

Lawrence Berkeley National Laboratory

Recent Work

Title

Determinants of Sediment Toxicity in San Francisco Bay

Permalink

<https://escholarship.org/uc/item/2fm0h1ps>

Authors

Hoffman, E.R.

Anderson, S.L.

Knezovich, J.P.

Publication Date

1994-12-01



Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

ENERGY & ENVIRONMENT DIVISION

**Determinants of Sediment Toxicity in
San Francisco Bay**

Final Report – December 1994

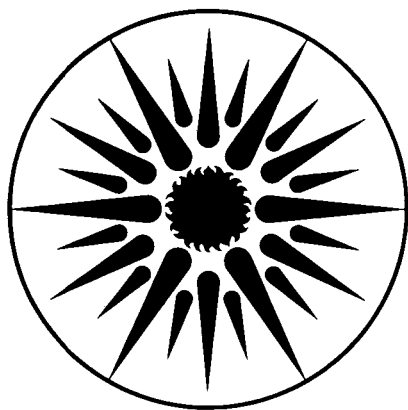
E.R. Hoffman, S.L. Anderson, J.P. Knezovich

December 1994

*U. C. Lawrence Berkeley Laboratory
Library, Berkeley*

FOR REFERENCE

Not to be taken from this room



**ENERGY
AND ENVIRONMENT
DIVISION**

REFERENCE COPY |
Does Not | Copy 1
Circulate |
Bldg. 50 Library.

LBL-36592

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

**DETERMINANTS OF SEDIMENT TOXICITY IN
SAN FRANCISCO BAY**

Final Report
December 1994

Erika R. Hoffman^{1,3}
Susan L. Anderson¹
John P. Knezovich²

¹ Energy and Environment Division
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

² Health and Ecological Assessment Division
Lawrence Livermore National Laboratory
Livermore, CA 94550

³ Wetlands and Sediment Management Section
U.S. Environmental Protection Agency, Region 9
San Francisco, CA 94105

Foreword

This research was performed under a contract (MIPR # E86-92-3026) between the Lawrence Berkeley Laboratory and the United States Army Corps of Engineers (USACOE) under a contract through the U.S. Department of Energy (No. DE-AC03-76SF00098). The project was administered by the USACOE with Mr. Tom Wakeman as contract officer. Additional oversight was provided by the interagency Long Term Management Strategy (LTMS) in-Bay studies work group, chaired by Mr. Tom Gandesbery. The contents of this document do not necessarily reflect the view and policies of the USACOE nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

TABLE OF CONTENTS

FOREWORD	i
EXECUTIVE SUMMARY	v
LIST OF TABLES	ix
LIST OF FIGURES	x
ACKNOWLEDGMENTS	xiii
1. INTRODUCTION	
Background	1-1
Goals	1-4
Scope	1-5
Study selection and site characteristics	1-5
2. PATTERNS OF TOXICITY IN FOUR GENERAL SITE CLASSES	
Introduction	2-1
Goals	2-1
Approach	2-1
Results	2-2
Discussion	2-5
Conclusions	2-7
3. RELATIONSHIP BETWEEN TOXICITY AND BULK SEDIMENT CHEMISTRY	
Introduction	3-1
Goals	3-1
Approach	3-2
Results	3-3
Discussion	3-16
Conclusions	3-19
4. SENSITIVITY OF TWO SPECIES OF INFAUNAL AMPHIPODS TO SEDIMENT GRAIN SIZE	
Introduction	4-1
Goals	4-2
Approach	4-3
Results	4-3
Discussion	4-5
Conclusions	4-7

Contents Continued

5. PATTERNS OF CONTAMINATION AND TOXICITY AT IN-BAY AND OCEAN REFERENCE SITES.

Introduction	5-1
Goals	5-2
Approach	5-2
Results	5-6
Discussion	5-9
Conclusions	5-12

6. AN EVALUATION OF EXISTING SEDIMENT QUALITY GUIDELINES USING DATA FROM SAN FRANCISCO BAY

Introduction	6-1
Goals	6-2
Background on sediment quality guideline derivation	6-2
Apparent Effects Threshold (AET)	6-2
Effects Range-Low (ERL) and Effects Range-Median values	6-5
EPA Sediment Quality Criteria (SQC)	6-7
Approach	6-9
Results	6-10
ERL Exceedences	6-15
ERM Exceedences	6-15
Chromium and ERL/ERM values	6-16
Mercury and ERL/ERM values	6-16
AET Exceedences	6-19
EPA Sediment Quality Criteria Exceedences	6-19
Discussion	6-20
Conclusions	6-23

7. BACKGROUND CONCENTRATIONS OF SELECTED TRACE METALS IN SAN FRANCISCO BAY SEDIMENT

Introduction	7-1
Goals	7-1
Approach	7-2
Results	7-5
Discussion	7-7
Conclusion	7-8

Contents Continued

8. ALTERNATIVE APPROACHES FOR EVALUATING SEDIMENT TOXICITY AND CHEMISTRY DATA

Introduction	8-1
Goals	8-2
Approach	
Summed PAH method	8-3
Critical Body Residue method	8-5
Results	
Summed PAH method	8-7
Critical Body Residue method	8-9
Discussion	
Summed PAH method	8-9
Critical Body Residue method	8-11
Conclusions	8-13

9. LITERATURE CITED 9-1

EXECUTIVE SUMMARY

San Francisco Bay has received anthropogenic waste for decades. Sediments, in particular, have served as a sink for numerous contaminants. Yet, the bioavailability of sediment-associated chemicals to benthic organisms and, thus, the potential for ecological impacts remains unclear. Various testing methodologies are currently used to assess the potential for impacts to the benthos associated with sediment contamination. However, increased development and application of these sediment quality tests has resulted in intense debate regarding their interpretation and reliability.

Currently, a wealth of data exist that can be used to develop a better understanding of the extent and impacts of sediment contamination in San Francisco Bay. Generated by numerous short- and long-term monitoring programs as well as dredged material testing, this data reflects both where we have been and where we are going in sediment testing. It was our goal to evaluate a subset of this new data with an eye on the evolving state of our understanding of chemical processes in sediments and their influence on the toxicity of contaminants.

In this report, several topical issues relating to sediment contamination in San Francisco Bay are examined. Each section addresses a discrete question and includes a discussion of the context for the analysis as well as description of the approach used, the results, and implications of our findings. Although sections are meant to stand alone, linkages do exist among them and these are discussed where appropriate.

In **Section 1** we provide general background on sediment toxicity testing methods and the issues surrounding their application in San Francisco Bay. We also outline the different component studies which comprise the San Francisco Bay data set used in subsequent analyses. This data set was primarily derived from harbor deepening studies and regional monitoring programs.

In **Section 2**, we examine the frequency and magnitude of toxicity observed in both

solid-phase and suspended-phase sediment bioassays from the San Francisco Bay data set. Significant toxicity using at least one bioassay test species was observed in numerous samples from different types of locations throughout the Bay. The results of multispecies toxicity testing provides a firmer basis for assessing the potential for toxic effects in these sediments. For any given location, patterns in toxic response generally reflected the relative influence of contaminant sources. The lack of concordance observed between the responses of different test species to the same sediments underscores the need for multispecies testing in order to comprehensively characterize the potential for ecological risk posed by sediment contaminants.

In **Section 3**, correlations between toxicity and bulk sediment chemistry are examined. The goal of this analysis was to determine the extent to which significant correlations between toxicity and sediment concentrations of organic compounds and trace metals are observable. Rank correlation coefficients were used to assess the strength of association between chemical and toxicological data. Many strong associations between toxicity and both contaminant and non-contaminant parameters were observed. However, univariate concentration-response analyses identified relatively few cases for which toxicity was clearly attributable to increasing bulk sediment concentrations of a particular chemical or class of chemicals. In particular, strong associations between contaminant and non-contaminant parameters confounded the identification of causes of toxicity. Based on these findings, we conclude that future testing programs should incorporate measures of the bioavailable fraction of sediment contaminants. Determination of acid volatile sulfide/simultaneously extracted metals (AVS/SEM) ratios and a greater focus on pore water testing would improve our ability to discern the relative contribution of various contaminant and non-contaminant parameters to observed sediment toxicity.

Section 4 addresses the relationship between sediment grain size and toxic response in two amphipod species (*Rhepoxynius abronius* or *Eohaustorius estuarius*). Both species are commonly used to evaluate toxicity of San Francisco Bay sediments. Data on grain size and amphipod survival were analyzed to determine whether increasing fines fraction and acute toxicity are positively associated. The presence of fine sediments was not consistently

associated with significant mortality for either amphipod species, indicating that grain size was neither a consistent nor significant determinant of amphipod toxicity in this data set. Further study on this topic should focus on assessing the contribution of other sediment properties such as shape and composition that may differentially contribute to the toxicity of similarly fine-grained sediments.

In **Section 5**, we compare and contrast patterns of contamination and toxicity from six reference sites located both within San Francisco Bay and in the ocean. The goal of this analysis was to assess levels of sediment contamination at all reference sites as well as to compare intersampling variability associated with chemical and toxicological measurements at in-Bay versus ocean references. Generally, we found that the concentrations of contaminants at these reference sites were relatively clean, although isolated instances of higher contamination were observed at several locations. Surprisingly, significant increases in toxicity were not observed in solid-phase bioassays as a result of exposure to these more contaminated sediments. Furthermore, intersampling variability over the course of one year was no greater at in-Bay than at ocean sites. Based on this analysis, we conclude that several of these reference sites appear to provide consistent points of comparison for continued use in testing programs.

Section 6 provides an evaluation of seven different types of sediment quality guidelines relative to their ability to accurately predict acute toxicity in San Francisco Bay sediments. An overview of the derivation approach for each guideline type is presented. We assessed the reliability of these different guidelines by evaluating where acute toxicity cooccurred with sediment contamination in excess of guideline values. For many of the chemicals measured, ERM and Puget Sound dry-weight amphipod AET values were better predictors of toxicity relative to the other guidelines. However, all of these sediment guidelines were limited in their ability to accurately predict toxicity in San Francisco Bay sediments. These results suggest that sediment guidelines, alone, are not adequate for evaluating the ecological hazards posed by contaminated sediments. However, used in conjunction with chemical and biological testing, region-specific guidelines may be valuable for the identification and prioritization of sites and/or sediments.

The goal of **Section 7** was to determine whether site-specific geochemical characteristics of sediments in this region could potentially influence the suitability of numerical sediment quality guidelines. To achieve this goal, we compared recently-collected USGS data measuring background concentrations of trace metals in San Francisco Bay sediments to a set of nationally-derived sediment quality guidelines (ERL values). We determined that ERL values for several trace metals (chromium, nickel, copper, and lead) are either equal to or lower than background concentrations, indicating that these particular guideline values cannot be used to predict sediment toxicity that would be associated with contamination. This analysis demonstrates that a consideration of naturally occurring concentrations of trace metals in addition to those that are indicative of contamination is critical in the development and/or application of numerical guidelines to San Francisco Bay sediments.

Finally, in **Section 8**, alternative approaches for evaluating the relationships between sediment contamination and toxicity are presented. Specifically, two methodologies (the Summed PAH model and the Critical Body Residue method) are applied to the data set in order to demonstrate their utility for predicting toxicity associated with specific classes of non-ionic organic compounds. Results using both methods suggest that organic compounds are not a significant determinant of toxicity in these samples. Our analysis illustrates that these models, when used in conjunction with results from bioassays, can provide important information for the identification of key contaminant classes of concern to sediment organisms.

The findings summarized above must be considered as preliminary since they were derived from only a fraction of the sediment data currently available for San Francisco Bay. They do indicate, however, that existing data do not provide us with a clear understanding of the causes of sediment toxicity in the Bay. Focused expansion of this data set is necessary to more accurately determine the utility of tools such as exposure models and sediment quality guidelines as well as to better characterize the general trends and potential causes of contamination and toxicity.

TABLES

Table 1-1.	Description of the eight San Francisco Bay studies used in this report. . . .	1-6
Table 1-2.	Physical and chemical parameters reported in San Francisco Bay sediment data sets	1-15
Table 1-3.	A summary of the test species and testing conditions used in this report.	1-17
Table 3-1.	Significant Spearman rank correlations for the Harbor studies.	3-4
Table 3-2.	Significant Spearman rank correlations for the Harbor studies.	3-6
Table 5-1.	Mean coefficients of variation for selected chemical parameters at in-Bay and ocean reference sites.	5-10
Table 6-1.	Sediment Quality Guidelines	6-2
Table 6-2.	ERL and ERM values and the frequency of their exceedence in San Francisco Bay sediment samples.	6-11
Table 6-3.	Puget Sound amphipod AET values and the frequency of their exceedence in San Francisco Bay sediment samples	6-12
Table 6-4.	Puget Sound bivalve AET values and the frequency of their exceedence in San Francisco Bay sediment samples.	6-13
Table 7-1.	Background concentrations for selected metals in San Francisco Bay sediments from Hornberger <i>et al.</i> (in review).	7-4
Table 7-2.	The percentage of sediment samples from San Francisco Bay studies with trace metal concentrations that were above the USGS background.	7-6
Table 8-1.	Neutral narcotic chemicals included in tissue residue calculation.	8-7
Table 8-2.	Toxic unit concentrations of PAHs in sediment, amphipod mortality, and the organic carbon content of selected samples from four San Francisco Bay studies.	8-8
Table 8-3.	Sediment and tissue concentrations of neutral narcotic chemicals at selected stations from the three Harbor studies.	8-10

FIGURES

Figure 1-1.	Location of composite samples taken for the Port of Oakland Berth Area Study during March 1991	1-7
Figure 1-2.	Location of samples taken for the Port of Oakland Inner Harbor Studies (Phase 3A and Phase 3A retest) during June and September 1990	1-8
Figure 1-3.	Location of samples taken for the Port of Oakland Outer Harbor Study (Phase 3B) during November 1990	1-9
Figure 1-4.	Location of samples taken for the Port of Richmond Deepening Study during June 1991	1-10
Figure 1-5.	Location of samples taken for the SedQual 3 Castro Cove Study during May 1991.	1-11
Figure 1-6.	Location of samples taken for the SedQual 3 Marsh Studies on multiple dates in 1991 and 1992	1-12
Figure 1-7.	Location of samples taken for the SedQual 3 Bay Studies #1 and #2 during August 1991 and April 1992	1-13
Figure 2-1.	Magnitude of toxicity observed in sediment samples from (A) harbor navigation channels and (B) mid-Bay station in San Francisco Bay.	2-3
Figure 2-2.	Magnitude of toxicity observed in sediment samples from (A) marshes and (B) Castro Cove	2-4
Figure 3-1	Bivalve larval mortality and abnormality in relationship to sediment concentrations of tributyltin from (A) Oakland Harbor berths using <i>Crassostrea gigas</i> and (B) Oakland Outer Harbor using <i>Mytilus galloprovincialis</i>	3-8
Figure 3-2	<i>Eohaustorius estuarius</i> mortality in relation to (A) %TOC and (B) %TON in 30 sediment samples from Castro Cove.	3-10
Figure 3-3	<i>Eohaustorius estuarius</i> mortality in relation to OC-normalized concentrations of (A) Methylphenanthrene and (B) total DDT in 30 sediment samples from Castro Cove.	3-11

Figures Continued

Figure 3-4.	The relationship between sediment TOC and dry-weight concentrations of Methylphenanthrene and total DDT in 30 sediment samples from Castro Cove.	3-12
Figure 3-5.	<i>Eohaustorius estuarius</i> mortality in relation to (A) zinc, (B) nickel, (C) lead, and copper in 30 sediment samples from Castro Cove.	3-13
Figure 3-6.	<i>Crassostrea gigas</i> mortality in relation to (A) copper and (B) chromium in 30 sediment samples from Castro Cove..	3-14
Figure 3-7.	<i>Eohaustorius estuarius</i> mortality in relationship to (A) TOC, (B) TON, (C) dibenzo(a,h)anthracene, and (D) nickel for 29 sediments from the Bay Run #1 (August 1991) and Bay Run #2 (April 1992) studies.	3-15
Figure 4-1.	Survival of <i>Rhepoxynius abronius</i> in relation to % fines in sediments from 69 sites in San Francisco Bay harbor studies..	4-4
Figure 4-2	Survival of <i>Eohaustorius estuarius</i> in relation to % fines in sediments from 98 sites in San Francisco Bay SedQual 3 study.	4-6
Figure 5-1.	Location of the Alcatraz mound (R-AC) and environs (R-AM) reference sites.	5-3
Figure 5-2	Location of the Bay Farm (R-BF) reference site.	5-4
Figure 5-3.	Location of the three ocean reference sites: the Point Reyes fine-grained (R-PF) and Point Reyes course-grained (R-PC) sites, and the Off-shelf (R-OS) site	5-5
Figure 5-4.	Sediment dry-weight concentrations of (A) total PAHs and (B) total PCBs at six reference sites tested in five harbor studies (March 1990 - June 1991).	5-7
Figure 5-5.	Mortality of (A) <i>Rhepoxynius abronius</i> and (B) <i>Nephtys caecoides</i> at six reference sites tested in four harbor studies (March 1990 - June 1991). ...	5-8
Figure 6-1.	A schematic representation of AET derivation based on results of toxicity testing.	6-4

Figures Continued

Figure 6-2 Frequency with which the following four dry-weight sediment guidelines were exceeded in San Francisco Bay sediments: Effects Range-Low (ERL) values for 24 chemicals; Effects Range-Median (ERM) values for 24 chemicals; Puget Sound Amphipod Apparent Effects Threshold (AAET) values for 27 chemicals; Puget Sound Bivalve Apparent Effects Threshold (BAET) values for 23 chemicals. . . . 6-14

Figure 6-3. Concentration-response relationship between total sediment chromium and mortality in the amphipod *Rhepoxynius abronius* and the polychaete *Nephtys caecoides*. 6-17

Figure 6-4. Mortality of (A) the amphipods, *Rhepoxynius abronius* and *Eohaustorius estuarius* and (B) bivalve larvae of *Crassostrea gigas* and *Mytilus galloprovincialis* as a function of total sediment mercury for samples above the Effects Range-Median value (ERM) 6-18

ACKNOWLEDGEMENTS

A project of this scope is possible only with the help of a number of persons who were actively involved in the provision of data and the preparation and review of the document. The efforts of each are greatly appreciated.

Principal investigators of the eight key sediment data sets were: Meg Pinza (Battelle), Jack Ward (Battelle), Nancy Kohn (Battelle), A. Russell Flegal (Institute of Marine Sciences, U.C. Santa Cruz), Robert Risebrough (Institute of Marine Sciences, U.C. Santa Cruz), John Oliver (Moss Landing Marine Laboratories) and Brian Anderson and John Hunt (Institute of Marine Sciences, U.C. Santa Cruz).

These data sets were graciously provided and deciphered by Karen Taberski (San Francisco Bay Regional Water Quality Control Board), Jack Word (Battelle), and Rick Packard (EcoAnalysis). Additional data and technical support were provided by Michelle Hornberger (U.S. Geological Survey) and Richard Swartz (U.S. Environmental Protection Agency, Newport).

Many reviewers provided substantive comments, greatly improving the content and style of this document. These people included Brian Ross (U.S. Environmental Protection Agency), Jack Word (Battelle), Donald Boesch, Tom Ginn (PTI Environmental Services), Robert Risebrough (Bodega Bay Institute), and Tom Dillon and David Moore (U.S. Army Corps of Engineers Waterways Experiment Station).

Finally, it is a pleasure to acknowledge the efforts of our contract managers, Tom Wakeman (formerly of the U.S. Army Corps of Engineers) and Tom Gandesbery and Michael Carlin of the San Francisco Bay Regional Water Quality Control Board, whose interest and patience helped sustain this project throughout its evolution.

1. INTRODUCTION

Background

Management of sediment quality, both locally and nationally, is important to maintain the beneficial uses of our nation's waters. Numerous chemicals enter marine and estuarine environments as a result of human activities, and a large fraction of these accumulate in the sediments, creating a potentially large and uncontrolled source of contamination in aquatic ecosystems. Sediment-associated contaminants may pose environmental risks if left in place or if disturbed through activities such as dredging. Dredging and the aquatic disposal of dredged material may cause the spread of contaminated sediments as well as potentially influence the bioavailability of the chemicals contained therein. Assessment of sediment quality is, therefore, important both to identify contamination as well as to determine appropriate remediation and/or disposal options for dredged sediments.

The state of the art of sediment quality assessment has advanced significantly in recent years, responding to the demand for robust approaches to identify potential biological impacts of sediment-associated contaminants. In particular, the development and proliferation of bioassay tests reflects an understanding that analytical chemistry alone can not provide sufficient information on the ecological risks of contaminants. Sediment bioassays have a distinctive advantage over chemical analysis because they integrate additive and interactive effects of the complex (often unmeasured) mixtures of contaminants which often characterize field sediments. Numerous reviews have been published describing methods for determining the toxicity of marine and estuarine sediments (Adams, *et al.*, 1992; Lamberson, *et al.*, 1992; Samoiloff, *et al.*, 1983).

Toxicity tests are recognized as powerful tools for studying sediment-related contamination, and they are applicable to a wide variety of regulatory and environmental management considerations. These include disposal of dredged material, monitoring of disposal sites, remediation and clean-up activities, and the determination of the temporal and spatial impacts associated with sediment contamination. Moreover, bioassay data have been

used to support the derivation of numerical sediment quality objectives (such as the Apparent Effects Threshold and the Effects Range-low and -median values) used in regulatory decision-making. In a research context, bioassays have been used successfully in combination with chemistry and benthic community surveys to predict and identify *in situ* sediment toxicity (Chapman, *et al.*, 1987; Swartz, *et al.*, 1994).

With increased development and application of chemical and biological tools for sediment toxicity assessment, debate regarding the appropriate application and reliability of these methods has also intensified (Chapman, *et al.*, 1991; Luoma & Carter, 1993; Spies, 1989). As a result, a major research goal has been to improve the predictive nature of these techniques, while at the same time enhance our understanding of the numerous processes that control contaminant bioavailability and toxicity.

Recent developments in environmental and analytical chemistry have included improved methods for determining the bioavailable fraction of contaminants in sediment such as acid volatile sulfide/simultaneously extracted metals (AVS/SEM) analysis (DiToro, *et al.*, 1992) and the extraction and analysis of sediment pore water (Kemp & Swartz, 1988). Research addressing concerns related to bioassay testing has included: determining the influence of non-contaminant related factors such as grain size and ammonia on test organism sensitivity (DeWitt, *et al.*, 1988; Kohn, *et al.*, 1994), demystifying the association between tissue concentrations of contaminants and toxic effects (McCarty & Mackay, 1993), development of bioassays that are based on sublethal endpoints such as genotoxic responses (Anderson, *et al.*, 1994) and enzymatic markers (Spies, *et al.*, 1988), and the refinement of bioassays which evaluate the effects of life-cycle or chronic exposures to sediment contaminants (Moore & Dillon, 1994; Nipper, *et al.*, 1989).

As a highly urbanized estuary, San Francisco Bay has proven to be a challenging region for sediment quality management. An estimated 5,000 to 40,000 metric tons of pollutants are disposed into the Bay each year by over 250 waste water treatment plants and

industrial dischargers (Davis, *et al.*, 1991). Furthermore, the contribution of contaminants associated with urban and agricultural non-point sources appears to be even greater than that from point sources. Because a large proportion of these contaminants are associated with particulate material, this chemical loading has a considerable influence on the quality of sediments in the estuary.

Sediments and their associated pollutants are transported throughout the Bay as a result of both natural and human activities. Overall, tide- and wind-driven resuspension account for the greatest amount of sediment movement, circulating approximately 130 million cubic yards of sediment each year. Dredging activities are another means by which contaminated sediments are redistributed in San Francisco Bay. Approximately 2-10 million cubic yards of material are dredged annually in order to maintain ports and navigation channels at necessary depths in an otherwise shallow estuary. Furthermore, these dredging volumes are expected to increase, as plans exist for substantial deepening of existing berths and channels to accommodate new deep-draft vessels. Because sediment testing is extensively used to determine the potential ecological risks posed by both in-Bay and ocean disposal of this dredged material, these tests have played a prominent role in regulatory decision making.

As testing methods have evolved and national-level debates have gained prominence, new concerns have arisen with respect to the use of these tools for sediment management in San Francisco Bay. A particular focus has been on the interpretation and application of sediment bioassay testing. Questions involving the use of sediment bioassays in San Francisco Bay that have been raised by current data include:

- What are the relationships between various bioassay endpoints and various measures of sediment contamination?
- How robust are the bioassays for identification of impacts due to chemical contamination in sediments?
- What influence do positive interferences have on the results of sediment testing?

- Can suitable reference sites be identified to serve as points of comparison for evaluating sediments tested as part of dredging and monitoring programs?
- Are existing sediment quality guidelines useful and appropriate as a supplement to bioassays for evaluating the potential for toxicity in Bay sediments?

Goals

Previous studies have characterized the extent and magnitude of toxicity associated with sediment contamination in San Francisco Bay, providing important syntheses of historical data from sediment testing programs (Davis, *et al.*, 1991; Long, *et al.*, 1988; Long & Markel, 1992; BTPC, 1993). While several well-defined "hot spots" exist at which consistent and clear causes of toxicity have been identified, a majority of the Bay falls into a "grey area" in which toxicity is sometimes observed but often with no apparent cause or relationship to *in-situ* impacts. Peripheral areas such as harbors and marshes often contain the highest concentrations of contaminants and sporadically high toxicity.

There remain, however, numerous unresolved concerns regarding not only the distribution and impacts of sediment contamination, but also the application and interpretation of sediment testing in San Francisco Bay. As new data are being generated by monitoring and regulatory programs, there is a large amount of information with which to examine, in greater detail, some of the current issues relating to sediment testing in the Bay.

The goals of this report are:

- to provide additional characterization of sediment toxicity and chemistry in San Francisco Bay using recently collected, high quality data,
- to evaluate the extent to which current research themes related to sediment can be applied to San Francisco Bay,
- and, in light of the above, to evaluate the certainty with which relationships between contamination and biological effects can be elucidated from this data.

Scope

Recently collected data from several different studies was used in order to examine topics of current concern with respect to sediment quality management in San Francisco Bay. Each chapter provides general and site-specific background on a particular topic, a description of the approach used for data analysis, results, and discussion of the key findings relative to what is known for San Francisco Bay. The criteria used for study selection, as well as a brief description of the different sites, is presented below.

Study selection and site characteristics

Sediment chemistry and toxicity data were taken from eight studies. A brief description and reference for each study are provided in Table 1-1. Studies fell into two general groups based on the source of the data.

- **Harbor Studies** - pre-dredging studies performed on sediments from the Ports of Oakland and Richmond between 1990-1991 by Battelle/Marine Sciences Laboratory (Figures 1-1 through 1-4).
- **SedQual 3 Studies** - data collected in the San Francisco Bay 1991-1992 Regional Monitoring Program sponsored by the San Francisco Bay Regional Water Quality Control Board (SFBRWQCB) (Figures 1-5 through 1-7). Measurement of trace metals, organics, and sediment toxicity were performed in separate laboratories.

Table 1-1. Description of the eight San Francisco Bay studies used in this analysis.

Location	Sampling Dates	Contracting Laboratory/Reference
Harbor Studies		
1. Oakland Harbor Berth Areas	March, 1990	Battelle/Pacific Northwest Laboratory for the Port of Oakland (Kohn et al., 1992a)
2. Oakland Inner Harbor (Phase 3A)	June, 1990	Battelle/Pacific Northwest Laboratory for the ACOE(a) (Ward et al., 1992)
3. Oakland Inner Harbor (Phase 3A retest)	September, 1990	Battelle/Pacific Northwest Laboratory for the ACOE(a) (Ward et al., 1992)
4. Oakland Outer Harbor (Phase 3B)	November, 1990	Battelle/Pacific Northwest Laboratory for the Port of Oakland (Kohn et al., 1992b)
5. Richmond Harbor	June, 1991	Battelle/Pacific Northwest Laboratory for the ACOE (Pinza et al., 1992)
SedQual 3 Studies		
6. Castro Cove	May, 1991	Various Laboratories (b) for the SFBRWQCB(c) (SFBRWQCB, 1994)
7. Critical Habitats (Marsh studies)	July, 1991 October, 1991 November, 1991 February, 1992	Various Laboratories for the SFBRWQCB (SFBRWQCB, 1994)
8. SF Bay Studies (Bay Run #1)	August, 1991	Various Laboratories for the SFBRWQCB (SFBRWQCB, 1994)
(Bay Run #2)	April, 1992	

(a) U.S. Army Corps of Engineers

(b) Participating laboratories in the Regional Monitoring Program include: UC Santa Cruz, UC Berkeley, Lawrence Berkeley Laboratories, CA Department of Fish and Game, Moss Landing Marine Laboratories, Granite Canyon Marine Laboratories, US Geological Survey, and Ecoanalysis.

(c) San Francisco Bay Regional Water Quality Control Board

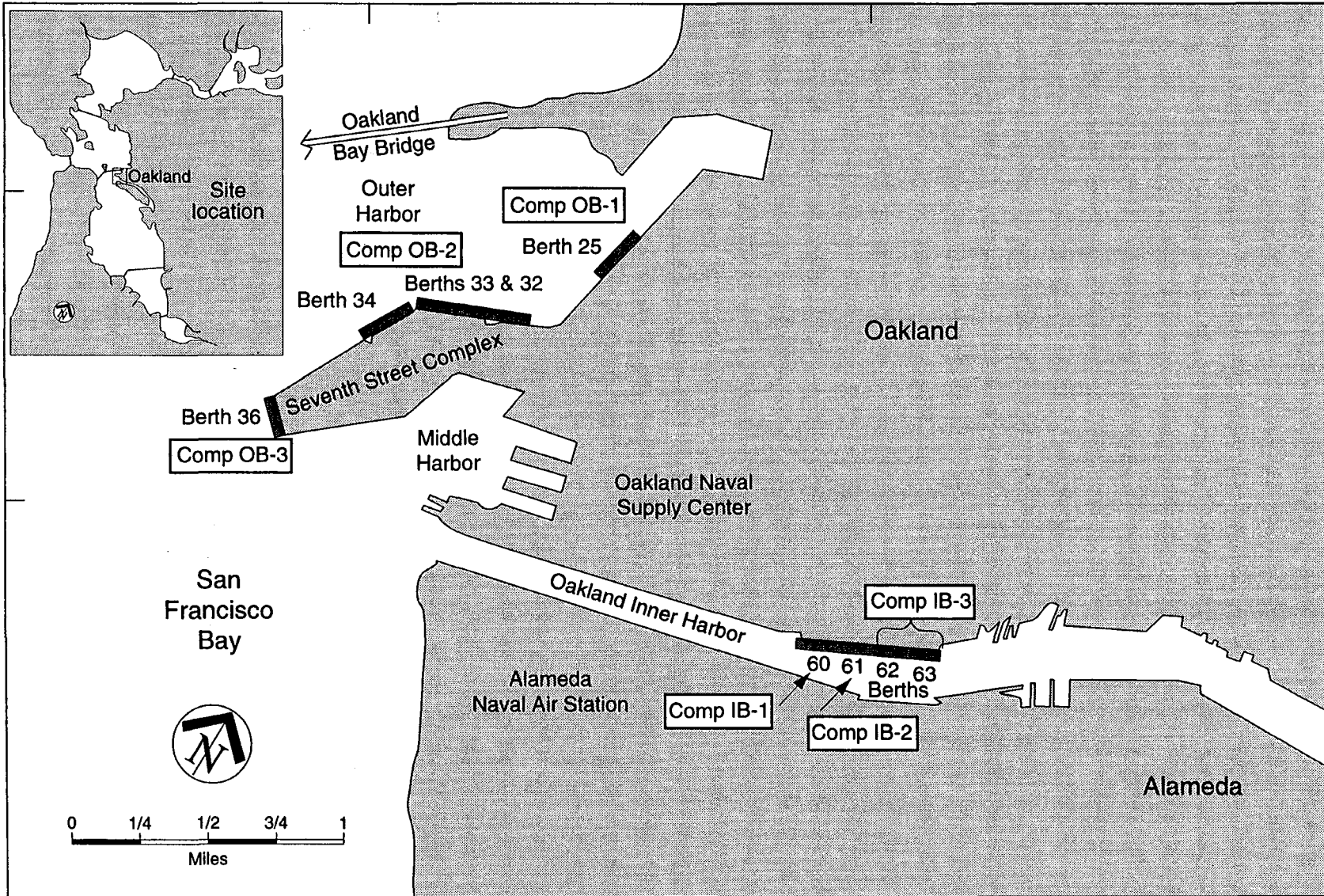


Figure 1-1. Location of composite samples taken for the Port of Oakland Berth Area study during March 1991 (adapted from Battelle, 1992a).

