle Harbor), local health departments have advised local residents to limit their consumption of seafood.

Pollution control programs in Puget Sound have traditionally focused on protecting water quality through effluent discharge limits and water quality standards. Such controls have generally not been effective in preventing sediment contamination. The control of sediment contamination is currently limited because no benchmarks in the form of guidelines or standards have been available to assess adverse biological impacts of contaminated sediments. In remedial action programs, such tools are needed to address the following specific regulatory needs:

- Identify problem chemicals
- Establish a link between contaminated sediments and sources
- Provide a predictive tool for cases in which site-specific biological testing results were 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
- Provide a reference point for establishing a cleanup goal.

In addition to remedial activities, several other regulatory programs require decision-making tools to enable characterization and management of contaminated sediments. For example, identification of problem areas and trend analyses is an integral part of ambient water quality monitoring. 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. Sediment quality values can provide useful tools for assessing the need for further chemical and biological testing and evaluation of sediment. The use of sediment quality values is a cost-effective alternative to extensive series of case-by-case biological testing or monitoring and is a complement to biological testing (e.g., enables source identification using chemical-specific values).

Source control efforts have also required consideration of sediment management tools. For example, chemical-specific sediment quality values have been identified as one useful tool in developing National Pollutant Discharge Evaluation System (NPDES) permit requirements associated with the maximum allowable concentration of contaminants in effluent particulate material. Such requirements are intended to ensure that the accumulation of effluent particulate material in sediments will not threaten the survival and reproductive potential of aquatic organisms. Sediment criteria must be translated from the ambient sediment quality goal into a source control limit.

SELECTION OF A SEDIMENT QUALITY VALUE APPROACH FOR PUGET SOUND

In the past decade, several federal, regional, and state agencies have developed numerical criteria or assessment methods for evaluating contamination in sediments and dredged material. Most early efforts at developing criteria were based on comparing chemical concentrations in contaminated areas to those in reference areas, and did not consider biological effects. More recently, approaches to evaluating sediment quality have focused on determining relationships between sediment contaminant levels and adverse biological impacts. Much of the information and analysis presented in this section is contained in recent reviews of approaches to sediment quality value development (e.g., Beller et al. 1986; Lyman et al. 1987; Battelle 1988; and Chapman in review). Based on these documents, the following approaches were reviewed for possible application in Puget Sound programs:

- Field-Based Approaches

  -Reference Area
  -Field-Collected Sediment Bioassay
  -Screening Level Concentration (SLC)
  -Sediment Quality Triad (Triad)
  -Apparent Effects Threshold (AET)

- Laboratory/Theoretically-Based Approaches

  -Water Quality Criteria/Interstitial Water (WQC)
  -Equilibrium Partitioning (sediment-water)
  -Equilibrium Partitioning (sediment-biota)
  -Spiked Sediment Bioassay.

Field-based approaches rely on empirical chemical and/or biological measurements of sediments to establish sediment quality values. Some of these approaches are purely chemical (reference area approach) or biological (field-collected sediment bioassay approach) in nature. Other approaches such as SLC, Triad, and AET associate biological responses (e.g., field-collected sediment bioassays, *in situ* biological effects observed in organisms living in or on sediments) and chemical concentrations measured in sediments to develop sediment quality values. Laboratory/theoretically-based approaches rely on extrapolation of water quality criteria to sediments, models of environmental interactions (e.g., sediment-water equilibrium partitioning), or extrapolation of laboratory cause-effect studies to develop sediment quality values.

As part of the evaluation of these approaches, emphasis was placed on the following management considerations relevant to the development of Puget Sound standards:

- Applicability to existing and planned sediment management programs in Puget Sound
- Feasibility of implementation in the near term
- Environmental protectiveness and cost effectiveness (i.e., reliability in predictions of adverse effects)
- Regulatory defensibility (i.e., supporting weight of evidence)
- Cost of initial sediment quality value development
- Cost of routine application as a regulatory tool.

In addition, each approach was evaluated according to the following technical considerations:

- Data requirements for initial sediment quality value development
- Data requirements for routine application as a regulatory tool
- Ability to develop chemical-specific sediment quality values
- Ability to develop sediment quality values for a wide range of chemicals
- Current availability of values for a wide range of Puget Sound problem chemicals
- Incorporation of influence of chemical mixtures in sediments
- Incorporation of a range of biological indicator organisms
- Incorporation of direct measurement of sediment biological effects
- Applicability of predictions to historical sediment chemistry data
- Ease and extent of field verification in Puget Sound.

These criteria include key features of most of the available approaches to developing sediment quality values (e.g., as advantages or limitations of the approaches). The AET approach scored favorably on these criteria except with respect to data requirements and cost of initial development. The AET approach has relatively extensive data requirements (e.g., sediment chemistry, one or more *in situ* biological effects measurements, and one or more laboratory bioassays) for sediment quality values development. However, the current additional cost of AET development is minimal because a large Puget Sound database has already been compiled and a wide range of AET values is available for use. The cost of routine application of the AET approach is comparable to that for other approaches because its implementation requires only the collection of sediment chemistry data for comparison to chemical-specific standards developed for the approach.

Field verification using diverse environmental samples was an important element of the evaluation of each approach for current use in Puget Sound because none of the

available approaches is fully capable of addressing all concerns over interactive effects among chemicals. Such effects may occur when organisms are exposed to multiple chemicals present in sediment. Using the three most promising approaches (Screening Level Concentration, AET, and sediment-water Equilibrium Partitioning), sediment quality values were generated and applied to biological effects data from Puget Sound to assess their ability to predict observed adverse impacts in actual environmental samples.

Based on consideration of the management and technical criteria and the results of the field verification exercise, the AET approach was selected by the Puget Sound regulatory agencies as the currently preferred method for developing sediment quality standards in Puget Sound (other approaches will still be considered as they are developed or tested). The AET approach can be used to provide sediment quality values for the greatest number and widest range of chemicals of concern in Puget Sound. The approach also incorporates the widest range of biological indicators that are directly applicable to sediment conditions. The approach taken for conducting the field verification exercise is described in the following section.

EVALUATION OF RELIABILITY

Environmental factors such as matrix effects or chemical interactions can affect the ability of chemical-specific criteria to correctly predict adverse biological effects in environmental samples. The only way to definitively test the chemical-specific predictions (i.e., whether particular biological effects are always observed above a given sediment quality value) is to conduct controlled laboratory spiking studies. However, the binary (impacted/nonimpacted) predictions of sediment quality values are most pertinent to sediment management and are testable using data sets that have matched chemical and biological data. Tests using such data are applicable to and recommended for assessing environmental predictions of any approach to developing sediment criteria as part of management performance objectives. The basic requirements of such tests of reliability are described in this section.

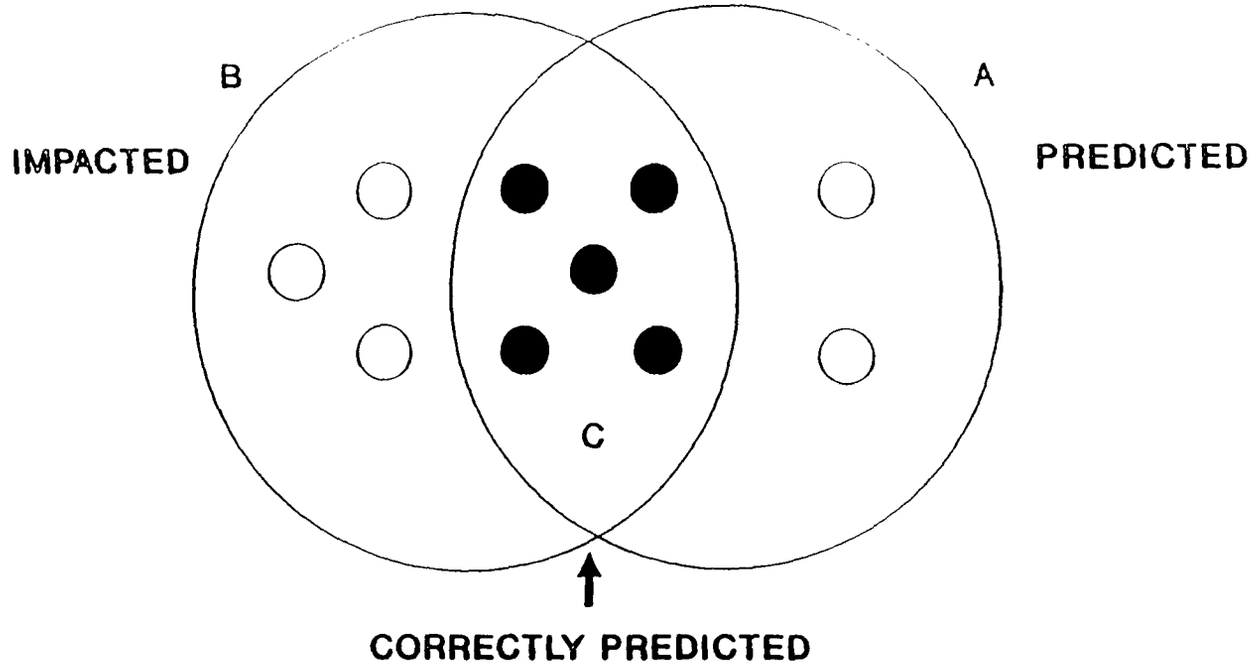
To meet the needs of most sediment quality management programs, an ideal sediment criteria approach would perform well on both of the following measures of reliability, which are evaluated with actual field data:

- **Sensitivity** in detecting environmental problems (i.e., are <u>all</u> biologically impacted sediments identified by the predictions of the chemical sediment criteria?)

- **Efficiency** in screening environmental problems (i.e., are <u>only</u> biologically impacted sediments identified by the predictions of the chemical sediment criteria?).

As a measure of reliability, sensitivity is defined as the proportion of all stations exhibiting adverse biological effects that are correctly predicted using sediment quality values. Efficiency is defined as the proportion of all stations predicted to have adverse biological effects that actually are impacted. The concepts of sensitivity and efficiency are illustrated in Figure 1.

Sensitivity and efficiency are independent measures of reliability. For example, a sediment criteria approach that sets values for a wide range of chemicals near their



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 1. Measures of reliability (sensitivity and efficiency).

analytical detection limits will probably be sensitive but inefficient. That is, it will predict a large percentage of sediments with biological effects but will also predict many biologically unimpacted sediments with only slightly elevated chemical concentrations. Such an approach may be environmentally protective but also may result in overregulation that would not be cost effective.

Conversely, a sediment criteria approach that sets values at the upper end of the range of environmental concentrations may be efficient but insensitive. That is, a high percentage of the stations with predicted impacts may indeed be biologically impacted, but the approach may fail to predict other biologically impacted stations with moderate to high chemical concentrations. Such an approach may be cost-effective and defensible in pursuing high priority remedial actions (i.e., would not overregulate) but would not be environmentally protective.

The overall reliability of any sediment criteria approach addresses both components of 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:

- **Overall reliability** = $\dfrac{\left[\begin{array}{c}\text{All stations correctly predicted as impacted}\\ +\\ \text{All stations correctly predicted as nonimpacted}\end{array}\right]}{\left[\text{Total number of stations evaluated}\right]}$

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).

These measures of reliability were important considerations in assessing the suitability for regulatory application of sediment quality values calculated by the AET approach. An assessment of reliability has recently been conducted using a large database comprising samples from 13 Puget Sound embayments (not all biological indicators are available in all embayments). In at least 85 percent of the available samples for each biological indicator, the approach either correctly classifies as impacted samples that exhibit adverse biological effects or correctly classifies as not impacted samples that do not exhibit adverse biological effects. Detailed results of the reliability evaluation for the AET approach are described in Section 3. The underlying concept of the AET approach is described in the following section.

2. THE CONCEPT OF AET

An AET is defined as the sediment concentration of a given chemical above which statistically significant ($P \leq 0.05$) biological effects (e.g., depressions in the abundances of indigenous benthic infauna) are always expected. If any chemical exceeds its AET for a particular biological indicator, an adverse biological effect is predicted for that indicator. If all chemical concentrations are below their AET for a particular biological indicator, then no adverse effect is predicted.

In this section, AET generation is described and the AET concept is discussed as it relates to the interpretation of chemical and biological data in field-collected sediments. AET generation is a conceptually simple process and incorporates the complexity of biological-chemical interrelationships in the environment without relying upon *a priori* assumptions as to the mechanistic nature of these interrelationships. The concept of the AET is presented in this section with little reference to specific chemicals, specific biological tests, or specific chemical normalizations, because the approach is not inherently limited to specific subsets of these variables. The specific use of the AET concept to generate AET values from Puget Sound data is described in Section 3.

DESCRIPTION OF THE AET APPROACH

The focus 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. The calculation of AET for each chemical and biological indicator is straightforward:

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)

2. **Identify "impacted" and "nonimpacted" stations**--Statistically test the significance of adverse biological effects relative to suitable reference conditions for each sediment sample and biological indicator; suitable reference conditions are established by sediments containing very low or undetectable concentrations of any toxic chemicals

3. **Identify AET using only "nonimpacted" stations**--For each chemical, the AET can be identified for a given biological indicator as the highest <u>detected</u> concentration among sediment samples that do not exhibit statistically significant effects (if the chemical is undetected in all nonimpacted samples, 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 is only a preliminary minimum estimate (or may not exist).

5. Repeat Steps 1-4 for each biological indicator.

A pictorial representation of the AET approach for two example chemicals is presented in Figure 2 based on results for a toxicity bioassay. Two subpopulations of all sediments analyzed for chemistry and subjected to a bioassay are represented by bars in the figure and include:

- Sediments that did not exhibit statistically significant ($P>0.05$) toxicity relative to reference conditions ("nonimpacted" stations)

- Sediments that exhibited statistically significant ($P\leq0.05$) toxicity in bioassays relative to reference conditions ("impacted" stations).

The horizontal axis in each figure represents sedimentary concentrations of chemicals (lead or 4-methylphenol) on a log scale. Dry weight normalized data are presented in Figure 2, although the AET approach is not limited to any particular normalization. For the toxicity bioassay under consideration, the AET for lead is the highest lead concentration corresponding to sediments that did not exhibit significant toxicity (the top bar for lead in Figure 2). Above this lead AET, significant toxicity was always observed in the data set. The AET for 4-methylphenol was determined analogously.

INTERPRETATION OF AET

An AET corresponds to the sediment concentration of a chemical above which all samples for a particular biological indicator were observed to have adverse effects. Thus, the AET is based on noncontradictory evidence of biological effects. 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.