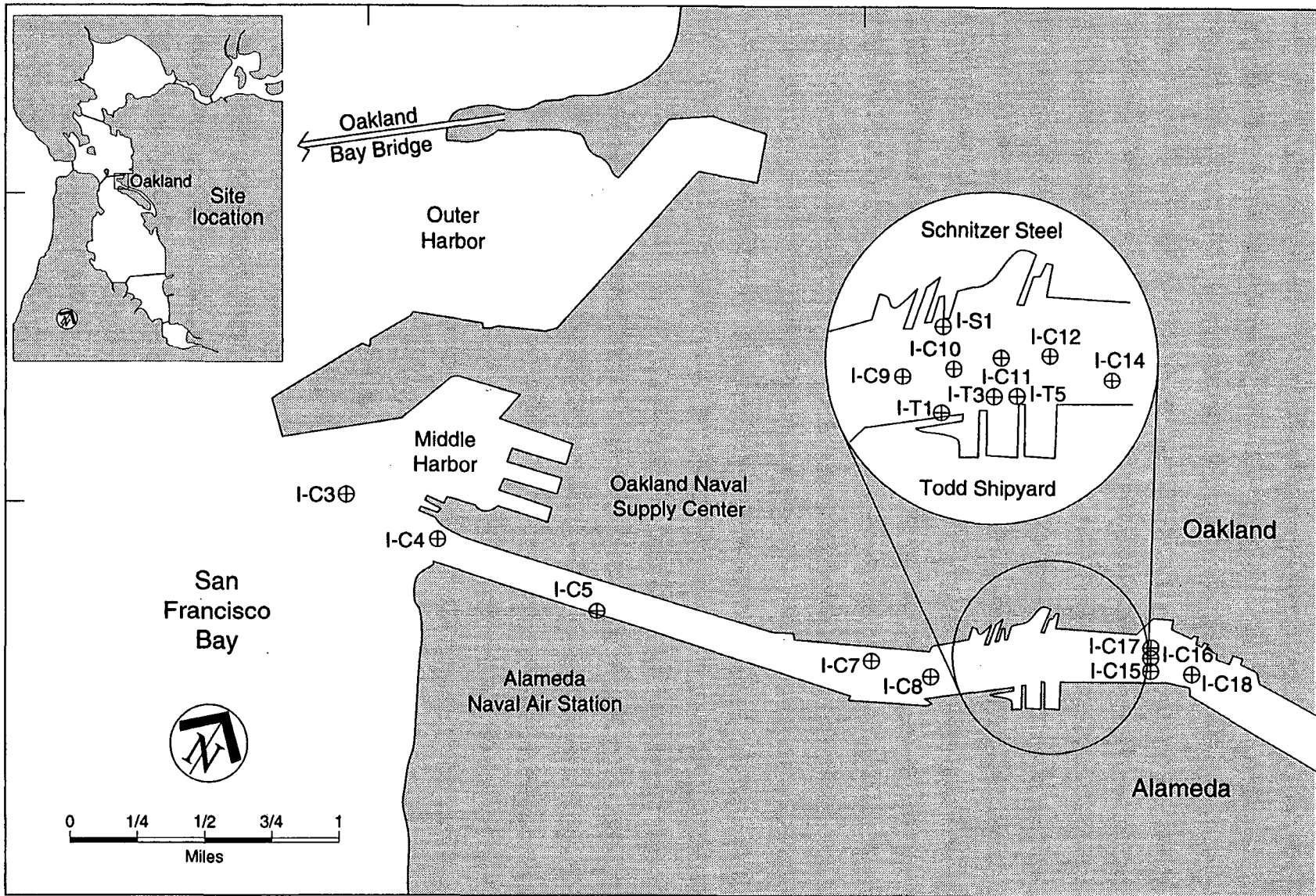


Figure 1-2. Location of samples taken for the Port of Oakland Inner Harbor studies (Phase 3A and Phase 3A retest) during June and September 1990 (adapted from Battelle, 1992c).

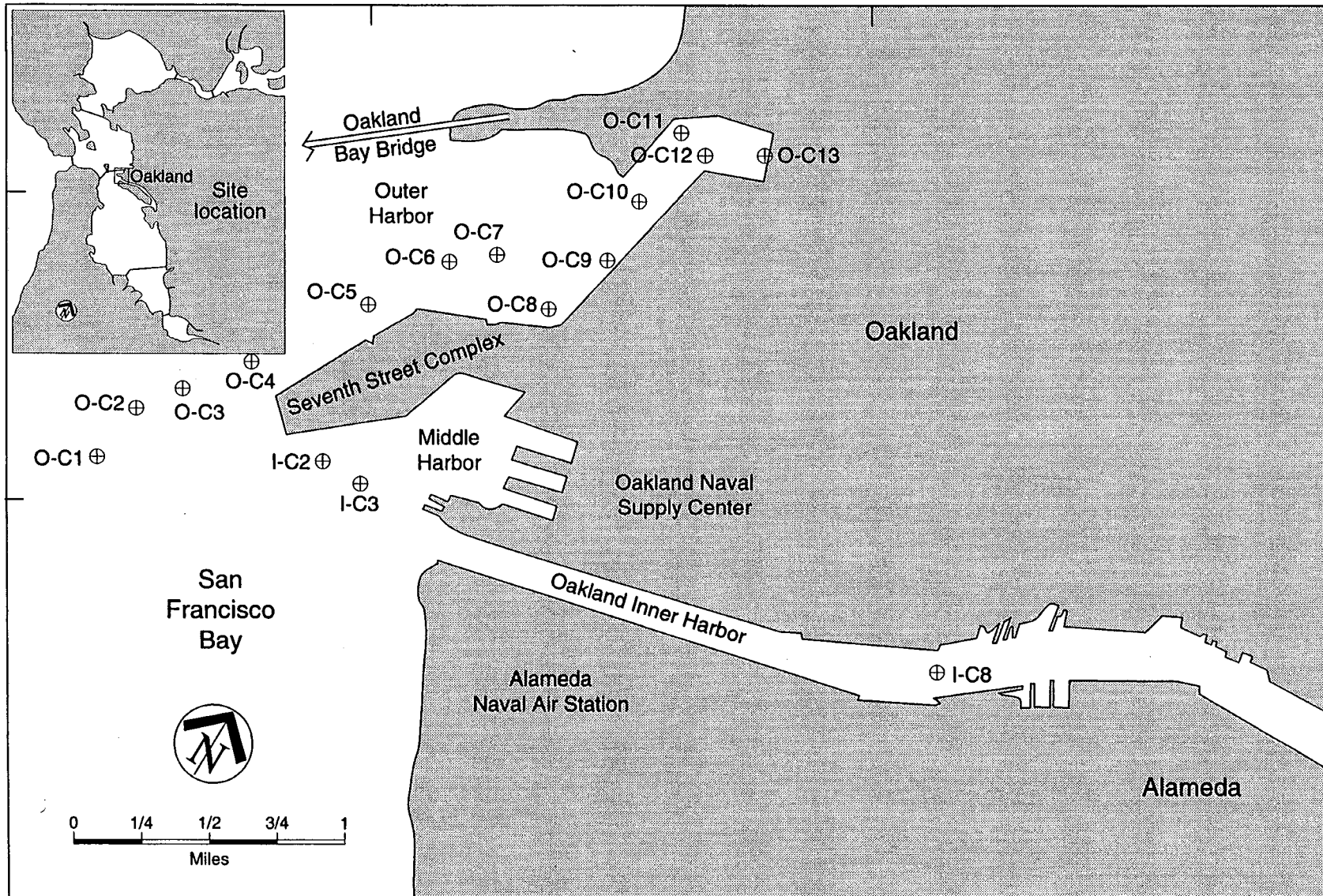


Figure 1-3. Location of samples taken for the Port of Oakland Outer Harbor study (Phase 3B) during November 1990 (adapted from Battelle, 1992b).

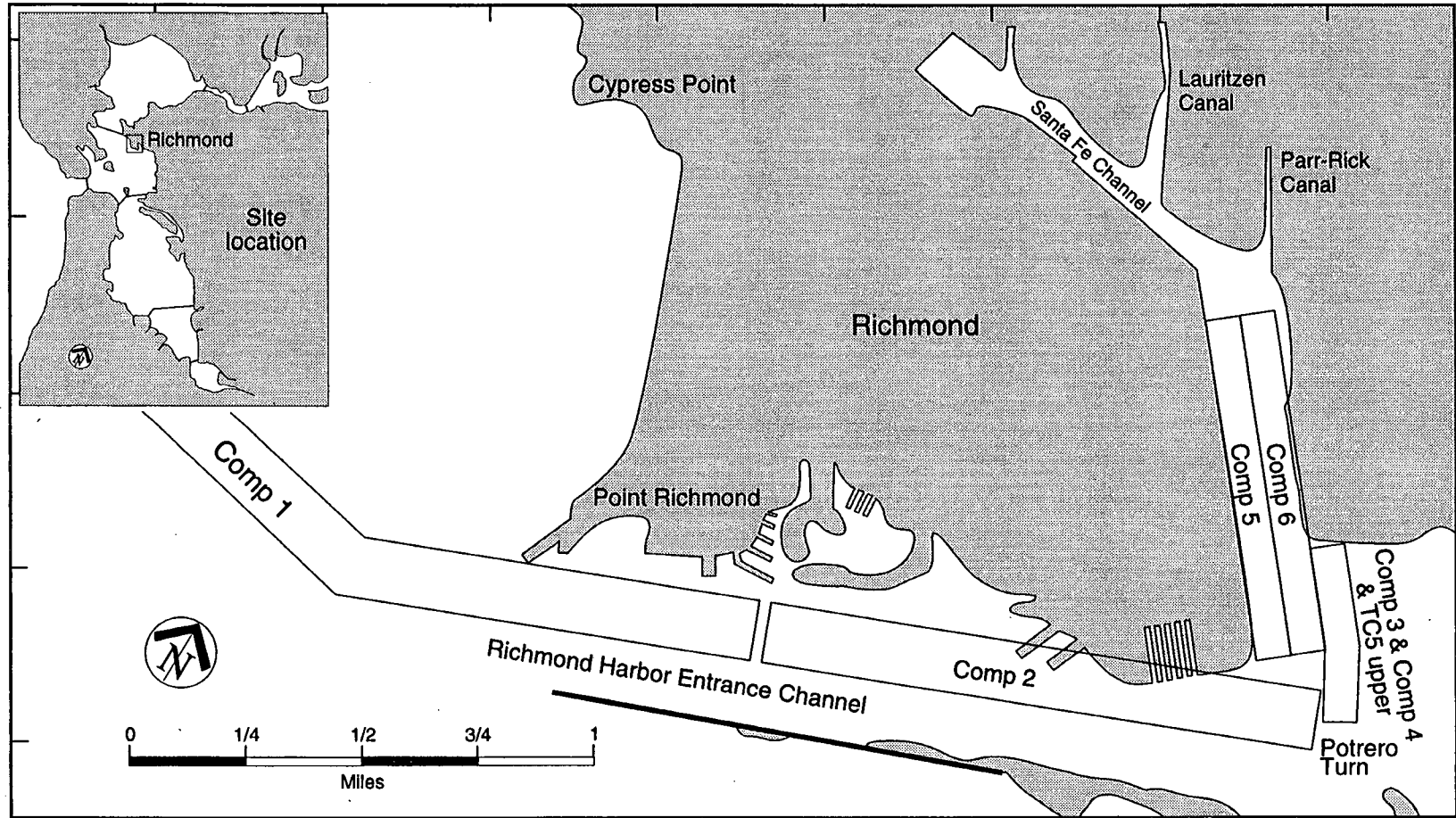


Figure 1-4. Location of samples taken for the Port of Richmond deepening study during June 1991 (adapted from Battelle, 1992d).

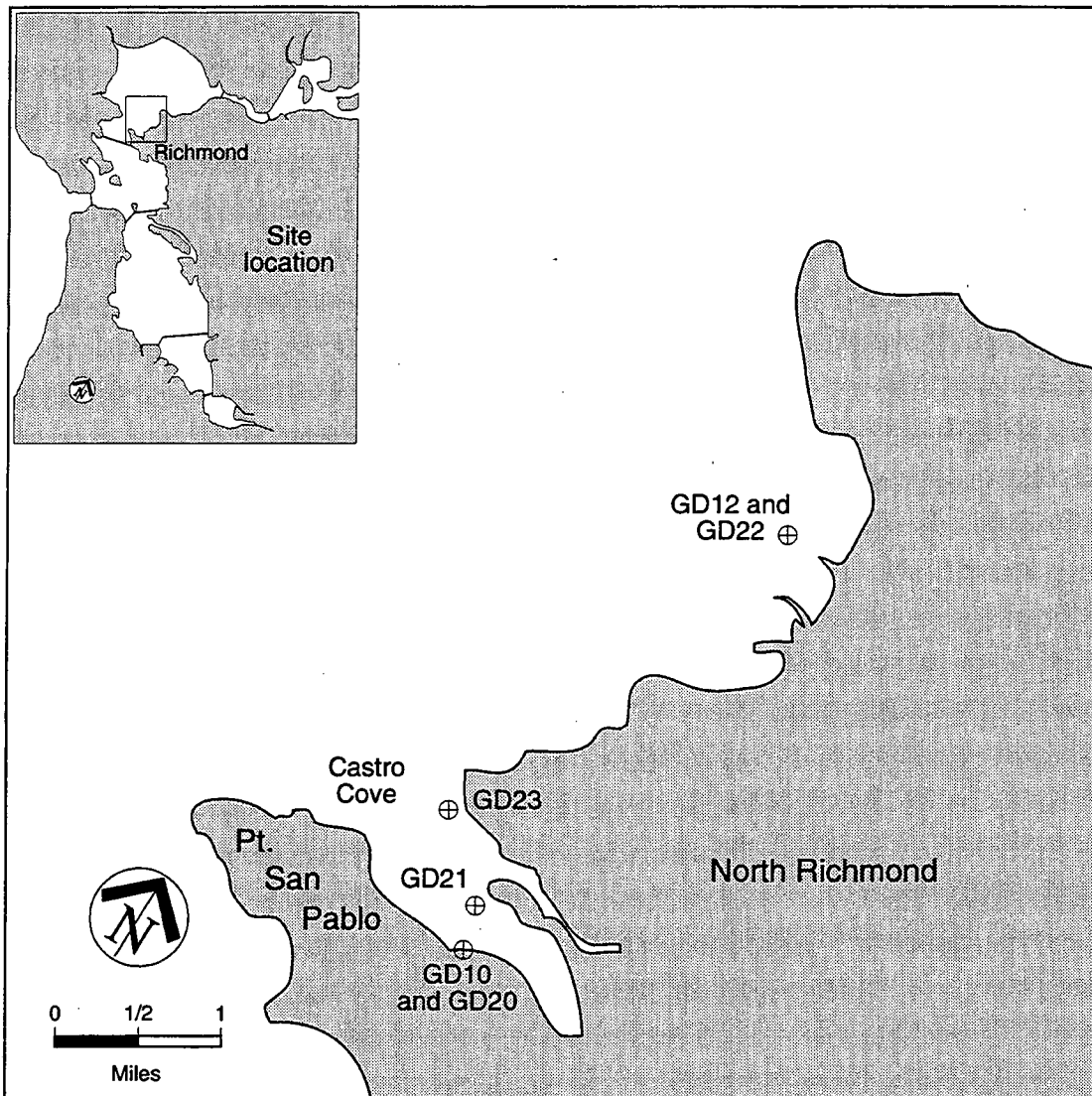


Figure 1-5. Location of samples taken for the SedQual 3 Castro Cove study during May 1991. Note that sites GD10/GD20 and GD22/GD12 indicate deep and depositional layer cores taken from the same location (adapted from SFBRWQCB, 1994).

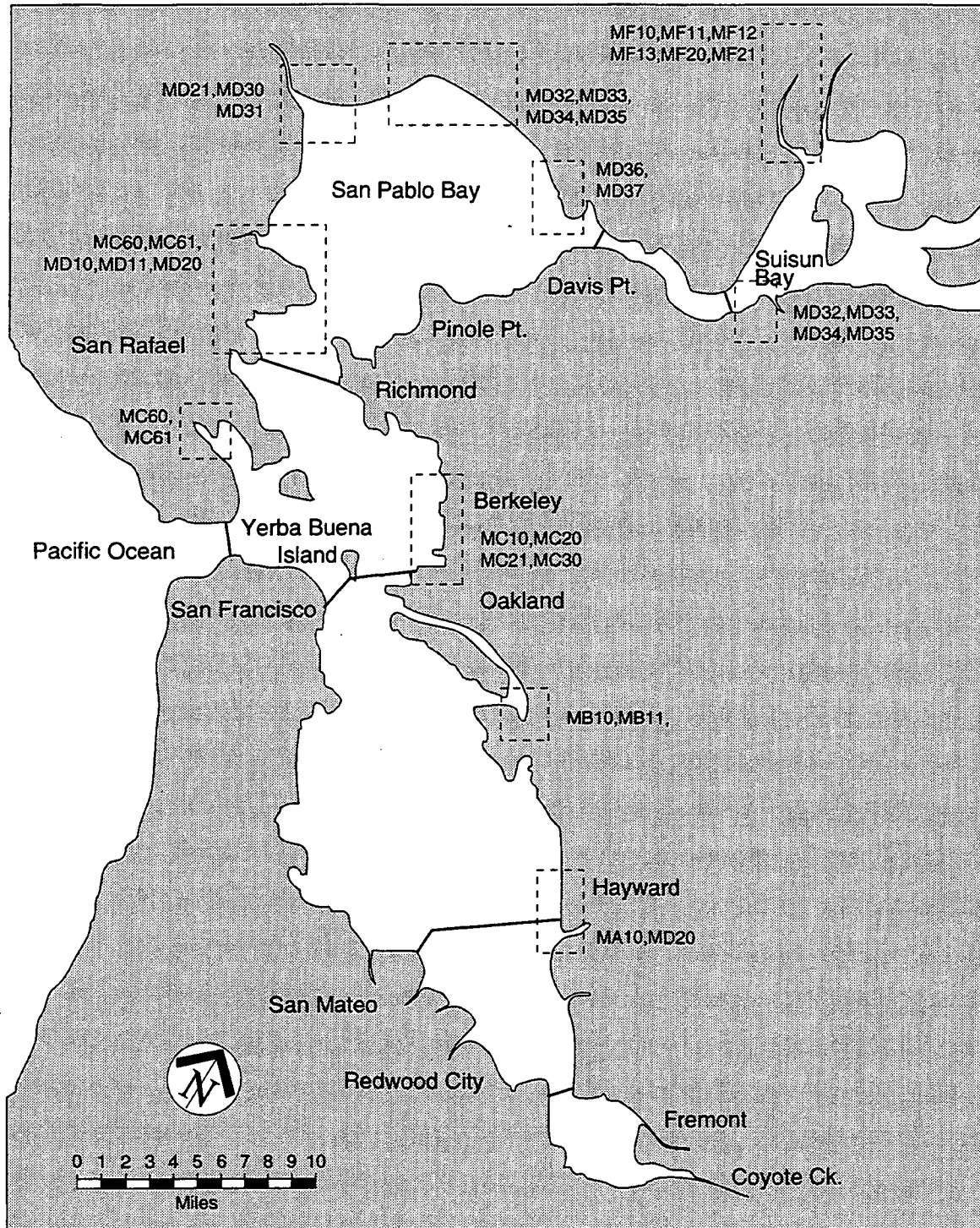


Figure 1-6. Location of samples taken for the SedQual 3 Marsh studies on multiple dates in 1991 and 1992 (adapted from SFBRWQCB, 1994).

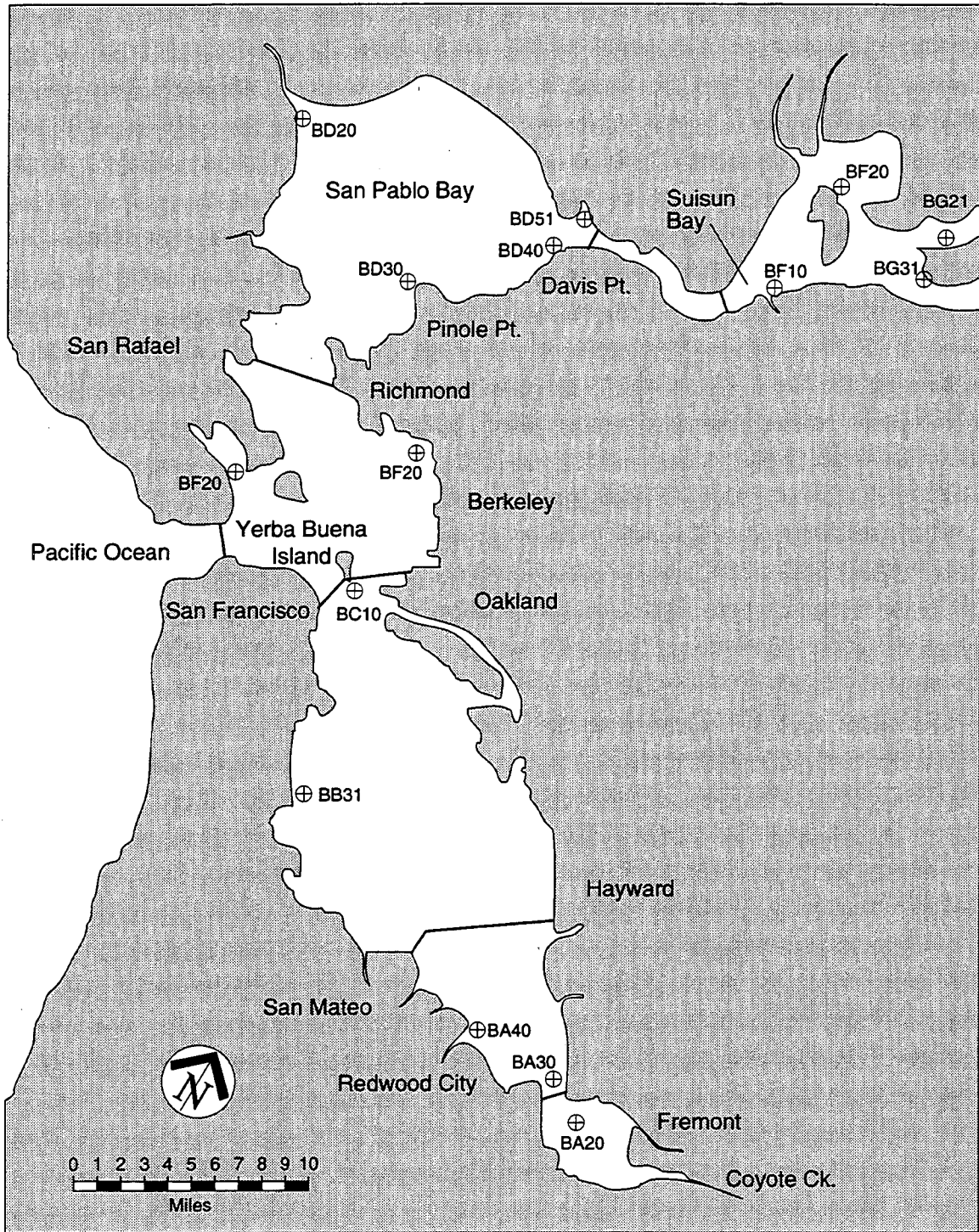


Figure 1-7. Location of samples taken for the SedQual 3 Bay Run studies #1 and #2 during August 1991 and April 1992 (adapted from SFBRWQCB, 1994).

The following were general criteria for study selection:

- recently conducted (1990-1992),
- paired bioassay and sediment chemistry testing,
- statistically significant toxicity observed with at least one test species in multiple samples,
- sediment bioassays include both solid- and elutriate-phase tests,
- sediment chemistry analysis include trace metals, polynuclear aromatic hydrocarbons (PAHs), pesticides, and polychlorinated biphenyls (PCBs).

The resulting data set includes sediments from a wide range of different locations (mid-Bay, marshes, mudflats, navigation channels, port berths, and in-Bay and offshore reference sites) representing areas both proximal to and removed from known sources of contamination.

The chemical and physical parameters measured were consistent within the Harbor and SedQual 3 groupings, but not between them. The largest differences in analytes measured were for the chlorinated pesticides for which 19 out of a total of 29 pesticides were unique to one of the two study groups. Furthermore, the components of chemical classes such as total low molecular weight PAHs (LPAH), total high molecular weight PAHs (HPAH), and total PCBs were not comparable for the two study groups. For example, several LPAHs (naphthalene, acenaphthylene, acenaphthene, and fluorene) were measured in the Harbor studies but not in SedQual 3, whereas several HPAHs (such as benzo(e)pyrene and methylphenanthrenes) were only measured in the SedQual 3 analysis. A complete list of the physical and chemical parameters analyzed in each grouping is provided in Table 1-2.

Although all the studies had solid-phase and elutriate toxicity testing components, the actual test species used and extent of testing were not consistent between studies. For example, all of solid-phase testing in the SedQual 3 studies was done using either the estuarine amphipod, *Eohaustorius estuarius* or the marine amphipod, *Rhepoxynius abronius*. In contrast, solid-phase testing in the Harbor studies was performed using several different

Table 1-2. Physical and chemical parameters reported in San Francisco Bay sediment data sets.

Measurements	In Common	SedQual 3 (unique)	Harbor Studies (unique)
Physical	% sand		% gravel % silt % clay
LPAH	Anthracene Phenanthrene		Acenaphthene Fluorene Phenanthrene Anthracene Naphthalene Acenaphthylene
HPAH	Fluoranthene Pyrene Benzo(a)anthracene Chrysene Benzo(b,k)fluoranthene Benzo(a)pyrene Dibenzo(a,h)anthracene Benzo(g,h,i)perylene Indeno(1,2,3-c,d)pyrene	Benzo(e)pyrene Methylphenanthrene	
Pesticides	Alpha-BHC Beta-BHC Delta-BHC Gamma-BHC Heptachlor epoxide Total chlordane p,p'-DDD p,p'-DDE p,p'-DDT	o,p'-DDD o,p'-DDE o,p'-DDT p,p'-DDMU Oxychlordane Gamma-Chlordane Alpha-Chlordane cis-Nonachlordane trans-Nonachlor Hexachlorobenzene	Endosulfan I & II Endosulfan sulfate Endrin Endrin aldehyde Dieldrin Methoxychlor Toxaphene Aldrin
Trace metals	Arsenic Cadmium Chromium Copper Lead Mercury Nickel Silver Selenium Zinc	Cobalt	
Other	Total organic carbon Total PCB Tributyltin	Total organic nitrogen	Monobutyltin Dibutyltin Oil and Grease Total petroleum hydrocarbon Total volatile solids

species: the amphipod (*R. abronius*), the infaunal polychaete (*Nephtys caecoides*), and the sanddab (*Citharichthys stigmaeus*). The elutriate tests in the SedQual 3 studies used either larvae from the pacific oyster (*Crassostrea gigas*) or the bay mussel (*Mytilus galloprovincialis*). In addition to oyster and mussel larvae, the mysid shrimp, *Holmesimysis sculpta* and the sanddab (*C. stigmaeus*) were used for elutriate-phase testing in the Harbor studies. A list of the test species considered in this report as well as the studies in which they were used is provided in Table 1-3.

Several of these studies had additional chemical and biological testing data that was used in this report. For example, each of the Harbor studies included a 28-day solid-phase bioaccumulation test using the bentnose clam (*Macoma nasuta*) and a polychaete (*N. caecoides*). The Castro Cove component of SedQual 3 included analyses of trace metal contamination and toxicity in pore water samples.

Generally, sediment handling and testing protocols were similar between the Harbor and SedQual 3 studies. However, because these studies represent the independent efforts of several different laboratories, the degree and nature of the quality assurance and quality control is not consistent among studies. Potential areas of interlaboratory variability include sediment holding times, sample preparation, chemical analysis methodologies, and test organism quality and holding.

Table 1-3. A summary of the test species and testing conditions used in this report.

Test Species	Lifestage	Endpoint	Exposure Period	Sediment Phase	Study (a) Reference
Amphipod <i>Rhepoxynius abronius</i>	adult	survival	10 day	solid/static (b)	1,2,4,5,8
Amphipod <i>Eohaustorius estuarius</i>	adult	survival	10 day	solid/static	6,7,8,9
Polychaete worm <i>Nephtys caecoides</i>	adult	survival bioaccumulation	10 day 28 day	solid/FT(c) solid/FT	1,3,4,5 1,3,4,5
Bentnose clam <i>Macoma nasuta</i>	adult	survival bioaccumulation	10 day 28 day	solid/FT solid/FT	1,3,4,5 1,3,4,5
Speckled sanddab <i>Citharichthys stigmaeus</i>	juvenile	survival	10 day 96 hrs	solid/FT elutriate	1,2,4,5 1,2,4,5
Bay mussel <i>Mytilus galloprovincialis</i>	embryo	survival development	48 hr	elutriate	4,5,7 4,5,7
Pacific oyster <i>Crassostrea gigas</i>	embryo	survival development	48 hr	elutriate	1,2,6,7,9 1,2,6,7,9
Mysid shrimp <i>Holmesimysis sculpta</i>	adult	survival	96 hr	elutriate	1,2,4,5
Purple sea urchin <i>Strongylocentrotus purpuratus</i>	sperm/egg embryo	fertilization development	1 hr 48 hr	pore water	6 6

(a) Key to study numbers found in Table 1-2.

(b) Static test system

(c) Flow-through test system

2. PATTERNS OF TOXICITY IN FOUR GENERAL SITE CLASSES

Introduction

Sediment bioassays are an important tool for sediment quality management. They are used extensively to monitor *in situ* sediment quality as well as to make regulatory decisions about the ecological hazards associated with the disposal of dredged material. In recent years, a large amount of toxicological data has been generated as part of monitoring and regulatory programs in San Francisco Bay. It follows that there is the need for a general characterization of the degree and extent of toxicity observed in this new data, because there is little known about the actual distribution of toxicity among various habitat types.

Goals

The goals of this analysis are to assess the frequency and magnitude of bioassay toxicity in this San Francisco data set as well as to determine the extent to which patterns of toxicity are predicated on the influences of contaminant sources at different sites. This analysis does not, however, include a formal comparison among the sites since the data are from different studies, each having independent sampling designs.

Approach

Bioassay data from eight of the studies in the San Francisco Bay data set (described in Section 1) were divided into four site groupings characterizing a range of sediment types and degrees of contamination:

- harbor navigation channels,
- marshes,
- mid-Bay,
- and an area known to be highly contaminated (ie., Castro Cove).

The number of bioassay species and type of exposures (solid-phase and elutriate) performed in a study were generally consistent within each site grouping¹.

The frequency and magnitude of toxic response was determined for each of these four types of sites by dividing the response endpoints (survival and normal development) into four impact categories: extremely high effects (75-100% mortality/abnormality), high effects (75-50% mortality/abnormality), moderate effects (50-25% mortality/abnormality), and low/no effects (25-0% mortality/abnormality)². We then identified the number of samples for which response of *at least one* test species fell into the extremely high, high, and moderate categories. All other samples were counted in the low/no toxic response category. Summed, the percentages for all four effects categories constitute unity. We also determined the number of samples for which responses in *at least two* species fell within the extremely high, high, and moderate categories. Excluded from this second calculation were samples in the low/no toxic effects category.

Results

Many of the samples in our data set elicited either moderate toxic effects (50-25% mortality/abnormality) or low/no effects (25-0% mortality/abnormality) in only one test species. Specifically, in the harbor navigation channels and mid-Bay groupings, 71-86% of the samples were of either moderate or low toxicity to one test species (Figure 2-1). For marsh samples (Figure 2-2a), the magnitude of toxicity was more evenly distributed over all the effects ranges. Only at the Castro Cove site, did a majority (67%) of the samples fall into the extremely toxic category (Figure 2-2b).

¹ One exception was the samples from the Oakland Inner Harbor Phase 3A retest study which were tested using only two species.

² It is important to emphasize that categorization of samples into impact categories was not contingent on statistical significance. Thus, there were sediments in the low/no effects category for which survival was significantly decreased relative to that observed in the control.

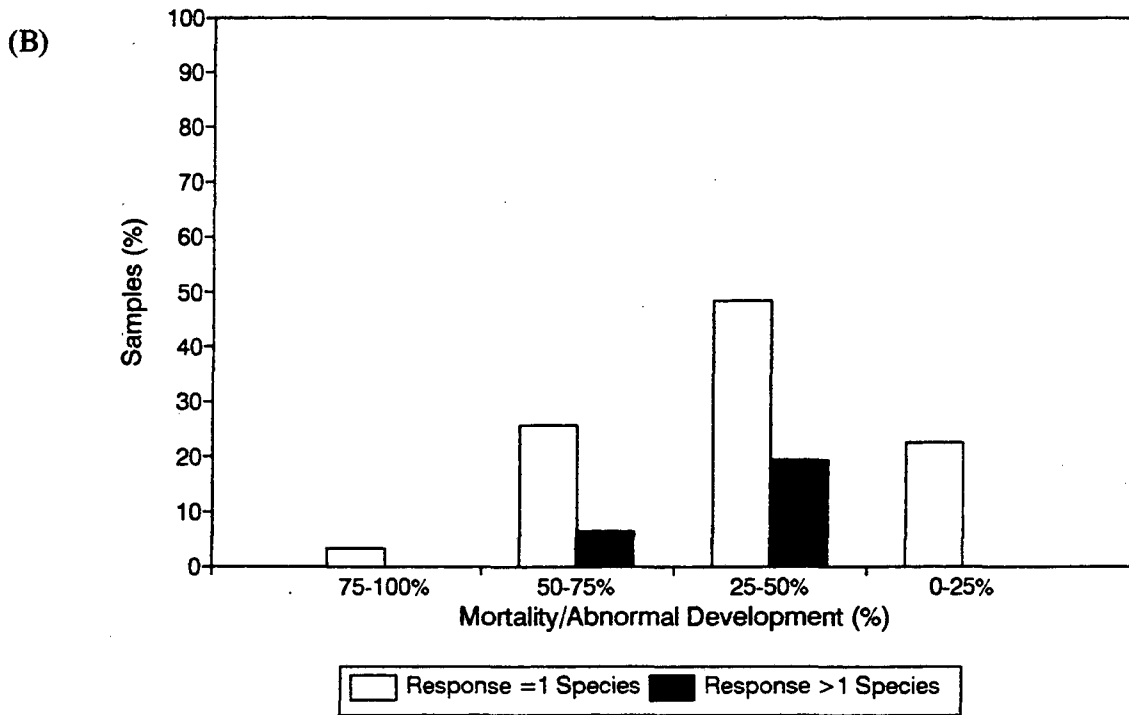
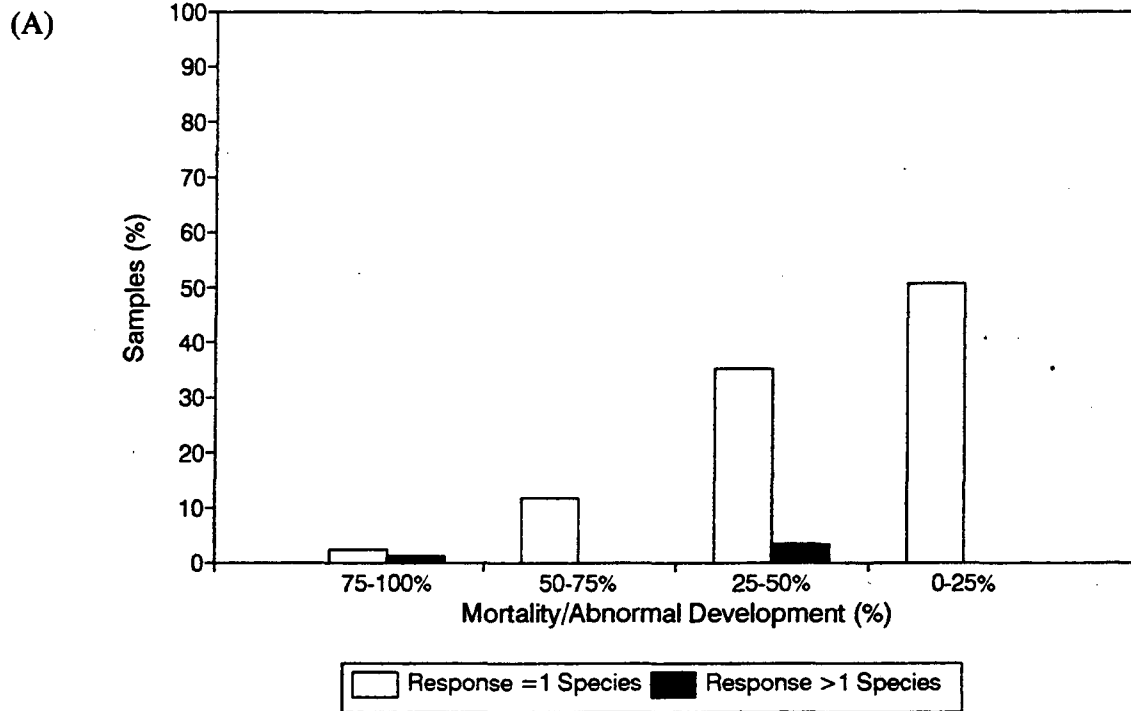


Figure 2-1. Magnitude of toxicity observed in sediment samples from (A) harbor navigation channels (n=69) and (B) mid-Bay stations (n=30) in San Francisco Bay.

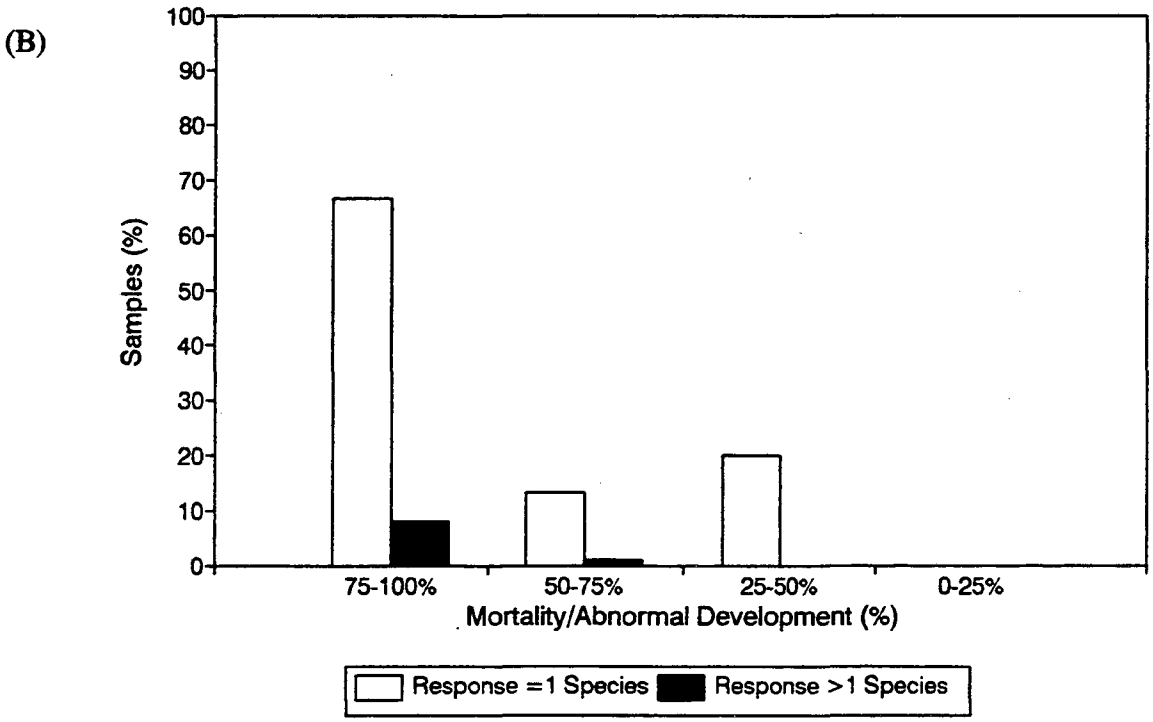
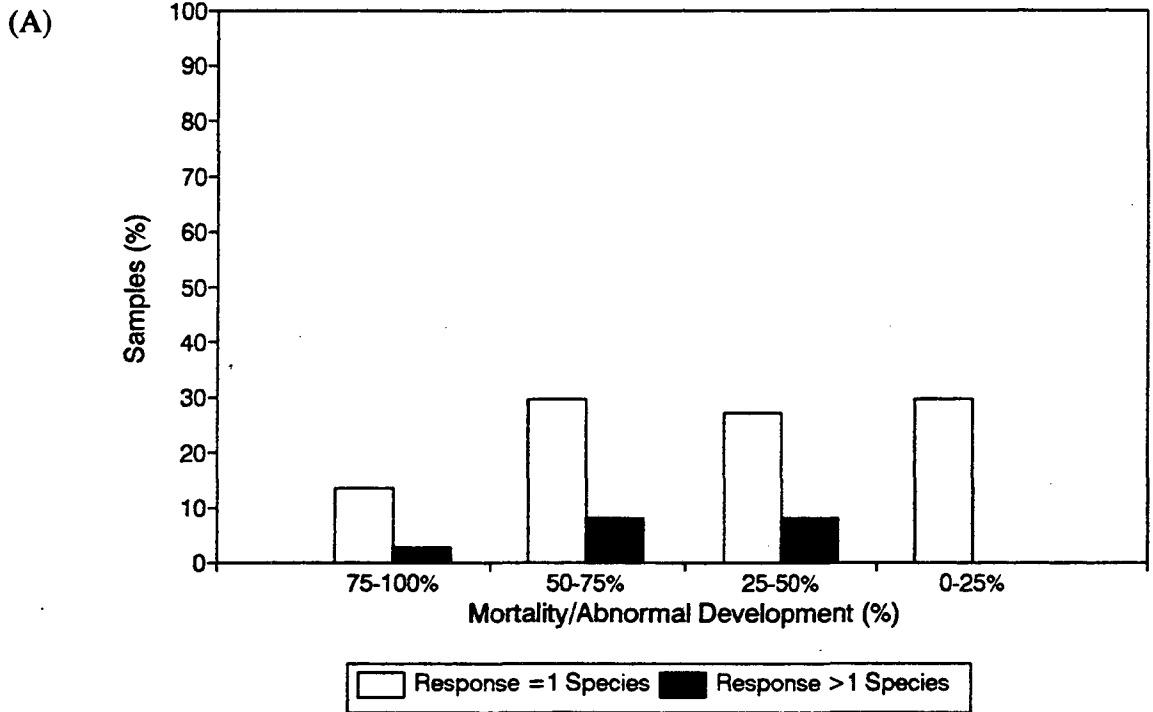


Figure 2-2. Magnitude of toxicity observed in sediment samples from (A) marshes (n=34) and (B) Castro Cove (n=30).

It is important to note that although the results above are expressed in terms of effects on a single test species, this was not consistently the same test species. The particular toxicity endpoint (e.g., bivalve abnormal development, amphipod and polychaete mortality) that determined the highest effects category of a given sample often differed between samples within a site grouping.

Overall, there was very little concordance among test species in the magnitude of their response to the sediment samples (represented by filled bars in Figures 2a-d). For all four site groups, the percentage of samples that were similarly toxic to more than two species was low (0-19%).

Discussion

Generally, we found that the distribution of toxic response in the San Francisco Bay sediment corresponded well with what might be expected, given the general characteristics of the four site groupings.

Harbor navigation channels: In general, sediment samples taken from the harbor navigation channels (Figures 1-2 through 1-4) are not expected to be highly toxic, despite their proximity to industrial activity. Because these navigation channels are frequently dredged, the near-surface sediments result from short-term deposition and are, therefore, less likely to be contaminated. The deeper sediments, when these have never been dredged or exposed, generally have low levels of contamination³. Of the 69 navigation channel samples considered in this analysis, 86% fell into either the low/no or moderate effects categories.

Results of toxicity testing from the Oakland Harbor Berth Areas (Figure 1-1) study were not included in this analysis, because these areas represented a substantially different

³Although all four Harbor studies used in this analysis were tests of material from deepening projects, the depths represented by these samples varied considerably between studies, with the Oakland Inner Harbor samples consisting of deeper sediments (38-44ft below water level) and most of the Oakland Outer Harbor and Richmond Harbor sediments consisting of both the shallow and deeper sections of the cores (mudline to 44ft below water level).

contamination regime. In general, sediments from berth areas are more likely to be contaminated and, therefore, toxic due to their close proximity and susceptibility to land-based activities. Because berths are often dredged less frequently than navigation channels, there is also a greater chance for contaminants to accumulate over time. The berth data was not analyzed as an individual site grouping, because the number of samples was too small (n=6).

Mid-Bay: Sediments sampled as part of the SedQual 3 Bay Studies #1 and #2 (n=30 per study) were primarily from mid-Bay areas off the main navigation channels (Figure 1-7). Many of the sampling stations were located in areas which are removed from known sources of contamination. Thus, we expected (and observed) toxicity to be normally distributed, with the majority of the samples having low to moderate toxic effects.

Marshes: The marsh samples from SedQual 3 were collected at 34 stations located in both mudflats and in marshes, some of which were proximal to point sources of contamination such as storm drains (Figure 1-6). The even distribution of toxicity that was observed may reflect the varied influence of anthropogenic factors at these marsh sites.

Castro Cove: The six stations sampled at Castro Cove were located on a presumed gradient of contamination moving away from the historic outfall of a petroleum refinery (Figure 1-5). High concentrations of both organic compounds and trace metals measured in sediments at several of these stations are reflected in the high to extremely high toxicity observed in 80% of the samples (n=30).

Although many sediment samples were toxic to more than one of the species tested, it was rare that the magnitude of toxic response was equal for any two species. This pattern is particularly evident in the Castro Cove study. Although a high percentage (80%) of samples elicited greater than a 50% toxic effect to at least one species, less than 10% of these samples were similarly toxic to any two test species. The extent of interspecies differences in the magnitude of their toxic response to sediments underscores the need for multispecies toxicity testing to comprehensively characterize the ecological hazards posed by sediment-associated

testing to comprehensively characterize the ecological hazards posed by sediment-associated contaminants.

Conclusions

Results of this analysis demonstrate that there are variations in the magnitude and distribution of toxicity between harbor navigation channels, mid-Bay sites, marshes, and a highly contaminated area. Patterns in toxic response generally reflected the relative influence of anthropogenic sources of contamination at these different types of locations. Toxicity testing with multiple species is essential to elucidate such patterns.

3. CORRELATIONS BETWEEN TOXICITY AND BULK SEDIMENT CHEMISTRY

Introduction

Previous analyses of San Francisco Bay data have come to differing conclusions regarding the influence of anthropogenically-derived sediment contamination on toxicity observed in sediment bioassays. Several studies have reported that there are few, if any, correlations between bulk measures of individual sediment contaminants and toxicity (Davis, *et al.*, 1990; Risebrough, 1994). These studies have suggested that non-contaminant parameters such as total organic carbon, total volatile solids, and grain size have a greater influence than do contaminants on toxicity observed in sediments. In the latter study, however, the identification of potential causes of toxicity was confounded by the high degree of correlation among the various chemical parameters themselves. While Long and Buchman (1989) also observed relatively strong relationships between the TOC content of sediments and their toxicity to bivalve larvae and amphipods, similarly strong correlations were observed for numerous contaminants. Long and Markel (1992), using information from a relatively large data base, observed strong and very significant correlations between various organic and inorganic sediment chemicals and toxicity, while correlations using TOC and sediment grain size, when significant, were generally weak.

Goals

Since the publication of the above studies, additional synoptic sediment chemistry and toxicity data have become available. The goal of this analysis was to examine recently-collected data to determine the extent to which correlations between toxicity and the bulk sediment concentrations of organic compounds and trace metals were observable. In addition, we evaluated the degree to which correlations between non-contaminant parameters and sediment contaminants obscure the identification of factor(s) responsible for toxicity.

Approach

To evaluate relationships between toxicity and chemical contamination, correlation analyses were conducted on data from the SedQual 3 and Harbor studies (described in Section 1) using Spearman's nonparametric rank correlation (r_s) (Daniel, 1990). Rank correlation coefficients were also used to measure the strength of association between concentrations of chemical parameters and the organic carbon content of the sediment. A non-parametric statistical procedure was used, because preliminary evaluations of the SedQual 3 data set indicated that these data do not meet the normality criteria required for use of parametric procedures (Risebrough, 1993). Correlation analyses were conducted using individual chemicals as well as various classes of organic compounds such as low molecular weight and high molecular weight polynuclear aromatic hydrocarbons (LPAH and HPAH) and total PCBs¹. Concentrations of all organic compounds were normalized to the organic carbon content of the sample (OC-normalized); whereas, trace metal concentrations are dry-weight values. For correlation analysis, studies were grouped in the following manner:

- Rank correlations were calculated on a per-study basis for the Harbor studies (Figures 1-1 through 1-4) and the Castro Cove study (Figure 1-5).
- Additional correlations were calculated for each of the Harbor studies using only those samples with greater than 0.2% total organic carbon (TOC). This was important because many samples had extremely low carbon content such that OC-normalization may have overestimated the potentially bioavailable fraction.
- Correlation analysis using data from the four Critical Habitat Studies (SedQual 3; Figure 1-6) was performed on a combined data set. These studies are hereafter collectively referred to as the Marsh study.

¹ Relationships between grain size variables and sediment toxicity samples are addressed separately in Section 5

- Correlations for the two Bay Run studies (SedQual 3; Figure 1-7) were performed on individual as well as the combined data sets. Sediment samples from Bay Run #1 and Bay Run #2 were collected from the same stations in August 1991 and April 1992.

All studies in the San Francisco Bay data set included sediment analysis of PAHs, PCBs, chlorinated pesticides, and trace metals. However, the component chemicals of these classes varied between the Harbor and SedQual 3 studies, particularly for the pesticides and LPAHs (See Table 1-2 in Section 1). For example, compounds that were unique to SedQual 3 include several chlordane pesticides, hexachlorobenzene, and total organic nitrogen. Compounds that were measured only in the Harbor studies include pesticides such as aldrin, dieldrin, toxaphenes, and endosulfan, as well as several LPAHs such as fluorene, acenaphthylene, naphthalene, and acenaphthene. These analytical differences must be taken into consideration when comparing the results of correlation analysis between SedQual 3 and Harbor studies.

Results

Overall, significant rank correlations with toxic effects were observed for numerous individual chemical parameters in both the Harbor and SedQual 3 studies (Tables 3-1 and 3-2). Upon closer examination, however, these correlations were rarely demonstrative of cause-effect relationships. Below, results of the correlation analyses are presented individually for the Harbor and SedQual 3 groups.

Harbor studies: The rank correlations between toxicity and concentrations of numerous chemical and nonchemical parameters were relatively strong and significant in samples from the Harbor studies (Table 3-1). Of the five studies, the Oakland Harbor berth areas had the highest number of significant rank correlations between sediment contaminants and toxicity. When all 12 samples were included in the analysis, 19 chemicals were strongly associated with toxicity in at least one test species ($p < 0.05$). Limiting the analysis to only

Table 3-1. Significant Spearman rank correlations for the Harbor studies

Study (sample size)	Chemical Parameter	Correlation Coefficients		
		<i>N. caecoides</i> (% mortality)	<i>R. abronius</i> (% mortality)	bivalve larvae [a] (% abnormal)
Oakland Harbor				
Outer Harbor (Phase 3B) Amphipod (n=22) Bivalve (n=7)	Tributyltin Lead Silver	NT NT NT	NS NS NS	0.90 ** 0.86 ** 0.82 **/****
Samples >0.2% TOC Amphipod (n=15) Bivalve (n=7)		NT	NS	NS
Inner Harbor (Phase 3A) Amphipod (n=22)		NT	NS	NT
Samples >0.2% TOC Amphipod (n=8)		NT	NS	NT
Inner Harbor (Phase 3A Repeat) Polychaete (n=22)		NS	NT	NT
Samples >0.2% TOC Polychaete (n=10)	Anthracene LPAH Total organic carbon Benzo(a)anthracene Phenanthrene HPAH Indeno(1,2,3-c,d)pyrene Benzo(b,k)fluoranthene Pyrene Total PAH	0.77 ** 0.73 ** 0.69 ** 0.68 ** 0.68 ** 0.67 ** 0.67 ** 0.67 ** 0.65 ** 0.65 **	NT	NT
Richmond Harbor polychaete/amphipod (n=12) bivalve (n=7)	Arsenic Lead Silver	0.64 ** NS NS	NS NS NS	NS NS NS
Samples >0.2% TOC polychaete/amphipod (n=11) bivalve (n=7)		NS	NS	NS

[a] Bivalve species used in Phase 3B was *Mytilus galloprovincialis*.

* p<0.01; ** p<0.05; *** Significant correlation (p<0.05) between chemical parameter and total organic carbon

NT not tested

NS no significant correlations

Table 3-1 continued. Significant Spearman rank correlations for the Harbor studies

Study (Sample size)	Chemical Parameter	Correlation Coefficients		
		<i>N. caecoides</i> (% mortality)	<i>R. abronius</i> (% mortality)	bivalve larvae [a] (% abnormal)
Oakland Harbor				
Berth Areas				
polychaete/amphipod (n=12)	Benzo(g,h,i)perylene	0.85 *	0.60 **	NS
bivalve (n=7)	Indeno(1,2,3-c,d)pyrene	0.85 *	NS	NS
	Acenaphthylene	0.81 *	NS	NS
	Dibenzo(a,h)anthracene	0.75 *	NS	NS
	Total PCB	0.73 **	NS	NS
	Lead	0.72 **	0.67 **	NS
	DDD	0.67 **	NS	NS
	Benzo(a)pyrene	0.61 **	NS	NS
	Naphthalene	0.61 **	NS	NS
	Copper	0.61 **/****	0.82 */****	NS ****
	Benzo(b,k)fluoranthene	0.61 **	NS	NS
	Mercury	0.60 **/****	0.83 */****	NS ****
	Total organic carbon	NS	0.60 **	NS
	Dieldrin	NS ****	0.60 **/****	NS ****
	Nickel	NS ****	0.86 */****	NS ****
	Silver	NS ****	0.80 */****	NS ****
	Zinc	NS ****	0.74 */****	NS ****
	Selenium	NS	0.61 **	NS
	Tributyltin	NS	NS	0.86 **
Samples >0.2% TOC [b]				
polychaete/amphipod (n=9)	Benzo(g,h,i)perylene	0.87 *	NS	1.0 *
bivalve (n=5)	Benzo(b,k)fluoranthene	0.87 *	NS	NS
	Indeno(1,2,3-c,d)pyrene	0.87 *	NS	NS
	Dibenzo(a,h)anthracene	0.86 *	0.71 **	1.0 *
	Acenaphthylene	0.86 *	NS	NS
	Total PCBs	0.86 *	NS	NS
	Benzo(a)anthracene	0.85 *	NS	NS
	Benzo(a)pyrene	0.85 *	NS	NS
	Fluoranthene	0.85 *	NS	NS
	Pyrene	0.85 *	NS	NS
	Total organic carbon	NS	0.71 **	NS

[a] Bivalve species used in Richmond Harbor was *Mytilus galloprovincialis*; *Crassostrea gigas* was used in Berthing areas.

[b] Representative significant rank correlations. Only representative values are presented because significant ($p < 0.05$) correlations existed between *N. caecoides* mortality and all chemical parameters measured with the exception of TOC, DDE, DDT, and Dieldrin. All chemicals, having significant correlations with *R. abronius* and bivalve larvae are listed.

* $p < 0.01$; ** $p < 0.05$; *** Significant correlation ($p < 0.05$) between chemical parameter and total organic carbon

NT not tested ; NS no significant correlations

Table 3-2. Significant Spearman rank correlations for SedQual 3 studies

Study (sample size)	Chemical Parameter	Correlation Coefficients	
		<i>E. estuarius</i> (% mortality)	<i>C. gigas</i> (% abnormal)
Bay Run #1 Amphipod (n=15) Bivalve (n=13)	Total organic nitrogen	0.67 *	NS
	Total organic carbon	0.62 ***	NS
Bay Run #2 Amphipod (n=14) Bivalve (n=14)	Nickel	0.73 */***	NS ***
	Cobalt	0.58 **/***	NS ***
Bay Runs #1&2 Amphipod (n=29) Bivalve (n=27)	Total organic nitrogen	0.56 *	NS
	Nickel	0.51 */***	NS ***
	Total organic carbon	0.50 *	NS
	Chromium	0.49 */***	NS ***
	Dibenzo(a,h)anthracene	0.41 *	NS
	Zinc	0.38 **/***	NS ***
Marshes Amphipod (n=36) Bivalve (n=35)	Total organic nitrogen	0.42 **/***	NS ***
Castro Cove [a] Amphipod (n=32) Bivalve (n=32)	Zinc	0.96 *	0.36 **
	Nickel	0.94 *	0.41 **
	Lead	0.92 *	
	Copper	0.91 *	0.36 **
	Silver	0.90 *	0.44 **
	Total organic carbon	0.89 *	0.43 **
	Total organic nitrogen	0.89 */***	0.43 **/***
	Chromium	0.84 *	0.37 **
	x-Methylphenanthrene	0.68 */***	NS ***
	Total DDT	0.54 */***	NS ***
	Beta HCH	0.44 **/***	0.45 */***
	Benzo(k)flouranthene	0.44 **/***	0.35 **/***
	Hexachlorobenzene	0.43 **/***	0.42 **/***
	Total PCBs	0.41 **	
	Total PAHs	0.40 **/***	0.36 **/***
	Lindane	0.37 **/***	0.41 **/***

* p<0.01; ** p<0.05

*** Significant correlation (p<0.05) between chemical parameter and total organic carbon.

NT not tested

NS no significant correlations

[a] Representative significant rank correlations. Only representative chemicals are presented because significant correlations (p<0.05) existed between amphipod mortality and all chemicals measured with the exception of fluoranthene, phenanthrene, and the pesticide Alpha HCH. All significant coefficients for bivalve development are included.

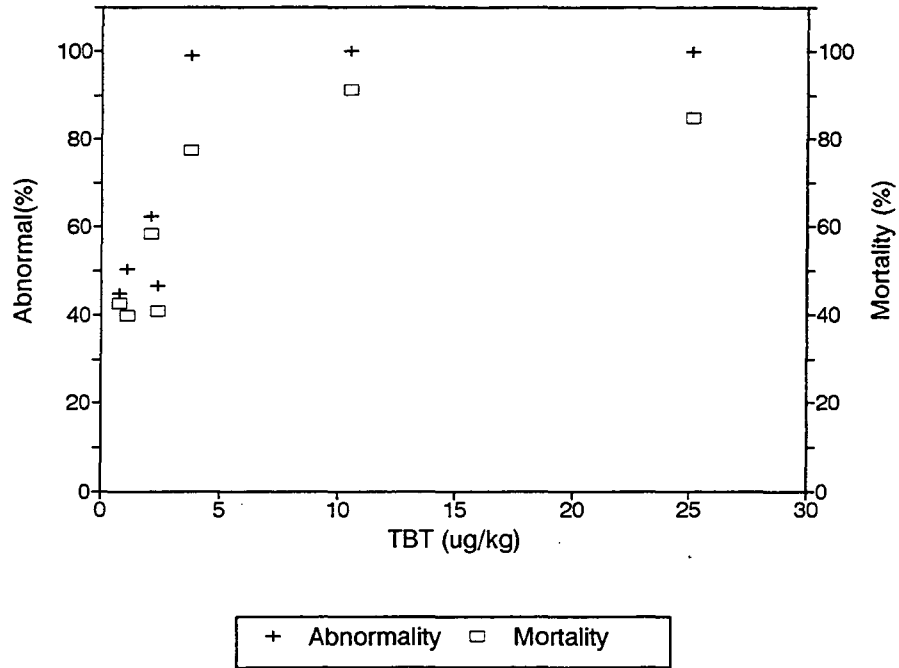
those nine berth samples with greater than 0.2% TOC, we observed significant associations between toxicity and every chemical parameter tested with the exception of three pesticides (DDT, DDE, and dieldrin). Significant rank correlations were also observed for ten chemicals in samples from the repeat testing of the Oakland Inner Harbor channel when analysis was limited to the ten samples with greater than 0.2% TOC. These correlations were generally weaker than those seen in the berth areas. In contrast, only four chemical parameters (three trace metals and tributyltin) were significantly associated with toxicity in the Oakland Outer Harbor and Richmond Harbor studies. Significant correlations between chemical parameters and TOC were infrequent and limited to six trace metals and one pesticide.

Most of the significant correlations observed in the Oakland Harbor berth areas and Oakland Inner Harbor studies were in relation to mortality of the polychaete, *Nephtys caecoides*. Concentration-response relationships using *N. caecoides*, however, were not particularly strong, as mortality never exceeded 31% in these data sets. Strong correlations were also seen in the Berthing Areas study between concentrations of nearly all the trace metals measured and mortality in the amphipod *Rhepoxynius abronius*, albeit over a very narrow range of mortality (4-28%) and contaminant concentrations. Strong correlations that are potentially demonstrative of a cause-effect relationship were observed between the bivalve abnormality endpoint and concentrations of the antifoulant, tributyltin, in both the Oakland Outer Harbor and Berthing areas studies² (Figure 3-1).

SedQual 3 Studies Nearly all the significant rank coefficients observed for the SedQual 3 grouping were limited to the Castro Cove study (Table 3-2). At Castro Cove, significant associations existed between mortality in the amphipod *Eohaustorius estuarius* and all parameters except two PAHs (fluoranthene and phenanthrene) and the pesticide, Alpha HCH. Significant correlations were also observed between concentrations of 12 chemicals and larval abnormality in *Crassostrea gigas*.

²Tributyltin was detected in only four samples from the SedQual 3 studies and thus was not included in rank correlation analysis for this data.

(A)



(B)

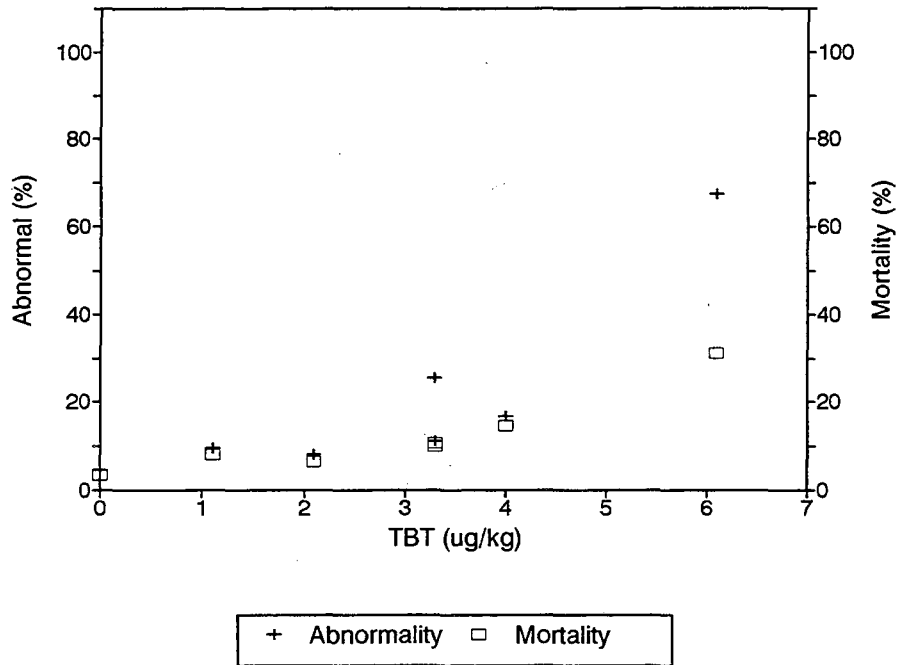


Figure 3-1 Bivalve larval mortality and abnormality in relationship to sediment concentrations of tributyltin from (A) Oakland Harbor berths using *Crassostrea gigas* ($r_s = 0.86$, $p < 0.05$) and (B) Oakland Outer Harbor using *Mytilus galloprovincialis* ($r_s = 0.90$, $p < 0.05$).

In all of the SedQual 3 studies, TOC and/or TON were among the chemical parameters most strongly associated with toxic effects³. Scatter plots for TOC-mortality and TON-mortality from the Castro Cove demonstrate a strong association (Figure 3-2). In the Castro Cove study, particularly strong associations were also observed between TOC and all of the organic chemicals which, themselves, were significantly associated with toxicity. A good example of this correlation pattern is seen for two pesticides, methylphenanthrene and total DDT. These two chemicals had the highest rank coefficients relative to toxicity (Figures 3-3a and 3-3b) and were significantly associated with TOC (Figure 3-4).