Relationships Among Chemical-Specific AET

Using Figure 2 as an example, sediment from Station SP-14 exhibited severe toxicity, potentially related to a greatly elevated level 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 2). Despite the toxic effects associated with the sample, sediments from many other stations with higher lead concentrations than 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. 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

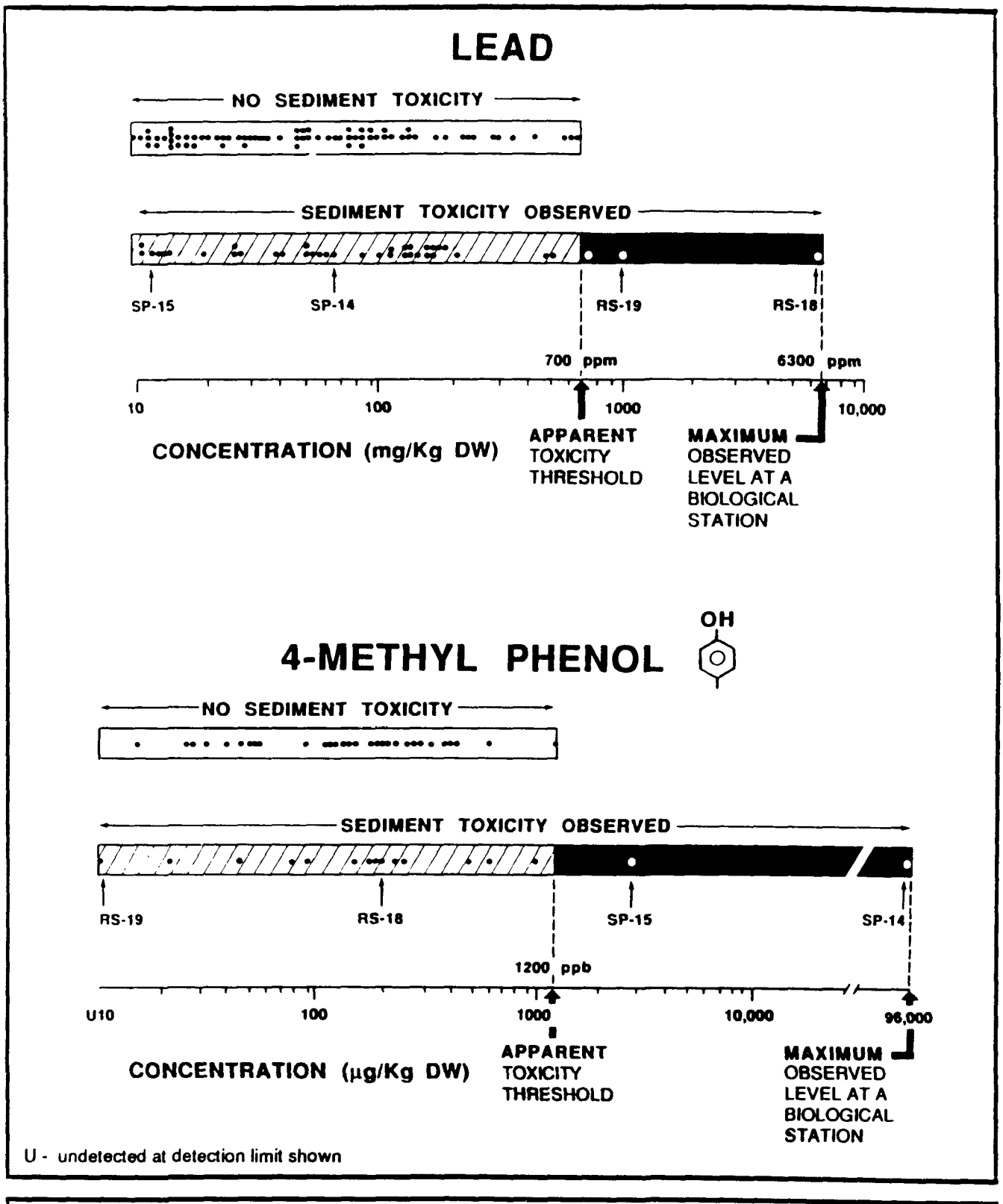


Figure 2. The AET approach applied to sediments tested for lead and 4-methyl phenol concentrations and toxicity response during bioassays.

other sites without associated biological effects). Based on the results for lead and 4-methylphenol, effects at 4 of the 28 impacted sites shown in the figures may be associated with elevated concentrations of 4-methylphenol, and effects at 7 other sites may be associated with elevated concentrations of lead (or similarly distributed contaminants).

These results illustrate that the occurrence of biologically impacted stations at concentrations below the AET of a single chemical does not imply that AET in general are not protective against biological effects, only that single chemicals may not account for all stations with biological effects. By developing AET for multiple chemicals, a high percentage of all stations with biological effects are accounted for with the AET approach (reliability results are presented in Section 3 of this briefing document).

AET 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) before generation of AET is appropriate.

Dose-Response Relationships and AET

The AET concept is consistent with empirical observations in the laboratory of dose-response relationships between increasing concentrations of individual toxic chemicals and increasing biological effects. A simple hypothetical example of such single-chemical relationships is shown for chemicals X and Y in Figure 3. In the example, data are shown for laboratory exposures of a test organism to sediment containing only increasing concentrations of chemical X, and independently, for exposures to sediment containing only increasing concentrations of chemical Y. The magnitude of toxic response in the example differs for the two chemicals and occurs over two different concentration ranges. It is assumed that at some level of response, for example >25 percent, the two different responses can be distinguished from reference conditions (i.e., responses resulting from exposure to sediments containing very low or undetectable concentrations of any toxic chemicals).

These single-chemical relationships cannot be proven in the field because organisms are exposed to complex mixtures of chemicals in environmental samples. In addition, unrelated discharges from different sources can result in uncorrelated distributions of chemicals in environmental samples. To demonstrate the potential effects of these distributions, response data are shown in Figure 4 for a random association of chemicals X and Y using the same concentration data as in Figure 3. The data have been plotted according to increasing concentrations of chemical X, and the same dose-response relationship observed independently for the two chemicals in the laboratory has been assumed. The contributions of chemicals X and Y to the toxic response shown for these simple mixtures is intended only for illustration purposes to enable direct comparison to the relationships shown in Figure 3, but are analogous to an additive toxic response. Other interactive effects are not considered in this example.

In Figure 4, a significant response relative to reference conditions would result whenever elevated concentrations of either chemical X or chemical Y occurred in a sample. Because of the random association of Y with X in these samples, the significant responses would appear to occur randomly over the lower concentration range of

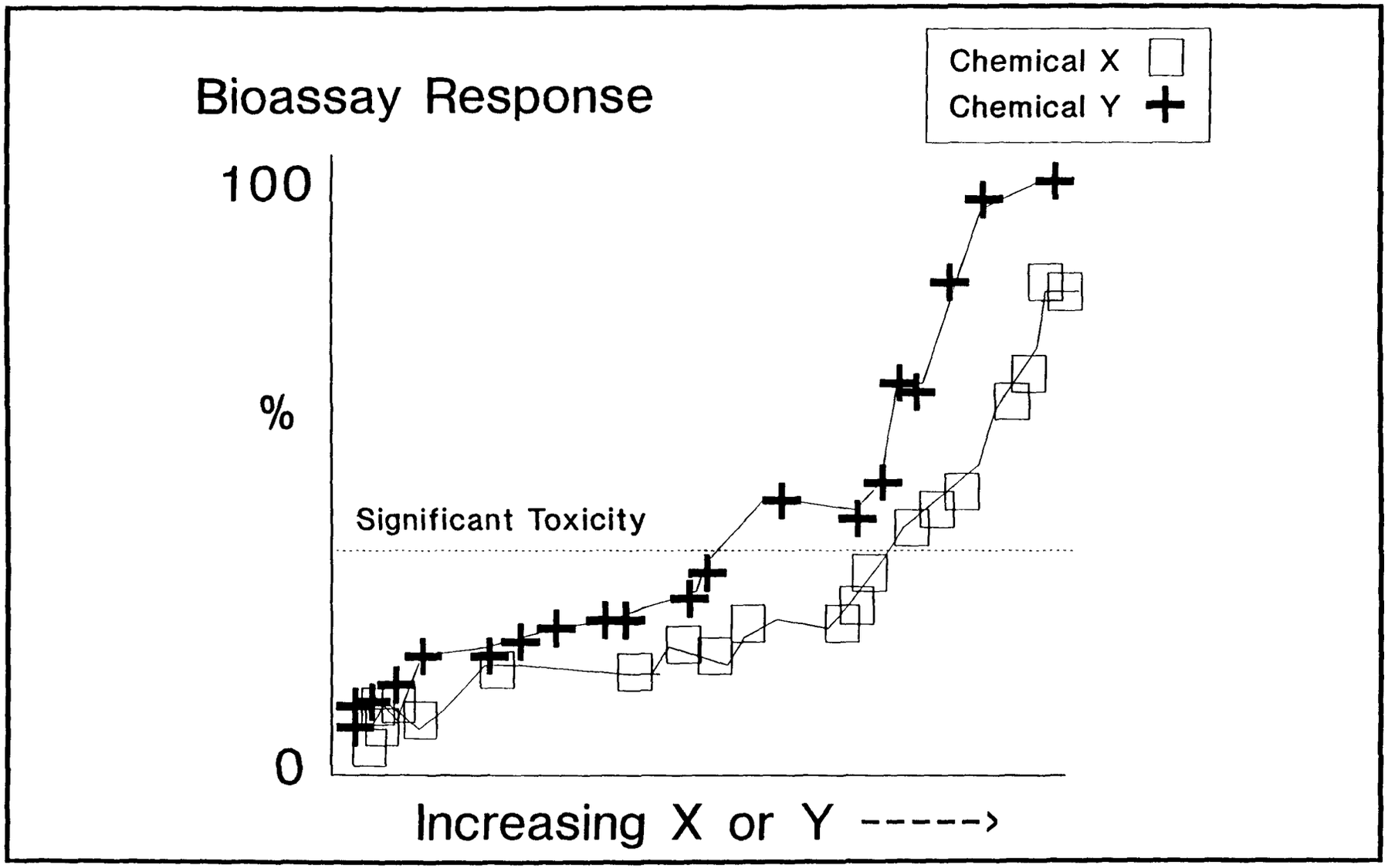


Figure 3. Hypothetical example of dose-response relationship resulting from laboratory exposure to single chemicals X and Y.

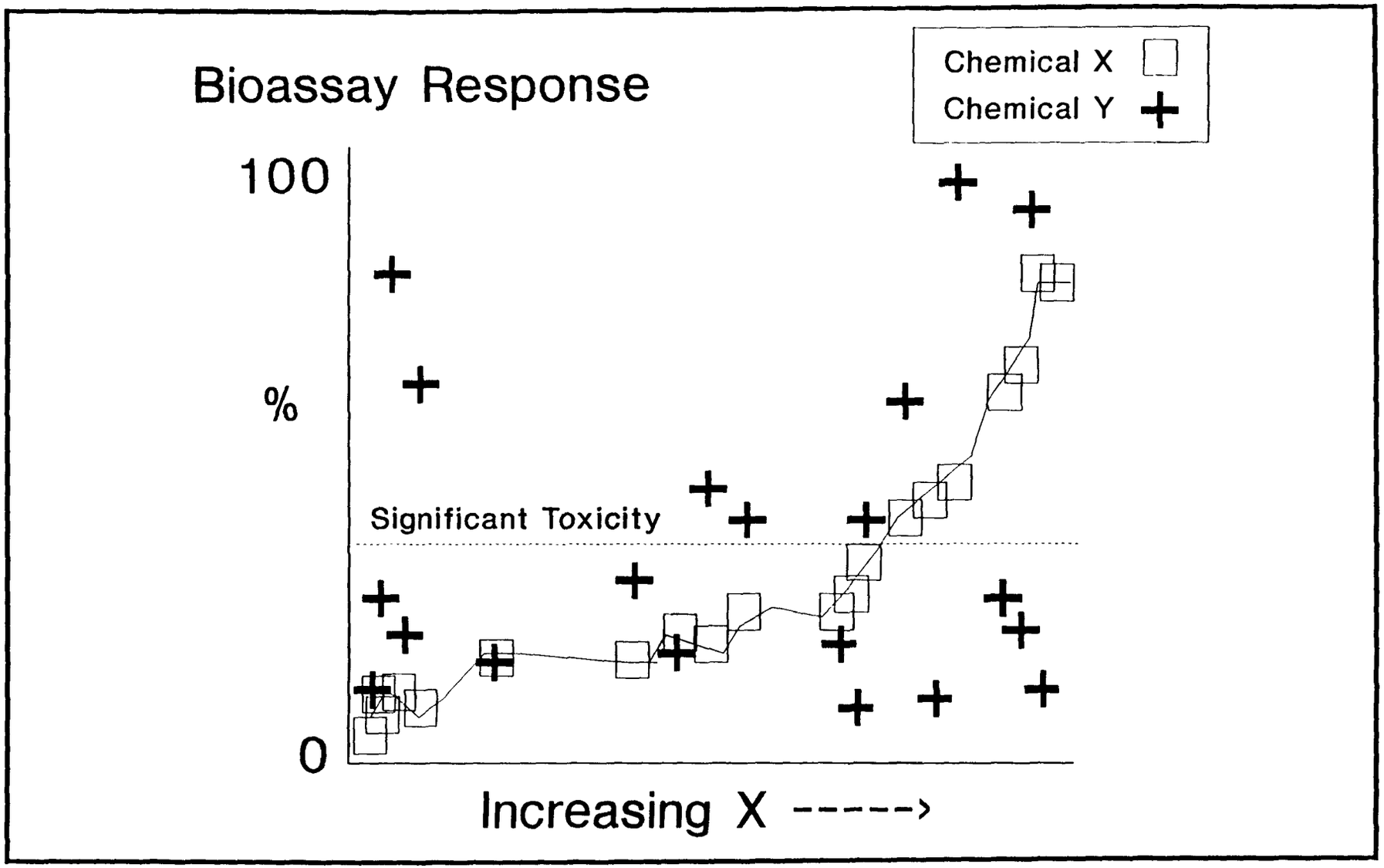


Figure 4. Hypothetical example of toxic response resulting from exposure to environmental samples of sediment contaminated with chemicals X and Y.

chemical X. The classification of the responses shown in Figure 4 into significant and nonsignificant groups (i.e., >25 percent response for either chemical) results in generation of Figure 5.

Figure 5 represents the appearance of the environmental results when ranked according to concentration of chemical X using these data. Below the AET for chemical X, significant toxicity is produced by elevated concentrations of chemical Y, which is randomly associated with the distribution of chemical X. Above the AET for chemical X, significant toxicity is always produced by elevated concentrations of chemical X, although in some samples, elevated concentrations of chemical Y also contribute to the overall toxicity. The AET for chemical X corresponds conceptually, in this simple example, to the concentration in Figure 3 at which a significant difference in response was observed in the laboratory for chemical X.

In environmental samples that contain complex mixtures of chemicals, a monotonic dose-response relationship such as in this simple two-chemical example may not always apply. For example, a consistently increasing biological response may not always occur at increasing concentrations of a chemical above its AET. Such observations could indicate that the AET is coincidental (i.e., that the observed toxicity in some or all samples above the AET is unrelated to the presence of that chemical), or that changing environmental factors in samples exceeding an AET obscure a monotonic dose-response relationship. Such factors are discussed in the following section.

Influence of Environmental Factors on AET Interpretation

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
- Matrix effects and bioavailability [i.e., phase associations between contaminants and sediments that affect bioavailability of the contaminants, such as the incorporation of polycyclic aromatic hydrocarbons (PAH) in soot particles].

The AET approach cannot distinguish and quantify the contributions of interactive effects, unmeasured chemicals, or matrix effects in environmental samples, but AET values may be influenced by these factors. To the extent that the samples used to generate AET are representative of samples for which AET are used to predict effects, the above environmental factors may not detract from the predictive reliability of AET. Alternatively, the infrequent occurrence of the above environmental factors in a data set used to generate AET could detract from the predictive reliability of those AET values. If confounding environmental factors render the AET approach unreliable, this should be evident from validation tests in which biological effects are predicted in environmental samples. Tests of AET values generated from Puget Sound data (see Section 3) indicate that the approach is relatively reliable in predicting biological effects despite the potential uncertainties of confounding environmental factors.

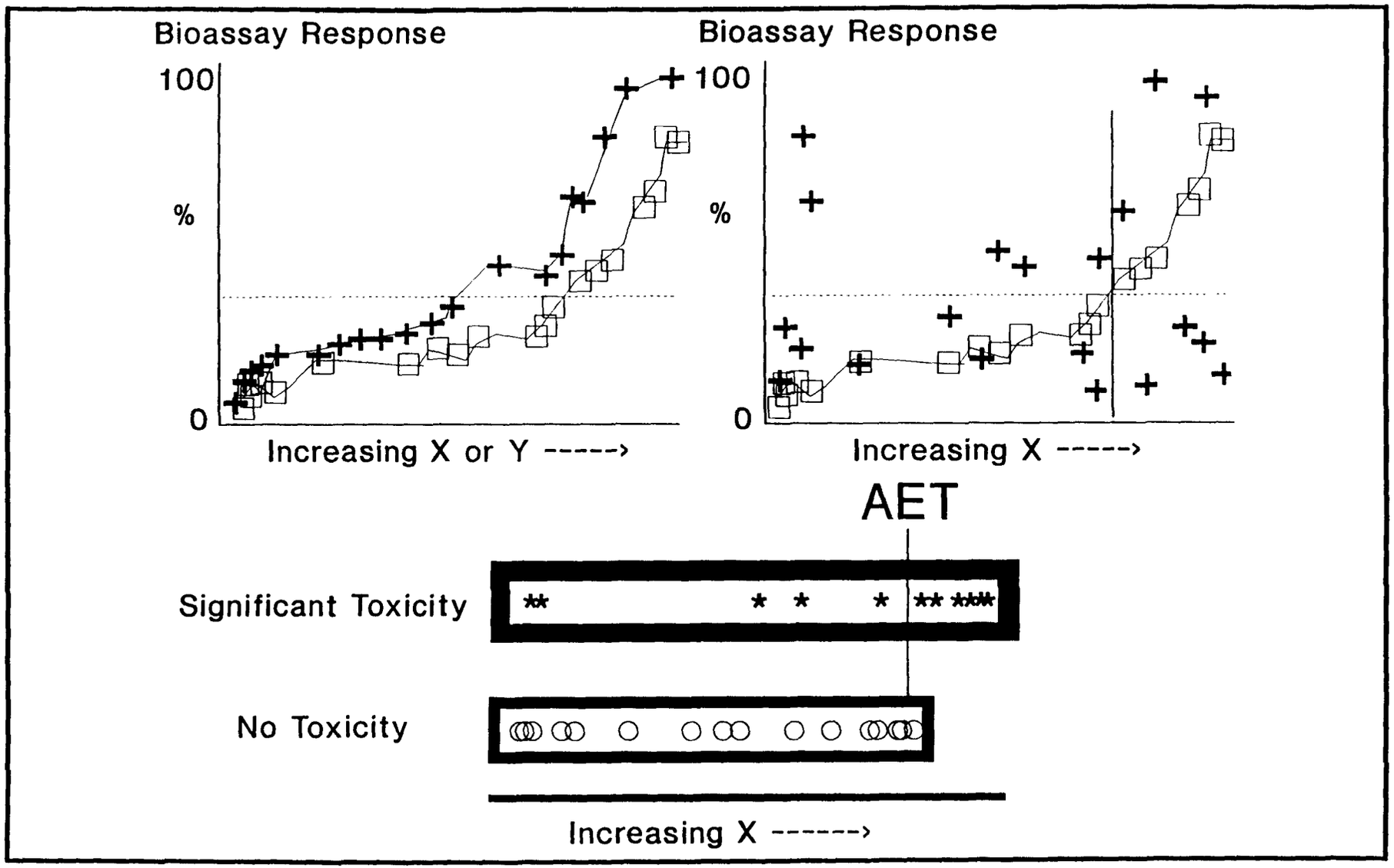


Figure 5. Hypothetical example of AET calculation for chemical X based on classification of significant and nonsignificant responses for environmental samples contaminated with both chemicals X and Y. [Previous Figures 3 and 4 are also shown for comparison; dashed line indicates level of significant toxicity]

Although the above environmental factors can influence the **generation** of field-based sediment quality values such as AET, they also may influence the **application** of all sediment quality value approaches for the prediction of adverse biological effects. For example, sediment quality values based on laboratory sediment bioassays spiked with single chemicals would not be susceptible to the effects of the environmental factors listed above. However, in applying such values to field-collected samples, predictions of biological effects could be less successful to the extent that interactive effects, unmeasured chemicals, and matrix effects occur in the environment.

The nature of the relationships between AET values and confounding environmental factors is discussed in the remainder of this section.

Interactive Effects and AET--AET uncertainty is increased by the possibility of interactive effects; the increase in uncertainty is expected to be less pronounced when large data sets collected from diverse areas are used to generate AET. Additivity and synergism can produce a comparatively low AET for a given chemical by causing impacts at concentrations that would not cause impacts in the absence of these interactive effects. This would effectively reduce the pool of nonimpacted stations used to generate AET. This effect should be reduced if a diverse database is used such that chemicals occur over a wide range of concentrations at stations where additivity and synergism are not operative. For chemicals that covary regularly in the environment (e.g., fluoranthene and pyrene), even a large, diverse database will not reduce the effects of additivity and/or synergism on AET generation. The resulting AET values for such chemicals may be reliable in predicting biological effects in environmental samples although not representative of the toxicities of the chemicals acting independently.

Antagonism will produce comparatively high AET values if (and only if) the AET is established at a station where antagonism occurs. A large, diverse database could not rectify this elevation of AET if the station at which antagonism occurred was the nonimpacted station with the highest concentration (i.e., the station setting an AET). An AET set by a station at which antagonism occurred would not be representative of the toxicity of the chemical acting independently. Hence, if antagonism did not occur widely, such antagonistic effects would cause the AET to be less sensitive in predicting adverse effects in the environment.

Empirical approaches such as the AET do not provide a means for characterizing interactive effects. Only laboratory-spiked sediment bioassays offer a systematic and reliable method for identifying and quantifying additivity, synergism, and antagonism. A great deal of research effort would be required to test the range of chemicals potentially occurring in the environment (both individually and in combination), a sufficiently wide range of organisms, and a wide range of sediment matrices to establish criteria. In addition, the applicability of bioassays conducted with laboratory-spiked sediments to environmentally-contaminated sediments requires further testing.

Unmeasured Chemicals and AET--Another source of uncertainty for AET and other field-based approaches is the possibility of effects being caused by unmeasured, covarying chemicals. Such chemicals would not be expected to substantially decrease the ability of AET to predict biologically impacted stations (excluding interactive effects discussed above). If an unmeasured chemical (or group of chemicals) varies consistently in the

environment with a measured chemical, then the AET established for the measured contaminant will indirectly apply to, or result in the management of, the unmeasured contaminant. In such cases, a measured contaminant would act as a surrogate for an unmeasured contaminant (or group of unmeasured contaminants). Because all potential contaminants cannot be measured routinely, management strategies must rely to some extent on "surrogate" chemicals.

If an unmeasured toxic chemical (or group of chemicals) does not always covary with a measured chemical (e.g., if a certain industry releases an unusual mixture of contaminants), the effect should be mitigated if a sufficiently large and diverse data set is used to establish AET. Use of a large data set comprising samples from a variety of areas with wide-ranging chemical concentrations would decrease the likelihood that an unrealistically low AET would be set. Because AET are set by the highest concentration of a given chemical in samples without observed biological effects, AET will not be affected by less contaminated samples in which unmeasured contaminants cause biological effects.

If an unmeasured toxic chemical does not covary with any of the measured chemicals, it is unlikely that the AET (or any other chemical-specific approach) could predict impacts at stations where the chemical is inducing toxic effects. The frequency of occurrence of stations with biological effects but no chemicals exceeding AET is the subject of validation tests (see Section 3 of this briefing document).

Matrix Effects and Bioavailability--Geochemical associations of contaminants with sediments that reduce bioavailability of those contaminants would affect AET analogously to antagonistic effects (i.e., they would increase AET relative to sediments in which this factor was not operative). Sediment matrices observed in Puget Sound that may reduce bioavailability of certain contaminants include slag material (containing high concentrations of various metals and metalloids, such as copper and arsenic) and coal or soot (which may contain high concentrations of largely unavailable PAH, as opposed to oil or creosote, in which PAH would be expected to be far more bioavailable; e.g., Farrington and Teal 1982). Many kinds of matrices may occur in the environment and a large proportion may be difficult to classify based upon appearance or routinely measured sediment variables. Hence, the use of matrix-specific data sets to generate AET, although desirable, would be difficult to implement.

To address this concern from a technical perspective (i.e., representativeness of data used in AET generation), the AET database could be screened for sediment with chemical concentrations that are anomalously high relative to those in other nonimpacted sediments from different geographic areas. From a management perspective, this guideline would generate more protective (sensitive) sediment quality standards that may also be less efficient in only identifying problem sediments. These sediments would be considered nonrepresentative and not used in AET generation unless and until additional data could substantiate that they are representative. Such data treatment methods are discussed in the following section.

3. GENERATION OF AET VALUES FOR PUGET SOUND

The AET approach was developed for projects conducted in the Puget Sound area and has thus far incorporated only Puget Sound data. An extensive database of biological and chemical data from numerous Puget Sound projects has been compiled for AET generation and validation tests (over 300 samples, four biological effects indicators, and over 125 chemicals although some chemicals were detected infrequently). The generation of AET values for Puget Sound has necessarily made use of available data, although the AET concept is not intrinsically limited to the biological effects indicators or chemicals used in Puget Sound. A brief history of the generation of AET values for Puget Sound projects is included below, followed by a description of the chemical and biological data in the existing Puget Sound database.

PUGET SOUND DATABASE

AET values were originally generated for a combined measure of sediment toxicity [i.e., either amphipod mortality (Swartz et al. 1985) or oyster larvae abnormality (Chapman and Morgan 1983)] and depressions in the abundance of benthic infauna (at phylum or class levels of taxonomic classification). These AET values were based on data from 50 to 60 stations sampled during the 1984-85 remedial investigation of Commencement Bay (a heavily industrialized embayment adjacent to Tacoma, Washington). In a 1986 project for the Puget Sound Dredged Disposal Analysis (PSDDA) and Puget Sound Estuary Program (PSEP), AET values were generated with a larger Puget Sound database consisting of 188 samples and including the previous Commencement Bay data. Biological indicators included individual measures of toxicity [i.e., amphipod mortality, oyster larvae abnormality, and Microtox bioassays (Williams et al. 1986)] and benthic infaunal depressions (at phylum or class levels). Matched biological and chemical data for 10 additional stations from a joint state and federal investigation of creosote contamination in Eagle Harbor (Barrick et al. 1986) have also been incorporated. Most recently, biological and chemical data for over 100 stations from Elliott Bay and Everett Harbor (Beller et al. 1988a; Pastorok et al. 1988) have been added to the database for generation of AET values. Elliott Bay is a highly urbanized embayment adjacent to Seattle, Washington. Everett Harbor is an urban embayment adjacent to Everett, Washington and is characterized by pulp industry activities.

Based on recent studies conducted for PSEP (see Barrick et al. 1988), the following strategy was recommended for developing reliable AET:

1. Collect chemical and biological effects data from preferably 50 stations or more

2. Bias the positioning of stations to ensure sampling of a variety of contaminant sources (e.g., an urban environment impacted by multiple contaminant sources and preferably multiple geographic areas at which AET predictions will be applied) over a range of contaminant concentrations (preferably over at least one to two orders of magnitude)

3. Conduct chemical tests for a wide range of chemical classes and ensure that <100 ppb detection limits (lower if possible) are attained for organic compounds (metals detection limits do not appear to be a problem).

The data included in the existing 334-sample Puget Sound data set are characterized in Table 1 and conform to these recommendations. The geographic distribution of samples in this data set is presented in Figure 6. A summary of concentration ranges of selected chemicals and key sediment characteristics [i.e., total organic carbon (TOC), fine-grained sediment content, sulfides] is provided in Table 2 for two groups of these samples: sediments sampled from reference areas (nonurban embayments) and from nonreference areas of Puget Sound (frequently from industrialized, urban embayments). Reference areas in Puget Sound are generally removed from the direct influence of contaminant sources in nonurbanized areas of the sound; almost all of these stations are nonimpacted. Most data for nonreference areas are from industrialized urban embayments and include both nonimpacted and impacted stations.

Biological Data

Two major kinds of biological tests are commonly used for environmental assessment of chemical contamination: sediment bioassays and evaluations of indigenous biota. Sediment bioassays involve the controlled exposure of test organisms (usually a single sensitive species) to test sediment for a fixed period of time. Although bioassays can be conducted *in situ*, most are conducted in the laboratory. Bioassays have at least two major advantages over evaluations of indigenous biota. First, because most experimental conditions can be controlled during bioassays (e.g., temperature, dissolved oxygen, lighting, sediment grain size, predation), measured effects can be attributed to chemical toxicity (i.e., the uncontrolled variable of interest) with reasonable confidence. Second, bioassays generally are considerably less expensive to conduct than evaluations of indigenous biota. A major disadvantage of most sediment bioassays is the lack of knowledge as to how the results correspond to potential impacts on diverse species under variable field conditions (Long and Chapman 1985; Swartz et al. 1985; Chapman et al. 1987). Part of this uncertainty can be evaluated by comparing bioassay responses with effects on indigenous biota.

Evaluations of sediment toxicity to indigenous biota can involve any kind of organism but usually are focused on benthic macroinvertebrates. These organisms are preferred because they live in close contact with bottom sediments, are relatively stationary, can be sampled quantitatively, and have been found to exhibit predictable patterns in response to environmental stress. Unlike bioassays, many environmental variables cannot be controlled during evaluations of indigenous biota. The relationship between measured effects on indigenous biota is therefore less certain than for bioassays. However, because effects on indigenous biota are measured in the field, there are no limitations encountered with extrapolating laboratory results to field situations. Specific analytical and statistical procedures for each of the biological tests used to develop AET are presented in the following sections.

Bioassay Tests--Currently, AET have been generated for three kinds of bioassays used in Puget Sound as acute lethal or sublethal indicators of sediment toxicity. The AET concept can be applied to other bioassays as they are developed.

TABLE 1. SUMMARY OF DATA SETS USED TO EVALUATE PUGET SOUND AET

Embayment	Survey Code[a]	Number/Kind of Bioeffect Samples[b]	Chemical Analyses Conducted[c]							
			Acid	Base	Neut.	PCB	Pest.	VOA	Metal	Misc
Bellingham Bay	EIGHTBAY	8 / A	x	x	x	x	x	x	x	
Carr Inlet	CBMSQS	4 / BAOM	x	x	x	x	x	x	x	x
Case Inlet	EIGHTBAY	4 / A	x	x	x	x	x	x	x	
Central Puget Sound Basin	ALKI	4 / B	x	x	x	x	x	x	x	
	EHCHEM	2 / BA	x	x	x	x	x	x	x	x
Commencement Bay	CBBLAIR	6 / BAO	x	x	x	x	x	x	x	x
	CBMSQS	42 / BAOM	x	x	x	x	x	x	x	x
		2 / AOM								
		2 / B OM								
Dabob Bay	EIGHTBAY	4 / A	x	x	x	x	x	x	x	
Eagle Harbor	EHCHEM	8 / BA	x	x	x	x	x	x	x	x
Elliott Bay	EBCHEM	71 / BA	x	x	x	x	x	x	x	x
		24 / A								
		4 / B								
	ALKI	7 / B	x	x	x	x	x	x	x	
	TPPS3AB	27 / B	x	x	x	x	x	x	x	
	DUWRIV1	8 / A			x	x	x	x	x	
	DUWRIV2	30 / A			x	x	x	x	x	
	EIGHTBAY	8 / A	x	x	x	x	x	x	x	
Everett Harbor	EVCHEM	13 / BA	x	x	x	x	x	x	x	x
		13 / A								
		3 / B								
	EVERETT1	6 / A		x	x		x	x		
	EIGHTBAY	8 / A	x	x	x	x	x	x	x	
Port Susan	EBCHEM	5 / BA	x	x	x	x	x	x	x	x
	EVCHEM	3 / BA	x	x	x	x	x	x	x	x
Samish Bay	EIGHTBAY	4 / A	x	x	x	x	x	x	x	
Sequim Bay	EIGHTBAY	4 / A	x	x	x	x	x	x	x	
	DUWRIV1	1 / A			x	x	x	x	x	
	DUWRIV2	1 / A			x	x	x	x	x	
Sinclair Inlet	EIGHTBAY	8 / A	x	x	x	x	x	x	x	

TABLE 1. **(Continued)**

[a] Puget Sound samples are derived from multiple surveys, which provided data for varying numbers of chemicals and biological indicators. The surveys include:

ALKI	Metro survey of Alki Point, Seattle (Osborn et al. 1985; Trial and Michaud 1985)
CBBLAIR	Port of Tacoma dredging survey of Blair Waterway in Commencement Bay (data analyzed integrated with CBMSQS data during the Superfund project; see Barrick et al. 1985)
CBMSQS	Commencement Bay Nearshore/Tideflats Superfund project; Carr Inlet reference area (Barrick et al. 1985)
DUWRIV1	PSDDA dredging study in the Duwamish River, Seattle (Phase I); Sequim Bay reference area (Chan et al. 1985)
DUWRIV2	PSDDA dredging study in the Duwamish River, Seattle (Phase II); Sequim Bay reference area (Chan et al. 1986)
EBCHEM	PSEP survey of Elliott Bay; Port Susan reference area (Beller et al. 1988a)
EHCHEM	Ecology Preliminary Investigation of Eagle Harbor; Blakely Harbor reference area in Central Puget Sound (Barrick et al. 1986)
EIGHTBAY	EPA survey of eight urban and nonurban embayments in Puget Sound (Battelle 1985)
EVCHEM	PSEP survey of Everett Harbor; Port Susan reference area (Pastorok et al. 1988)
EVERETT1	U.S. Navy preliminary dredging study in Everett Harbor (U.S. Navy 1985)
TPPS3AB	Toxic Pretreatment Planning Study conducted in Central Puget Sound and Elliott Bay by Metro (Romberg et al. 1984).