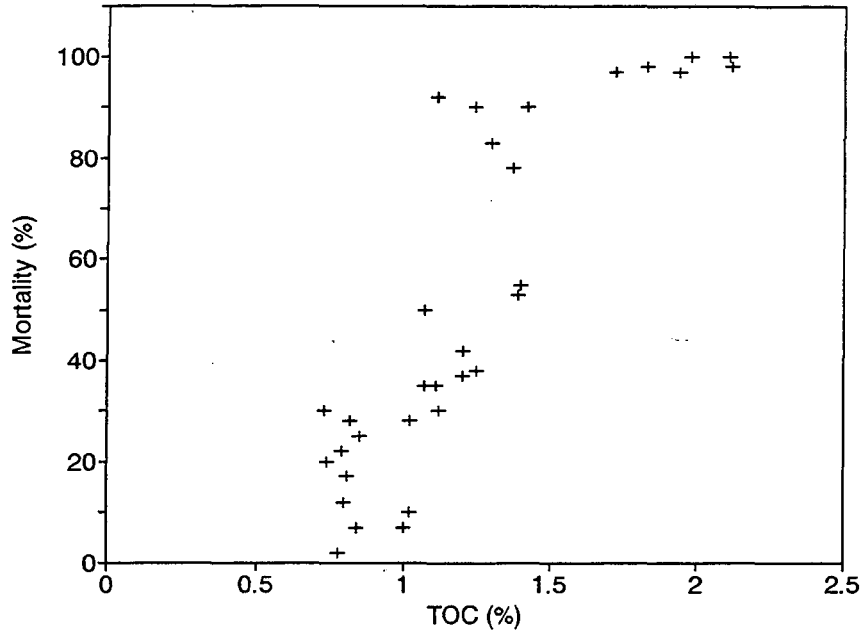
The highest rank correlation coefficients evident in the Castro Cove data were between concentrations of trace metals (zinc, nickel, lead, copper, and silver) and amphipod mortality. The concentration-response relationships for these metals are presented in Figure 3-5. Effects Range-Low (ERL) and, where possible, Effects Range-Median (ERM) values for each metal are included on these figures. These values indicate the low end of the range of sediment concentrations in which toxic effects were observed in a national data base (Long, *et al.*, in press). The highest *E. estuarius* mortality was observed in four samples from one station (GD20) which consistently exhibited the highest sediment trace metal concentrations.

Rank correlation coefficients between nearly all trace metals measured and oyster larvae abnormality were also significant in Castro Cove ($p < 0.05$). The concentration-response plots for these metals, however, show that determination of significant association was highly influenced by a small number of samples in the high concentration ranges (exemplified in Figure 3-6).

Combining the data from the Bay Run #1 and #2 studies resulted in a higher number of significant associations with amphipod mortality than was observed using either of the Bay Run studies alone (Table 3-1). While the scatter plots for TOC and TON (Figure 3-7a and 3-7b) show a strong concentration-response relationship, those for the chemical parameters

³ TON not measured in any of the five Harbor studies.

(A)



(B)

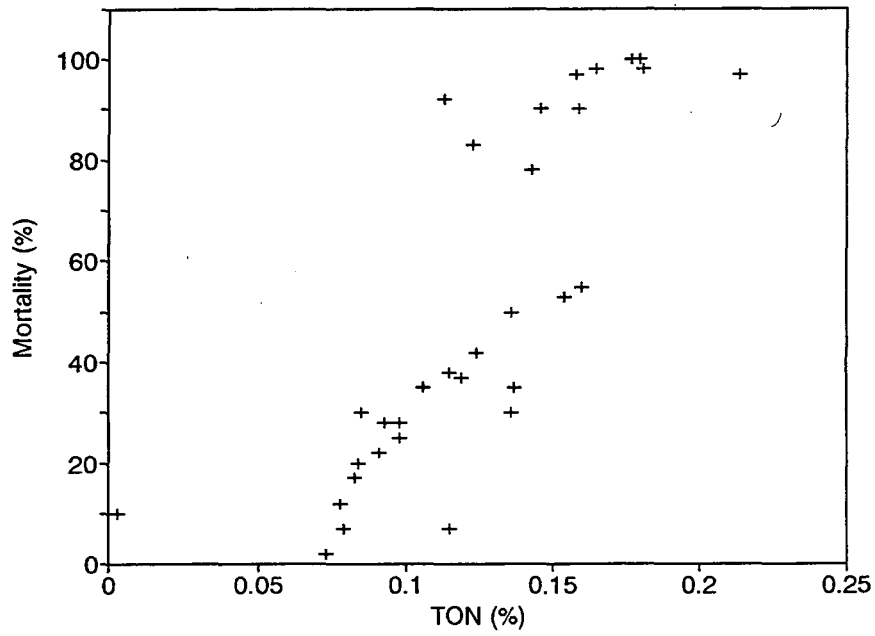
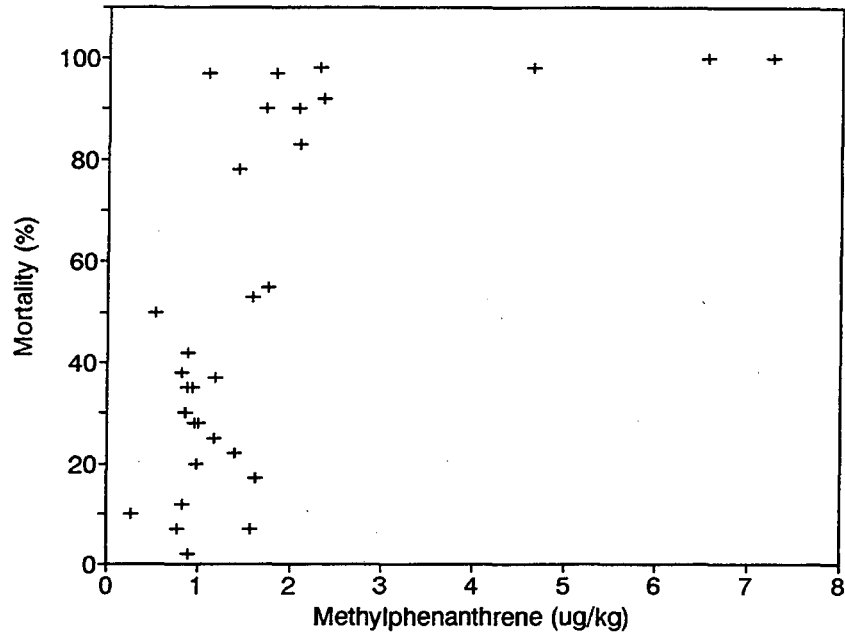


Figure 3-2 *Eohaustorius estuarius* mortality in relation to (A) %TOC ($r_s = 0.89, p < 0.01$) and (B) %TON ($r_s = 0.89, p < 0.01$) in 30 sediment samples from Castro Cove.

(A)



(B)

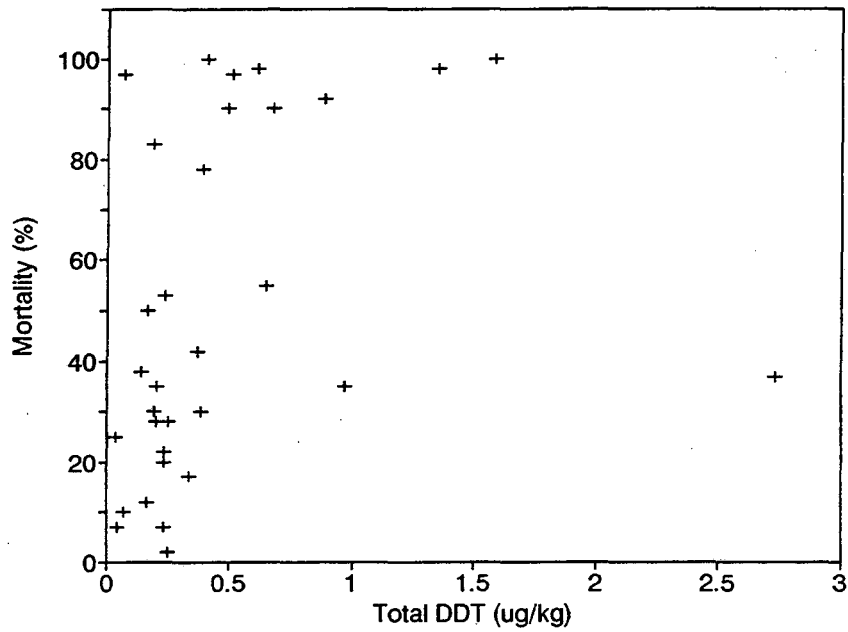


Figure 3-3 *Eohaustorius estuarius* mortality in relation to OC-normalized concentrations of (A) Methylphenanthrene ($r_s = 0.68$, $p < 0.01$) and (B) total DDT ($r_s = 0.54$, $p < 0.01$) in 30 sediment samples from Castro Cove.

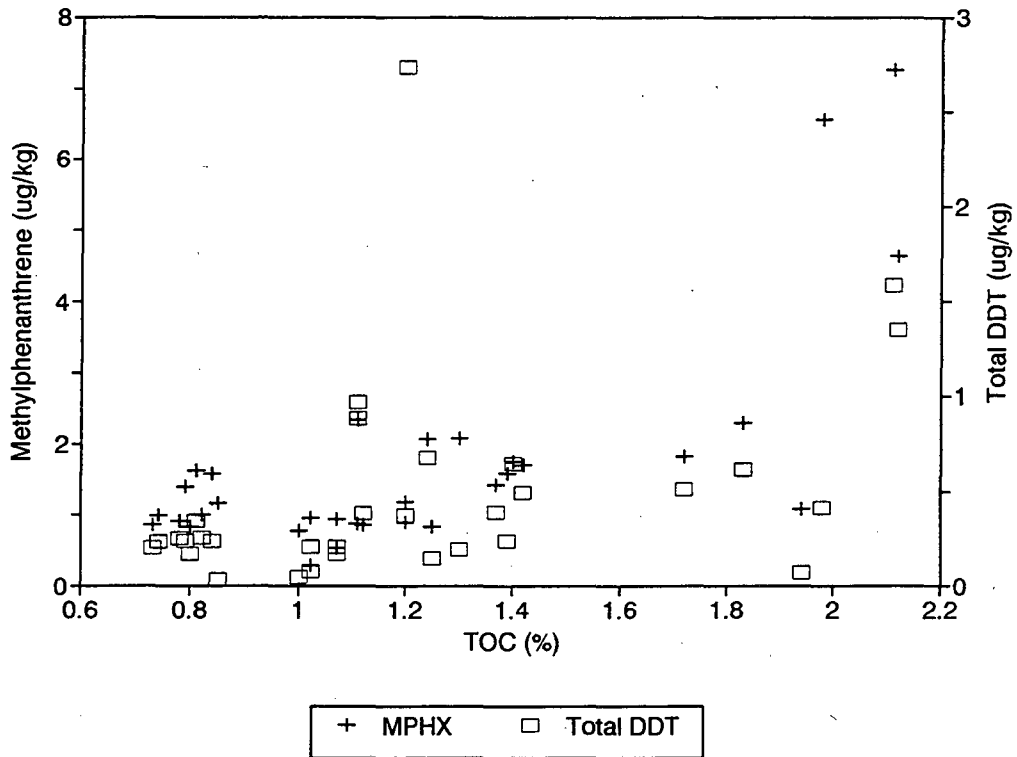


Figure 3-4. The relationship between sediment TOC and dry-weight concentrations of Methylphenanthrene ($r_s = 0.60$, $p < 0.01$) and total DDT ($r_s = 0.48$, $p < 0.01$) in 30 sediment samples from Castro Cove.

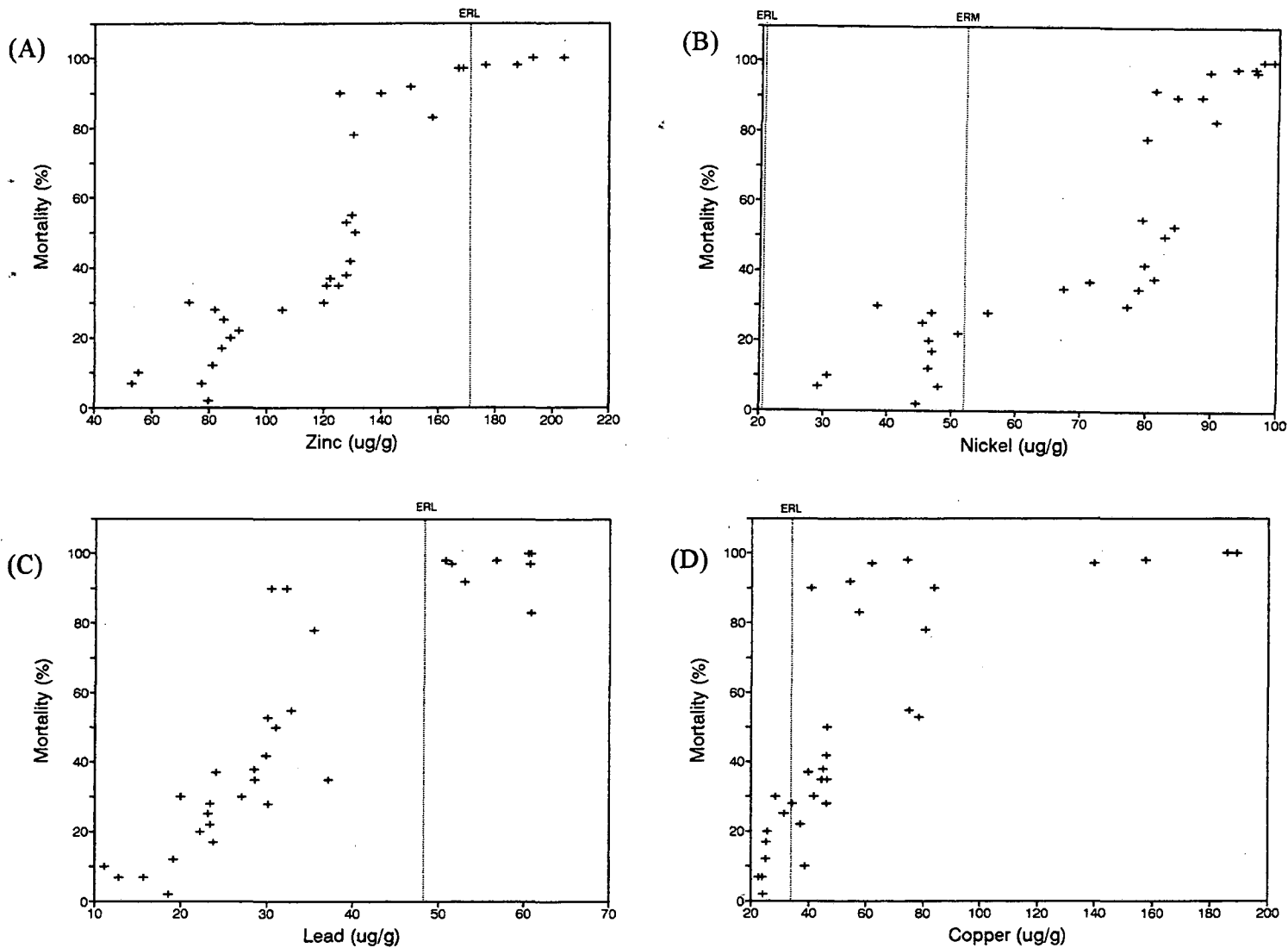


Figure 3-5. *Eohaustorius estuarius* mortality in relation to (A) zinc ($r_s=0.96$, $p<0.01$), (B) nickel ($r_s=0.92$, $p<0.01$), (C) lead ($r_s=0.92$, $p<0.01$), and copper ($r_s=0.91$, $p<0.01$) in 30 sediment samples from Castro Cove. For comparison, Effects Range-Low (ERL) and Effects Range-Median (ERM) values are indicated with vertical dotted lines (Long *et al.*, in press).

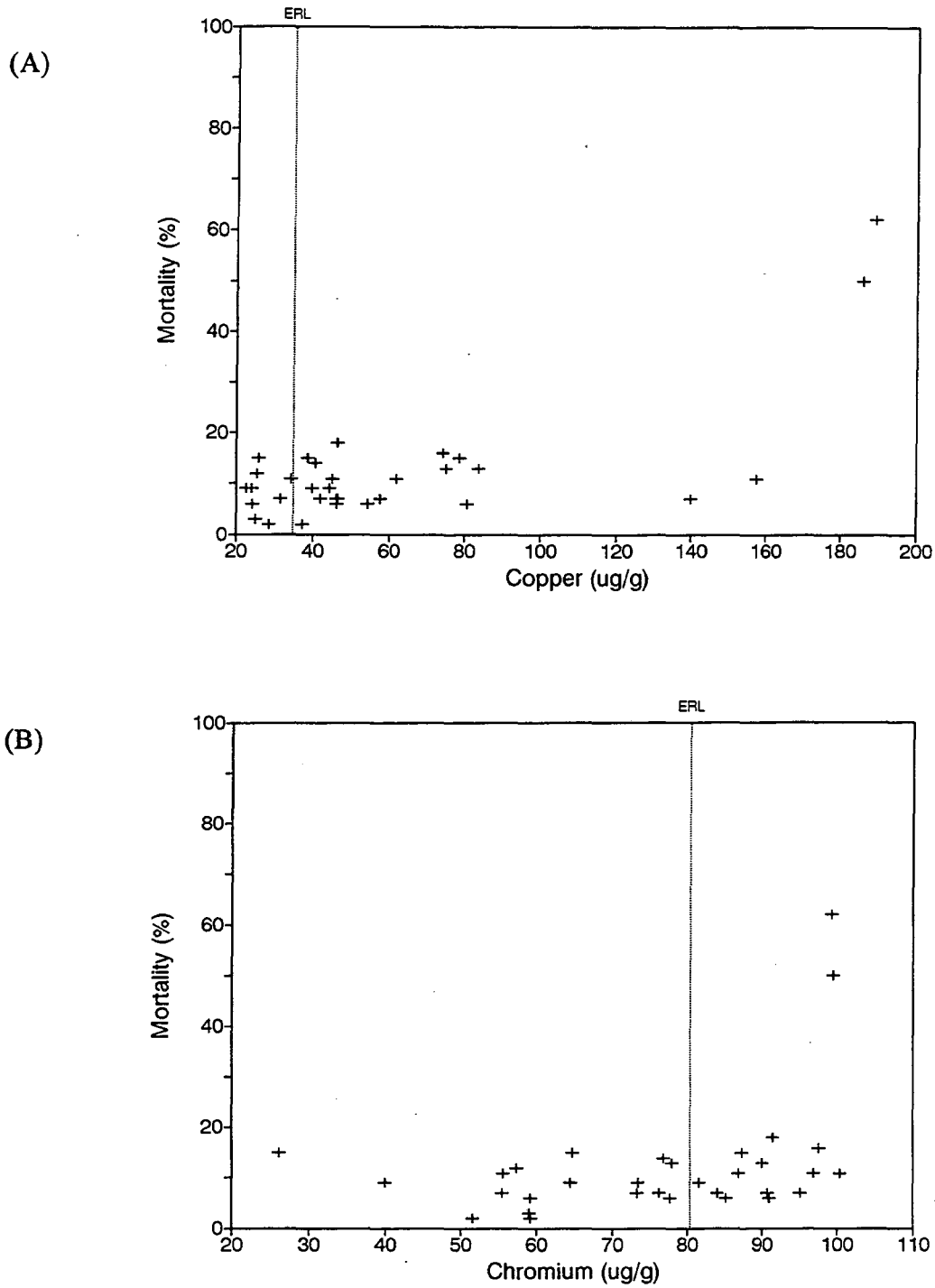


Figure 3-6. *Crassostrea gigas* mortality in relation to (A) copper ($r_s = 0.36$, $p < 0.05$) and (B) chromium ($r_s = 0.37$, $p < 0.05$) in 30 sediment samples from Castro Cove. For comparison, ERL values are indicated with vertical dotted lines.

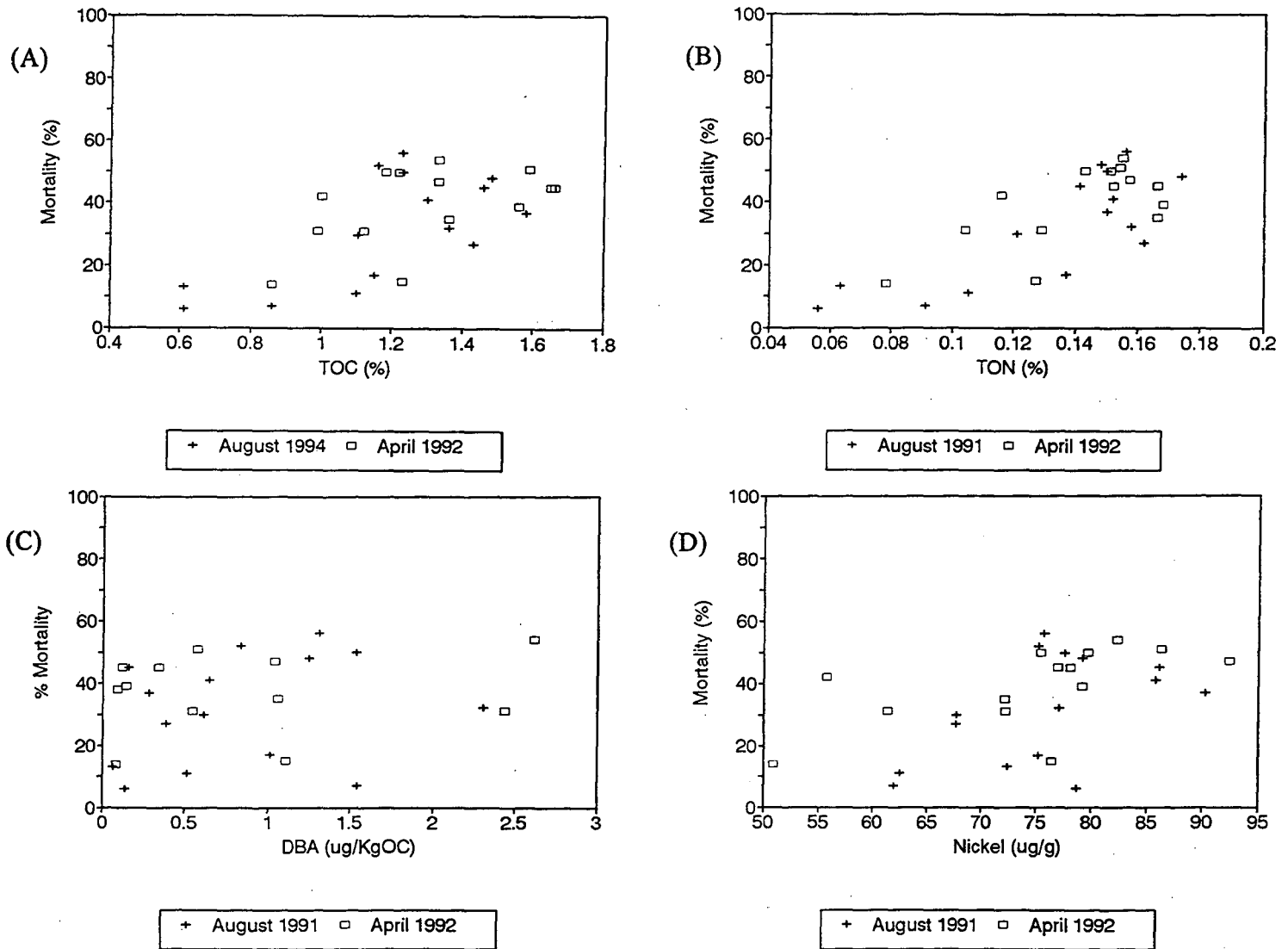


Figure 3-7. *Eohaustorius estuarius* mortality in relationship to (A) TOC (for Bay Run #1 $r_s = 0.62$; for Bay Runs combined $r_s = 0.50$, $p < 0.01$), (B) TON (for Bay Run #1 $r_s = 0.67$; for Bay Runs combined $r_s = 0.56$, $p < 0.01$), (C) dibenzo(a,h)anthracene (for Bay Runs combined $r_s = 0.41$, $p < 0.01$), and (D) nickel (for Bay Run #2 $r_s = 0.73$, $p < 0.01$; for Bay Runs combined $r_s = 0.51$, $p < 0.01$) for 29 sediments from the Bay Run #1 (August 1991) and Bay Run #2 (April 1992) studies.

(exemplified in Figure 3-7c and 3-7d by nickel and dibenzo(a,h)anthracene) showed weak relationships over concentration ranges that were lower than those expected to cause mortality in *E. estuarius*. Significant associations were observed between TOC and three trace metals in the combined data set.

In both the Harbor and SedQual 3 studies, a higher number of the significant rank correlations were associated with the solid phase toxicity tests as compared to the bivalve elutriate tests. For example, in the Castro Cove study, significant correlations with *E. estuarius* mortality were observed for 45 chemical parameters (Table 3-2). In contrast, there were only 12 chemical parameters that were significantly correlated with the oyster larval toxicity endpoint. Likewise, in the Oakland Harbor Berth Area study (Table 3-1), significant coefficients associated with polychaete or amphipod mortality existed for anywhere between 10 to 21 chemical parameters, depending on the test species and sample size used; whereas, significant correlations with bivalve larval abnormality were evident for only four chemicals.

Discussion

Spearman Rank analysis was chosen as a first-order approximation of the potential for concentration-response correlations between chemical parameters and toxicity endpoints. The rank correlation coefficient (r_s) gives an indication of the strength of association between any two variables. However, these correlations only demonstrate covariance and are not necessarily an indication of causality. Clearly, there are significant limitations to using univariate chemical measures to explain biological response in a multivariate exposure scenario. Further information such as synoptic measurements of chemistry and toxicity in pore water and the AVS/SEM ratios of sediments are necessary to better estimate exposures as well as to discriminate between the relative contributions of co-occurring chemicals to toxicity. Most of the studies in our data set lacked such information; however we attempted to compensate for this by including chemical classes (such as PAHs and total PCBs) as

variables in the rank correlation analysis as well as OC-normalizing concentrations of all organic compounds.

In both the Harbor and SedQual 3 groupings, stronger associations between contamination and toxicity were observed in studies for which samples were taken proximal to intensive industrial activity and/or point sources of contamination. Samples from outer harbor channels (Oakland Outer Harbor and Richmond Harbor) or generally removed from direct influence of point discharges (Bay Runs #1 and #2 and Marsh Studies) had fewer significant rank correlation coefficients between chemicals and toxic endpoints.

It is unclear whether the strong correlations observed between organic carbon content and biological effects are an indication that sediment toxicity is related to non-contaminant parameters or whether these correlations are merely a consequence of the strong association often observed between sediment contaminants and TOC. Particularly at Castro Cove, strong correlations between numerous organic compounds and TOC made it difficult to distinguish which contaminant or non-contaminant components were responsible for the toxicity observed. On one hand, TOC was significantly correlated with at least one toxicity endpoint in nearly all of the SedQual 3 and Harbor studies; and, in several studies, it was one of the parameters for which the strongest correlations were observed. On the other hand, the strong correlations often observed between TOC and organic compounds and trace metals suggest that these TOC-associated contaminants (and not the TOC itself) may be responsible for the toxicity observed. The fact that fine sediment grain size (often associated with higher levels of TOC) is not a significant determinant of amphipod toxicity in either Harbor or SedQual 3 sediments⁴ suggests that contaminants associated with organic carbon, rather than natural sediment characteristics, are a likely cause of the toxicity observed in these samples.

In several studies, strong associations were observed between bulk concentrations of trace metals and toxicity. Upon closer examination, however, the concentration-response

⁴ See results of Section 5 for a further discussion.

relationships for these trace metals were often weak or occurred over narrow concentrations and/or response ranges. For example, although clear concentration-response relationships were observed for all trace metals measured in Castro Cove sediments, a majority of the samples had trace metal concentrations that were approximately equal to or lower than the highly conservative ERL values⁵ (Figure 3-5). Indeed, concentrations of individual trace metals in many of these samples are lower than levels expected to be of concern for toxicity in San Francisco Bay⁶ (Ross, personal communication). This suggests that trace metals may be additive in their toxic effects or that they are simply co-occurring with a more probable cause of toxicity such as petroleum hydrocarbons which are abundant at this site (Taberski, personal communication).

Many of the samples in the Harbor studies had a very low organic carbon content, limiting the utility of OC-normalization for these samples and potentially skewing the outcome of the rank correlation analysis. For this reason, a second analysis was performed for each of the five Harbor studies using reduced data set consisting only of those samples that had TOC greater than 0.2%. While a larger number of significant associations were often observed for some of these reduced data sets, low ranges of contamination and toxic response generally precluded identification of factors responsible for toxicity.

A higher number of significant rank associations between chemical parameters and toxicity were observed using solid phase tests (amphipod and polychaete) as compared to the

⁵ The predictive accuracy for ERLs and ERMs for toxicity in San Francisco Bay sediments is discussed in detail in Section 7.

⁶ We performed a supplemental analysis of pore water chemistry and toxicity data from Castro Cove, the results of which demonstrate that trace metal contamination is an unlikely cause of toxicity at this location. Concentrations of individual trace metals in pore water were generally lower than EPA marine water quality criteria values. Furthermore, summation of pore water toxic units calculated based on the oyster larval abnormality endpoint identified only one sample in which trace metal concentrations would be expected to cause toxicity in this test species. A pore water toxic unit for a given chemical is defined as the pore water concentration divided by the water-only LC50 for a given species. Further information on pore water toxicity units is given in Section 4. This analysis was limited to trace metals, because pore water measurements of organic compounds were not available.

elutriate phase tests using bivalve larvae. Based on a somewhat larger data base, Long and Markel (1992) observed just the opposite trend, namely that linear correlation coefficients (r^2 values) for bivalve larvae were generally higher than those for amphipods. One possible explanation for this difference is the smaller amount of bivalve data used in the present analysis relative to amphipod data. Specifically, the number of Harbor samples tested using the elutriate bioassays was always less than half that tested with solid phase bioassays, substantially increasing the critical value necessary for a significant rank correlation with the bivalve tests. It is noteworthy, however, that although we observed fewer significant associations between toxicity endpoints in the elutriate bioassays and chemical parameters, the concentration-response plots for these were often more suggestive of a potential cause-effect relationship.

Conclusions

Significant associations were observed between toxicity endpoints and bulk sediment concentrations of numerous chemicals in areas proximal to contaminant inputs such as Castro Cove and ship berths in Oakland Harbor. In contrast, relatively few significant associations were seen between toxicity endpoints and chemical concentrations in samples taken from navigation channels, marshes, and mid-Bay areas. Furthermore, significant correlations were observed more frequently using the mortality endpoints from solid phase tests (amphipod and polychaete) than using the bivalve larval development endpoint in the elutriate tests.

The significance of the rank correlation coefficients alone, however, was often an unfaithful translation of the actual relationship between chemical parameters and toxicity. An examination of the individual scatter plots was required to accurately characterize the concentration-response relationships. Studies for which strong associations were observed between bulk sediment chemistry and toxicity generally fell into one of three categories:

- Significant associations observed but inconclusive concentration-response relationships because of very low levels of contamination and/or toxicity.

- Significant associations observed but attributable to a few individual data points that had a strong effect on the strength of correlation.
- Significant associations and strong concentration-response relationships observed but strongest associations were for non-contaminant parameters such as TOC or TON.
- Significant associations and strong concentration-response relationships observed but for contaminants which were, themselves, significantly correlated with TOC.

The general absence of cases in which toxicity was clearly attributable to increasing bulk sediment concentrations of an individual chemical or chemical class coupled with the frequency with which toxicity was associated with non-contaminant parameters is by no means an indication that sediment toxicity is unrelated to contamination. Rather, these results demonstrate that paired sediment chemistry and toxicity testing are suitable tools for determining the overall toxicity and contamination associated with a sample but not for identifying the cause(s) of the toxicity observed.

Determination of the actual concentrations to which a test organism is exposed is critical to relating cause to effect in toxicity testing. Quantification of exposure, however, is impossible based on bulk sediment chemistry alone. There are numerous sediment factors which influence the bioavailability of contaminants in sediments which must be taken into consideration. Abiotic factors include the chemical form of the compound, its sorption to particulate or dissolved organic material, the presence of other ions (AVS, salinity, iron and manganese oxides), and the oxidation state of the sediment. Interaction among these factors may further modify the bioavailability of a contaminant in the sediment. Biotic factors influencing bioavailability and, therefore, toxicity include the mode of existence or habitat of an organism (interstitial, tube-dwelling, epibenthic) and route of contaminant uptake (feeding, respiration).

Nevertheless, bulk sediment chemistry and bioassay testing results can provide useful information on potential causes of toxicity that can then be used to direct further questioning and investigation. For example, strong concentration-response relationships with sediment

TOC are an indication that contaminant or non-contaminant factors that are associated with the organic carbon fraction are likely determinants of the toxicity observed. Sediments that are relatively rich in organic carbon are likely to have higher levels of dissolved organic carbon (DOC) in their pore water. Contaminants associated with DOC may then constitute a significant exposure to those infaunal test organisms for which sediment pore water is a primary route of exposure. Likewise, contaminants that are bound to the organic carbon fraction of sediments may constitute a significant additional exposure for test organisms that directly ingest sediment particles. The finer grain size of organically rich sediments may also, independently, influence organism mortality if it exceeds the grain size tolerance limits of the test species. While strong correlations observed between TOC, bulk sediment chemicals, and toxicity confound cause-effect determinations, they do provide positive identification of a highly contaminated location where toxicity is likely to result from exposure to a mixture of different chemicals or chemical classes.