Station locations for each survey are summarized in Appendix B of Barrick et al. (1988).

[b] 334 distinct samples (including 12 repeated samplings) at a total of 322 locations:
(B) 201 benthic infaunal analyses; (A) 287 amphipod mortality bioassays; (O) 56 oyster larvae abnormality bioassays; (M) 50 Microtox (saline extract) bioassays. The seven amphipod bioassay stations excluded as biological anomalies and the three benthic infauna and eight amphipod bioassay stations excluded as chemical anomolies (see text) are <u>not</u> included in these totals.

[c] Chemical analyses conducted for U.S. EPA priority pollutant acid, base, neutral, PCB, pesticide, and volatile organic compounds, metals, and miscellaneous compounds not recognized as EPA priority pollutants (e.g., resin acid compound data for the EVCHEM survey, and tentatively identified organic compounds).

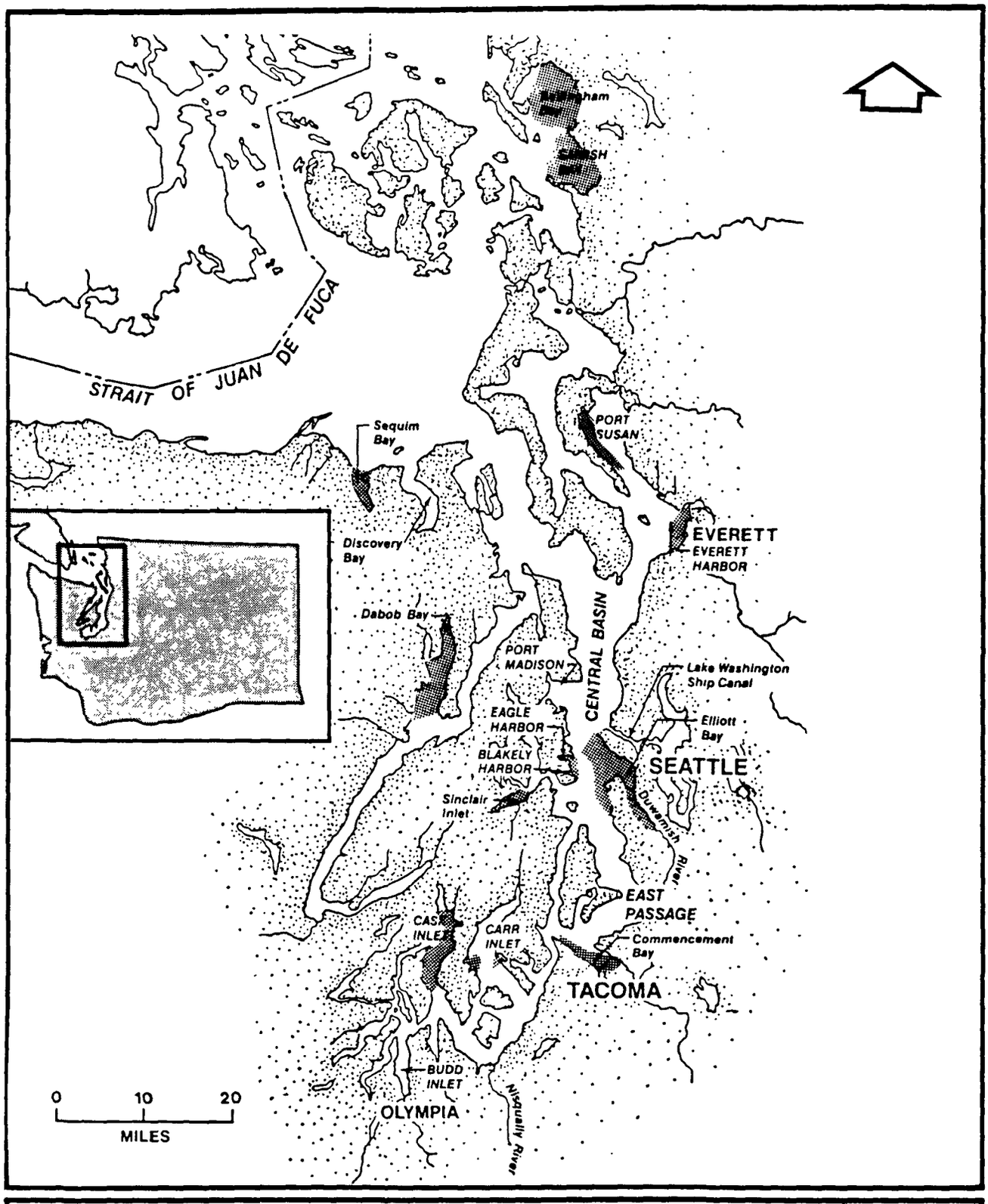


Figure 6. Location of sampling sites for Apparent Effects Threshold data sets in Puget Sound.

TABLE 2. SUMMARY OF SELECTED CHEMICAL CONCENTRATIONS AND KEY SEDIMENT CHARACTERISTICS FOR BIOLOGICAL STATIONS FROM REFERENCE AND NONREFERENCE AREAS OF PUGET SOUND[a]

Chemical	Reference Area[b]		Non-Reference Area[c]	
	Minimum	Maximum	Minimum	Maximum
Metals (mg/kg dry weight; ppm)				
Antimony	0.13	2.86	0.1	1,370
Arsenic	1.9	22	0.34	9,700
Cadmium	0.047	1.9	0.3	184
Chromium	9.6	255	5.4	555
Copper	4.9	74	3.6	11,400
Lead	0.4	23	0.79	71,100
Mercury	0.016	0.134	0.06	52
Nickel	11	141	6.9	118
Silver	0.02	0.78	0.013	8.27
Zinc	15	102	18.4	6,010
Organic Compounds (ug/kg dry weight; ppb)				
Low molecular weight PAH	2.5	55	1.3	630,000
High molecular weight PAH	22	140	7.5	3,200,000
1,4-Dichlorobenzene	U5	U160	1	31,000
Hexachlorobenzene (HCB)	U0.3	U800	0.05	730
Hexachlorobutadiene	U1.3	U820	1	730
N-Nitrosodiphenylamine	U5	U2,500	4	610
4-Methylphenol	2	290	3	100,000
Phenol	13	560	0.9	2,900
Total PCBs	2.7	37	0.5	9,600
Conventional Sediment Variables (dry weight)				
Total organic carbon (%)	0.19	2.69	0.06	29.4
Fine-grained material (%)	7.4	88	2.7	95.5
Sulfides (mg/kg; ppm)	2.2	31	1	7,610

[a] Ranges are based only on biological stations used in the evaluation of Puget Sound AET. Higher concentrations for some chemicals have been detected in sediment samples for which there are no biological data, including several samples collected at depth. "U" indicates undetected in all samples over the detection limit range shown; otherwise the minimum value is the lowest detected concentration and the maximum value is the highest detected concentration (detection frequencies are not indicated but are less that 100 percent for most chemicals except metals).

[b] Reference areas in Puget Sound are generally removed from the direct influence of contaminant sources in nonurbanized areas of the sound; almost all of these stations are nonimpacted.

[c] Most data for nonreference areas are from industrialized urban embayments; nonreference data include both nonimpacted and impacted stations.

The amphipod mortality bioassay (Swartz et al. 1985) is an indicator of acute lethal toxicity in whole sediments. Significant mortalities of the adult amphipod *Rhepoxynius abronius* were determined by statistically comparing results of tests on sediments from potentially impacted sites with those from a reference area. All comparisons were made within studies (i.e., data were not compared among studies).

The oyster larvae abnormality bioassay (modified for sediments; Chapman and Morgan 1983) is an indicator of acute sublethal toxicity in sediment elutriates. Significant abnormalities in the larvae of the Pacific oyster *Crassostrea gigas* were determined by statistically comparing results of tests on sediment elutriate samples from potentially impacted sites with those from a reference area. Oyster larvae bioassay data were available for 56 stations from Commencement Bay and Carr Inlet in Puget Sound.

Change in bacterial luminescence is an indicator of acute sublethal effects in sediment elutriates (Williams et al. 1986), although the response may also reflect acute lethal effects. The Microtox bioassay (Beckman Instruments 1982; Bulich et al. 1981) using *Photobacterium phosphoreum* was applied in these tests. Significant Microtox toxicity for samples from Commencement Bay was assessed by statistically comparing the predicted decrease in luminescence in the presence of a 15-g sediment sample to that observed for a 15-g sample from a reference area (Carr Inlet).

The following statistical procedures were applied to test results from the amphipod bioassay:

- 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
- An 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, a t-test was used to compare the two means
- If variances were not homogenous, an approximate t-test (Sokal and Rohlf 1969) was used to compare means
- Statistical significance was tested with a pairwise error rate of 0.05 to ensure consistency among studies of differing sample sizes.

Oyster larvae bioassay results were treated similarly to the amphipod test results, except that a t-test was always used to compare mean results. The following procedure was used for Microtox results (Williams et al. 1986):

- For each sample, decrease in luminescence for a 15-g sample was predicted with a least-squares regression of the percent decrease in luminescence versus the logarithm of the standardized sample dilution,

where five serial dilutions of supernatants from samples of 13.0-26.4 g were used as values for the independent variable

- Statistical significance of the difference between the predicted luminescent response and the response of control sediment was determined using a t-test with a comparison-wise error rate of 0.001, which yielded an experiment-wise error rate of 0.05 (Zar 1974) for the single Commencement Bay study that was available.

Benthic Infauna Analyses--AET were also developed using field data on benthic infauna abundances. Depressions in the abundances of indigenous benthic infauna, unlike laboratory bioassays, are *in situ* measurements of chronic and/or acute effects in sediments.

Significant depressions of the abundances of polychaetes, molluscs, crustaceans, and total benthic infauna were determined separately by comparing values from potentially impacted sites with those from reference areas. Comparisons were made within respective studies unless appropriate reference data were not available for a particular study. In those cases, comparisons were made among studies. 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 um) as those of the potentially impacted site. In this manner, comparisons were stratified by three of the major natural variables known to influence the abundance and distribution of benthic macroinvertebrates.

The AET approach does not require the analysis of benthic infaunal data at any particular taxonomic level, but the data discussed in this report apply to higher level taxa only. As discussed by Beller et al. (1986; Appendix H), higher level taxa were used initially in the development of AET values for two major reasons. First, because the AET approach is based on pair-wise statistical comparisons with reference conditions, the benthic taxa must either be abundant enough or have a low enough variance to allow major depressions in abundance to be discriminated statistically. These criteria were best met by use of higher level taxa abundances. Second, comparisons with bioassay results (i.e., amphipod mortality and oyster larvae abnormality) suggested that benthic comparisons based on higher taxa were as sensitive as the bioassays in identifying problem sediments, although different species may differ widely in their sensitivity to individual chemicals present as a complex mixture in contaminated sediments.

Use of higher level taxa does not allow compensatory shifts in species abundances to be evaluated explicitly within each higher level taxon, but does allow an evaluation of the net effects of such shifts on the total abundance of each higher level taxon. Use of higher level taxa also does not allow impacts on pollution-sensitive indicator species to be evaluated explicitly. Such species are frequently characterized by relatively low abundances and high variability and, therefore, are not always amenable to the determination of statistically significant reductions in abundance even when they are considered explicitly.

In a limited test conducted using species-level data from Commencement Bay in Puget Sound (Beller et al. 1988), 10 species or genera were sufficiently abundant to enable statistical analysis of the benthic data. These species-level benthic AET were

similar in magnitude to higher taxa benthic AET (generally within a factor of two) although they were based on considerably fewer data.

Stations tested for benthic infaunal abundances were evaluated for statistically significant benthic depressions as follows:

- All abundances were $\log_{10}$-transformed
- All replicates from each set of reference conditions were pooled, and a mean and standard deviation were calculated for each of the four benthic groups (i.e., total benthos, Polychaeta, Mollusca, and Crustacea)
- The $\log_{10}$-transformed mean abundance of each benthic group at each potentially impacted site was then treated as described for the amphipod bioassay results (i.e., pairwise comparisons by the F_{max}-test followed by the t-test or approximate t-test).

Chemical Data

Chemical measurements of sediments are particularly useful for identifying specific problem chemicals, for comparing historical sediment concentrations with existing sediment quality values in the absence of biological data or for developing new sediment quality values, and for identifying the potential sources of contaminating chemicals. The chemical data available in Puget Sound for different studies have not all been generated using the same analytical method or laboratory, but a general quality assurance review has been conducted to assess their applicability for development of sediment quality values. This general review is described by Beller et al. (1986) and was carried out in three basic steps:

- All available data sets were reviewed for synoptic collection of data (i.e., matched chemical and biological measurements at each station), and only matched chemical and biological data sets were considered further. [Note: a matched data set was defined as one for which toxicity data were collected on the same sediment homogenate used for sediment chemistry, and replicate benthic infaunal samples were collected at the identical station location and time, or at nearly the same time, as sediment chemistry samples.]
- Each data set was reviewed for documentation of quality assurance methods and summaries of quality assurance review (such documentation was typically provided in the reports in which the data were presented)
- Data were subjected to a more detailed review that focused on issues related to data comparability.

Overall, matched data were considered to provide the most reliable basis for deriving or validating sediment quality values with site-specific biological field data. Matched data sets were 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) were assumed to be affected by direct association with contaminants in the surrounding environment, it was considered essential that chemical and biological data be collected from nearly identical subsamples from a given station.

In the quality assurance review of chemical data, analytical techniques, detection limits, and the chemical scope of contaminants analyzed (e.g., ionizable and nonionic semivolatile organic compounds, metals, volatile organic compounds) were assessed and summarized. 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.

AET were developed for over 60 chemicals frequently detected in the environment, including 16 PAH; several alkylated PAH 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; and 11 metals and metalloids. Data for other miscellaneous chemicals that were less frequently detected or analyzed for were also evaluated for their potential use in developing AET (e.g., resin acids and chlorinated phenols that have recently been measured in selected sediments from Everett Harbor, an area influenced by pulp and paper mill activity).

AET were 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, AET were also developed for data normalized to fine-grained particle content (expressed as the percent of silt and clay, or <64-um 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) and will not be discussed here.

Guidelines for Data Treatment

Before generating AET values, options were developed to address the following factors that can affect AET uncertainty (in addition to factors discussed previously, such as interactive effects and unmeasured chemicals):

- Low statistical power and Type I statistical error in biological tests
- Distributions of chemical concentrations (i.e., ranges and continuity of concentrations)
- Chemical analysis variability
- Anomalous stations with relatively high chemical concentrations but without statistically significant biological effects (possibly relating to bioavailability and matrix effects).

Discussion of these options for biological or chemical analyses is presented in the following sections.

Bioassay Data--Several modifications were adopted to improve the consistency of the results among the various studies pooled for amphipod bioassay results. Similar modifications were not made for the oyster larvae abnormality bioassay and the Microtox bioassay because each of these indicators is represented by only a single study area in the database used to generate AET.

The first modification addressed the level of significance at which pairwise comparisons between impacted and reference sites were judged significant. A level of $P \leq 0.05$ pairwise was consistently used for all comparisons between stations, instead of using an experiment-wise error rate that would result in variable *alpha* levels for pairwise station comparisons depending on the sample size of different studies [see Appendix C of Barrick et al. (1988)].

The second modification considered screening criteria and power analyses for the amphipod tests conducted in the Elliott Bay and Everett Harbor surveys conducted by PSEP [discussed in detail in Appendix C of Barrick et al. (1988)]. This modification resulted in the exclusion of seven potentially nonimpacted stations for which there was inadequate power to distinguish significant effects relative to reference conditions. The excluded stations were Stations DR-05, DR-08, EW-03, and NS-04 in Elliott Bay, and Stations EW-14, SD-02, and SR-07 in Everett Harbor. Impacted stations were defined as those that exhibited statistically significant mortality ($P \leq 0.05$) relative to reference conditions **and** exceeded 25 percent mortality (used for this study as a minimum level of concern).

Benthic Infauna Analyses--Power analyses have not been conducted for benthic infaunal data used to develop AET in Puget Sound. In lieu of a power analysis, and analogous to the amphipod bioassay guideline, a guideline was developed to ensure that benthic effects were of sufficient magnitude to be of regulatory concern as adverse impacts and to be discriminated statistically in most cases. Thus, only significant effects ($P \leq 0.05$) that also exceeded a 50-percent reduction in major taxa abundance were considered impacts. This guideline was derived partly from consideration of the natural variability of benthic infauna in relatively undisturbed environments of Puget Sound. Based on a summary of data from Lie (1968), Nichols (1975), and Word et al. (1984) in Tetra Tech (1987), the abundances of selected major taxa (Polychaeta, Mollusca, Crustacea) and total infauna may vary seasonally by roughly a factor of two (i.e., lowest mean abundances are roughly 50 percent of the highest mean abundances). In most cases, ≥ 50 percent reductions in mean abundance can be detected statistically ($P<0.05$), whereas ≤ 30 percent reductions cannot be detected ($P>0.05$). Finally, the guideline of 50 percent reduction in benthic infauna abundances provided a level of environmental protectiveness with a reasonable balance between underprotection due to tolerance of major effects and overprotectiveness due to misclassification of nonimpacted sites as impacted.

Chemical Data--All detected chemical data entered in the sediment quality database after quality assurance review were included in AET calculations. These calculations were based on biological effect stations that passed the biological screening criteria summarized in the previous section.

Anomalous stations with relatively high chemical concentrations but without statistically significant biological effects were considered to indicate the possibility of unusual bioavailability or matrix effects. AET values based on such anomalous stations

are not necessarily unreliable, but may be unrepresentative of Puget Sound conditions. The inclusion of these anomalous stations in the generation of AET may increase AET values solely in response to a localized condition. Alternatively, such stations may simply indicate the need for additional data to confirm an increase in AET values that is representative of regional environmental conditions.

In recent updates of AET, an option was implemented to exclude anomalous chemical data from AET calculations until they could be confirmed (i.e., by another nonimpacted station with chemical concentrations within a factor of 3 of the currently anomalous station) [for more detailed discussion see Appendix C of Barrick et al. (1988)]. The purpose of this option is to reduce the possibility that Puget Sound AET might be set based on nonrepresentative data for exceptional chemical matrices (e.g., slag, coal) or unusual biological conditions (e.g., extremely tolerant species under localized conditions).

For the amphipod bioassay, this procedure affected 8 of the 296 amphipod stations that and resulted in changes to the AET for nine chemicals where the ratio of the anomalous station to the nonimpacted station with the next highest concentration ranged from 3.2 to 14. For benthic infauna, this procedure affected 4 of 206 benthic stations and resulted in changes to the AET for eight chemicals [including high molecular weight PAH (HPAH) as a class) where the ratio of the anomalous station to the nonimpacted station with the next highest concentration ranged from 3.0 to 20. By implementing this option, the sensitivity of the amphipod bioassay and benthic infauna AET each increased by 11 percent, yielding values of 58 and 75 percent, respectively. Data currently identified as anomalous will be reevaluated if and when confirming data become available.

VALIDATION TEST METHODS

The reliability of AET generated from Puget Sound data was evaluated with tests of sensitivity and efficiency (defined in Section 2). 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 of each group were classified as impacted or nonimpacted based on the appropriate statistical criteria (i.e., F_{max}-tests and t-tests at alpha = 0.05).**

- **Several tests of reliability were conducted at this point:**

Test 1 - AET (dry weight) were generated with the entire Puget Sound database available in 1988 (see Table 1 for the distribution of samples for individual indicators) and were then tested against the same database for

each biological indicator. Stations predicted to have a particular biological effect (i.e., stations with one or more chemicals exceeding AET for the biological indicator) were compared to those stations known to have the predicted biological effect. This test provides a useful assessment of sensitivity but results in 100 percent efficiency by definition (i.e., AET are defined such that all stations with chemical concentrations above AET are biologically impacted). This test directly assesses, for data available in 1988, the number of biologically impacted stations that the AET approach could not account for with any chemicals (e.g., as a result of interactive effects or unmeasured chemicals; see Section 2, Interpretation of AET).

Test 2 - The test described above was repeated in two parts: (a) using TOC-normalized AET for nonionic organic compounds and dry weight-normalized AET 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, AET were recalculated for the remaining data set, and biological effects were predicted for the single deleted station. The predictive efficiency was then 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 AET to confirm the sensitivity and efficiency measurements in Tests 1 and 3. AET (dry weight) generated in 1986 with 188 stations from diverse geographic regions in Puget Sound were later tested with an independent set of 146 Puget Sound stations (from Blakely Harbor, Eagle Harbor, Elliott Bay, Everett Harbor, and Port Susan). These 1986 AET are distinct from the 1988 AET discussed in Test 1. Results were also compiled for comparisons of the 1986 AET to the entire database (i.e., including stations used to calculate these AET; hence, this latter test is not completely independent but provides a comparison for the identical total number of stations used in Test 1).

Details of these validation tests and other tests that have been conducted are reported by Barrick et al. (1988) and Beller et al. (1986).

VALIDATION RESULTS AND DISCUSSION

Selected AET generated from the 334-station database are presented in Table 3 (dry weight-normalized) and Table 4 (organic carbon-normalized). The chemicals presented in Tables 3 and 4 are among the most commonly measured and detected chemicals in Puget Sound sediments. A complete list of 1988 AET used in validation tests is provided in Appendix F in Barrick et al. (1988).

TABLE 3. 1988 PUGET SOUND AET FOR SELECTED CHEMICALS (normalized to dry weight)[a]

Chemical	Amphipod AET[b]	Oyster AET[c]	Benthic AET[d]	Microtox AET[e]
Metals (mg/kg dry weight; ppm)				
Antimony	200[g]	--	150[g]	--
Arsenic	93	700	57[h]	700
Cadmium	6.7	9.6	5.1[h]	9.6
Chromium	270[g]	--	260[g]	--
Copper	1300[fg]	390	530[g]	390
Lead	660	660	450[g]	530
Mercury	2.1[f]	0.59	2.1[g]	0.41
Nickel	>140[fg]	--	>140[g]	--
Silver	6.1[fg]	>0.56	>6.1[g]	>0.56
Zinc	960[fg]	1,600	410[g]	1,600
Organic Compounds (ug/kg dry weight; ppb)				
Low molecular weight PAH	24,000[g]	5,200	13,000[fg]	5,200
Naphthalene	2,400[g]	2,100	2,700[g]	2,100
Acenaphthylene	1,300[g]	>560	1,300[fg]	>560
Acenaphthene	2,000[fg]	500	730[g]	500
Fluorene	3,600[g]	540	1,000[fg]	540
Phenanthrene	6,900[fg]	1,500	5,400[fg]	1,500
Anthracene	13,000[fg]	960	4,400[fg]	960
2-Methylnaphthalene	1,900[g]	670	1,400[g]	670
High molecular weight PAH	69,000[fg]	17,000	69,000[fg]	12,000
Fluoranthene	30,000[fg]	2,500	24,000[fg]	1,700
Pyrene	16,000[fg]	3,300	16,000[fg]	2,600
Benz(a)anthracene	5,100[fg]	1,600	5,100[fg]	1,300
Chrysene	9,200[fg]	2,800	9,200[fg]	1,400
Benzofluoranthenes	7,800[g]	3,600	9,900[fg]	3,200
Benzo(a)pyrene	3,000[g]	1,600	3,600[fh]	1,600
Indeno(1,2,3-c,d)pyrene	1,800[fg]	690	2,600[fh]	600
Dibenzo(a,h)anthracene	540[fg]	230	970[fh]	230
Benzo(g,h,i)perylene	1,400[fg]	720	2,600[fh]	670
Chlorinated benzenes				
1,3-Dichlorobenzene	>170	>170	>170	>170
1,4-Dichlorobenzene	120[h]	120	110[h]	110
1,2-Dichlorobenzene	>110[h]	50	50	35
1,2,4-Trichlorobenzene	51	64	--	31
Hexachlorobenzene (HCB)	130	230	22[h]	70
Total PCBs	3,100[fg]	1,100	1,000[h]	130

TABLE 3. (Continued)

Chemical	Amphipod AET[b]	Oyster AET[c]	Benthic AET[d]	Microtox AET[e]
Phthalates				
Dimethyl phthalate	>1,400[g]	160	>1,400[g]	71
Diethyl phthalate	>1,200[g]	>73	200[g]	>48
Di-n-butyl phthalate	1,400[h]	1,400	>5,100	1,400
Butyl benzyl phthalate	900[g]	>470	900[g]	63
Bis(2-ethylhexyl)phthalate	>3,100	1,900	1,300[h]	1,900
Di-n-octyl phthalate	>2,100[g]	>420	6,200[h]	--
Phenols				
Phenol	1,200[fg]	420	1,200	1,200
2-Methylphenol	63	63	72[g]	>72
4-Methylphenol	3,600[g]	670	1,800[g]	670
2,4-Dimethyl phenol	72[g]	29	210[g]	29
Pentachlorophenol	360[g]	>140	690[g]	>140
Miscellaneous Extractables				
Benzyl alcohol	870[g]	73	870[g]	57
Benzoic acid	760[g]	650	650	650
Dibenzofuran	1,700[g]	540	700[g]	540
Hexachlorobutadiene	180[h]	270	11[h]	120
N-Nitrosodiphenylamine	48[h]	130	28[h]	40
Volatile Organics				
Tetrachloroethene	>210	140	57[h]	140
Ethylbenzene	>50	37	10[h]	33
Total xylenes	>160	120	40[h]	100
Pesticides				
p,p'-DDE	15	--	9	--
p,p'-DDD	43	--	16[g]	--
p,p'-DDT	>270[g]	>6	34[g]	--

[a] ">" indicates that a defined AET could not be established because there were no "effects" stations with chemical concentrations above the highest concentration among "no effects" stations. "--" indicates AET data not available.

[b] Based on 287 stations (including recent surveys in Eagle Harbor, Elliott Bay, and Everett Harbor not included in the previous generation of 1986 AET).

[c] Based on 56 stations (all from Commencement Bay Remedial Investigation and Blair Waterway dredging study); no additional stations added since 1986.

TABLE 3. (Continued)

[d] Based on 201 stations (including recent surveys in Eagle Harbor, Elliott Bay, and Everett Harbor not included in the previous generation of 1986 AET).

[e] Based on 50 stations (all from Commencement Bay Remedial Investigation); no additional stations added since 1986.

[f] The value shown exceeds AET established from Commencement Bay Remedial Investigation data (Barrick et al. 1985) because of addition of Puget Sound data presented in Beller et al. (1986).

[g] The value shown exceeds AET presented in Beller et al. (1986) because of addition of Puget Sound data from the Eagle Harbor, Elliott Bay, or Everett Harbor surveys.

[h] The value shown is less than AET presented in Beller et al. (1986) because of the exclusion of chemically anomalous stations from the AET dataset and the application of power analyses to assess significantly impacted stations ($P<.05$) (see text and Barrick et al. 1988).

TABLE 4. 1988 PUGET SOUND AET FOR SELECTED CHEMICALS (normalized to total organic carbon)[a]

Chemical	Amphipod AET[b]	Oyster AET[c]	Benthic AET[d]	Microtox AET[e]
Nonionic Organic Compounds (mg/kg organic carbon; ppm)				
Low molecular weight PAH	2,200	370	780	>530
Naphthalene	220	99	170	>170
Acenaphthylene	66	>27	66	>27
Acenaphthene	200	16	57	>57
Fluorene	360	23	79	>71
Phenanthrene	690	120	480	>160
Anthracene	1,200	>79	220	>79
2-Methylnaphthalene	>120	---	64	---
High molecular weight PAH	5,300	960	7,600	1,500
Fluoranthene	3,000	160	1,200	>190
Pyrene	1,000	>210	1,400	>210
Benz(a)anthracene	270	110	650	>160
Chrysene	460	110	850	>200
Benzofluoranthenes	450	230	1,500	>430
Benzo(a)pyrene	210	99	>1,300	>140
Indeno(1,2,3-c,d)pyrene	88	33	900	>87
Dibenzo(a,h)anthracene	47	120	89	33
Benzo(g,h,i)perylene	78	31	>1,200	>67
Chlorinated benzenes				
1,3-Dichlorobenzene	>15	>15	>15	>15
1,4-Dichlorobenzene	9	3.1	16	>16
1,2-Dichlorobenzene	>5.8	2.3	2.3	2.3
1,2,4-Trichlorobenzene	1.8	2.7	---	0.81
Hexachlorobenzene (HCB)	4.5	9.6	0.38	2.3
Total PCBs	190	>46	65	12
Phthalates				
Dimethyl phthalate	53	>22	53	>19
Diethyl phthalate	>110	>5.3	61	>5.3
Di-n-butyl phthalate	260	260	1,700	220
Butyl benzyl phthalate	42	>9.2	64	4.9
Bis(2-ethylhexyl)phthalate	78	60	60	47
Di-n-octyl phthalate	58	>57	4,500	--

TABLE 4. (Continued)

Chemical	Amphipod AET[b]	Oyster AET[c]	Benthic AET[d]	Microtox AET[e]
Miscellaneous Extractables				
Dibenzofuran	>170	15	58	>58
Hexachlorobutadiene	6.2	11	6.9	3.9
N-nitrosodiphenylamine	>11	>11	11	>11
Volatile Organics				
Tetrachloroethene	>22	>22	>22	>22
Ethylbenzene	>3.8	>3.8	>3.8	>3.8
Total xylenes	>12	>12	>12	>12
Pesticides				
p,p'-DDE	0.81	--	0.31	--
p,p'-DDD	2.2	--	1.0	--
p,p'-DDT	>16	--	3.7	--
Ionizable Organic Compounds (mg/kg organic carbon; ppm)				
Phenols and Miscellaneous Extractables				
Phenol	440	>39	>140	33
2-Methylphenol	3.1	3.1	10	>10
4-Methylphenol	780	37	250	81
2,4-Dimethyl phenol	6.5	>1.3	2.6	0.63
Pentachlorophenol	24	>11	66	>11
Benzyl alcohol	73	5.0	>73	5.0
Benzoic acid	>170	>170	>170	>170
Metals (mg/kg organic carbon; ppm)				
Antimony	>55,000	3,300	5,500	3,300
Arsenic	32,000	88,000	4,400	88,000
Cadmium	1,100	1,200	580	1,200
Chromium	>150,000	---	65,000	---
Copper	100,000	49,000	13,000	48,000
Lead	110,000	66,000	18,000	66,000
Mercury	210	210	120	77
Nickel	>41,000	---	31,000	---
Silver	170	>100	490	100
Zinc	210,000	>200,000	48,000	>200,000

TABLE 4. (Continued)

[a] ">" indicates that a defined AET could not be established because there were no "effects" stations with chemical concentrations above the highest concentration among "no effects" stations (normalized to TOC). "--" indicates AET data not available.

[b] Based on 287 stations (including recent surveys in Eagle Harbor, Elliott Bay, and Everett Harbor not included in the previous generation of 1986 AET).

[c] Based on 56 stations (all from Commencement Bay Remedial Investigation and Blair Waterway dredging study); no additional stations added since 1986.

[d] Based on 201 stations (including recent surveys in Eagle Harbor, Elliott Bay, and Everett Harbor not included in the previous generation of 1986 AET).

[e] Based on 50 stations (all from Commencement Bay Remedial Investigation); no additional stations added since 1986.

Results of validation tests 1 through 4 are summarized in Table 5. Using dry weight-normalized AET calculated from all available data (Test 1), sensitivity ranged from 58 to 93 percent, depending on the biological indicator, and efficiency was 100 percent (100 percent efficiency is attained by definition when AET are generated and tested with the same data set of nonimpacted stations). The sensitivity for dry-weight AET in Test 1 was largely confirmed by Tests 3 and 4, which relied on independent data sets (Table 5; sensitivity ranged from 57 to 93 percent). For independent data sets, efficiency of dry weight AET generally ranged from 55 to 75 percent (except Microtox AET Test 3, which was 37 percent efficient).