4. SENSITIVITY OF TWO SPECIES OF INFAUNAL AMPHIPODS TO SEDIMENT GRAIN SIZE

Introduction

The infaunal amphipods, *Rhepoxynius abronius* and *Eohaustorius estuarius* are widely used on the West Coast to evaluate toxicity of sediments of varying grain size and composition. Nevertheless, concern that grain size characteristics alone may influence the outcome of toxicity tests has not been fully resolved (DeWitt, *et al.*, 1988; Spies, 1989). Responding to these concerns, EPA and the U.S. Army Corps of Engineers have recently issued joint technical recommendations which discuss the optimal grain size ranges within which adverse effects in different amphipod species are not expected to occur (Davies, *et al.*, 1993). However, the effects of very fine grain size on amphipod survival remains controversial in the San Francisco Bay area where, to date, this issue has not been comprehensively addressed using the available data.

The basis of concern lies in the fact that these amphipod species are often tested in sediments which are much finer than those in which they are found in nature. The native sediment of both the marine *R. abronius* and the estuarine *E. estuarius* has been described as fine to very fine sands (approximately 70-200 μm in diameter) (DeWitt, *et al.*, 1989; Oakden, 1984). In comparison, the particle size of fine sediments in San Francisco Bay ranges from 62.5 μm to less than 3.9 μm . Data from laboratory experiments as well as field collection have demonstrated that *R. abronius*, in particular, has a strong preference for fine sands over sediments that are coarser or finer (Oakden, 1984).

The effects of sediment particle size on the survival of both amphipod species have been evaluated in detail by DeWitt *et al.* (1988; 1989). In bioassays using laboratory manipulated and clean field sediments, survival of *R. abronius* was reduced by as much as 15% in finer uncontaminated sediments as compared to native sediments; although the actual cause of mortality was not identified (DeWitt, *et al.*, 1988). While several mechanisms by which fine sediment may be deleterious to *R. abronius* are described, DeWitt and his co-

workers suggest that sediment grain size is actually a "super variable" that is correlated with the actual cause of mortality rather than grain size itself being the cause. Using the regression relationship of *R. abronius* survival to % fines in clean sediments, they proposed a model for calculating the 95% prediction limits as a means of estimating the effect that grain size alone should have on amphipod survival in field-collected sediments. In a later study comparing sensitivities of *R. abronius* and *E. estuarius*, DeWitt *et al.* (1989) found that mean survival of *E. estuarius* declined somewhat as grain size decreased but that, generally, *E. estuarius* showed little sensitivity to sediment grain size and was slightly more tolerant of fine, uncontaminated sediments than *R. abronius*.

While the findings described above indicate that neither *R. abronius* nor *E. estuarius* are highly sensitive to uncontaminated fine-grained sediments, they do leave open the potential for positive interferences to occur that are directly or indirectly linked to sediment grain size. Moreover, it is difficult to use information based on clean uncontaminated sediments to draw definitive conclusions on the relative toxicological influence of grain size in contaminated field sediments. Strong correlations among % fines, organic carbon, and concentrations of contaminants often complicate the identification of cause and effect in field-collected sediments. Because San Francisco Bay sediments are frequently characterized by a high fines fraction ($\geq 80\%$ silt-clay), there remains controversy as to whether moderate levels of amphipod mortality (such as that described in Section 2 of this report) may be partially attributable to grain size effects.

Goals

The purpose of this analysis was twofold: 1) to characterize the relationship between grain size and the toxic responses of two amphipod species (*R. abronius* and *E. estuarius*) for sediments from our data set, and 2) to determine whether mortality due to grain size can be differentiated from that due to chemical contaminants in these sediments.

Approach

Grain size data were compared to amphipod survival data from sediment bioassays to elucidate potential relationships between decreasing grain size and mortality. Complete grain size information (% gravel, % sand, % silt, and % clay) and mortality data were available from four of the Harbor studies (Oakland Harbor Phase 3A, Phase 3B, Berth Areas, and Richmond Harbor). The fines content, comprised of the silt (3.9-62.5 μm) and clay (<3.9 μm) fractions, was calculated for each sample and is expressed as % fines (dry-weight). Toxicity testing in the Harbor studies was conducted using *R. abronius*.

Grain size data available for the SedQual 3 studies was limited to % sand. For these studies, the fines content of each sample was estimated by assuming that the gravel content was negligibly small and using the relationship: % fines = 100 - % sand. This data was then compared to the results of toxicity testing using *E. estuarius*.

Spearman Rank correlation analysis (Daniel, 1990) was performed on both the Harbor and SedQual 3 data groups to determine the association (r_s) between % fines, % clay, % silt and amphipod survival. Relationships between various grain size measures and amphipod survival were further analyzed by regression analysis. Those samples which had a grain size distribution of $\geq 80\%$ fines and for which greater than 20% amphipod mortality was observed were examined in greater detail to determine if explanations other than grain size could account for the mortality observed.

Results

Survival of *R. abronius* was negatively associated with the % fines ($r_s = -0.394$, $p < 0.01$), % clay ($r_s = 0.348$, $p < 0.01$), and % silt ($r_s = 0.425$, $p < 0.01$) content of sediments from the Harbor studies. The relationship between *R. abronius* survival and % fines is depicted in Figure 4-1. The regression equation as well as mean survival in native sediment and fine-grained controls are also indicated. Although the linear regression of amphipod survival with % fines shows a slight negative trend that is highly significant ($p = 0.002$), the low r^2 value

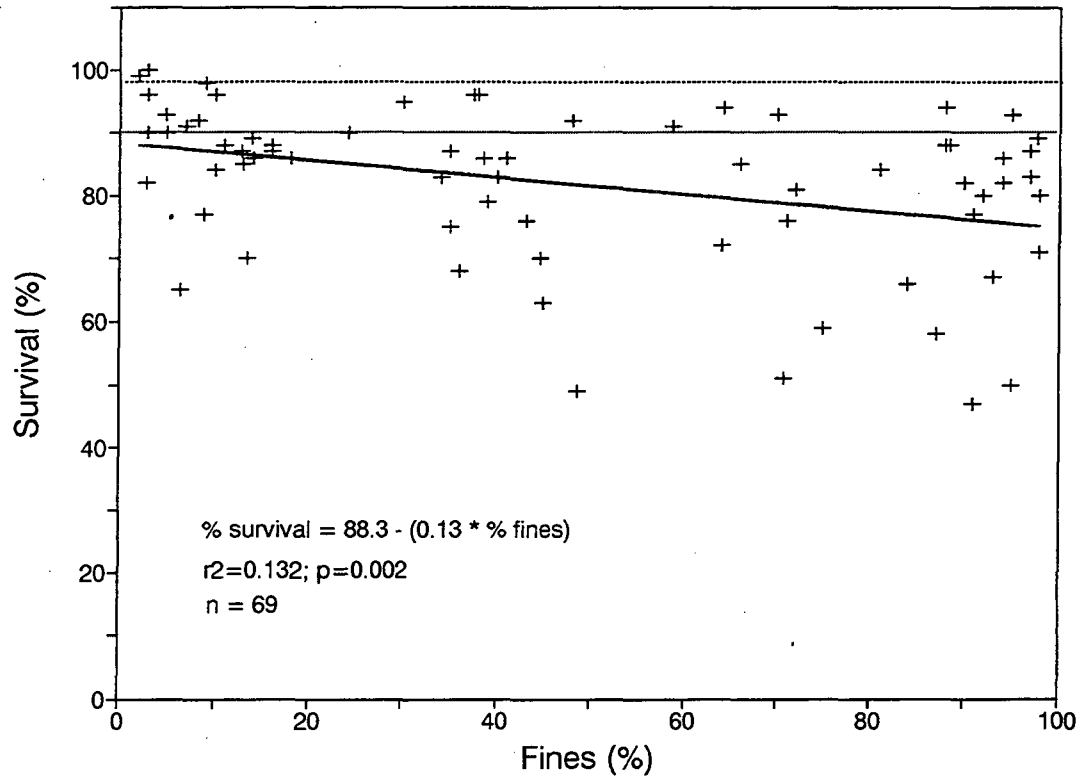


Figure 4-1. Survival of *Rhexopynius abronius* in relation to % fines (% fines = % silt + % clay) in sediments from 69 sites in San Francisco Bay Harbor studies. The regression line (solid) overlies the original data. Mean survival in the native control (98% in 3% fines) is indicated by the upper dashed line. Mean survival in the fine-grained control (90% in 67% fines) is indicated by the lower dotted line.

(0.13) indicates a poor fit of this data to the linear model¹. In fact, variability in amphipod response increased with increasing % fines. Of the 20 samples with a fines content greater than 80%, 18 samples had elicited amphipod survivorship that was less than the fine-grained control average of 90%, and only six samples had amphipod survival of less than 80%. While two of these six samples contained somewhat elevated concentrations of mercury (1.8 µg/g) and nickel (133µg/g), neither trace metals nor organic compounds were noticeably elevated in any of the other four samples.

While there was a negative association between survival of *E. estuarius* and estimated % fines content of the samples from the SedQual 3 studies ($r_s = -0.490$ $p < 0.01$), no significant linear relationship was observed ($r^2 = 0.10$) (Figure 4-2). Amphipod survival in the high fines sediments ($\geq 80\%$ fines) was highly variable, ranging from 0 to 95%.

Discussion

Survival of both species of amphipod in very fine ($\geq 80\%$ silt-clay) San Francisco Bay sediments was highly variable (50-100% for *R. abronius*; 0-95% for *E. estuarius*), indicating that grain size alone cannot fully account for the mortality observed in some of these samples. It seems plausible that increased mortality observed in the finer-grained sediments could be attributable to contaminant partitioning associated with the higher organic carbon content of these sediments. However, examination of the few fine-grained samples for which amphipod mortality exceeded 20% rarely yielded a definitive individual contaminant which might explain toxicity. Therefore, we cannot rule out the possibility that variable amphipod survival in the finer sediments is attributable to unmeasured properties of fine sediments such as their composition and/or shape. The specific characteristics that distinguish similarly fine sediments from different areas, and the toxicological relevance of these characteristics are difficult to quantify and, as such, have received little attention in sediment research.

¹ The linear regressions of % silt and % clay with *R. abronius* survival showed a negative trend similar to that seen using % fines. Their r^2 values were also very low (0.16 for % silt and 0.08 for % clay).

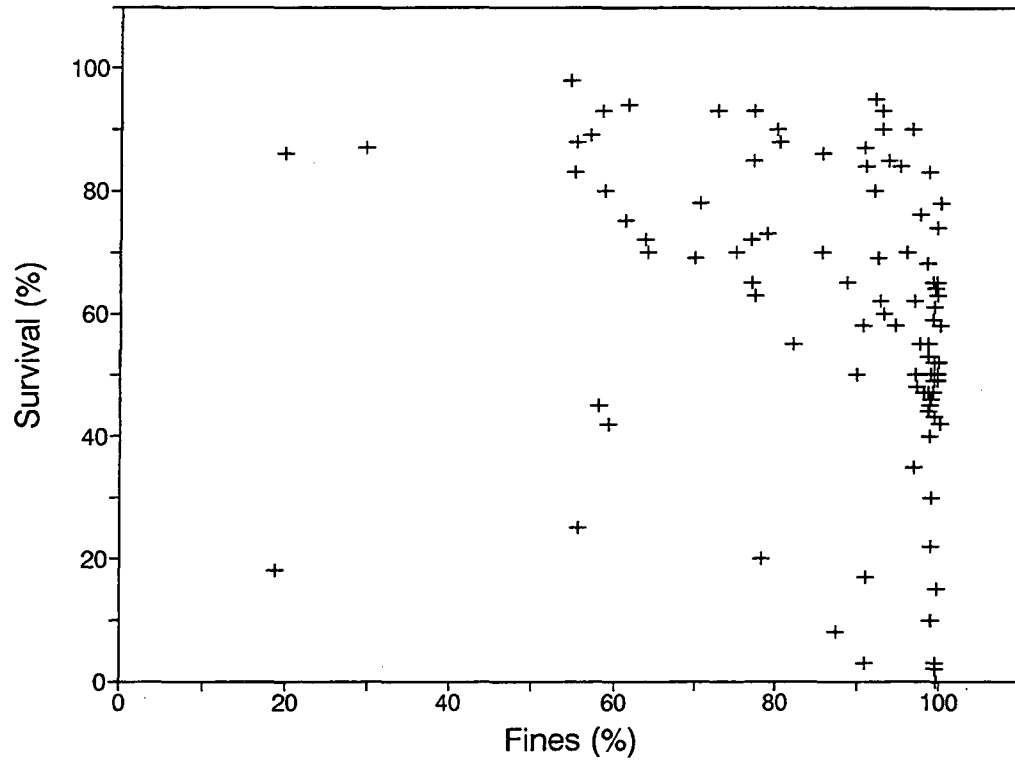


Figure 4-2. Survival of *Eohaustorius estuarius* in relation to % fines (100 - % sand) in sediments from 98 sites in the San Francisco Bay SedQual 3 study.

The relationships between grain size and *R. abronius* and *E. estuarius* survival observed in samples from our data set (Figures 4-1 and 4-2) were strikingly similar to those which have been reported previously for sediments from uncontaminated sites in Oregon and Washington by DeWitt *et al.* (1989; Figure 3, pp. 1041). The slopes of the regression lines for *R. abronius* are nearly equal (San Francisco Bay $b = -0.13$; Oregon/Washington $b = -0.12$) and, in both studies, variability of *R. abronius* survival increased with decreasing grain size. DeWitt *et al.* also reported an absence of a correlation between decreasing sediment grain size and *E. estuarius* mortality that is similar to our findings. Thus, the results of our analysis corroborate those of another study that has evaluated the tolerance of these species to uncontaminated sediments of different grain sizes.

Conclusions

For the majority of the Harbor and SedQual 3 samples, direct particle-size induced mortality does not appear to be a significant determinant of amphipod mortality. Rather, indirect grain size effects associated with the partitioning of contaminants in fine-grained sediments is the most likely explanation for the weak positive relationship between mortality and increasing fines for *R. abronius* in the San Francisco Bay data. In contrast, no such relationship was observed for *E. estuarius*. An unequivocal test of the association between grain size and amphipod mortality in field-collected sediments would be difficult to perform because of confounding factors such as contamination which covaries with grain size and unquantified differences in the physical properties of similarly fine sediments.

5. PATTERNS OF CONTAMINATION AND TOXICITY AT IN-BAY AND OCEAN REFERENCE SITES

Introduction

Toxicological and chemical evaluation of sediments from reference sites is an important component of sediment quality management. In regional monitoring programs, reference sites are used as negative controls to evaluate organism response to "background" conditions. The ideal sediment for this purpose would be uncontaminated, yet otherwise representative of local sediment characteristics. Reference sites are also used for comparative purposes in dredged material testing programs to represent the chemical and toxicological background conditions in the vicinity of a disposal site. In this case, the particular disposal history of a site will influence the contamination and toxicity associated with background. While there is no substantial agreement on how "clean" a reference site should be, it is clearly important that intersampling variability at a given reference site should be low enough so that it may serve as a consistent basis of comparison.

There are numerous factors that can potentially affect intersampling variability in both chemistry and in toxicity determinations at reference sites. Listed below are factors that may be pertinent to reference sites used in San Francisco Bay sediment testing programs:

- patchy distribution of contamination and/or sediment types,
- variable QA/QC within and between analytical laboratories,
- inconsistent positioning during sampling (especially a problem at deep water sites),
- temporal variability in factors controlling contaminant partitioning,
- temporal variability in test species sensitivity,
- influence of spills and/or discharge events (important for sites located at or near disposal areas).

Currently, there are efforts underway to identify and quantify the different parameters influencing reference site characterization in San Francisco Bay (Anderson, *et al.*, in preparation; Taberski, personal communication).

Goals

In this section, we examine the variation associated with chemical and biological measurements at six reference sites over the course of one year. Our goal was to assess the absolute levels of contamination as well as the magnitude of intersampling variability for both chemical and toxicological measurements. We compared intersampling variability in chemistry at the three reference sites located within San Francisco Bay with that observed at the three ocean reference sites. Specifically, we tested the hypothesis that intersampling variability for individual compounds would be greater at in-Bay reference stations (due to their closer proximity to sources of contamination) than at the offshore reference stations.

Approach

Intersampling variability was evaluated at each of six reference sites sampled as part of the Harbor studies (Figures 5-1 through 5-3). Three of the references are from locations within San Francisco Bay: Alcatraz mound, Alcatraz environs (Figure 5-1), and the Bay Farm Borrow Pit site (Figure 5-2). The other three sites are located off-shore of the Golden Gate Bridge: the Off-shelf, Point Reyes course-grained, and Point Reyes fine-grained sites (Figure 5-3). These reference sites were sampled on five different occasions starting in March 1990 and ending June 1991 with the exception of the two reference sites at Pt. Reyes. Comparable data for the Pt. Reyes fine- and course-grained references were only available from four and three sampling events, respectively, because of changes in sampling location made after March 1990.

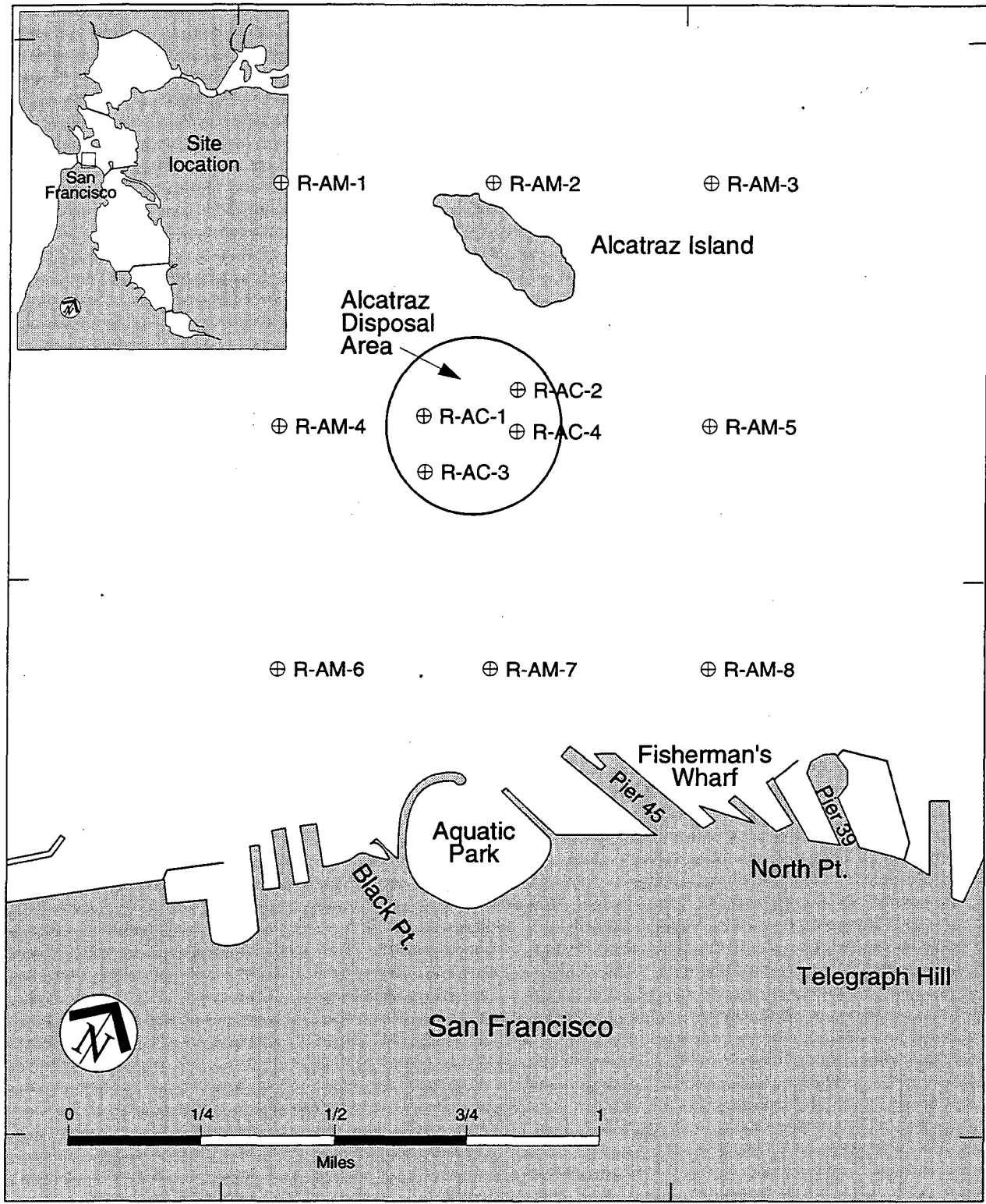


Figure 5-1. Location of the Alcatraz mound and environs reference sites. The Alcatraz mound reference is a composite of sediment taken from four locations (R-AC1 through R-AC4). The Alcatraz environs reference is a composite of sediment taken from eight locations (R-AM1 through R-AM8) (adapted from Battelle, 1992c).

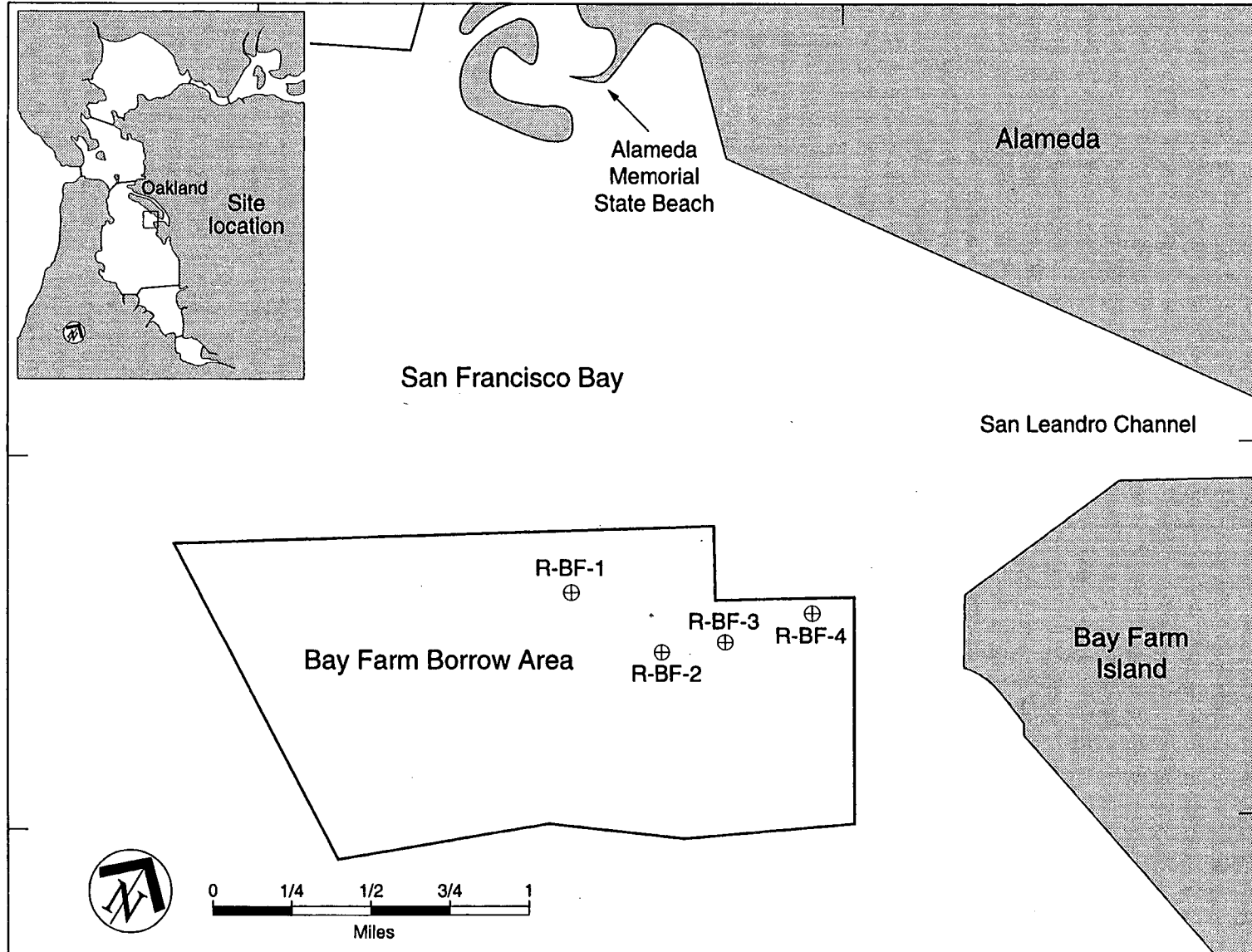


Figure 5-2. Location of the Bay Farm reference site. The Bay Farm reference is a composite of sediment taken from four locations (R-BF1 through R-BF4) (adapted from Battelle, 1992a).

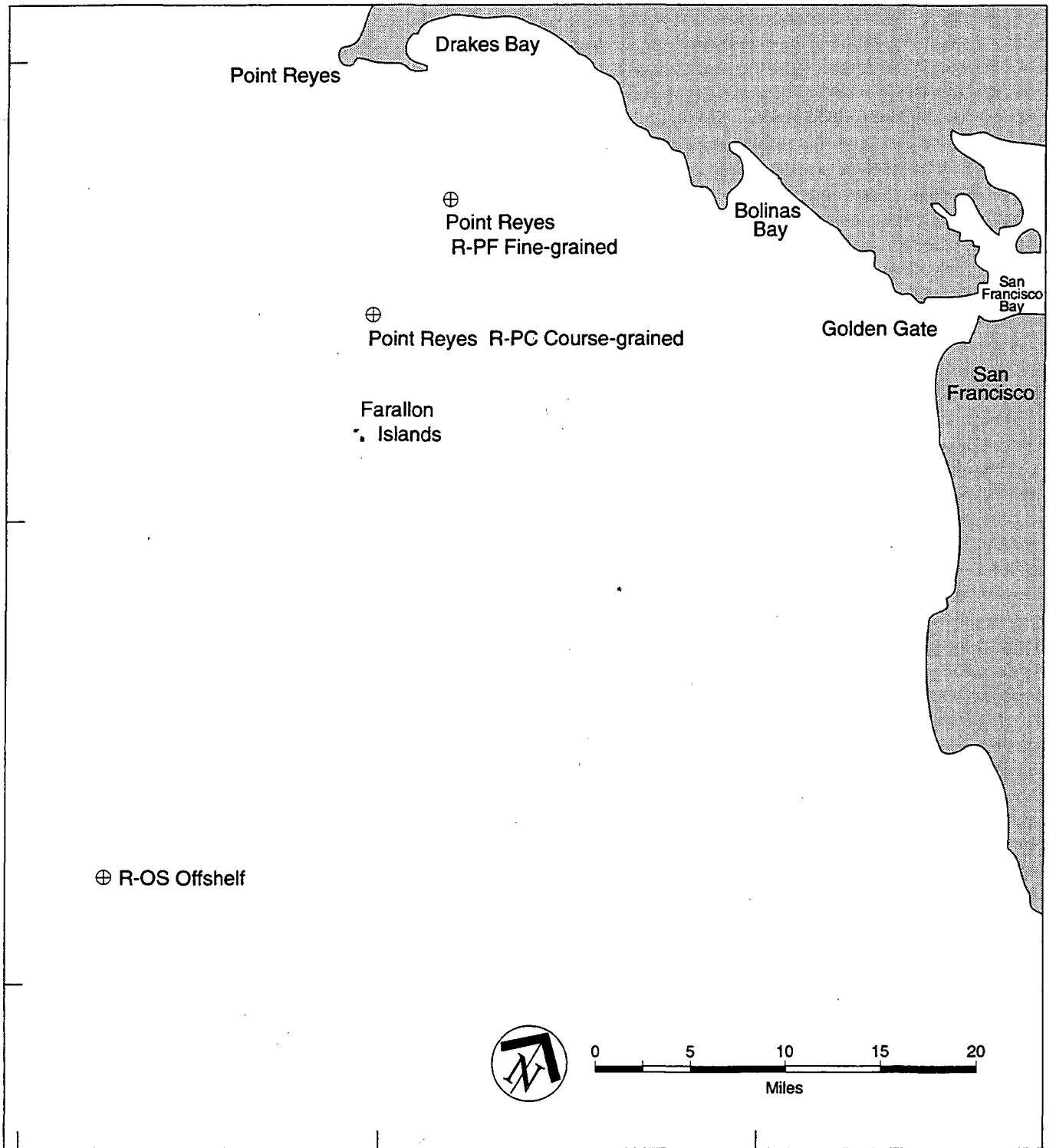


Figure 5-3. Location of the three ocean reference sites: Point Reyes fine-grained (R-PF) and Point Reyes course-grained (R-PC) sites, and Off-shelf (R-OS) (adapted from Battelle, 1992d).

For each reference site, an intersampling coefficient of variation (CV) was calculated for each of 26 chemical parameters¹ (grain size, total organic carbon, organic compounds, and trace metals). CV values were also calculated using the survivorship results from the solid-phase toxicity tests using the amphipod, *Rhepoxynius abronius*, and the polychaete, *Nephtys caecoides*. Differences in intersampling variability between in-Bay and ocean reference sites were evaluated for each chemical parameter by comparing the mean CV values for each reference grouping (in-Bay vs ocean). Note that CV values were used expressly as a means for comparing the intersampling variability associated with particular sites and/or contaminants. For the purposes of this analysis, the magnitude of any given CV is less meaningful than its ranking relative to the CV values from other reference sites.

Results

Generally, measured concentrations of sediment contaminants were low at the six reference sites during this one year period. Although values for the coefficients of variation were often high (CV values ranged from 2.2 to 224%), this variation occurred over very low concentration ranges. Overall, consistently higher intersampling variability was observed for sediment organic compounds (mean CV = 150%) as compared to physical parameters (mean CV = 70%) or trace metals (mean CV = 40%).

There were several instances in which unusually high concentrations of a particular class of compounds were detected at a particular site. For example, at the Alcatraz mound reference site, a 36 µg/g spike of total polynuclear aromatic hydrocarbons (PAH) was seen in the March 1990 Oakland Harbor Berth Areas study (Figure 5-4a). Moderate PAH contamination was observed at the other two in-Bay reference sites, but at much lower concentrations (i.e., 1.6 to 2.3 µg/g). Likewise, a spike of total PCB (260 µg/kg) was detected at the Off-shelf reference site during the Oakland Inner Harbor (Phase 3A repeat)

¹ CV values were only calculated for those chemical parameters that were consistently present above detection limits during the 1990-91 sampling period.