For AET sets in which nonionic organic chemical concentrations were normalized to TOC (and other chemicals normalized to dry weight; see Test 2), sensitivity ranged from 55 to 77 percent and efficiency was 100 percent by constraint. A more detailed analysis of the results of these and other validation tests is presented by Barrick et al. (1988). Only two aspects of these results are discussed in detail below:

- Characteristics of biologically nonimpacted stations not predicted by AET
- The performance of dry weight AET vs. that of TOC-normalized AET (Test 1 vs. Test 2).

Characteristics of Biologically Impacted Stations Predicted as Nonimpacted by AET

The purpose of this section is to identify the characteristics of biologically impacted stations for which no identified chemical exceeded the AET for that biological indicator. In this case, the results from Test 1 are being considered. In the 334-station database, biological effects were found at 49 amphipod bioassay or benthic infauna stations at which no chemical concentration exceeded a 1988 AET. These incorrect predictions included 29 of 287 stations for the amphipod mortality bioassay (Table 6) and 25 of 201 stations for effects on benthic macroinvertebrate communities (Table 7). Both kinds of effects were found at five of the impacted stations that were not predicted as impacted by AET. Chemical and biological data for each of these stations were reviewed to assess possible reasons for the incorrect predictions including the following:

- The indicated adverse biological effect resulted from factors other than toxic chemical exposures (e.g., physical disruption, grain size distribution)
- Unidentified chemicals accounted for the sediment toxicity
- High chemical detection limits precluded an assessment of whether the AET was exceeded
- The biological effect was incorrectly classified as significant because of a Type I statistical error.

Amphipod Bioassay Stations--DeWitt et al. (1986) have shown that elevated mortality can result in the amphipod mortality bioassay by exposing test organisms to sediments having a high percentage of fine-grained material. These responses were found to occur in the absence of apparent chemical contamination and were thought to result from the physical characteristics of the sediment or some other natural variable that

TABLE 5. SENSITIVITY AND EFFICIENCY RESULTS FOR VALIDATION TESTS CONDUCTED WITH PUGET SOUND AET
[*normalized to dry weight (DW) or total organic carbon (TOC)*]

Validation Test[a] and AET Dataset	Number of Stations Evaluated	Sensitivity	Efficiency	Overall Reliability
Benthic Infaunal Abundance				
1: 1988 AET (DW)	201	75% (81/108)	100% (81/81)[b]	87% (174/201)
2a: 1988 AET (mixed DW/TOC)	201	77% (83/108)	100% (83/83)[b]	88% (176/201)
2b: 1988 AET (TOC)	201	76% (82/108)	100% (82/82)[b]	87% (175/201)
3: Independent AET[c] (DW)	201	75% (81/108)	72% (81/112)	71% (143/201)
4a: 1986 AET (DW); independent data	109	83% (59/71)	75% (59/79)	71% (77/109)
4b: 1986 AET (DW); all data	201	76% (82/108)	82% (82/100)	78% (157/201)
Amphipod Mortality Bioassay				
1: 1988 AET (DW only)	287	58% (62/106)	100% (62/62)[b]	85% (243/287)
2a: 1988 AET (mixed DW/TOC)	287	55% (58/106)	100% (58/58)[b]	83% (239/287)
2b: 1988 AET (TOC only)	287	45% (48/106)	100% (48/48)[b]	80% (229/287)
3: Independent AET[c] (DW)	287	57% (60/106)	67% (60/90)	74% (211/287)
4a: 1986 AET (DW); independent data	146	68% (36/53)	55% (36/66)	68% (99/146)
4b: 1986 AET (DW); all data	287	56% (59/106)	69% (59/86)	74% (213/287)
Microtox Bioassay				
1: 1986 AET[d] (DW only)	50	93% (27/29)	100% (27/27)[b]	96% (48/50)
2a: 1986 AET (mixed DW/TOC)	50	93% (27/29)	100% (27/27)[b]	96% (48/50)
2b: 1986 AET (TOC only)	50	83% (24/29)	100% (24/24)[b]	90% (45/50)
3: Independent AET[c] (DW)	50	93% (27/29)	61% (27/44)	62% (31/50)
4: not applicable[d]				
Oyster Larvae Abnormality				
1: 1986 AET[d] (DW only)	56	88% (15/17)	100% (15/15)[b]	96% (54/56)
2a: 1986 AET[d] (mixed DW/TOC)	56	88% (15/17)	100% (15/15)[b]	96% (54/56)
2b: 1986 AET (TOC only)	56	71% (12/17)	100% (12/12)[b]	91% (51/56)
3: Independent AET[c] (DW)	56	88% (15/17)	37% (15/41)	50% (28/56)
4: not applicable[d]				

[a] See text (Validation Test Methods) for a complete description of the following tests:

Test 1: Dry weight-normalized AET generated and tested using the same database of nonimpacted and impacted stations (note: the concentration at which an AET value is set is determined only by nonimpacted stations).

TABLE 5. (Continued)

Test 2: TOC-normalized AET generated for nonpolar, nonionic organic compounds and dry weight-normalized AET generated for the remaining chemicals (Test 2a) or TOC-normalized AET calculated for all chemicals (Test 2b). Predictions tested against the same database of nonimpacted and impacted that was used to generate the AET.

Test 3: Dry weight-normalized AET generated independently of each station used to test predictions (see footnote c and text for description of iterative procedure used).

Test 4: Dry weight-normalized AET generated in 1986 using approximately half of the amphipod bioassay and benthic infauna stations now available in 1988. Predictions tested using the remaining independent stations only (Test 4a) or all stations (i.e., the stations used to calculate AET plus the remaining independent stations).

[b] By definition, efficiency is 100 percent because all Puget Sound stations for each indicator were included in the calculation of these AET.

[c] Cumulative results for (1) deleting a station from the AET database; (2) recalculating AET using remaining stations; (3) predicting effects at deleted station; (4) restoring the deleted station to the AET database; (5) repeating for each station in the database (see text for discussion of why this test affects efficiency more than sensitivity).

[d] No new data were available to update the 1986 AET for oyster larvae and Microtox bioassays; validation tests were conducted using the 1986 AET although more chemicals (e.g., including tentatively identified organic compounds and conventional sediment variables such as TOC, grain size, sulfides) were used for these tests than used previously by Beller et al. (1986).

TABLE 6. CHARACTERISTICS OF STATIONS AT WHICH SIGNIFICANT AMPHIPOD MORTALITY WAS FOUND BUT NO EFFECTS WERE PREDICTED

Embayment	Survey[a]	Station[a]	Percent Mortality[b]	Key Distinguishing Factors[c,d]
Elliott Bay	EBCHEM	DR-14	32	% fines = 80.6
		EW-10	58	% fines = 80.3; DL
		KG-02	37	DL (6% fines)
		KG-03	27	DL (49% fines)
		KG-09[e]	33	DL (51% fines)
		KG-11	32	DL (8% fines)
		NH-02	45	DL (31% fines)
		NH-09	58	DL (11% fines)
		NH-11	55	no obvious factors (29% fines)
		NS-08[e]	82	% fines = 83.9; DL
		WW-08[e]	41	DL (58% fines)
Sinclair Inlet	EIGHTBAY	SC-08	38	% fines = 89.8; DL
		SC-17	26	% fines = 78.2; DL
		SC-18	32	% fines = 77.8; DL
Case Inlet		CS-01	28	% fines = 89.4; DL
		CS-11	28	DL (39% fines)
		CS-15	43	% fines = 81.3; DL
		CS-17	41	% fines = 78.1; DL
Dabob Bay		DB-07	26	DL (49% fines)
		DB-15	36	% fines = 89.7; DL
Samish Bay		SM-03	47	% fines = 80.9; DL
		SM-07	25	% fines = 84.9; DL
		SM-20	31	% fines = 87.2; DL
Everett Harbor		EV-02	29	% fines = 83.1; DL
		EV-11	26	no obvious factors (69% fines)
Bellingham Bay		BH-05	34	% fines = 96.6; DL
		BH-23	58	% fines = 95.3; DL
Everett Harbor	EVCHEM	NG-06	43	no obvious factors (7% fines)
Everett Harbor	EVERETT1	EV-24	40	few chemicals measured (63% fines)

[a] Surveys listed in Table 1; see Appendix B in Barrick et al. (1988) for station locations.

[b] Percent mortality observed at each anomalous station.

[c] Percent fine-grained material (i.e., <63 um).

[d] DL = Detection limits of at least one chemical exceeded the 1988 amphipod bioassay AET for that chemical.

[e] Station predicted by 1986 Puget Sound amphipod AET but not by 1988 AET.

TABLE 7. CHARACTERISTICS OF STATIONS AT WHICH SIGNIFICANT BENTHIC EFFECTS WERE FOUND BUT NO EFFECTS WERE PREDICTED

Embayment	Survey[a]	Station[a]	Taxon[b]	Key Distinguishing Factors[c,d]
Elliott Bay	EBCHEM	AB-03	P	% coarse = 95.3; DL
		KG-07[d]	M	% coarse = 83.3; DL
		KG-08[e]	M	% coarse = 90.8; DL
		KG-11	M	% coarse = 92.5; DL
		MG-01	P	DL (25% coarse)
		MG-02	P	% coarse = 95.5; DL
		MG-03	P	% coarse = 77.5; DL
		MG-04	P	DL (14% coarse)
		NH-01[e]	M	% coarse = 81.1; DL
		NH-02[e]	M	DL (69% coarse)
		NH-11	P	% coarse = 71.3; DL
		NS-08[e,f]	T,M,C	DL; (16% coarse)
		WW-01	M	no obvious factors (60% coarse)
		WW-03	T,P,M,C	% coarse = 93.4
		WW-08[e]	M,C	no obvious factors (42% coarse)
		WW-17[e]	M	% coarse = 94.7
Everett Harbor	EVCHEM	SD-01	T,P,M,C	% coarse = 95.5
Eagle Harbor	EHCHEM	EH-02	M	% coarse = 89.3; DL
		EH-03	T,M	% coarse = 92.1; DL
Alki Pt.	ALKI	AP-04	M	% coarse = 95.9
Elliott Bay	TPPS3AB	EB-38 (3/15/82)	M	no obvious factors (13% coarse)
		EB-38 (7/15/82)	T,P,M	no obvious factors (38% coarse)
		WP-03 (7/15/82)	P	% coarse = 92.4; DL
		WP-08 (7/15/82)	M	% coarse = 92.6
		WP-16 (3/15/82)	P	no obvious factors (4% coarse)

[a] Surveys listed in Table 1; see Appendix B in Barrick et al. (1988) for station locations.

[b] Taxa showing significant depressions. T = total taxa, P = Polychaeta, M = Mollusca, C = Crustacea.

[c] Percent coarse-grained material (i.e., ≥63 um).

[d] DL = Detection limits of at least one chemical exceeded proposed benthic infauna AET for that chemical.

[e] Station predicted by 1986 Puget Sound amphipod AET but not 1988 AET.

[f] Station NS-08 is located near the Pier 91 naval dock. Tributyltin contamination is possible in this area but has not been tested.

correlated with the physical characteristics. In the Puget Sound database, significant amphipod mortality was associated with 35 of 74 stations (47 percent) with a relatively high percentage (approximated as ≥75 percent) of fine-grained sediment. In contrast, significant mortality was associated with less than 30 percent of the remaining stations. Sixteen of the 29 impacted amphipod bioassay stations (55 percent) that were not predicted by AET were characterized by sediments having a relatively high percentage of fine-grained material. In the Battelle (1985) survey (Table 6), 81 percent of the impacted stations that were not predicted were consistent with this factor. A significant depression in benthic infaunal abundance was found at only one of the 29 fine-grained stations [Station NS-08 (Beller et al. 1988a)]. However, adequate benthic infaunal data were not available at all amphipod bioassay stations [e.g., only amphipod bioassay data were included from the Battelle (1985) survey].

Elevated detection limits precluded an assessment of whether one or more AET were exceeded at nearly all of the impacted amphipod bioassay stations that were not predicted. Aside from detection limit concerns or the possible presence of unmeasured chemicals, no obvious factors account for the apparent incorrect predictions at 12 of the 29 stations (41 percent). Most of these amphipod bioassay stations were found in or near the mouth of the Duwamish River in Seattle, an industrial area contaminated with a variety of chemical classes, suggesting that the biological effects may have resulted from unmeasured chemicals or from interactive effects among the numerous chemicals present at most of those stations.

Benthic Infauna Stations--The characteristics of benthic macroinvertebrate assemblages are influenced, in part, by sediment grain size distribution. Coarse-grained sediments may be suboptimal habitats for many benthic species because they are generally low in organic content and are indicative of high-energy environments. Thus, the food supply for benthic infauna may be limited, and organisms may have difficulty maintaining burrows, tubes, or position in the shifting sediment. In the present study, 16 of the 25 impacted benthic infauna stations (64 percent) that were not predicted by AET were characterized by sediments having a relatively high percentage (i.e., ≥70 percent) of coarse-grained sediment. For example, the observed depression of benthic infaunal abundance at Station SD-01 in Everett Harbor is potentially attributable to the location of this station in a current-swept channel on the Snohomish River delta, rather than toxic effects. Elevated detection limits precluded an assessment of whether one or more AET were exceeded at nearly all of the impacted stations that were not predicted (Table 7).

Aside from potential concerns over detection limits, there was no other obvious factor that might explain apparent incorrect predictions at 9 of the 25 stations (36 percent). As discussed above for amphipod bioassay stations, most of these potentially unexplained stations occurred in or near the mouth of the lower Duwamish River in Seattle. Various dredging projects have occurred in this area and may have disturbed benthic assemblages sufficiently to result in depressions of major taxa. However, detection limits, unmeasured chemicals, or chemical synergism are the most likely factors at these stations, especially considering that four of the five incorrectly predicted stations at which both the amphipod bioassay and benthic infauna indicators were significant were from this area [i.e., Stations KG-11, NH-02, NH-11, and WW-08 from Beller et al. (1988a)]. The fifth such station (Station NS-08) is located near Pier 91 in Elliott Bay. Potential (as yet untested) tributyltin contamination from historical naval ship operations

may account for the substantial amphipod mortality and benthic infaunal depressions observed at this station.

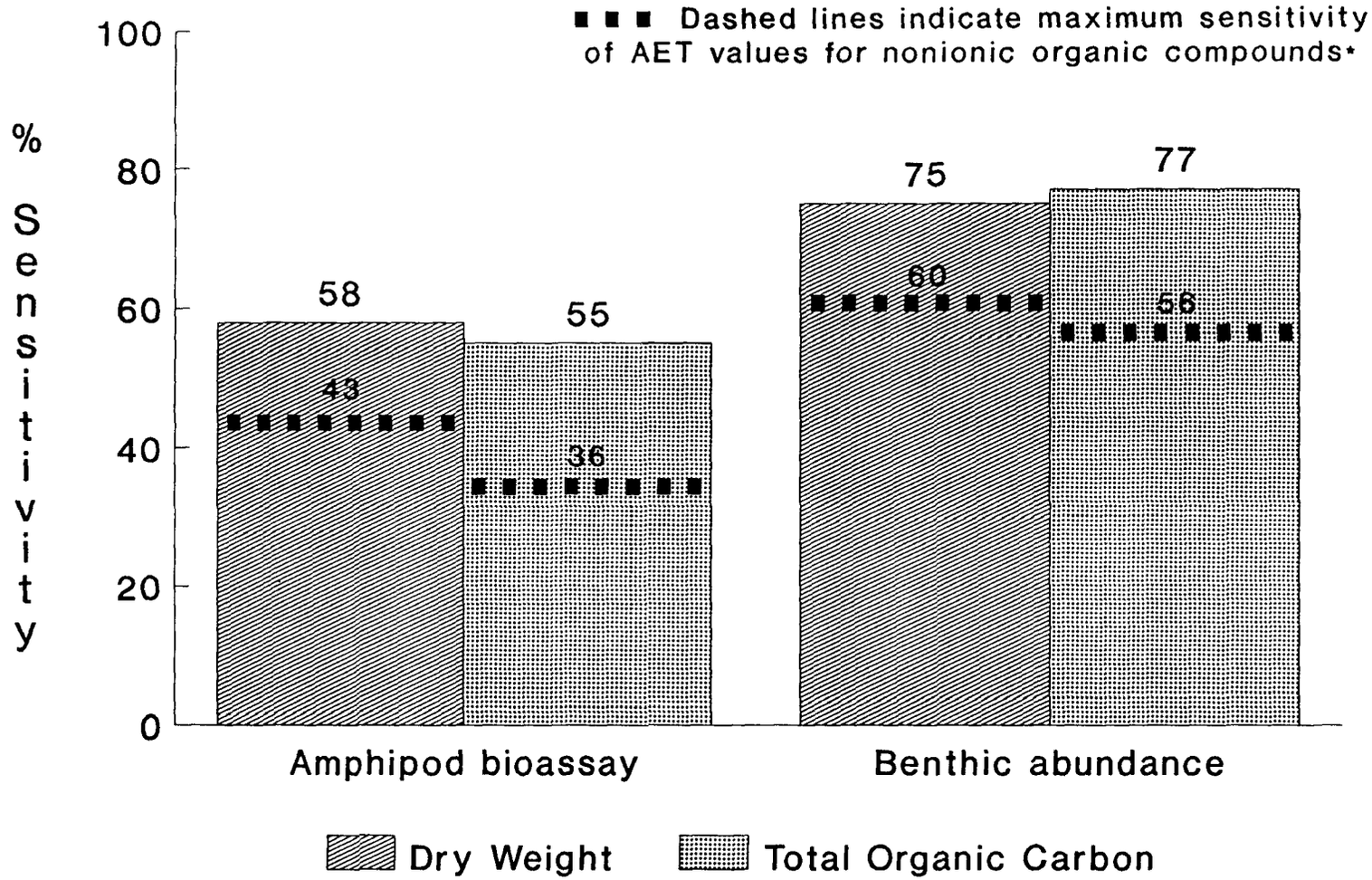
Although unmeasured chemicals or other factors may account for some of the impacted stations that were not predicted by AET, sediment grain size appears to be a dominant factor at the majority of these stations. The fact that fine-grained sediments dominate such amphipod bioassay stations and that coarse-grained sediments dominate such benthic infauna stations reinforces this conclusion. Only 17 percent of the impacted amphipod bioassay stations that were not predicted had a relatively high percentage of coarse-grained material, and only 16 percent of the impacted benthic infauna stations that were not predicted had a relatively high percentage of fine-grained sediments. No correction factor is recommended for grain size, other than matching reference area sediments as closely as possible with sediments from the study site. The 1988 AET are efficient with respect to not predicting effects for stations at which grain size or associated factors may dominate over toxic effects.