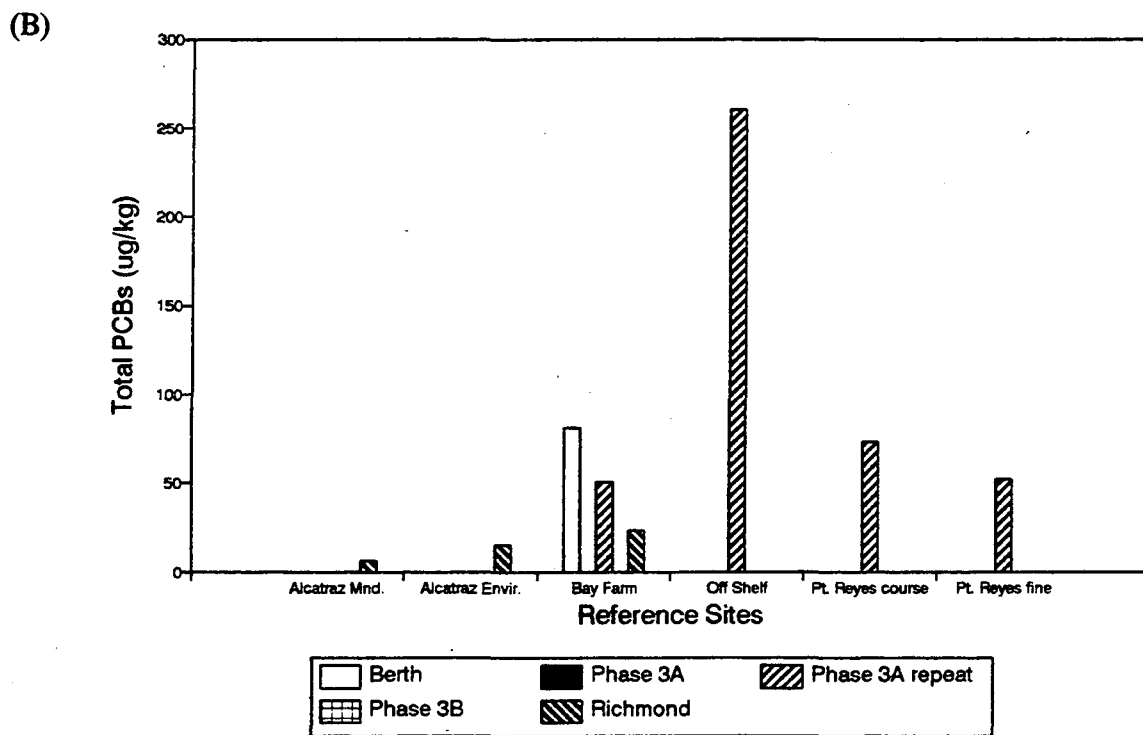
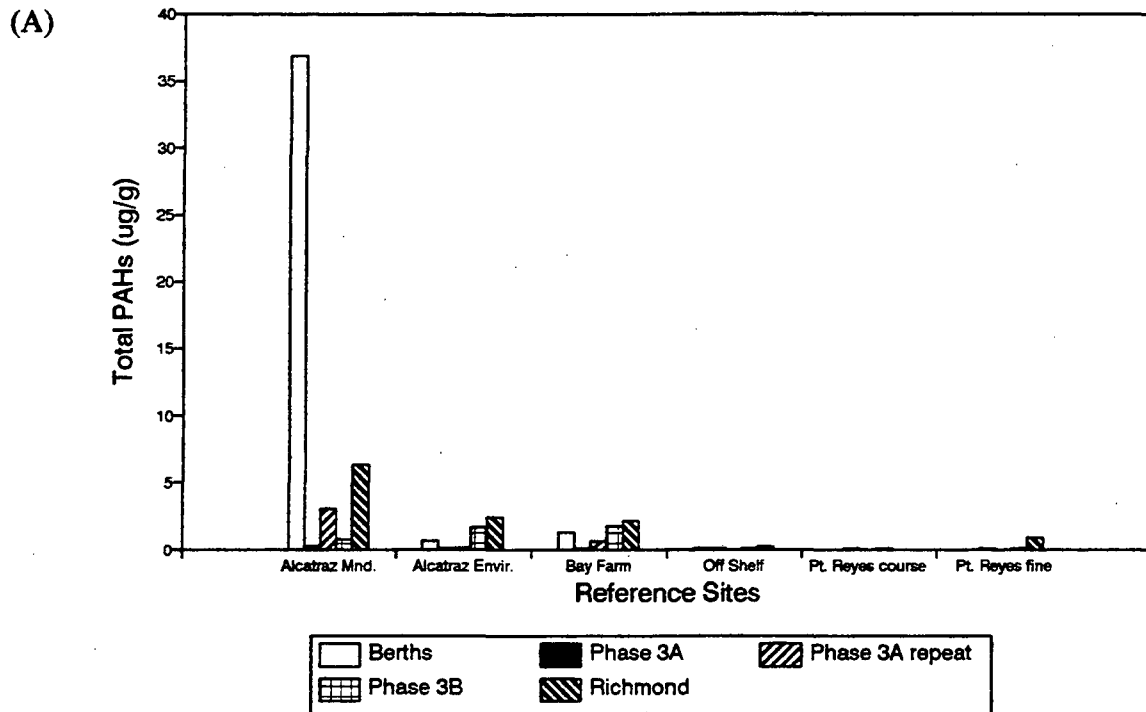


Figure 5-4. Sediment dry-weight concentrations of (A) total PAH and (B) total PCB at six reference sites tested in five harbor studies (March 1990 - June 1991).

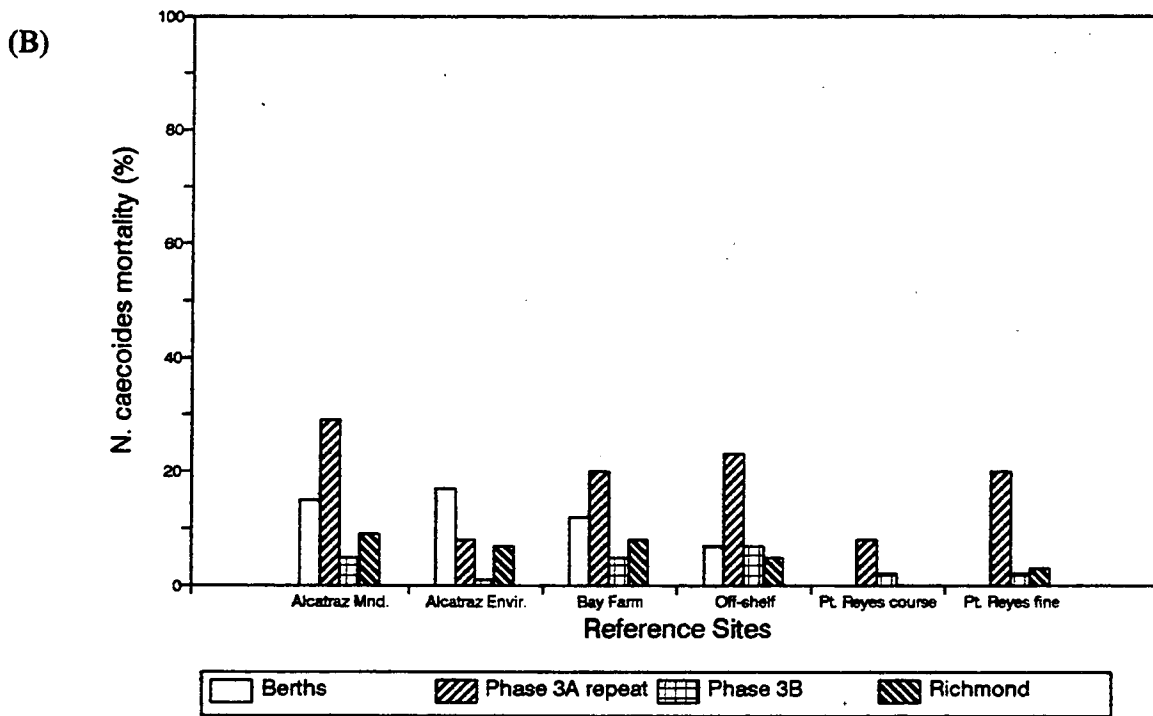
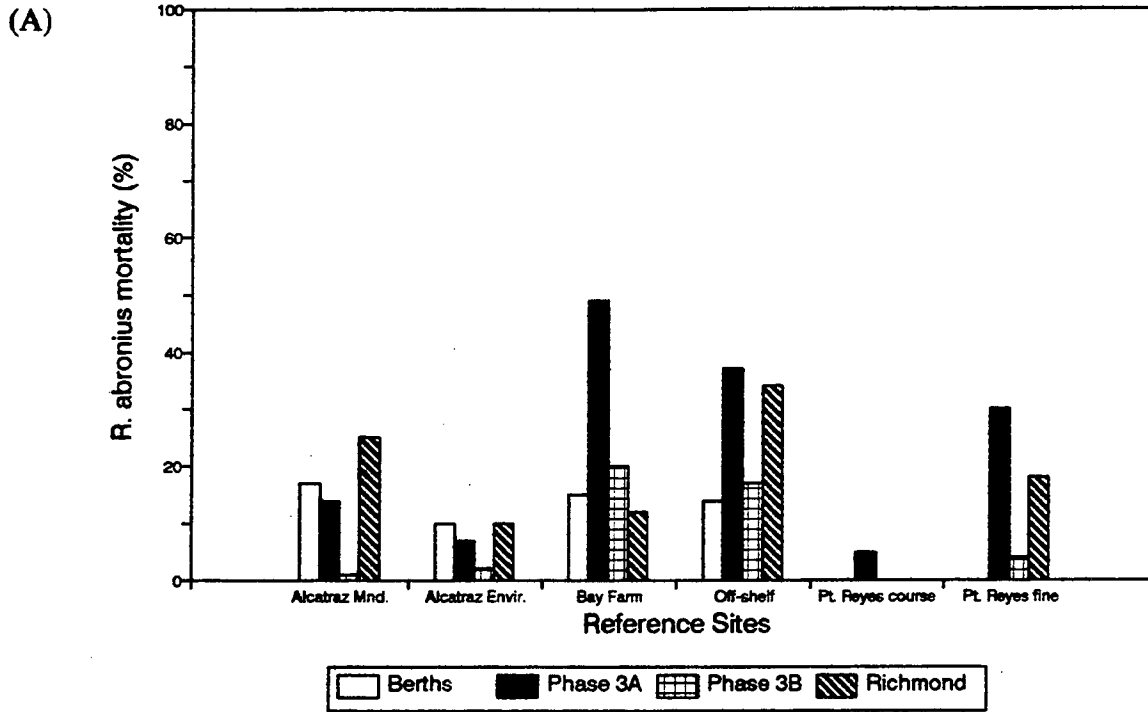


Figure 5-5. Mortality of (A) *Rhepoxynius abronius* and (B) *Nephtys caecoides* at six reference sites tested in four harbor studies (March 1990 - June 1991).

sampling in September 1990 (Figure 5-4b). In this study, elevated but lower levels of total PCB was also observed at the other two ocean reference sites (Pt. Reyes fine and coarse). Notably, neither major nor minor variations in sediment contamination were mirrored in the responses of either the amphipod (*R. abronius*) or the polychaete (*N. caecoides*) (Figure 5-5a and b). *N. caecoides* mortality was generally less than 20%, with the exception of results from the Oakland Inner Harbor study (Phase 3A repeat), for which mortality reached 29% at the Alcatraz mound site. *R. abronius* mortality exceeded 20% on several occasions at both in-Bay and ocean reference sites, with the highest mortality (49%) observed in a sample from the Bay Farm site. Nevertheless, significant mortality was never observed in sediment samples containing unusually high PAH or PCB contamination.

Intersampling variability in chemical measurements over a one-year sampling time was not obviously higher at in-Bay reference sites as compared to ocean sites (Table 5-1). For physical parameters and most trace metals, higher mean CV values were observed in the in-Bay reference grouping as compared to the ocean reference sites. In contrast, ocean sites consistently had the highest mean CV values for organic compounds such as PAHs and PCBs.

Discussion

The high intersampling variability observed for many chemical and physical parameters at the Alcatraz mound station is not surprising given that this station has served as a disposal site for dredged material for the last 15 years². What is surprising about the data from this site is the general absence of a toxic response to instances of elevated contamination. For example, the low molecular weight component of the PAH spike seen at the Alcatraz mound site in Figure 5-4a had an OC-normalized concentration of 2,375 µg/gOC, which is well within the range of concentrations that are expected to be toxic to *R. abronius*

² Responding to frequent observations of elevated contamination and toxicity at this location, the EPA and Army Corps of Engineers have suspended its use as a reference site for the dredged material testing program substituting instead the Alcatraz environs reference. Alcatraz environs is a composite of stations outside of but contiguous to the mound itself (U.S. Army Corps of Engineers, 1993).

Table 5-1. Mean coefficients of variation for selected chemical parameters* at in-Bay and ocean reference sites

Parameter	In-Bay (a) CV (%)	Ocean (b) CV (%)	Parameter	In-Bay CV (%)	Ocean CV (%)
Sand	<i>42</i>	22	Napthalene	100	<i>160</i>
Silt	<i>61</i>	51	Fluorene	128	<i>212</i>
Clay	<i>61</i>	31	Phenanthrene	117	<i>130</i>
TOC	<i>66</i>	34	Anthracene	121	<i>212</i>
Arsenic	<i>27</i>	16	LPAH	118	<i>133</i>
Chromium	<i>43</i>	11	Pyrene	108	<i>158</i>
Copper	<i>34</i>	15	Benzo(a)anthracene	118	<i>179</i>
Lead	<i>22</i>	17	Chrysene	107	<i>173</i>
Mercury	<i>94</i>	44	Benzo(b,k)fluoranthene	126	<i>154</i>
Nickel	<i>27</i>	17	Benzo(a)pyrene	103	<i>188</i>
Selenium	<i>119</i>	<i>132</i>	Indeno(1,2,3-c,d)pyrene	99	<i>157</i>
Silver	<i>51</i>	<i>54</i>	Benzo(g,h,i)perylene	96	<i>173</i>
Zinc	<i>26</i>	11	Total PCB	119	<i>154</i>

For each parameter, the higher of the mean CV values is italicized.

* Acenaphthylene, Acenaphthene, Dibenzo(a,h)anthracene, DDT, DDD, DDE, Endrin, and Dieldrin were below detection limits in most reference samples and, therefore, are not included in this analysis.

(a) In-Bay reference sites include: Alcatraz mound (R-AC), Alcatraz environs (R-AM), and Bay Farm Borrow Pit (R-BF).

(b) Ocean reference sites include: Point Reyes fine-grained (R-PF), Point Reyes course-grained (R-PC), and Off-shelf (R-OS).

(Swartz, personal communication). It appears that factors such as sediment organic carbon or a non-homogenous distribution of contaminant in the sediment decreased the bioavailability of total PAH to amphipods in this sample.

The incidence of PCB contamination at the ocean reference sites was unexpected. The one-time peak observed at the Off-shelf site (Figure 5-4b) appears to reflect the difficulty of accurately locating a site located in nearly 1,300 meters of water. The area in the immediate vicinity of the Off-shelf reference site has been used previously for the disposal of industrial and military waste, which might explain the presence of PCB-contaminated sediment in this sample. At the time of the September 1990 sampling, there was a suspicion that the sampling position was not accurate (Word, personal communication). Subsequently, greater care has been taken in positioning during sampling, and no appreciable PCB has been detected in sediment samples at this site. Although PCB concentrations at this site were elevated, the OC-normalized concentration ($26\mu\text{g/gOC}$) was well below levels considered toxic to *R. abronius*.

A comparison of the mean CV values for three in-Bay and three ocean reference sites was used to test the hypothesis that sediment chemistry at in-Bay reference sites was more variable than at ocean sites. There are several reasons why the in-Bay references were expected to be more variable than ocean references. The Alcatraz and Bay Farm reference sites were expected to be more variable than the ocean sites, because these in-Bay sites are relatively shallow and located in areas that are subject to either the direct influence of dredged material disposal or the indirect influence of resuspended sediments. In contrast, the ocean reference sites are located in deeper water that, at least theoretically, should be more distant from sources of contamination.

Although episodically high levels of contamination were more frequently observed at in-Bay sites, overall intersampling variability at these reference locations did not differ significantly from that at ocean sites. There are several possible explanations for why intersampling variability at the ocean sites was higher than expected. Of the three ocean

sites, the Pt. Reyes fine-grained (R-PF) reference is the most influenced by the discharge from San Francisco Bay and is, by definition, a deposition zone for finer material. Higher deposition rates at this site of material originating from the Bay may explain the higher variability observed for several chemical parameters. The difficulty in accurately sampling the Off-shelf (R-OS) reference site combined with what appears to be patchy contamination in the vicinity of this site due to historical dumping activities may account, in part, for the observed spikes of PCB and other organic compounds.

The pattern of higher intersampling variability seen for measurements of organics relative to physical and trace metal components may be a by-product of the higher recovery associated with trace metals as compared to those for organic compounds.

Conclusions

Based on this examination of trends in contamination and toxicity at several in-Bay and ocean reference sites, the following conclusions can be drawn:

- Sediment concentrations of trace metals and organic compounds at the six in-Bay and ocean reference sites in this analysis were generally low, with occasional elevated levels of organic contaminants (PAH or PCB) observed at both in-Bay and ocean sites.
- Amphipod and polychaete survival in solid-phase bioassays did not appear to respond to these fluctuations in sediment contamination. However, there were several reference samples for which amphipod survival was lower than expected without any apparent chemical cause.
- Intersampling variability in concentrations of physical and chemical parameters was no greater at in-Bay than at ocean reference locations.

- Consistently higher intersampling variability was observed at all reference sites for organic compounds as compared to trace metals and physical parameters.

Our analysis considered only a small fraction of the chemistry and toxicity data available with which to evaluate these reference sites. Although these results demonstrate that sediment contamination at the majority of existing reference sites is generally low, an analysis of the larger data set would be necessary to determine which reference sites can provide a consistent basis of comparison for use in sediment monitoring programs. These results do demonstrate the need to closely evaluate conditions at selected in-Bay and ocean sites to determine the extent and cause of episodic PAH and PCB contamination. Likewise, instances of high amphipod mortality should be examined in light of the potential for non-contaminant influences.

6. AN EVALUATION OF EXISTING SEDIMENT QUALITY GUIDELINES USING DATA FROM SAN FRANCISCO BAY

Introduction

In the absence of federally-mandated sediment quality standards, many scientists and regulators have advocated the development of sediment quality guidelines to aid in the identification and prioritization of potential sediment contamination areas (DiToro, *et al.*, 1991; Long, *et al.*, 1994 in press). In some regions of the United States, the utility of guidelines has been extended beyond informal use as screening tools to regulatory use, particularly in the suitability-determination process for the unconfined disposal of dredged material.

There are many debates in the scientific and regulatory communities as to the best methods for developing sediment quality guidelines, the appropriate use of guidelines in the regulatory process, and the applicability of guidelines developed for one region to other regions. Many individuals question the utility of basing any sort of guidelines or standards on measurements of bulk sediment concentrations, because these are a poor predictor of the actual exposure of an organism. Until factors that govern bioavailability of contaminants in different sediments are better understood, and methods for normalizing the influence of these factors in different sediment types are fully developed, sediment guidelines must be carefully scrutinized to determine their applicability to individual regions.

In the following section, we evaluate the extent to which seven different sediment quality guidelines predict toxicity in sediment samples from San Francisco Bay (Table 6-1). Three different approaches to guideline development are represented: the Apparent Effects Threshold (AET), the Effects Range-Low/Median (ERL/ERM), and sediment-water Equilibrium Partitioning (EqP). These particular guidelines were chosen, because they are among those most often referenced in evaluations of the ecological risks associated with sediment contamination.

Table 6-1. Sediment quality guidelines

Indicator	Normalization	Target Species	Source
Effects Range-low (ERL) Effects Range-medium (ERM)	dry-weight dry-weight	Marine benthos Marine benthos	Long <i>et al.</i> , in press
Puget Sound Amphipod AET Puget Sound Amphipod AET	dry-weight organic carbon	<i>Rhepoxynius abronius</i>	Barrick <i>et al.</i> , 1988
Puget Sound Bivalve AET Puget Sound Bivalve AET	dry-weight organic carbon	<i>Crassostrea gigas</i>	Barrick <i>et al.</i> , 1988
EPA Sediment Quality Criteria	organic carbon	Marine benthos	EPA, 1993

Goals

The goal of our analysis was to evaluate the reliability of these sediment guidelines for San Francisco Bay by determining the percent incidence of acute toxicity associated with guideline exceedence. An overview of each guideline derivation approach is given including the assumptions involved, sources of uncertainty, and (where appropriate) the context in which it is commonly used. This is followed by an evaluation of the reliability of different guidelines by comparing them to sediment chemistry and toxicity data from the San Francisco Bay data set.

Background on sediment quality guideline derivation

Apparent Effects Threshold (AET)

An Apparent Effects Threshold is defined as the bulk sediment concentration of a given chemical above which statistically significant ($p \leq 0.05$) biological effects (e.g., mortality, developmental abnormality, infaunal abundance) are always observed in the data set used to generate the AET. Synoptic measurements of sediment chemistry and biological effects from field and laboratory studies are compared to a "clean" reference site to determine impacted and non-impacted samples. AETs for each chemical are then determined on an species-specific basis as the highest concentration that does not elicit

statistically significant effects. AET values are based on a correlation between chemical concentrations and biological effects, rather than definitive proof of a causal relationship.

As an example, a schematic representation of AET determination is provided in Figure 6-1. Each point in the top bar is a sediment sample that, for a given concentration of this contaminant, was not significantly toxic for a particular test organism. Likewise, each point in the bottom bar is a sample in which significant toxicity was observed. The AET is then set at the highest detected concentration that did not exhibit statistically significant toxic effects. AETs are calculated in this manner to compensate for the fact that it is difficult to determine the cause of toxicity in sediment samples containing a mixture of different contaminants. By definition, no-effects data give a good indication of the potentially non-toxic concentrations of a given chemical. Using both the no-effects and effects data to set the threshold at a point above which all studies have recorded significant toxicity, increases the likelihood that a guideline reflects an actual threshold of toxicity.

AETs for organic compounds are expressed as both dry-weight and organic carbon normalized (OC-normalized) values. Various forms of chemical normalization are often performed to account for variable bioavailability of contaminants in samples due to variable sediment types. OC-normalization of bulk sediment organic concentrations is considered important for non-ionic organic compounds, because the sorption capacity of many sediments for this class of compounds is primarily determined by the mass fraction of organic carbon in the sediment. While AETs for trace metals are currently available only as dry-weight values, recent findings have demonstrated that the bioavailability of some metals is strongly influenced by sulfide binding, suggesting the need for normalizing bulk metal concentrations to the acid volatile sulfide (AVS) content of sediments.

There are several drawbacks to the AET approach that may lower the predictive value of these guidelines. Although data from field collected sediments is used, the AET approach does not take into account possible synergistic and antagonistic interactions of the chemical mixtures found in these sediments. Furthermore, periodic updating of AETs can

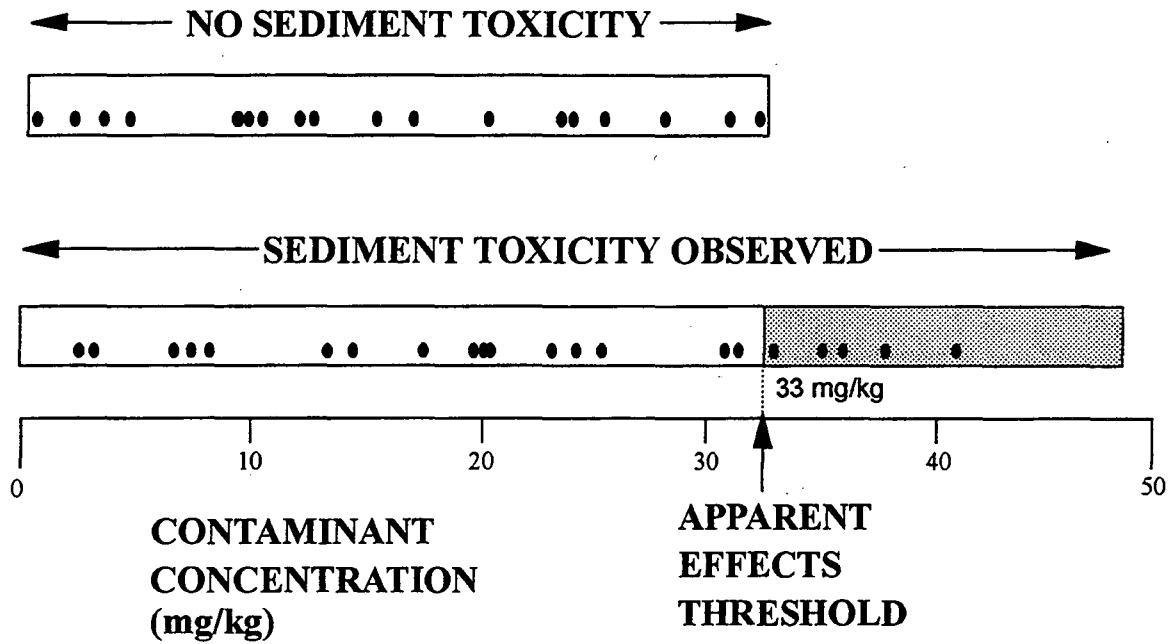


Figure 6-1. A schematic representation of AET derivation based on results of toxicity testing. Dots represent individual sediment samples. Generally, the AET value is set at the sediment concentration above which toxicity is always observed in the data base used to generate the AET (adapted from Becker *et al.* 1990).

only serve to make the thresholds less stringent (i.e., higher) but never more stringent. In a sense, threshold determination is driven by the data points on the top bar (no sediment toxicity) in Figure 6-1. This makes AET determination potentially sensitive to false negatives in the data set.

AETs have been developed for the Washington State Department of Ecology for use in the identification and classification of sediments potentially causing adverse impacts on biological resources (Barrick, *et al.*, 1988). Puget Sound AETs were developed for four biological indicators: amphipod mortality in *Rhepoxynius abronius*, bivalve larval abnormality using the Pacific oyster (*Crassostrea gigas*), benthic invertebrate abundance, and bacterial luminescence using Microtox™. The Puget Sound Dredge Disposal Analysis (PSDDA) program uses AETs as the basis for setting screening and maximum levels (SLs and MLs) for the evaluation of dredged material proposed for aquatic disposal. The PSDDA program uses SLs and MLs as a first tier screening of dredged material to determine the potential for adverse biological effects (Ginn & Pastorok, 1992).

California AETs have also been developed using data from both northern and southern parts of the state (Becker, *et al.*, 1990). These AETs, however, were derived from a relatively small data base as compared to the Puget Sound AETs. Consequently, the uncertainty associated with these values is high. As part of its Regional Monitoring Program, the San Francisco Bay Regional Water Quality Control Board (SFBRWQCB) is currently updating and standardizing its sediment quality database in order to facilitate recalculation of AETs for San Francisco Bay (Taberski, personal communication).

Effects Range-Low (ERL) and Effects Range-Median values

Long and Morgan first calculated ERL and ERM values for the National Oceanic and Atmospheric Administration (NOAA) as a non-regulatory, screening tool for identifying and prioritizing National Status and Trends Program sites with respect to the potential for adverse biological effects due to sediment contamination (Long & Morgan, 1990). Recently, these guidelines were recalculated based on an expanded data base, as well as

revisions in data selection criteria (Long, *et al.*, 1994 in press). While these revised values are not recommended for regulatory use, they are a valuable tool for screening sediment chemistry.

ERL and ERM values have been calculated for nine trace metals, 16 organic compounds, and three classes of PAHs by pooling information from laboratory spiking and field studies as well as existing guideline values from other derivation methods (including AETs and values based on the EqP approach). All test data used in the calculation of ERL and ERM values were screened for the presence of linear concentration-response relationships showing a discernable chemical gradient associated with observed toxic effects. The data and guidelines for a given compound were then sorted according to increasing chemical concentration. The ERL is calculated from this sorted data as the lower 10th percentile concentration at which some form of significantly toxic effects have been observed. Stated otherwise, ERL values are equivalent to the sediment concentrations where adverse effects may begin. The ERM is the 50th percentile concentration or the concentration approximately midway in the range of reported values associated with biological effects. It is the concentration at which effects are frequently observed.

Although ERL and ERM values were intended for the specific non-regulatory purposes of the NOAA Status and Trends Program, they have been used both informally and officially as guidelines for sediment quality screening in many areas of the country (MacDonald, 1992; Wolfenden & Carlin, 1992). Their use by the San Francisco RWQCB in developing sediment screening criteria for wetland creation has prompted criticism on both their derivation and their specific applicability to the sediments of the San Francisco Bay (Ginn, 1993).

Perhaps the most fundamental criticism of the ERL and ERM guidelines focuses on the absence of no-effects data in the derivation of these values. Much of the data used to generate ERL and ERM values are from field studies of sediments that are contaminated with a mixture of chemicals. In such cases, definitive cause-effect relationships are difficult

to determine, and the information provided by no-effects data is particularly important to distinguish co-occurring, but potentially nontoxic chemicals, from those chemicals which are likely to be causing toxicity.

Additionally, the inclusion of other sediment quality guidelines for ERL and ERM derivation appears to be both an asset and a drawback for the reliability of these values. Clearly, the use of sediment guidelines, which are, themselves, based on large data sets, can add to the preponderance of evidence indicating adverse effects above a given sediment concentration. However, the specific derivation of ERL and ERM values (i.e., the calculation of 10th and 50th percentile concentrations) gives equal weight to effects data from a single study and concentrations from guideline values. Thus, there is a large potential for the calculated ERL or ERM to be disproportionately biased by the results of one study.

EPA Sediment Quality Criteria (SQC)

The sediment-water equilibrium partitioning approach (EqP) is the basis for the development of EPA Sediment Quality Criteria for the protection of benthic aquatic life. Thus far, SQC have been proposed for six non-ionic organic compounds (flouranthene, acenaphthylene, phenanthrene, dieldrin, and endrin). SQC values are defined as the sediment concentrations of a particular chemical at which EqP-predicted pore water concentrations are not expected to exceed EPA water quality criteria. The following is a simplified version of the equation used to derive SQC:

$$SQC_{oc} = K_{oc} FCV$$

where SQC_{oc} is the OC-normalized criteria value ($\mu\text{g}/\text{kgOC}$), K_{oc} is the partition coefficient for sediment organic carbon, and FCV is the EPA water quality criteria Final Chronic

Value¹ (µg/L) for the compound of interest. Conceptually, one can understand SQC as a water quality criteria value (FCV) adapted to a sediment context based on the expected partitioning of a chemical between the sediment and soluble phases (defined by K_{oc}). A complete explanation of the derivation and assumptions of applying the EqP approach to SQC development is given by DiToro *et al.* (1991).

EqP theory is used to predict contaminant concentrations that would occur in sediment pore water under conditions of thermodynamic equilibrium with the sediment. Several important assumptions are involved in the application of EqP theory to developing SQC:

- partitioning of a contaminant between the sediment and soluble phases is at equilibrium such that K_{oc} is static and quantifiable,
- contaminant partitioning at equilibrium occurs primarily between two phases (pore water and sediment),
- exposures occurring in sediment pore water are equivalent to those in a water-only system.

Considering the dynamic nature of the estuarine environment, in which factors such as salinity, pH, and redox potential of the sediments are in constant flux, it is unlikely (especially for chronic exposures), that steady-state or equilibrium conditions between sediment and water are ever achieved. Furthermore, other unconsidered variables such as dissolved organic carbon appear to have a significant influence on the partitioning and bioavailability of contaminants in pore water. Nevertheless, data from laboratory spiking studies have demonstrated that EqP theory gives a good approximation of bioavailable contaminants for many types of organic compounds and sediments (DiToro, *et al.*, 1991).

¹ EPA Final Chronic Values are calculated based on chronic toxicity data as part of water quality criteria development.

Approach

Data from the 197 samples which comprise our San Francisco Bay data set² were used to evaluate the application of various sediment quality guidelines to San Francisco Bay sediments. Sediment chemistry data were compared to the guideline values for 31 chemicals (eight trace metals and 23 organic compounds including low and high molecular weight polynuclear aromatic hydrocarbons (LPAH/HPAH), pesticides, and PCBs). These chemicals are among the parameters that are most commonly measured in sediment monitoring programs. Samples for which significant acute toxicity, using at least one test species, co-occurred with contaminant concentrations exceeding a guideline for a particular chemical were designated toxic exceedences (TE). The percentage of the total samples tested for which guideline exceedences were also toxic (%TE) was calculated for each chemical within a particular guideline type as a measure of that value's predictive potential for biological effects in these samples.

Toxic exceedences were determined for the ERL, ERM, and proposed EPA Sediment Quality Criteria values using all of the samples, regardless of the species used for toxicity testing. Exceedences for the species-specific Puget Sound AETs were determined using only samples in which the appropriate species (amphipod or bivalve) were tested.

Organic carbon in the samples examined in this study ranged from 0.04 to 4.3% in the SedQual 3 studies and from 0.01 to 1.4% in the Harbor studies. Samples with organic carbon content less than 0.2% were not included in the exceedence analysis of OC-normalized guidelines. In sediments with very low organic carbon fractions, other factors (e.g., particle size and sorption to mineral fractions) appear to be relatively more influential in the partitioning of organic chemicals (DiToro, *et al.*, 1991).

The opportunity to conduct more detailed evaluation of guideline efficacy for predicting sediment toxicity was limited due to the low number of guideline exceedences for

² A summary of the San Francisco Bay studies used in this report is found in Table 1-2 of Section 1.

many chemicals and small size of the data set used. We did, however, examine the relationship between the magnitude of guideline exceedence and the degree of toxic response for chromium and mercury, for which there were a relatively large number of exceedences over a broad range of sediment concentrations and degrees of toxic response.

Results

The complete results of the guideline exceedence analysis as well as the guideline values themselves are presented in Tables 6-2 through 6-4. The frequency of exceedences is schematically summarized for the four dry-weight guidelines in Figure 6-2. Guideline values for many of the chemicals were exceeded in only a few (generally less than five) of the 197 sediment samples. The most notable exception to this observation was for the ERL values, all of which were exceeded in at least one sediment sample, and 20 of which were exceeded by more than 10 samples. Generally, more samples exceeded the guidelines for trace metals than for organic compounds.

Although these sediment guidelines cover a wide range of chemicals, there are still commonly measured contaminants that are missing from their ranks. For example, the antifoulant tributyltin and its breakdown products are regarded as potentially significant contaminants in harbors especially near boat maintenance and construction areas, but there are as yet no effects-based sediment guidelines available for butyltins. In 1989, the Puget Sound Dredged Disposal Analysis (PSDDA) agencies set an interim screening level for tributyltin at 30 µg/kg. However, the PSDDA Screening levels are designed specifically as regulatory triggers, indicating the need for biological testing, rather than delineating a threshold of ecological risk in the manner of the sediment quality indicators in Table 6-1³.

The predictive ability of each of these guidelines with respect to toxicity in San Francisco Bay sediments is described below. In addition, the relationship between the

³ Of the 197 samples included in our analysis, there were only five that exceeded this screening level for tributyltin.

Table 6-2. ERL and ERM values and the frequency of their exceedence in San Francisco Bay sediment samples

Chemical	ERL (a) (ug/kg)	#Exceedences	% TE (b)	ERM (a) (ug/kg)	#Exceedences	% TE
HPAH	1,700	19	47.4	9,600	[1] (c)	[0]
Fluoranthene	600	10	40.0	5,100		
Pyrene	665	18	66.7	2,600	4	75.0
Benzo(a)anthracene	261	10	60.0	1,600	3	66.7
Chrysene	384	13	61.5	2,800	[1]	
Benzo(b,k)fluoranthene	-	-	-	-	-	-
Benzo(a)pyrene	430	16	75.0	1,600	3	100.0
Dibenzo(a,h)anthracene	63.4	16	75.0	260	11	100.0
Benzo(g,h,i)perylene	-	-	-	-	-	-
Indeno(1,2,3-c,d)pyrene	-	-	-	-	-	-
LPAH	552	3	66.7	3,160	[1]	[0]
Napthalene	160	1	0.0	2,100		
Acenaphthylene	44	2	0.0	640		
Acenaphthene	16	11	63.6	500	[1]	[0]
Fluorene	19	13	53.8	540	[1]	[0]
Phenanthrene	240	15	46.7	1,500	[1]	[0]
Anthracene	85.3	14	57.1	1,100	[1]	[0]
DDE	2.2	24	70.8	27		
DDD	-	-	-	-	-	-
DDT	-	-	-	-	-	-
Dieldrin	-	-	-	-	-	-
Endrin	-	-	-	-	-	-
Total PCB	22.7	54	51.9	180	4	50.0
	(ug/g)			(ug/g)		
Silver	1	3	66.7	3.7		
Cadmium	1.2	12	50.0	9.6	1	100.0
Chromium	81	144	61.1	370	29	41.4
Copper	34	112	66.1	270	1	100.0
Mercury	0.15	113	59.3	0.71	21	59.1
Nickel	20.9	195	55.9	51.6	159	63.5
Lead	46.7	30	56.7	218	1	0.0
Zinc	150	27	63.0	410	1	100.0
Average %TE			55			71 [50] (d)
Std. dev.			19			30

(a) Effects Range-Low (ERL) and Effects Range-Median (ERM) values from Long et. al. (1993).

(b) %TE is percent of the total samples exceeding a guideline that were toxic to at least one test species.

(c) Square brackets denote that an exceedence or %TE was from the Alcatraz mound reference site (Oakland Harbor Berthing Areas study).

(d) Value in square brackets is average %TE including %TE values from ERMs exceeded only by the single Alcatraz mound reference site. See text for further discussion of this sample.

Table 6-3. Puget Sound amphipod AET values and the frequency of their exceedence in San Francisco Bay sediment samples

Chemical	Amph AET (a) ($\mu\text{g}/\text{kg}$)	#Exceedences	% TE (b)	Amph AET ($\mu\text{g}/\text{gOC}$)	#Exceedences	% TE
HPAH	69,000			5,300	[1] (c)	[0]
Fluoranthene	30,000			3,000		
Pyrene	16,000			1,000	[1]	[0]
Benzo(a)anthracene	5,100			270		
Chrysene	9,200			460		
Benzo(b,k)fluoranthene	7,800	3	67	450	1	0
Benzo(a)pyrene	3,000	1	100	210	3	67
Dibenzo(a,h)anthracene	540	8	100	47	5	100
Benzo(g,h,i)perylene	1,400	6	100	78	9	88
Indeno(1,2,3-c,d)pyrene	1,800	1	100	88	3	67
LPAH	24,000			2,200	[1]	[0]
Napthalene	2,400			220		
Acenaphthylene	1,300			66	1	0
Acenaphthene	2,000			200	1	0
Fluorene	3,600			360		
Phenanthrene	6,900			690		
Anthracene	13,000			1,200		
DDE	15			0.81	5	20
DDD	43	2	0	2.2	4	0
DDT	-	-	-	-	-	-
Dieldrin	-	-	-	-	-	-
Endrin	-	-	-	-	-	-
Total PCB	3,000			190		
	($\mu\text{g}/\text{g}$)					
Silver	5.9					
Cadmium	6.7	1	100			
Chromium	270	29	21			
Copper	1,300	1	100			
Mercury	2.1	1	0			
Nickel						
Lead	660					
Zinc	960	1	100			
Average %TE			72			38 [29] (d)
Std. dev.			41			39

(a) Puget Sound Amphipod AETs from Barrick et al. (1988).

(b) %TE is percent of the total samples exceeding a guideline that were toxic to at least one test species.

(c) Square brackets denote that an exceedence or %TE was from the Alcatraz mound reference site.

(d) Value in square brackets is average %TE including %TE values from ERMs exceeded only by the single Alcatraz mound reference site (Port of Oakland Berth Area Study); See text for further discussion of this sample.

Table 6-4. Puget Sound bivalve AET values and the frequency of their exceedence in San Francisco Bay sediment samples

Chemical	Biv AET (a) ($\mu\text{g}/\text{kg}$)	#Exceedences	% TE (b)	Biv AET ($\mu\text{g}/\text{gOC}$)	#Exceedences	% TE
HPAH	17,000			960		
Fluoranthene	2,500			160		
Pyrene	3,300	3	67	-	-	-
Benzo(a)anthracene	1,600	2	50	110	1	100
Chrysene	2,800	1	100	110	1	100
Benzo(b,k)fluoranthene	3,600	3	67	230	3	67
Benzo(a)pyrene	1,600	3	67	99	4	50
Dibenzo(a,h)anthracene	230	11	18	120		
Benzo(g,h,i)perylene	720	10	20	31	16	19
Indeno(1,2,3-c,d)pyrene	690	5	40	33	12	25
LPAH	5,200			370		
Napthalene	2,100			99		
Acenaphthylene	-	-	-	-	-	-
Acenaphthene	500			16	1	0
Fluorene	540			23		
Phenanthrene	1,500			120	1	0
Anthracene	960					
DDE	-	-	-	-	-	-
DDD	-	-	-	-	-	-
DDT	-	-	-	-	-	-
Dieldrin	-	-	-	-	-	-
Endrin	-	-	-	-	-	-
Total PCB	1,100					
	($\mu\text{g}/\text{g}$)					
Silver	0.56					
Cadmium	9.6	1	0			
Chromium	-	-	-			
Copper	390	1	0			
Mercury	0.59	5	20			
Nickel	-	-	-			
Lead	660					
Zinc	1600	1	0			
Average %TE			37			45
Std. dev.			31			38

(a) Puget Sound Bivalve AETs from Barrick et al. (1988).

(b) %TE is percent of the total samples exceeding a guideline that were toxic to at least one test species.

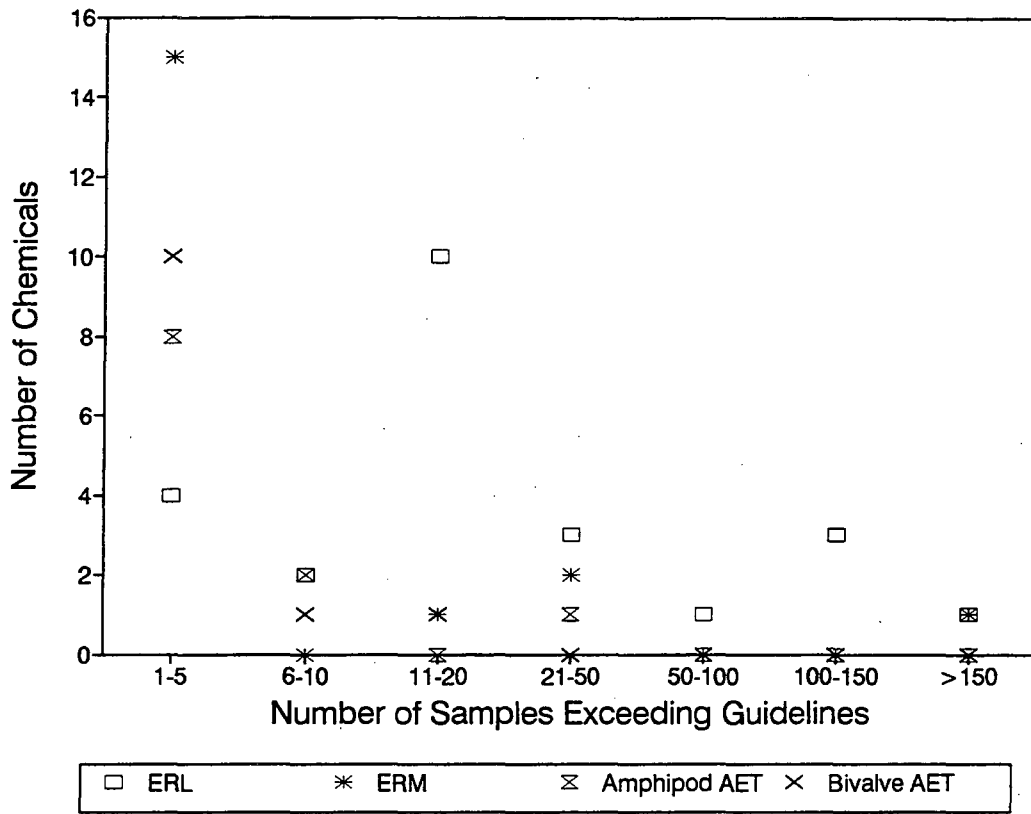


Figure 6-2. Frequency with which the following four dry-weight sediment guidelines were exceeded in San Francisco Bay sediments: Effects Range-Low (ERL) values for 24 chemicals; Effects Range-Median (ERM) values for 24 chemicals; Puget Sound Amphipod Apparent Effects Threshold (AAET) values for 27 chemicals; Puget Sound Bivalve Apparent Effects Threshold (BAET) values for 23 chemicals.

magnitude of guideline exceedence and toxicity is examined in greater detail for mercury and chromium, because the data for these two trace metals contained a relatively large number of exceedences over a wide range of concentrations and toxic responses.

ERL Exceedences

The number of San Francisco Bay samples exceeding trace metal ERL values was highly variable depending on the metal, ranging from 3 to 195 samples (Table 6-2). The ERL for nickel was exceeded in nearly all (195) of the samples tested. Nickel was followed in frequency by chromium, mercury, and copper which were exceeded by 144, 113, and 112 samples, respectively. In contrast, ERL values for organic compounds were exceeded by fewer samples, ranging from 1 to 56 samples depending on the compound.

The frequency with which significant toxic effects were associated with sediment concentrations greater than an ERL was similar for trace metals and organics. On average, 55% of the samples which exceeded an ERL were toxic to at least one test species, indicating that there was effectively a random chance that exceeding the ERL for a particular chemical would accurately predict significant toxicity in these sediments.

ERM Exceedences

ERM values were exceeded less frequently than ERLs (Table 6-2). The highest number of samples (159) exceeded the ERM for nickel. Moreover, ERMs for nickel, mercury, and chromium together accounted for 85% of all the exceedences of ERM values.

An average of 50% of the samples that exceeded an ERM were also toxic, implying that ERM values were no more predictive of acute toxicity than were ERL values. However, this average is highly skewed by a single sample from the Alcatraz mound reference site (Port of Oakland Berth Area study). This sample accounted for the sole exceedence of the ERM values for LPAH, HPAH, chrysene, acenaphthene, fluorene, anthracene, and phenanthrene while not being acutely toxic to any of the test species. If the guideline exceedences due to this individual sample are excluded from the calculations, the

average %TE for ERMs increases from 50% to 71%. The percentage of ERM exceedences that were toxic was higher than the %TEs for ERLs for eight of the 12 remaining chemicals. However, this was not the case for chromium or mercury, as discussed below. Although these two metals had the second and third highest number of ERM exceedences, the %TEs were either equal to or lower than those for the ERLs.

Chromium and ERL/ERM values

Despite a four-fold difference between the ERL (81 µg/g) and ERM (370 µg/g) values for chromium, 61% of the samples exceeding the ERL were toxic as compared to 41% that exceeded the ERM. The concentration-response relationship for the 26 samples containing chromium concentrations that exceeded the ERM is presented in Figure 6-3. There was no linear correlation between the increasing concentrations of total chromium and survival in either the amphipod, *R. abronius* ($r^2 = 0.024$, $p=0.29$) or the polychaete *N. caecoides* ($r^2 = 0.062$ $p=0.45$). Notably, one sample from an area of the Oakland Inner Harbor channel located near a metal recycling facility contained concentrations of total chromium in excess of 900 µg/g, yet mortality in both test species was less than 10%.

Mercury and ERL/ERM values

The ERM value for mercury (0.71 µg/g) was exceeded by 21 samples, 59% of which were significantly toxic to at least one test organism (Table 6-2). The frequency of mercury ERM exceedences is essentially equal to that seen for the ERL value. The concentration-response relationships for these samples are presented in Figure 6-4 relative to mortality in two amphipod species and two bivalve species. Consistently high mortality was observed in samples tested with *E. estuarius*, however, all five data points were derived from a single highly contaminated area at Castro Cove where mercury co-occurred with elevated sediment concentrations of several other trace metals and organic compounds. No significant linear relationship was observed between mortality in either *R. abronius* ($r^2=0.012$, $p=0.74$) or

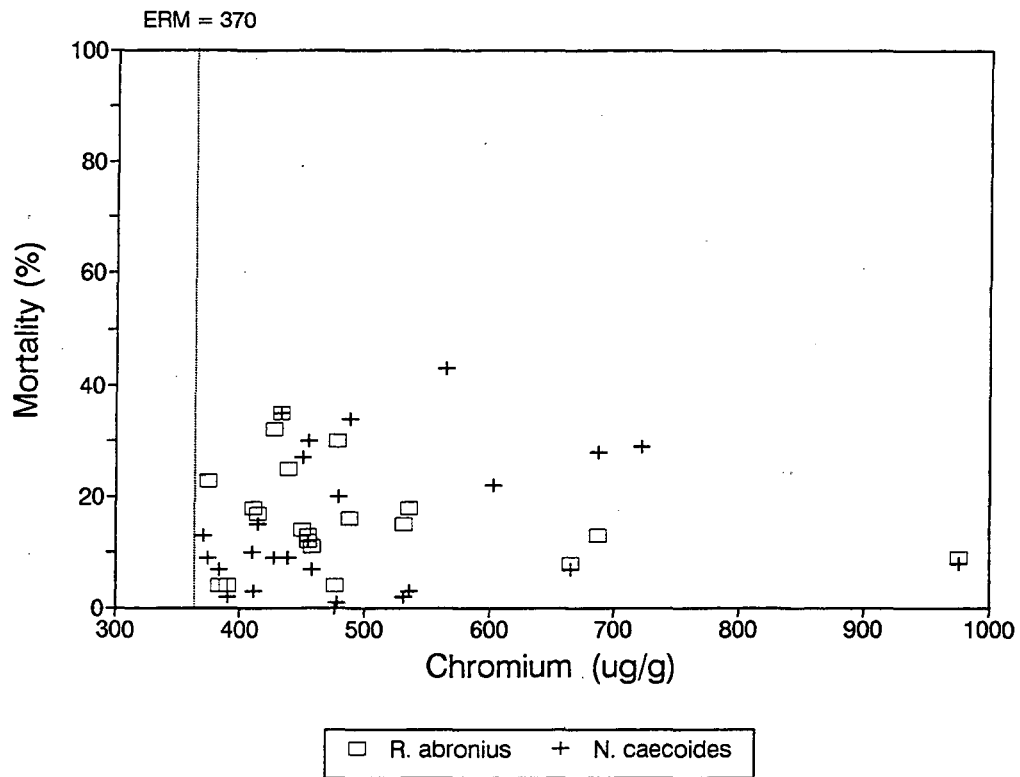
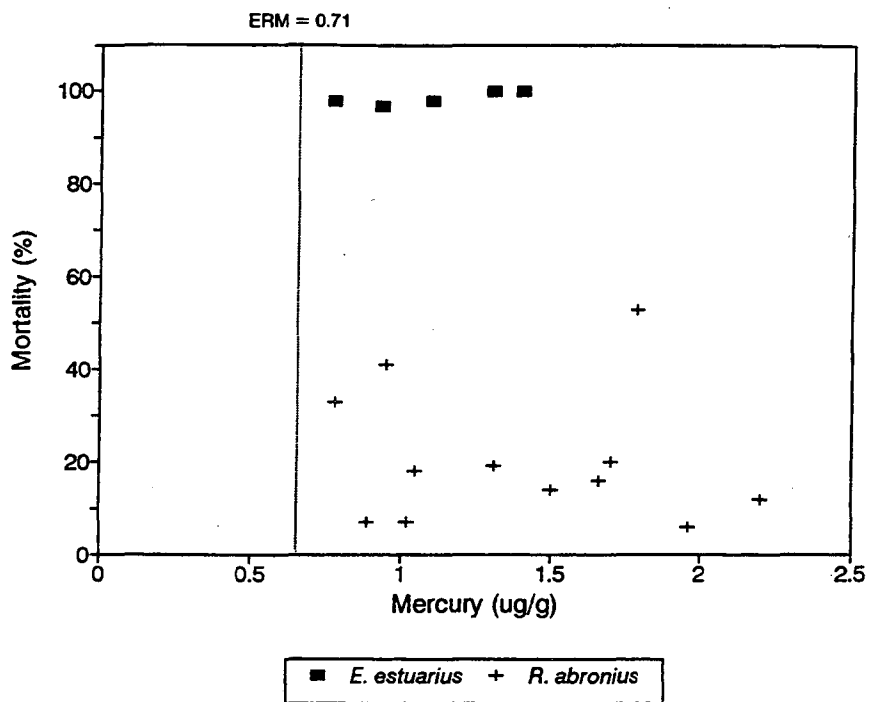


Figure 6-3. Concentration-response relationship between total sediment chromium and mortality in the amphipod *Rhepoxynius abronius* ($r^2=0.062$; $p=0.29$) and the polychaete *Nephtys caecoides* ($r^2=0.024$; $p=0.45$). Data shown are for samples with sediment chromium in excess of the Effects Range-Median value (ERM).

(A)



(B)

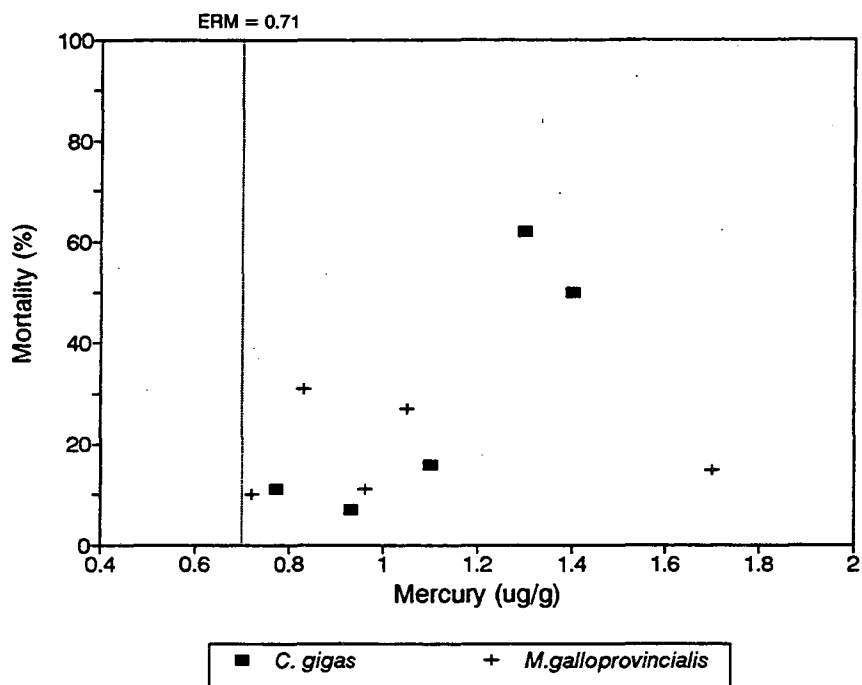


Figure 6-4. Mortality of (A) the amphipods, *Rhepoxynius abronius* and *Eohaustorius estuarius* and (B) bivalve larvae of *Crassostrea gigas* and *Mytilus galloprovincialis* as a function of total sediment mercury values that were above the Effects Range-Median value (ERM).

M. galloprovincialis ($r^2=0.011$, $p=0.87$) and sediment concentrations of mercury. A significant correlation between sediment mercury and toxicity was seen, however, using *C. gigas* ($r^2=0.76$, $p=0.05$) although the total number of samples was very small ($n=5$).

AET Exceedences

Relatively few of these San Francisco Bay samples exceeded the Puget Sound amphipod and bivalve AET values as compared to the ERL and ERM values (Tables 6-3 and 6-4). The highest number of exceedences was observed for the dry-weight amphipod AET for chromium which was exceeded in 29 samples. For organic compounds there were no substantial differences in the number of samples exceeding dry-weight vs OC-normalized AET values. Although the amphipod AET values were almost always higher than those for the bivalve, the total number of exceedences was larger for the amphipod AET, because more samples were tested using this species.

The percentage of samples that exceeded AETs and were acutely toxic averaged from a low of 28% (dry-weight bivalve AET) to a high of 71% (dry-weight amphipod AET). As was the case for the chromium ERM, exceedence of the chromium AET was a poor predictor of amphipod toxicity (%TE = 21%). For a large proportion of the other chemicals, the %TEs are based on a small number of exceedences. Furthermore, samples from the Castro Cove study make up 98% of the guideline exceedences for the dry-weight bivalve AETs and 64% of the exceedences for the OC-normalized bivalve AET values. Significant bivalve larval toxicity was observed in only two of the thirty samples tested from Castro Cove, considerably influencing the %TE calculation for many of the AETs.

EPA Sediment Quality Criteria Exceedences

The EPA's proposed sediment quality criteria (SQC) for three PAHs (fluoranthene, acenaphthylene, and phenanthrene) and two pesticides (dieldrin and endrin) were exceeded by only two of the samples in our data set. Each of the SQC PAH values were exceeded by a single Alcatraz mound reference sample (R-AC from the Port of Oakland Berth Area Study).

The SQC value for endrin was also exceeded in only a single sample. Neither of these samples, however, was significantly toxic to any of the test species.

Discussion

Generally, the small size of the San Francisco Bay data set we evaluated (197 samples) and the small number of exceedences seen for many of the sediment quality guidelines precludes drawing firm conclusions about their accuracy for predicting toxicity in San Francisco Bay sediments. Nevertheless, there were notable differences between different types of guidelines in the frequency with which they were exceeded and the degree of correlation between significant toxicity and guideline value exceedences.

The ERL and ERM values for organics and trace metals were generally the most conservative of the dry-weight guidelines tested and, therefore, were the most frequently exceeded. We would expect that while an ERM value would be exceeded less often than an ERL, any exceedences would be more predictive of toxicity. In fact, this is what was observed for most chemicals with the overall average %TE for ERLs being 55% while that for ERMs was 71%⁴. There were, however, some notable exceptions for individual chemicals. Although a large percentage of all our samples exceeded the ERM values for chromium, nickel, and mercury, these values were either similarly or less predictive of toxicity than the ERLs. Therefore, it is likely that factors other than increasing concentrations of these trace metals are responsible for the toxicity observed in many of the samples. ERL and ERM values appear to be of little use for evaluating risks associated with chromium, nickel, and mercury in San Francisco Bay sediments.

Based on the results of this analysis, we are unable to make definitive conclusions regarding the accuracy of Puget Sound AETs for predicting toxic effects in San Francisco Bay sediments. Although the number of samples exceeding AET values for organic compounds

⁴ The Alcatraz mound reference site from the Port of Oakland Berthing Area study was excluded from this %TE calculation.

was similar to those exceeding ERMs, many of these samples were clustered within a particular study. For example, a relatively large number of samples exceeding AET values for several HPAH compounds (dibenzo(a,h)anthracene, benzo(g,h,i)perylene, and indeno(1,2,3-c,d)pyrene) came from the Castro Cove study. There are two potential problems with drawing conclusions about AET values based on samples primarily from Castro Cove. First, Castro Cove is a relatively small and highly contaminated area. With the available chemical information it is difficult to discern which of the co-occurring chemicals are likely cause(s) of the toxicity observed here. Second, because samples from the Castro Cove study made up a large number of the exceedences for both the bivalve and amphipod AETs, the calculated %TE for many of the chemicals was highly influenced by the extreme differences in toxic response between bivalves and amphipods observed at this site. Specifically, two-thirds of the Castro Cove samples were significantly toxic in the amphipod bioassay, whereas only two out of 30 samples were toxic in the bivalve test.

There are several factors influencing the predictive accuracy of individual sediment quality guidelines for toxicity in San Francisco Bay samples. A detailed discussion of these factors, which include the biological basis for and derivation of the guidelines themselves, their regional specificity (or lack of specificity), and the characteristics of the data set to which they are being compared, is presented below.

By definition, ERL and ERM values delineate the concentration ranges at which toxic effects "occasionally" and "frequently" occur in the data base used to generate these guidelines. Thus, we would not expect that all cases in which ERLs were exceeded would necessarily be correlated with significant toxicity. Furthermore, it is important to remember that ERLs and ERMs were derived from a multistudy and multispecies data base which represents a wide range of species sensitivity and sediment types. For example, in the data base used to generate the revised chromium ERL, toxic effects occurring below the ERL value were derived from four sources: two observations of low to moderate benthic species richness, one observation of increased burrowing time in a clam (*Macoma balthica*), and a chronic marine threshold based on equilibrium partitioning theory. Acute toxicity to

amphipod and polychaete species was not observed until chromium concentrations were at or above the ERM value (370 µg/g). Therefore, it follows that the chromium ERL would not be highly predictive of acute toxicity for the test species used in the San Francisco Bay data set.

Another important consideration that may influence the utility of both Puget Sound AET and ERL/ERM values for San Francisco Bay sediments is the environmental specificity of these sediment guidelines. Marine-specific or generalized guidelines may not account for the variable characteristics that govern contaminant bioavailability in an estuarine environment. For example, the influence of freshwater inputs (particularly important in a shallow estuary such as San Francisco Bay) would not be reflected in AET guidelines developed specifically for the marine benthos of Puget Sound. In addition, AET values based on the response of a marine amphipod (*R. abronius*) may not accurately predict acute toxicity for an estuarine amphipod species such as *E. estuarius*. Likewise, nationally-derived criteria such as the ERL/ERM values and EPA Sediment Quality Criteria may be less likely to account for the specific background geological conditions of a particular region. Thus, high background levels of some trace metals (discussed further in Section 7) and local conditions affecting chemical speciation may affect the applicability of national guidelines to San Francisco Bay. These considerations underscore the importance of evaluating the biological basis of sediment quality guidelines as well as the site-specific sediment chemistry of the area in question before applying them to regional data sets.

The potential for sediment guidelines to accurately predict sediment toxicity is further limited by the accuracy and specificity of the sediment data set in question. For example, bulk sediment measurements of total chromium make no distinction between soluble ionic species in pore water and insoluble complexes associated with the sediment - although the latter are generally non-toxic and the former have differing toxicities to aquatic organisms. While trivalent chromium (Cr^{+3}) is the more toxic ion, hexavalent chromium (Cr^{+6}) is considered to be the more toxicologically significant because it is more reactive and able to cross cell membranes where it is then reduced to Cr^{+3} (Vonburg & Liu, 1993). The speciation of soluble chromium is highly dependant on the oxidation state of the medium. Under the anoxic conditions of most sediments, Cr^{+3} predominates over Cr^{+6} . Therefore, one possible

explanation for why the ERM and Puget Sound amphipod dry-weight AET for chromium are poor predictors of toxicity in these San Francisco Bay sediments may be because elevated levels of total chromium in these sediments are predominantly Cr^{+3} and/or insoluble complexed forms that are relatively non-toxic to the test species.

While our data set did cover a broad range of habitats, sediment types, and degrees of contamination, the incidence of elevated contamination and acute toxicity was not evenly distributed over the studies. Considerable expansion of the data set would enable an evaluation of these guidelines that would be less biased by site-specific or study-specific conditions.

Conclusions

In this section, we examined the utility of several commonly-used sediment quality guidelines for predicting acute toxicity in a San Francisco Bay data set. We found that some guidelines appear to be better predictors of toxicity than others. Specifically, sediment contamination in excess of ERM values was more often correlated with an observation of toxicity than was exceedence of an ERL value. Puget Sound dry-weight amphipod AET values were more predictive of toxicity than were bivalve AETs or OC-normalized AETs; however, exceedences of AET values were limited to a small fraction of the entire data set. The proposed EPA Sediment Quality Guidelines are only available for a limited number of chemicals and were not exceeded in enough samples to permit conclusions to be drawn on their predictive potential.

Perhaps the most limiting factor in our analysis was the size of the data set used. Although a broad range of concentrations measured for many of the chemicals, higher concentrations were frequently represented by an individual "hot" sample or measurements made in a single, highly contaminated area. A larger set of sediment chemistry and toxicity data, representing a wider range of chemical concentrations from multiple sites would

facilitate efforts to determine the applicability of these guidelines to San Francisco Bay sediments.

Considering the numerous (and largely unmeasured) site-specific factors that govern contaminant partitioning and bioavailability in San Francisco Bay, it is unlikely that sediment guidelines alone could ever be adequate for the determination of risks posed by contaminants in sediments. Indeed, the present analysis demonstrates that a significant proportion of the acutely toxic sediments would have been missed had guidelines not been used in conjunction with bioassays. However, sediment quality guidelines can be valuable in both regulatory and monitoring programs for: 1) determination of potential impact areas, 2) prioritizing areas for examination or remediation, and 3) use as relatively inexpensive triggers for additional biological and chemical testing.

Instead of applying of sediment guidelines as a substitute for biological testing, it would be more constructive to integrate guidelines in a tiered manner with biological and chemical testing into sediment monitoring programs. For example, conservative guidelines such as the ERLs appear to be a good indication of the upper threshold of concentrations that are not likely to be toxic and therefore might be useful for preliminary screening purposes. Furthermore, ERL and/or ERM values may be more useful for the identification of areas with significant contamination (ie., where sediments exceed guidelines for numerous chemicals) rather than to determine the specific chemical(s) responsible for toxicity (Berry, personal communication). Ultimately, guidelines such as ERMs, which are more predictive of acute toxicity in San Francisco Bay sediments, could be used as the basis for developing site-specific sediment screening guidelines for regulatory use. In addition, site-specific guidelines could be validated using techniques such as spiked pore water bioassays.

7. BACKGROUND CONCENTRATIONS OF SELECTED TRACE METALS IN SAN FRANCISCO BAY SEDIMENT

Introduction

Sediment quality guidelines have come into greater use as a management tool for the identification (and prioritization) of contaminated sediments that may be toxic to aquatic communities. However, as was discussed in Section 6, many of the existing guidelines have been developed from national data bases and, as such, do not account for region-specific sediment characteristics. With this in mind, an important step in evaluating the applicability of sediment guidelines is to identify whether these values are representative of contamination above the background levels associated with the sediments of a particular region. A sediment guideline that is lower than the naturally occurring background concentration is not useful for the identification of risks associated with anthropogenic contamination.

Earlier surveys of San Francisco Bay sediments such as those performed by Long *et al.* (1988) and Long and Morgan (1992) have reported elevated levels of trace metals at stations throughout the Bay. Sediment levels of mercury, lead, and chromium often exceeded nationally derived Effects Range-low (ERL) and Effects Range-medium (ERM) values established by Long and Morgan (1990) for the National Status and Trends program¹. Sediments with elevated concentrations of nickel and chromium are also frequently observed in harbors and navigation channels tested as part of the dredged material testing program. However, these sediments are frequently not toxic in both solid-phase and suspended-phase bioassays (Ross, personal communication).