Relative Performance of Dry Weight- and TOC-Normalized AET

An interesting result of validation tests with Puget Sound data was the roughly equivalent predictive success of dry weight-normalized AET and organic carbon-normalized AET (Table 5 and Figure 7). This finding was unexpected in light of a body of laboratory bioassay and sorption data that strongly supports the use of TOC normalization for nonionic organic chemicals. Possible explanations for the relative performance of dry weight- and organic carbon-normalized AET are presented in this section.

Dry weight-normalization simply assumes that the overall burden of a contaminant in sediment is a predominant factor influencing toxicity to exposed organisms (although organic carbon content may be a secondary factor). Because dry weight-normalization does not focus on any specific solid fraction of the sediment (e.g., organic matter, fine particles), it essentially averages among sediment fractions in terms of sorptive affinity and relationship to bioavailability. If specific sediment fractions consistently mediate toxicity in most samples, then averaging across all sediment fractions potentially reduces the representativeness of dry weight-normalized AET. Alternatively, if specific sediment fractions do not consistently mediate toxicity in most samples, normalization to dry weight may better account for the variability of contaminant-toxicity relationships among environmental samples than normalization to these sediment fractions.

Organic carbon normalization assumes that organic matter in sediments is a generic sink for nonionic organic contaminants and is a predominant factor influencing the bioavailability and toxicity of these compounds to exposed organisms. Laboratory bioassay and sorption studies provide a mechanistic rationale for organic carbon normalization. Simply stated, bioavailability and toxicity of nonionic organic chemicals in sediments appear to correspond to interstitial water concentrations, and, under equilibrium conditions, the distribution of nonionic organic compounds between sedimentary organic matter (represented by organic carbon content) and interstitial water should be constant (and can be expressed as K_{oc}). For sediments with a given bulk concentration of a nonionic organic chemical, increases in organic carbon content should correspond to proportional decreases in interstitial water concentrations of that chemical. Hence, as sediment organic carbon content increases, toxicity "threshold" values expressed per gram of bulk sediment should decrease. If contaminant concentrations are normalized to organic carbon content, threshold values should be constant for that contaminant in



*Note: Dashed lines do not necessarily represent unique contributions of nonionic organic compounds (i.e., AET for metals or ionizable organic compounds may also be exceeded at some of the same stations)

Figure 7. Sensitivity of dry weight- and total organic carbon-normalized AET. [Efficiency is 100 percent by constraint for both tests.]

all sediments regardless of organic carbon content. Thus, use of TOC-normalized data could be expected to result in a more consistent estimation of sediment toxicity thresholds than use of dry weight-normalized data for generation of AET for nonionic organic compounds.

A study by Adams et al. (1985) illustrates the rationale for organic carbon normalization. Adams et al. (1985) conducted a series of bioassays in which freshwater midges *Chironomus tentans)* were exposed to water, sediments (with various levels of organic carbon content), and food contaminated with Kepone. No-effect concentrations based on total sediment Kepone concentrations increased in proportion to total organic carbon content of sediments, whereas no-effect levels based on interstitial water Kepone concentrations were relatively constant regardless of sediment concentration. The authors suggested that no-effect concentrations should be based on interstitial water concentrations and sediment organic carbon content, not on bulk sediment weight.

The AET approach is empirical and does not favor one mechanistic explanation over any other but can operate whether one or a combination of assumptions is appropriate. The interpretation of operative mechanisms is, for the AET approach, an *a posteriori* inference based on the results of validation tests. For contaminated sediments in the environment, organic carbon normalization might not offer advantages in predictive success over dry weight normalization if the following hypotheses are applicable to environmental sediment samples:

- Sediment/interstitial water systems are not typically at equilibrium in the environment as a result of impeded contaminant exchange

- Sediment organic matter occurring in the environment does not have uniform affinity for nonionic organic compounds.

These hypotheses are discussed further in the following sections.

Sediment-Water Contaminant Exchange--Based on the mechanistic rationale for organic carbon normalization, an assumed advantage of TOC-normalized AET is the constant relationship between sedimentary organic matter, interstitial water concentration, and sediment toxicity that should exist under equilibrium conditions. Based in part on evidence from laboratory studies, it is plausible that equilibrium could be difficult to attain in the environment because of kinetic aspects of sorption/desorption processes. The attainment of equilibrium requires a relatively rapid transfer of a contaminant between various phases in a system. Studies of sorption/desorption have demonstrated that attainment of equilibrium of a nonionic organic compound between sediment and aqueous phases can take weeks, months, or longer (e.g., Karickhoff 1984, Karickhoff and Morris 1985b). The existence of reversible (rapidly exchangeable) and irreversible (highly retarded) components of contaminant loadings in sediments has been postulated by several investigators based on hysteresis in sorption/desorption isotherms or direct observation of desorption kinetics (e.g., DiToro and Horzempa 1982; Karickhoff and Morris 1985b). In laboratory-spiked sediments studied by Karickhoff and Morris (1985b), one half or more of the total sorbed contaminant concentration was irreversible. For highly hydrophobic compounds and systems with high solids concentrations, the irreversible fraction increased to ≥ 90 percent in some cases (Karickhoff and Morris 1985b). For the irreversible component, the attainment of equilibrium distributions could be very slow, on the order of years for highly hydrophobic compounds (i.e., relatively widespread

contaminants such as PAH with five or more rings and PCB congeners with six or more chlorine atoms) (Karickhoff and Morris 1985b).

Very few studies have been conducted with field-collected sediments to examine the kinetic aspects of contaminant exchange between sediment and ambient water. However, in diffusion studies with field-contaminated sediments, Fisher et al. (1983) reported apparent diffusion coefficients for trichloro-, tetrachloro-, and pentachlorobiphenyls that were one to three orders of magnitude lower than observed for laboratory-spiked hexachlorobiphenyl (DiToro et al. 1985); the laboratory-spiked PCB congener appeared to exhibit fully reversible behavior. DiToro et al. (1985) proposed that the low apparent diffusion of PCB congeners observed by Fisher et al. (1983) could be explained by the fact that 90 to 99.9 percent of the sediment loading of PCBs was irreversibly sorbed in those field-collected sediments. Such a proportion of irreversibly sorbed contaminants in the environment could result in considerable deviations from equilibrium conditions. Relatively large and variable deviations from equilibrium in environmental samples would result in considerable variability in the relationship between organic carbon content and bioavailability or toxicity on a sample-by-sample basis, reducing the potential advantage of TOC normalization.

Physical explanations for retarded sediment/water exchange of contaminants and deviations from equilibrium include entrainment or trapping of contaminants in refractory matrices such as fecal pellets (Karickhoff and Morris 1985a) and humic matter (Freeman and Cheung 1981). In a study of sediments collected in Puget Sound, Prahl and Carpenter (1983) observed that PAH were disproportionately concentrated in certain fractions of refractory sedimentary organic matter (e.g., charcoal fragments and vascular plant detritus, such as lignin). This disproportionality suggests that PAH may not have been at equilibrium within the sediment phase or, alternatively, that different kinds of organic matter may have different affinities for PAH.

The Uniformity of Organic Matter--For a given nonionic organic chemical, consistent partitioning between organic matter and interstitial water in sediments from different areas requires that organic matter have a consistent affinity for the chemical. A number of studies indicate that organic matter can be considered uniform to a first approximation, based on the derivation of single K_{oc} values for ranges of sediments and soils (e.g., Karickhoff 1984 and references therein). However, direct studies of the associations of nonionic organic compounds with dissolved (Gauthier et al. 1987; Carter and Suffet 1985) and particulate (Diachenko 1981) humic materials indicate that affinities can vary considerably as a function of the source and properties of the organic matter. For example, K_{oc} values for pyrene in dissolved humic materials from various sources differed by as much as an order of magnitude (Gauthier et al. 1987). The aromatic vs. aliphatic character of organic matter has been related to binding affinity (Diachenko 1981; Gauthier et al. 1987) and is likely an important structural difference between humic materials derived from terrestrial vs. marine sources (e.g., Hatcher et al. 1980, 1981) in estuarine sediments. Variability in organic matter alone may not account for the roughly equivalent performance of dry weight- and organic carbon-normalized AET, as variability within the sedimentary pool of organic matter does not necessarily favor normalization to bulk sediment rather than bulk organic carbon.

4. APPLICATION OF AET IN PUGET SOUND SEDIMENT MANAGEMENT PROGRAMS

The reliability of the AET approach (particularly AET normalized to dry weight for a broad range of chemicals) at predicting biological effects indicates its potential utility as a tool for sediment quality management. Uses for which the AET approach is well suited include:

- Determination of the extent and relative priority of potential problem areas to be managed
- Identification of potential problem chemicals in impacted sediments and, as a result, potential sources of contaminants
- Prioritization of laboratory studies for determining cause-effect relationships
- With appropriate safety factors or other modifications, use in regulatory programs and in screening decisions on the need for further chemical or biological testing of sediments.

Proposed regulations for sediment contamination are currently under review in Puget Sound and may include use of AET to develop sediment standards. These regulations are the culmination of cooperative planning and scientific investigations that were initiated by several federal and state agencies in the early and mid-1980s, including:

- Commencement Bay Superfund Investigations
- Puget Sound Dredged Disposal Analysis (PSDDA)
- Urban Bay Toxics Action Program
- Puget Sound Water Quality Authority (Authority) Management Plan.

Each program and particular application of the AET approach is described further below. A key decision in many of these programs was to develop two sets of sediment quality values, focusing separately on the sensitivity and efficiency concepts of reliability discussed in Section 1 (see Figure 1). 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 <u>and</u> 100 percent efficient values. For these programs, direct biological testing is used to resolve the differences in predictions of the two sets of sediment quality values (i.e., prediction of adverse biological effects by highly sensitive sediment quality values, which at lower chemical concentrations are not predicted by highly efficient sediment quality values).

COMMENCEMENT BAY NEARSHORE/TIDEFLATS SUPERFUND INVESTIGATION

Commencement Bay is a heavily industrialized harbor in Tacoma, Washington. Recent surveys have indicated over 281 industrial activities in the nearshore/tideflats area. Comprehensive shoreline surveys have identified over 429 point and nonpoint source discharges in the study area, consisting primarily of seeps, storm drains, and open channels. Only 27 of the point sources were identified as NPDES-permitted discharges. A remedial investigation under Superfund, started in 1983, revealed 25 major sources contributing to sediment contamination. The magnitude of sediment contamination was characterized using chemical analyses and biological effects indicators. Comparisons of chemical concentrations in various waterways of Commencement Bay with concentrations in relatively uncontaminated areas, along with the presence of measurable biological effects provided a basis for ranking problem areas. No single chemical accounts for this contamination. Concentrations of several different chemicals were found at greater than 1,000 times reference area concentrations, primarily in sediment adjacent to major chemical manufacturing, pulp mill, shipbuilding and repair, or smelter plant operations. Adverse biological effects were found in each of these areas.

During the course of the remedial investigation, the AET approach was developed to assign values to sediment quality. For the Commencement Bay study, biological effects included depressions in the number of individual benthic taxa, presence of tumors and other abnormalities in bottomfish, and several laboratory toxicity tests (amphipod mortality, oyster larvae abnormality, bacterial bioluminescence). At the Commencement Bay site, AET are currently being used to evaluate cleanup alternatives. Optional biological testing can be used by potentially responsible parties to appeal site-specific predictions of adverse biological effects.

PUGET SOUND DREDGED DISPOSAL ANALYSIS

In 1985, PSDDA was initiated to develop environmentally safe and publicly acceptable options for open-water unconfined disposal of dredged material. PSDDA is a cooperative program conducted under the direction of the U.S. Army Corps of Engineers Seattle District (Corps), EPA Region 10, Ecology, and the Washington Department of Natural Resources (WDNR).

In 1988, PSDDA produced a management plan and an environmental impact statement specifying procedures for disposal site management and evaluation of dredged material and identifying recommended open-water unconfined disposal sites in central Puget Sound. Ongoing activities of PSDDA are focused on designating sites for open-water, unconfined disposal in south and north Puget Sound. AET were used by PSDDA as tools to develop chemical-specific guidelines for evaluating the need for biological testing of contaminated dredged material.

In PSDDA, a chemical screening level was established above which biological testing would be required to determine the suitability of dredged material for unconfined, open-water disposal. Above a higher maximum level of contaminant concentrations, additional biological testing was considered unnecessary to determine that the dredged material was unsuitable, although an extensive series of biological tests could be conducted to demonstrate the suitability for unconfined, open-water disposal.

URBAN BAY TOXICS ACTION PROGRAM

The national program for estuarine studies and pollution abatement is implemented through EPA regional offices under the guidance of the Office of Marine and Estuarine Protection (OMEP). Region 10, through the Office of Puget Sound, is responsible for the development and implementation of the Puget Sound Estuary Program (PSEP). The Urban Bay Toxics Action Program is a major component of PSEP, and was begun in 1984 by EPA's Office of Puget Sound and Ecology. Substantial participation has also been provided by the Authority and other state agencies and local government. Major funding and overall guidance for the program is provided by EPA/OMEP. The Urban Bay Toxics Action Program consist of the identification of problem areas in urban bays (predominantly based on sediment contamination), identification of potential sources, development of an action plan for source control, and formation of an action team for plan implementation.

In PSEP urban bay programs, AET are used in conjunction with site-specific biological tests during the assessment of sediment contamination to define and rank problem areas. Actions to date have focused on Elliott Bay, Everett Harbor, and Budd Inlet adjacent to the cities of Seattle, Everett, and Olympia. Source control actions are well underway, but sediment remediation has not yet begun at any of the sites.

PUGET SOUND WATER QUALITY MANAGEMENT PLAN

The Puget Sound Water Quality Management Plan (Plan) was published by the Authority in 1987. The Plan presents programs and plans for 12 issue areas. The Plan identifies the following goal for the contaminated sediments program: "To reduce and ultimately eliminate adverse effects on biological resources and humans from sediment contamination throughout the sound by reducing or eliminating discharges of toxic contaminants and by capping, treating, or removing contaminated sediments." The strategy to achieve this goal is:

- To establish a quality standard that will classify sediments that cause adverse biological effects
- To implement soundwide controls on sources of contaminants causing sediments to fail the classification criteria
- To provide rules and sites for disposal of dredged material
- To conduct remedial actions for existing areas of high sediment contamination levels.

The strategy to develop classification criteria requires Ecology to develop and adopt regulations establishing criteria for identifying and designating sediments that have adverse effects on biological resources or pose a significant health risk to humans. Ongoing efforts to develop sediment quality standards are focused by the requirements of the Authority's Plan toward producing legally enforceable tools for preventing sediment contamination and managing contaminated sediment. Development of state

sediment standards by Ecology in response to the Authority's Plan is described in the following section.

PROPOSED STATE SEDIMENT QUALITY STANDARDS

Ecology is currently developing multiple categories of sediment management standards, based in part on AET values. Development of these standards relies heavily on the past and ongoing efforts described above and involves active participation by Ecology, EPA, the Authority, WDNR, the Corps, and a variety of public interest groups. The draft regulation currently under development affects only sediment in Puget Sound, although the adopted regulation will be broadened, and modified as necessary, to include the entire state at a future date.

Ecology has convened several public workshops in 1988 for interested parties such as environmental and public interest groups, ports, industry, state and federal agencies, local governments, and Indian tribes. In addition, a Sediments Advisory Group provides technical and policy review of the standards development. Ecology will adopt the regulations following formal public review processes beginning in January 1989. The current schedule for adoption of standards and guidelines is outlined below:

■ Sediment Quality Standards	Draft by December 31, 1988 Final by June 30, 1989
■ Effluent Particulate Limits	Interim by July 30, 1988 Final by June 30, 1989
■ Standards for Unconfined Disposal of Dredged Material	Central Puget Sound by September 1988; north and south Puget Sound by September 1989
■ Standards for Confined Disposal of Dredged Material	Interim by September 1989 Final by July 1990
■ Sediment Remedial Action Guidelines	Final by January 1991
■ Method for Ranking Contaminated Sediment Sites	By June 1, 1990
■ Inventory of Sites with Adverse Effects	Initial inventory by October 1, 1990
■ Priority List of Contaminated Sites and Investigation Schedule	By December 31, 1990

5. SUMMARY

Ideally, sediment quality values would be supported by a complete understanding of the physical, chemical, and biological interactions that comprise a dynamic ecosystem. However, extensive research is required to understand the complex relationships that drive such a system. The near-term needs for enhanced regulation and environmental protection in Puget Sound are not compatible with the probable long-term development of sediment criteria based on such complete understanding. In Washington state, the development of sediment quality values is viewed as an iterative process in which the best available information is integrated into an appropriate management framework.

The sediment quality values that have been developed for Puget Sound using the AET approach provide decision tools that have the following characteristics:

- Developed empirically from field data in Puget Sound
- Developed to provide chemical-specific values
- Supported by a variety of biological indicators including acute lethal and sublethal bioassays and *in situ* benthic infaunal analyses reflecting acute and/or chronic effects
- Driven by statistically significant adverse effects relative to Puget Sound reference conditions
- Supported by noncontradictory evidence of adverse effects within a database incorporating approximately 300 samples from 13 embayments in Puget Sound (including 287 amphipod bioassay stations, 201 benthic infauna stations, 56 oyster larvae bioassay stations, and 50 Microtox bioassay stations).

For this database, AET are from 85 to 96 percent reliable in predicting adverse effects when they do occur and in not predicting adverse effects when none are observed. Viewed as an environmental risk management tool, the application of AET in sediment programs enhances the ability to characterize and clean up existing sediment contamination and to prevent future contamination from occurring. The AET approach is currently used in several regional programs and is proposed as the basis for developing state sediment quality standards.

6. 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. American Society for Testing and Materials, Philadelphia, PA.

Barrick, R., S. Becker, D. Weston, and T. Ginn. 1985. Commencement Bay nearshore/ tideflats remedial investigation. Final Report. Prepared by Tetra Tech, Inc. for Washington Department of Ecology and U.S. Environmental Protection Agency. Tetra Tech, Inc., Bellevue, WA.

Barrick, R., H. Beller, and M. Meredith. 1986. Eagle Harbor Preliminary Investigation. Final Report. Prepared by Tetra Tech, Inc. for Black & Veatch/Washington Department of Ecology. Tetra Tech, Inc., Bellevue, WA.

Barrick, R., S. Becker, R. Pastorok, L. Brown, and H. Beller. 1988. Sediment quality values refinement: 1988 update and evaluation of Puget Sound AET. Prepared by PTI Environmental Services for Tetra Tech/U.S. Environmental Protection Agency Region 10, Office of Puget Sound. PTI Environmental Services, Bellevue, WA.

Battelle 1985. Detailed chemical and biological analyses of selected sediment from Puget Sound. Draft Final Report. U.S. Environmental Protection Agency Region 10, Seattle, WA. 300 pp.

Battelle. 1988. Overview of methods for assessing and managing sediment quality. Prepared for U.S. Environmental Protection Agency, Office of Water, Office of Marine and Estuarine Protection, Washington DC. Battelle Ocean Sciences, Duxbury, MA. 30 pp.

Beckman Instruments, Inc. 1982. Microtox system operating manual. Carlsbad, CA.

Beller, H., R. Barrick, and S. Becker. 1986. Development of sediment quality values for Puget Sound. Prepared by Tetra Tech, Inc. for Resource Planning Associates/U.S. Army Corps of Engineers, Seattle District for the Puget Sound Dredged Disposal Analysis program. Tetra Tech, Inc., Bellevue, WA.

Beller, H., R. Pastorok, S. Becker, G. Braun, G. Bilyard, and P. Chapman. 1988a. Elliott Bay action program: analysis of toxic problem areas. Final Report. Prepared for U.S. Environmental Protection Agency Region 10, Office of Puget Sound. PTI Environmental Services and Tetra Tech, Inc., Bellevue, WA. July, 1988.

Beller, H., R. Barrick, L. Jacobs, and S. Becker. 1988b. Commencement Bay Nearshore/ Tideflats Feasibility Study: Development of sediment cleanup goals. Public Review Draft. Prepared by PTI Environmental Services for Tetra Tech, Inc./Washington Department of Ecology and U.S. Environmental Protection Agency, Region 10. PTI Environmental Services, Bellevue, WA.

Bulich, A.A., M.W. Greene, and D.L. Isenberg. 1981. Reliability of the bacterial luminescence assay for determination of the toxicity of pure compounds and complex effluent. pp. 338-347. In: Aquatic Toxicology and Hazard Assessment: Proceedings of the Fourth Annual Symposium. D.R. Branson and K.L. Dickson (eds). ASTM STP 737. American Society for Testing and Materials, Philadelphia, PA.

Carter C.W., and I.H. Suffet. 1985. Quantitative measurements of pollutant binding to dissolved humic materials compared with bulk properties of humic materials. Org. Geochem. 8:145-146.

Chan, S., M. Schiewe, and D. Brown. 1985. Analysis of sediment samples for U.S. Army Corps of Engineers Seattle harbor navigation project, operations and maintenance sampling and testing of Duwamish River sediments. Draft Report. 15 pp. + appendices.

Chan, S. M. Schiewe, K. Grams, A. Friedman, R. Bogar, U. Varanasi, W. Reichert, P.D. Plesha, S.J. Demuth, and D.W. Brown. 1986. East, West, and Duwamish Waterways Navigation Improvement Project: physical/chemical/biological analyses of sediments proposed for dredging. Final Report. Prepared by NOAA, National Marine Fisheries Service for U.S. Army Corps of Engineers, Seattle District. 106 pp.

Chapman, P.M. In Review. Current approaches to developing sediment quality criteria. Submitted to Environmental Toxicology & Chemistry.

Chapman, P.M., and J.D. Morgan. 1983. Sediment bioassays with oyster larvae. Bull. Environ. Contam. Toxicol. 31:438-444.

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 survival of the phoxocephalid amphipod *Rhepoxynius abronius*. Mar. Environ. Res. 25:99-124.

Diachenko, G.W. 1981. Sorptive interactions of selected volatile hydrocarbons with humic acids from different environments. PhD Dissertation, University of Maryland.

Dexter, R.N., D.E. Anderson, E.A. Quinlan, L.S. Goldstein, R.M. Strickland, S.P. Pavlou, J.R. Clayton, Jr., R.M. Kocan, M. Landolt. 1982. A summary of knowledge of Puget Sound related to chemical contaminants. NOAA Technical Memorandum. OMPA-13. NOAA/OMPA, Boulder, CO, 435 pp.

DiToro, D.M. and L.M. Horzempa. 1982. Reversible and resistant components of PCB adsorption-desorption: isotherms. Environ. Sci. Technol. 16:594-602.

DiToro, D.M., J.S. Jeris, and D. Ciarcia. 1985. Diffusion and partitioning of hexachlorobiphenyl in sediments. Environ. Sci. Technol. 19:1169-1176.

Farrington, J.W. and J.M. Teal. 1982. Pollutant chemical measurements and biological availability: a lesson to be learned from fossil fuel hydrocarbons. Trans. Am. Geophys. Union. 53:86.

Fisher, J.B., R.L. Petty, and W. Lick. 1983. Release of polychlorinated biphenyls from contaminated lake sediments: flux and apparent diffusivities of four individual PCBs. Environ. Pollut. (Ser. B) 5:121-132.

Freeman, D.H. and L.S. Cheung. 1981. A gel partition model for organic desorption from a pond sediment. Science (Wash., DC) 214:790-792.

Gahler, A.R., J.M. Cummins, J.N. Blazevich, R.H. Riech, R.L. Arp, C.E. Gangmark, S.V.W. Pope, and S. Filip. 1982. Chemical contaminants in edible non-salmonid fish and crabs from Commencement Bay, WA. U.S. Environmental Protection Agency 910/9-82-093. EPA Region 10, Seattle, WA. 117 pp.

Gauthier, T.D., W.R. Seitz, and C.L. Grant. 1987. Effects of structural and compositional variations of dissolved humic materials on pyrene K_{oc} values. Environ. Sci. Technol. 21:243-248.

Ginn, T.C., and R.C. Barrick. 1988. Bioaccumulation of toxic substances in Puget Sound organisms. Proc. Fifth International Ocean Disposal Symposium. pp. 157-168. In: Oceanic Processes in Marine Pollution, Volume 5: Urban Wastes in Coastal Marine Environments. D.A. Wolfe and T.P. O'Conner (eds). Robert E. Krieger Publishing Co., Malabar, FL.

Hatcher, P.G., R. Rowan, and M.A. Mattingly. 1980. ^{1}H and ^{13}C NMR of marine humic acids. Org. Geochem. 2:77-85.

Hatcher, P.G., G.E. Maciel, and L.W. Dennis. 1981. Aliphatic structure of humic acids; a clue to their origin. Org. Geochem. 3:43-48.

Karickhoff, S.W. 1984. Organic pollutant sorption in aquatic systems. J. Hydraulic Engin. 110:707-735.

Karickhoff, S.W., and K.R. Morris. 1985a. Impact of tubificid oligochaetes on pollutant transport in bottom sediments. Environ. Sci. Technol. 19:51-56.

Karickhoff, S.W., and K.R. Morris. 1985b. Sorption dynamics of hydrophobic pollutants in sediment suspensions. Environ. Toxicol. Chem. 4:469-479.

Lie, U. 1968. A quantitative study of benthic infauna in Puget Sound, Washington, U.S.A., in 1963-64. FiskDir. Skr., (Ser Havunders.). 14:229-556.

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. Pollut. Bull. 16:405-415.

Lyman, W.J., A.E. Glazer, J.H. Ong, and S.F. Coons. 1987. An overview of sediment quality in the United States. Final Report. Prepared for U.S. Environmental Protection Agency, Office of Water Regulations and Standards, Monitoring and Data Support Division, Washington, D.C. Arthur D. Little, Inc., Cambridge MA. 112 pp. w/appendices.

Malins, D.C., B.B. McCain, D.W. Brown, A.K. Sparks, H.O. Hodgins, and S.L. Chan. 1982. Chemical contaminants and abnormalities in fish and invertebrates from Puget

Sound. Technical Memorandum OMPA-19. National Oceanic and Atmospheric Administration, Seattle, WA.

Malins, D.C., B.B. McCain, D.W. Brown, S.L. Chan, M.S. Myers, J.T. Landahl, P.G. Prohaska, A.J. Friedman, L.D. Rhodes, D.G. Burrows, W.D., Gronlund, and H.O. Hodgins. 1984. Chemical pollutants in sediments and diseases of bottom-dwelling fish in Puget Sound, Washington. Environ. Sci. Technol. 18:705-713.

Nichols, F.H. 1975. Dynamics and energetics of three deposit-feeding benthic invertebrate populations in Puget Sound, Washington. Ecol. Monogr. 45:57-82.

Osborn, J.G., D.E. Weitkamp, and T.H. Schadt. 1985. Alki wastewater treatment plant outfall improvements predesign study. Tech. Rep. No. 6.0, Marine Biology. Municipality of Metropolitan Seattle. 50 pp.

Pastorok, R., H. Beller, S. Becker, T. Ginn, L. Jacobs, N. Musgrove, G. Braun, and P. Chapman. 1988. Everett Harbor action program: analysis of toxic problem areas. Final Report. Prepared for U.S. Environmental Protection Agency Region 10, Office of Puget Sound. PTI Environmental Services and Tetra Tech, Inc., Bellevue, WA. September, 1988.

Prahl, F.G., and R. Carpenter. 1983. PAH-phase associations in Washington coastal sediment. Geochim. Cosmochim. Acta 47:1013-1023.

Romberg, G.P., S.P. Pavlou, and E.A. Crecelius. 1984. Presence, distribution, and fate of toxicants in Puget Sound and Lake Washington. METRO Toxicant Program Report No. 6A. Toxicant Pretreatment Planning Study Technical Report C1. Municipality of Metropolitan Seattle, Seattle, WA. 231 pp.

Sokal, R.R., and F.J. Rohlf. 1969. Biometry. W.H. Freeman and Company, San Francisco, CA. 776 pp.

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, U.S.A. Mar. Pollut. Bull. 13:359-364.

Swartz, R.C., W.A. DeBen, J.K.P. Jones, J.O. Lamberson, and F.A. Cole. 1985. Phoxocephalid amphipod bioassay for marine sediment toxicity. pp. 284-307. In: Aquatic Toxicology and Hazard Assessment: 7th Symposium, ASTM STP 854. R.D. Cardwell, R. Purdy, and R.C. Bahner (eds). American Society for Testing and Materials, Philadelphia, PA.

Tetra Tech. 1987. Recommended season for sampling subtidal benthic macroinvertebrates during Puget Sound monitoring. Draft Report. Prepared for U.S. Environmental Protection Agency, Region 10. Tetra Tech, Bellevue, WA. 22 pp.

Trial, W., and J. Michaud. 1985. Alki wastewater treatment plant outfall improvements predesign study. Tech. Rep. No. 8.3, Water Quality. Municipality of Metropolitan Seattle. 89 pp.

U.S. Department of the Navy. 1985. Final environmental impact statement. Carrier Battle Group Puget Sound region ship homeporting project. Tech. Appendix. Vol. 2.

Prepared for U.S. Department of the Navy, Western Division, Naval Facilities Engineering Command, San Bruno, CA.

Williams, L.G., P.M. Chapman, and T.C. Ginn. 1986. A comparative evaluation of marine sediment toxicity using bacterial luminescence, oyster embryo, and amphipod sediment bioassays. Mar. Environ. Res. 19:225-249.

Word, J.Q., P.L. Striplin, K. Keeley, et al. 1984. Renton sewage treatment plant project. Seahurst baseline study. Volume V. Section 6. Subtidal benthic ecology. University of Washington Fisheries Research Institute, Seattle, WA. 461 pp.

Yake, W., and D. Norton. 1986. Memorandum, July 1, 1986: Chemical contamination of groundwater, intertidal seepage, and sediments on and near Wyckoff Company property; Eagle Harbor, Bainbridge Island. Water Quality Investigations Section, Washington Department of Ecology, Olympia, WA 60 pp.

Zar, J.H. 1974. Biostatistical analysis. Prentice-Hall, Inc. Englewood Cliffs, NJ.

Environmental Protection Agency
Regional Program Office
Library