Goals

The purpose of this section is to use recent information on background levels of trace metals in San Francisco Bay sediments to assess the suitability of a set of nationally-derived sediment guidelines for this region. We also used these background levels to determine the

¹ See Section 6 for a description of ERL and ERM values.

degree to which measured concentrations of trace metals from our data set represent contamination as opposed to concentrations that might be naturally occurring.

Approach

There have been few geological studies which have focused specifically on San Francisco Bay. Consequently, trace metal concentrations typical of shale sediment are often used as an approximation of expected background levels (Krauskopf, 1967). However, the U.S. Geological Survey (USGS) has recently completed a survey evaluating vertical and horizontal profiles of trace metals in sediment cores taken in central and north Bay locations (Hornberger, *et al.*, in review). This data constitutes the first study specific to San Francisco Bay which attempts to distinguish enriched from background concentrations

We used this USGS data as an approximation of background in order to assess the applicability of the newly revised ERL and ERM values for trace metals to San Francisco Bay sediments (Long, *et al.*, 1994). ERL and ERM values were singled-out for comparison, because they are the lowest of all the available national guidelines and, for several metals, exceedences of these guidelines were often not associated with toxicity in our San Francisco data set (discussed in Section 7). We also compared the USGS background data from the Richardson Bay core² to the metals concentrations of the Harbor and SedQual 3 samples to assess the degree of their contamination.

The USGS study included coring at four locations in the central and north Bay with subsequent analysis of seven trace metals (chromium, copper, lead, mercury, nickel, silver, and zinc). Analyses were conducted on sediment samples taken at ten-centimeter intervals for the full length of each core, and cores ranged from 1.5 to 2.5m in length. Trace metals concentrations reported in Table 7-1 represent averages from measurements made on replicate cores.

² The Richardson Bay core was used because it was the only core for which Pb 210 dating confirmed the age of the sediments.

Background levels of a given metal were defined in two ways. "True background" was determined using Pb-210 radiometric analysis which is used to date sediments that are less than 100 years old. However, Pb-210 analysis was only conducted on cores from the San Pablo Bay and Richardson Bay locations. Although the metals concentrations given for Richardson Bay have been confirmed as true background by Pb-210 dating, cores collected in the San Pablo Bay require additional radiometric analysis. Therefore, metals concentrations from this core may not represent true background. In Table 7-1, we have also classified metal concentrations as "approximate background" based on the methods described below.

A second estimation of background was made based on concentrations measured at depth. For all cores (with the exception of those from Richardson Bay), background is "qualitatively" defined as that concentration of a given metal which remains unchanged over 60-70 cm intervals in the deepest sections of the core. Nearly uniform concentrations of chromium and nickel were measured throughout all the cores. Consequently, there were no significant differences from which to qualitatively delineate background. For this reason, the range of chromium and nickel concentrations measured over the length of the entire core are reported under the "no background detected" heading in Table 7-1. One possible explanation for these uniform concentrations may be that naturally high background concentrations have overwhelmed any more recent signal which may be due to anthropogenic inputs.

We made additional verification of both radiometric "true" and "qualitative" background designations³ using recently published information from a sediment budget study for

³ It is important to note that the all of the radiometric and qualitative designations of background in Table 5-1 are not necessarily synonymous with pre-human or crustal concentrations. Rather, these designations give a good indication of pre-industrial concentrations. The Richardson Bay core, in which additional Carbon-14 dating suggests that sediment at 150 cm was deposited several hundred years ago, appears to give a complete record of pre- and post-anthropogenic activities. Trace metal profiles in this core, however, show two points where identifiable changes in enrichment are observed. Slight enrichment occurs approximately 100 years ago and may be attributable to runoff from hydraulic mining in the Sierras. A second and much more significant amount of enrichment starts 50 years ago, coinciding with the intensification of industrialized activity in the vicinity of this estuary. In nearly all the cores, the onset of this second phase of enrichment is clearly detectable and indicative of the relatively higher contribution of industrialization to trace metal enrichment of sediments in San Francisco Bay (Hornberger, personal communication).

Table 7-1. Background concentrations for selected metals in San Francisco Bay sediments from Hornberger et al. (in review)

Location (Lat/Long)	Accretion/Erosion potential (c)	Core length (cm)	Dating Info	Analysis Method	Approx. Background (a)					No Background (b)	
					Ag (µg/g)	Cu (µg/g)	Pb (µg/g)	Zn (µg/g)	Hg (µg/g)	Cr (µg/g)	Ni (µg/g)
Grizzly Bay 38 06 07 122 01 49	Slight erosion	250	nm (d)	Near total	nm	41	15	100	nm	102-119	100
San Pablo Bay 38 02 04 122 19 22	3-6' moderate accretion	250	might not be true background @2.5m	Total	0.1	55	40	60	nm	150-170	100
				Near total	0.1	40	20	70	nm	100-120	100
Central Bay 122 21 40 37 51 23	Slight accretion/erosion	200	nm	Total	nm	25	25	70	nm	115-150	70-90
				Near total	nm	30	10	70	nm	90-120	70-90
Richardson Bay 37 51 38 122 28 21	Slight erosion	150	"true" background from Pb 210	Total	0.1	20	20	70	nm	110-130	70-80
				Near total	0.1	20	10	50	0.05	70-100	50-70
ERL (e)					1.0	34	46.7	150	0.15	81	20.9
ERM					3.7	270	218	410	0.71	370	51

(a) "Approx. Background" indicates background defined qualitatively according to vertical profile of contamination. See Text for further explanation.

(b) "No background" indicates that ranges were seen throughout core (i.e., uniform concentrations). See text for further explanation.

(c) Accretion/Erosion potential from U.S. Army Corps of Engineers (1992)

(d) not measured

(e) ERL and ERM are Effects Range-low and -median values from Long et al. (in press)

Cadmium was measured in the San Pablo Bay and Richardson Bay cores but was below detection limits at both sites.

Cd concentrations: <0.7 µg/g for total extraction; <0.3 µg/g for near total extraction.

All sediment samples sieved to < 64 microns.

All trace metals except silver measured by Inductively Coupled Argon Plasma Emission Spectroscopy (ICAPES).

Silver measured by Graphite Furnace Atomic Absorption Spectrophotometry (GFAAS).

San Francisco Bay (U.S. Army Corps of Engineers, 1992). Each of the four USGS coring locations were plotted on bathymetric difference plates that represent the accretion/erosion rates over the 35 year period from 1955-1990. As indicated in Table 7-1, nearly all of the coring locations were in areas with only slight (0-1 ft. over 35 years) accretion or erosion. Only the core from San Pablo Bay was located in an area of moderate (3-6 ft. over 35 years) accretion. Higher accretion rates at this site corroborate with preliminary radiometric dating results indicating that sediment in the deeper sections of this core are less than 100 years old

The sample preparation and analysis methods used the USGS study were comparable to those used in both the Harbor and SedQual 3 studies with the following qualifications:

- All USGS sediment samples were sieved and only the silt-clay (<64-micron) fraction was used for trace metal analysis. Consequently, the trace metal concentrations given in Table 7-1 may be 10-20 $\mu\text{g}/\text{kg}$ higher than if entire samples had been analyzed (Hornberger, personal communication).
- Cores from three of the four USGS locations were prepared for analysis using both a total hydrofluoric acid (HF) digestion and a near total (concentrated nitric by reflux) digestion. The fourth core, from Grizzly Bay, was analyzed using only a near total digestion. Results using both extraction methods are supplied in Table 7-1 to facilitate comparisons with Harbor samples (analyzed using HF digestion) and SedQual 3 samples (analyzed using near total digestion).

Results

A comparison of ERL and ERM values to the USGS background ranges demonstrates that ERLs for four of the seven trace metals are either nearly equal to or lower than background (Table 7-1). ERLs for copper and lead fell within the range of background values observed at the USGS coring locations. Chromium and nickel ERLs were lower than the background ranges for these metals, neither of which had a vertical concentration gradient in

the USGS cores. Moreover, the ERM value for nickel was lower than background. In contrast, ERLs for silver, zinc and mercury were substantially higher than background.

A comparison of bulk sediment concentrations of selected metals from the SedQual 3 and Harbor studies to the USGS background levels from the Richardson Bay coring site is presented in Table 7-2. Concentrations of trace metals in SedQual 3 samples were generally higher than background while those in the Harbor sediments were often lower than background. The number of samples from the SedQual 3 studies that had metal

Table 7-2. The percentage of sediment samples from San Francisco Bay studies with trace metal concentrations that were above the USGS background (a)

Trace Metal	SedQual 3 Studies		Harbor Studies	
	Background (b) ($\mu\text{g/g}$)	% Above Background	Background (c) ($\mu\text{g/g}$)	% Above Background
Silver	0.1	97	0.1	54
Copper	20	98	20	59
Lead	10	95	20	28
Zinc	50	99	70	47
Mercury	0.05	100	na	-
Nickel	50 - 70	86 - 69	70 - 80	44 - 35
Chromium	70 - 110	73 - 7	110 - 130	100 - 98

- (a) Background concentrations measured at the USGS Richardson Bay coring location (Hornberger *et al.*, in review).
- (b) Samples from the SedQual 3 studies were compared to background concentrations based on the near total extraction method.
- (c) Samples from the Harbor studies were compared to background concentrations based on the HF extraction method.
- na = not available

concentrations higher than background ranged from 69 to 99% of the total, depending on the metal. Chromium was the only metal that was often lower than background. In contrast, trace metal concentrations in excess of background were observed less frequently in the Harbor samples (28-59% depending on the metal). Once again, chromium was the exception but unlike the SedQual 3 samples, 98-100% of those from the Harbor studies had chromium concentrations which were higher than background.

Discussion

If sediment guidelines are to be a useful tool in the identification of anthropogenically contaminated sediments in San Francisco Bay, they must reflect the site-specific conditions of enrichment. Marginal exceedence of guidelines such as ERLs, which are equal to or nearly equal to background levels of several trace metals, would be expected to be a poor predictor of toxicity related to contamination (see discussion in Section 7). This exercise suggests that for guidelines to be locally robust, they must be developed using regional data sets and with a consideration of site-specific variables such as background trace metal chemical concentrations.

Comparison of USGS trace metal background values to the SedQual 3 and Harbor data sets is a very general means of answering the question, "What is the extent of trace metal contamination in San Francisco Bay?". Combining the chemistry data from the individual studies that comprise the SedQual 3 and Harbor data, however, substantially limits the chemical- and site-specificity of the conclusions that can be drawn from this analysis. Sources of samples in the SedQual 3 studies ranged from the highly contaminated Castro Cove to mid-Bay sites that are physically removed from direct influence of discharges. Harbor study samples were predominantly from frequently-dredged navigation channels, although a fraction were from berth areas adjacent to intensive industrial activity. Furthermore, the background metal concentrations represented by the USGS Richardson Bay core were generally among the lowest of all the USGS coring sites, making this a conservative point of comparison. Thus, our observation of generally higher trace metal

contamination in the SedQual 3 samples as compared to those from the Harbor studies should not be interpreted as indicating that harbors are generally less contaminated with trace metals than non-harbor areas. Rather, this result suggests that trace metal contamination is not limited to (nor guaranteed to be a problem in) harbors.

Conclusions

In this section, we have demonstrated that ERL values for several trace metals (chromium, nickel, copper, and lead) are equal to or lower than background concentrations of these elements in San Francisco Bay sediments. Based on this comparison, we conclude that these particular guideline values can not be used to accurately predict sediment toxicity that would be associated with contamination. These results highlight the importance of using available site-specific information on background chemistry in determining the regional/local applicability of nationally-derived sediment quality guidelines.

8. ALTERNATIVE APPROACHES FOR EVALUATING SEDIMENT TOXICITY AND CHEMISTRY DATA.

Introduction

The results discussed in Section 3 demonstrate that the toxicity of field-collected sediments is frequently not explained by elevated bulk sediment concentrations of individual chemicals and/or chemical groups. Rather than indicating that chemical contamination is not the cause of toxicity, these results confirm a general concern that bulk concentrations alone do not provide adequate information to predict the bioavailability of sediment contaminants. Evaluations of sediment data are further complicated by the fact that field sediments are usually comprised of mixtures of contaminants the interactions of which can potentiate, antagonize, or have no effect on the toxicity of other co-occurring chemicals. Typical concentration-response analyses that are limited to bulk concentrations of individual chemicals or chemical classes do not account for these interactions.

The toxicity of sediments is governed by the interactions of numerous factors including the physical and chemical properties of the sediment, the chemical properties of the contaminants, and the exposure conditions of the organism. In order to determine why samples with relatively high contamination were not toxic to test species or, conversely, why samples with relatively low contamination were toxic, we would need to know more about such factors (e.g., acid volatile sulfide (AVS), organic and inorganic ligands, dissolved organic carbon, oxidation state, and pH). We would also need to understand the mode of toxic action as well as the route(s) of exposure to the test organism in question.

There have been several methods proposed to determine the "bioavailable" fraction of trace metals in sediment samples. Evaluation of the ratio of AVS to simultaneously extracted metals (SEM) has proven to be a good estimator of the bioavailability of several divalent metals (cadmium, copper, lead, nickel, and zinc) in sediments (Casas & Crecelius, 1994; DiToro, *et al.*, 1992; DiToro, *et al.*, 1990). Measurements of SEM and AVS are only slowly

being incorporated into sediment testing programs in San Francisco Bay and were not available in either the Harbor or SedQual 3 studies.

Additionally, numerous models have been developed for predicting the toxicological effects of mixtures of organic compounds in sediments (Adams, *et al.*, 1992; Burton, 1992). Assumptions about critical pathways of exposure and the interactions among contaminants in mixtures differ between models, as do the data requirements for their use. Yet, they all estimate (or require data regarding) the bioavailable fraction of sediment contaminants. Those models that employ standard measurements (bulk sediment chemistry, acute toxicity, bioaccumulation) are finding greater application in ongoing efforts to identify dominant ecotoxicological factors in San Francisco Bay (e.g., Swartz, *et al.*, 1994).

Goals

In this section, we apply the following two models to the San Francisco Bay data set to examine their utility as alternative methods for predicting the toxicity associated with specific classes of non-ionic organic compounds that commonly occur in the Bay.

- The Summed Polynuclear Aromatic Hydrocarbon (Σ PAH) model was recently developed by Swartz and his co-workers to predict whether sediment PAH levels are expected to be acutely toxic to the amphipod *Rhepoxynius abronius*.
- Critical Body Residue Analysis is a method for determining the potential toxicity of chemical groups with similar modes of toxic action, based on bioaccumulation of these chemicals in tissue. CBR analysis relies on the premise that whole-body residues of certain chemical classes are a good first approximation of the amount of chemical present at the sites of toxic action.

Approach

The Summed PAH model

We used the Σ PAH model recently developed by Swartz *et al.* (in review) to evaluate the potential toxicity of PAH concentrations in sediment samples from four of the Harbor Studies (Oakland Harbor: Phase 3A, Phase 3B, Berth Areas and the Port of Richmond). The model results were then compared to the actual amphipod toxicity observed for these samples.

The Σ PAH model uses standard chemistry data to predict the amphipod toxicity of field-collected sediments contaminated by PAHs. The model combines Toxic Units (TU), Equilibrium Partitioning Theory, Quantitative Structure Activity Relationships (QSAR), and information from laboratory concentration-response experiments to estimate the influence of sediment PAHs on *R. abronius* mortality. The Σ PAH model is particularly useful, because the data requirements (dry weight concentrations of 13 PAHs from the EPA priority pollutant list and sediment organic carbon) are standard measurements in many sediment monitoring programs. Furthermore, *R. abronius* is one of the amphipod species most commonly used in West coast sediment toxicity testing programs. A brief description of the model and its application follow.

For a given sediment sample, pore water TU values were calculated for each of the 13 PAHs. A toxic unit (TU) is a dimensionless measure of the potential for a contaminant to cause adverse affects to biota; it is defined as the concentration of chemical in the water divided by the water-only LC50. Theoretically, 50% mortality will occur in a sample with a TU = 1. Applied in the context of sediment toxicity, a TU for a given chemical is defined as $[PW]/[PW\ LC50]$, where [PW] is the chemical concentration in pore water ($\mu\text{g/L}$) and [PW LC50] is the pore water LC50 for that chemical ($\mu\text{g/L}$). Pore water concentrations and LC50 values are estimated as follows:

- Pore water concentrations of each PAH compound were estimated from the OC-normalized sediment concentration and the sediment-water partitioning coefficient using the following equilibrium relationship:

$$K_{oc} = [\text{OC-PAH}] / [\text{PW PAH}]$$

where K_{oc} is the organic carbon normalized sediment-water partition coefficient, OC-PAH is the PAH concentration per kilogram sediment organic carbon ($\mu\text{g}/\text{kgOC}$), and PW PAH is the concentration of PAH in the sediment pore water ($\mu\text{g}/\text{L}$). The K_{oc} values used were either derived experimentally or computed using laboratory measurements of K_{ow} and a K_{oc}/K_{ow} regression model (U.S. Environmental Protection Agency, 1993).

- Water-only LC50 values (substituted here for pore water LC50s) used in the TU calculation were either derived experimentally or estimated using the QSAR method¹, because amphipod LC50 values for most of these PAHs are unavailable. LC50s were derived experimentally for three PAH compounds (flouranthene, acenaphthylene, phenanthrene) using two amphipod species and plotted relative to their octanol/water partition coefficients (K_{ow}) (DeWitt, *et al.*, 1992). Although based on a relatively small amount of data, the linear relationship between the LC50 and K_{ow} values was highly significant ($r^2 = -0.98$, $p < 0.05$) (Swartz, personal communication). The pore water LC50s for the remaining PAH compounds were then estimated from this toxicity QSAR based on their individual K_{ow} values. The LC50 values used in this model will presumably change as more data from PAH spiking experiments becomes available precluding the need for QSAR estimations.

¹ QSARs describe relationships between the physico-chemical properties of chemicals and the biological responses these chemicals elicit. Comprehensive discussions of QSARs and their application in aquatic toxicology are given by Hermens *et al.* (1985) and McCarty *et al.* (1985).

The TU's for each of the PAH compounds were then summed for each sample². The probability that the total PAH content of a given sample would be acutely toxic to the amphipod, *R. abronius*, was estimated by comparing the Σ TU values to an empirically-derived concentration-response model. This model is based on data from spiked sediment studies using flouranthene, acenaphthylene, phenanthrene and describes the relationship between the percentage of samples with amphipod mortality greater than 24% and the value for Σ TU.

Critical Body Residue Analysis

Arguably, the most direct indication of the bioavailability (and in some cases, toxicity) of certain classes of chemicals, is the degree to which they bioaccumulate in the tissues of sediment-dwelling organisms. McCarty and Mackay (1993) have discussed the utility of the Critical Body Residue (CBR) method for evaluating the ecological risk of various groups of chemicals.

Numerous studies have demonstrated that organic chemicals with similar modes of toxic action also have comparable tissue residue thresholds for acute toxic response in invertebrates (Donkin, *et al.*, 1989; Landrum, *et al.*, 1991) and fish (Mackay & Hughes, 1984) (McCarty (1993) provides a thorough review). Furthermore, these thresholds appear to be fairly consistent across a variety of organisms (McCarty, 1986; McCarty, *et al.*, 1992). In other words, interspecies differences in sensitivity to many chemicals appear to be governed more by toxicokinetics (how a chemical gets to the site of toxic action) than by toxicodynamics (the mode of toxic action once at the target site). Unfortunately, there have been relatively few studies which have either estimated or measured tissue residues associated

² Toxic unit summation is frequently performed when analyzing concentration-response relationships for field samples containing mixtures of chemicals. While the assumption of additive toxicological effects may not apply in every situation, as a first approximation it appears to be reasonably accurate for mixtures of chemical classes having both similar and dissimilar modes of toxic action (McCarty & Mackay, 1993). Thus, even if an individual chemical is present at a concentration which was lower than its threshold for a specific mode of toxicity, it can still contribute to the overall non-specific (narcotic) toxicity of the mixture.

with acute and chronic toxicity endpoints in fish and aquatic invertebrates. The data which are available apply primarily to freshwater organisms. McCarty *et al.* (1992) estimated a residue range from 200 to 800 $\mu\text{M}/\text{kg}$ as producing acute narcosis³ in early life stages of fathead minnows, while residues associated with chronic effects appear to be an order of magnitude lower (McCarty, 1986). Landrum *et al.* (1991) have reported a 26 day PAH LD50 at body residues of 6100 $\mu\text{M}/\text{kg}$ for the freshwater amphipod, *Diporeia* sp.

One important feature of the CBR method, is that it is possible to group chemicals according to a common mode of toxic action, thus considering the effects of a chemical class rather than those of the individual compounds. This is particularly useful for real-world situations, in which sediments often contain mixtures of contaminants. Within classes of similarly-acting compounds (such as neutral narcotics), individual chemicals appear to have additive effects, with each chemical contributing to the overall toxicity of the mixture (Hermens, *et al.*, 1985).

In the present study, body residues of the neutrally narcotic organic compounds cited in Table 8-1 were calculated using tissue chemistry data from 69 sediment samples from four of the Harbor studies (Oakland Harbor Phase 3A retest, Phase 3B, Berth areas and Richmond Harbor). Bioaccumulation and 28-day toxicity tests were performed using the bentnose clam (*Macoma nasuta*) and an infaunal polychaete (*Nephtys caecoides*). Tissue samples from individuals of both species surviving at the end of 28 days⁴ were analyzed for the same suite of organic compounds as was measured in the bulk sediment. For each sample, wet weight tissue concentrations of these neutral narcotic compounds were converted to molar concentrations and summed⁵. Those samples from each of the studies with the highest

³ Acute narcosis describes toxicity resulting from non-specific modes of action.

⁴ Differences in bioaccumulation between individuals of the same species do not appear to be significant as evidenced by the low inter-replicate variability in tissue residues. Therefore, the fact that only tissues of surviving organisms were used is not expected to underestimate the residues which might be associated with acute toxicity.

⁵ Tissue residues were not normalized to the lipid content of the organism, because this information was not available.

summed tissue residues were then evaluated with respect to the results of the 28-day toxicity testing using the same species.

Table 8-1. Neutral narcotic chemicals included in tissue residue calculation

HPAHs	LPAHs	PCBs
Fluoranthene	Naphthalene	Aroclor 1242
Pyrene	Acenaphthylene	Aroclor 1248
Benzo(a)anthracene	Acenaphthene	Aroclor 1254
Chrysene	Fluorene	Aroclor 1260
Benzo(b,k)fluoranthene	Phenanthrene	
Benzo(a)pyrene	Anthracene	
Indeno(1,2,3,-c,d)pyrene		
Dibenzo(a,h)anthracene		
Benzo(g,h,i)perylene		

Results

Summed PAH model

Summary data from the Σ PAH analysis for selected samples from each of the four studies are provided in Table 8-2. Summed TU values, % *R. abronius* mortality, and % organic carbon for the three stations with the highest Σ TU values from each study are found on the left side of the Table 8-2. The same information from the three stations with the highest amphipod mortality is on the right side of the table. With two exceptions, all the Harbor samples tested had Σ TU values that were less than 1, ranging from 0.002 to 0.67. The highest Σ TU values were observed in samples from the Alcatraz disposal mound (4.52) and Alcatraz environs (1.64) reference stations. However, *R. abronius* mortality in these samples was generally low, 17% and 10%, respectively.

Table 8-2. Toxic unit concentrations of PAHs in sediment, amphipod mortality, and the organic carbon content of selected samples from four San Francisco Bay studies

<u>Highest Summed TUs (a)</u>					<u>Highest Mortality (b)</u>			
Study	Station	Sum TUs (c)	% Mort. (d)	% TOC	Station	Sum TUs	% Mort.	% TOC
Oakland Harbor Phase 3B	Alcatraz Mound Ref.	0.58	1	0.05	OC-1 Outer Harbor	0.06	53	0.81
	Alcatraz Environs Ref.	0.35	2	0.19	I-C2 Middle Harbor	0.10	41	0.87
	I-C8 Outer Harbor	0.29	11	0.10	O-C2 Middle Harbor	0.02	33	0.75
Oakland Harbor Berth Areas	Alcatraz Mound Ref.	4.52	17	0.43	Comp IB-1 Berth 60	0.19	28	0.79
	Comp OB-1 Berth 25	0.38	12	0.05	Alcatraz Mound Ref.	4.52	17	0.43
	Comp IB-3 Berths 62&63	0.27	13	0.07	Bay Farm Ref.	0.07	15	0.72
Oakland Harbor Phase 3A	I-C10 Inner Harbor	0.35	9	0.04	I-C17 Inner Harbor	0.02	51	0.09
	I-T3 Todd Shipyard	0.35	13	0.05	I-C9 Inner harbor	0.01	50	0.69
	I-T5 Todd Shipyard	0.30	12	0.09	Bay Farm Ref.	0.04	49	0.60
Richmond Harbor	Alcatraz Environs Ref.	1.64	10	0.06	Comp III Turning Basin	0.02	42	0.68
	Alcatraz Mound Ref.	0.67	25	0.38	Off Shelf Ocean Ref.	0.01	34	1.44
	Comp II Inner Harbor chnl	0.13	20	0.73	Alcatraz Mound Ref.	0.67	25	0.38

(a) Listed are three samples from each study with the highest summed PAH TUs.

(b) Listed are three samples from each study with the highest amphipod mortality.

(c) Summed toxic units for: naphthalene, acenaphthylene, acenaphthene, fluorene, phenanthrene, anthracene, fluoranthene, pyrene, benzo(a,h)anthracene, chrysene, benzo(b)fluoranthene, benzo(k)fluoranthene, benzo(a)pyrene.

(d) Mortality endpoint using *Rhepoxynius abronius*.

Samples with summed TUs >1 are indicated by shading.

Critical Body Residue Analysis

Body residues of neutral narcotics in both test species were very low for all of the samples from the Harbor studies. Tissue and sediment chemistry information from three of the four studies⁶ is presented in Table 8-3 for those samples having the highest summed concentrations of neutral narcotics. The highest tissue concentration of summed neutral narcotics, 7.2 $\mu\text{M}/\text{kg}$ in *M. nasuta*, was observed in reference sample from the Alcatraz disposal mound (R-AC from the Oakland Harbor Phase 3A retest). The Alcatraz mound sample from the Richmond Harbor study was the only sample associated with the highest bioaccumulation in both test species (3.4 $\mu\text{M}/\text{kg}$ in *M. nasuta* and 7.0 $\mu\text{M}/\text{kg}$ in *N. caecoides*). Overall, *N. caecoides* and *M. nasuta* were comparable in terms of their propensity toward accumulation of these neutral narcotic compounds. There were no cases in which the samples with the highest sediment concentration of neutral narcotics (indicated by (c) in Table 8-3) corresponded to the sample(s) in which the highest tissue residues were observed.

Nearly all of the samples having the highest tissue or sediment concentrations of neutral narcotics were non-toxic in 28-day tests for both species (Table 8-3). Overall survival for *N. caecoides* ranged between 31 to 98% and that for *M. nasuta* ranged from 84 to 98%. Slightly reduced survival (77%) was associated with a tissue concentration of 1.6 $\mu\text{M}/\text{kg}$ in *N. caecoides*.

Discussion

Summed PAH model

The ΣTU values were frequently higher at the Alcatraz reference stations and the highest ΣTU values overall were observed in samples from Alcatraz mound ($\Sigma\text{TU} = 4.52$) and Alcatraz environs ($\Sigma\text{TU} = 1.64$). However, amphipod mortality actually observed in these two samples (17% at the mound and 10% at the environs) was much lower than would

⁶ Results from the fourth study (Oakland Harbor Phase 3B) are not included because the highest tissue concentrations of neutral narcotic chemicals were extremely low ($<0.5\mu\text{M}/\text{kg}$ wet weight) in comparison with the other studies.

Table 8-3. Sediment and wet-weight tissue concentrations of neutral narcotic chemicals at selected stations from three harbor studies

Study	Station	Sediment NN (a) ($\mu\text{g}/\text{kg}$ OC)	% TOC	% fines	Σ NN <i>N. caecoides</i> ($\mu\text{M}/\text{kg}$)(b)	% Surv. <i>N. caecoides</i> (28-day)	Σ NN <i>M. nasuta</i> ($\mu\text{M}/\text{kg}$)	% Surv. <i>M. nasuta</i> (28-day)
Oakland Harbor	Comp IB-1 (c)	468	0.79	64	1.63	77	1.17	98
Berths	Comp IB-2 (d)	605	0.18	18	0.70	90	1.19	92
	R-AC (e)	8565	0.43	34	0.71	91	0.98	94
Oakland Harbor	I-S1 (c)	558	1.16	96	5.50	82	1.05	97
Phase 3A	R-AC (d)	317	0.94	85	0.76	90	7.17	94
Retest	I-C12 (e)	1012	0.17	18	1.62	81	0.05	99
Richmond	R-AC (c/d)	1659	0.38	35	6.95	92	3.09	91
Harbor	R-AM (e)*	2380*	0.06	5	0.59	95	0.82	90

(a) Neutral Narcotic chemicals calculated as LPAHs + HPAHs + total PCBs.

(b) All tissue concentrations are expressed as $\mu\text{M}/\text{kg}$ wet-weight

(c) station with highest summed neutral narcotics in *N. caecoides* tissue.

(d) station with highest summed neutral narcotics in *M. nasuta* tissue.

(e) station with the highest concentrations of OC-normalized neutral narcotics.

* Low total organic carbon at this station precludes OC-normalization; Sediment concentration of neutral narcotics given as $\mu\text{g}/\text{kg}$ dry weight.

For each study, tissues with the highest summed concentrations of neutral narcotics are shaded for each species.

be predicted from the TU concentration-response model for ΣPAHs, indicating that elevated PAHs at these sites were not a likely cause of the marginal toxicity observed (Swartz, personal communication). All of the samples with relatively higher amphipod mortality (>30%) had very low ΣTU values for PAHs ranging from 0.01 to 0.1.

Notably, 43% of all the samples and 75% of those listed in Table 8-2 as having the highest ΣTU values had low organic carbon content (< 0.2%). The utility of using equilibrium partitioning theory to predict the partitioning of non-polar organics between pore water and sediment decreases substantially when the organic content of a sample falls under 0.3% (DeWitt, *et al.*, 1992; DiToro, *et al.*, 1991). Thus, the use of EqP theory with samples having low organic carbon content may have resulted in an overestimation of pore water PAH concentrations and, consequently, pore water TU values.

Critical Body Residue Analysis

Overall, bioaccumulation of neutral narcotics was very low for both test species⁷. The highest tissue concentration observed (7.2 μM/kg in *M. nasuta*) falls significantly short of both the estimated acute narcosis residue range of 200-800 μM/kg in fish (McCarty & Mackay, 1993) and 6100 μM/kg in freshwater crustaceans (Landrum, *et al.*, 1991). Even if bivalves and polychaetes are more sensitive to these compounds than fish and amphipods, we would not expect this difference to be much more than an order of magnitude. The absence of significant toxicity in these samples lends further support to the conclusion that neutral narcotics are not present in tissue concentrations that would be expected to cause toxicity.

⁷ Over a four week exposure period, no noticeable differences were observed between tissue residues of neutral narcotics in the polychaete (*N. caecoides*) and those in the bivalve (*M. nasuta*). Mixed function oxidases such as aryl hydrocarbon hydroxylases which are capable of transforming various hydrocarbons into water-soluble degradation products, have been found in several species of polychaete (Lee, *et al.*, 1977) but are rare in bivalves (Payne, 1977). Biotransformation and excretion of accumulated compounds due to the activity of inducible enzyme systems are considered important mechanisms influencing tissue residues over longer exposure periods.

Furthermore, our findings provide confirmation that bulk sediment concentrations of contaminants are not indicative of their bioavailability for these test species. Samples having the highest OC-normalized sediment concentrations of neutral narcotics never had the highest bioaccumulation of these compounds.

There have been numerous studies demonstrating that interstitial water concentrations of non-ionic hydrophobic chemicals (such as PAHs) are the primary determinant of acute toxicity to benthic organisms (Swartz, *et al.*, 1988; Swartz, *et al.*, 1989). However, the relative contribution of the sediment-sorbed versus soluble fraction of contaminants to tissue residues has not been well defined for deposit feeding organisms and appears to be strongly dependent on both the partitioning characteristics of the chemical and the exposure conditions of the organism (Knezovich, *et al.*, 1987). Determination of the significance of different exposure pathways has been complicated by variability in the lipophilicity between chemicals as well as in their elimination half-lives for different organisms ⁸.

On one hand, if exposure to interstitial concentrations of neutral narcotics was the primary source of bioaccumulation in these species, then we would not expect bulk sediment concentrations to be a good predictor of tissue residues. The relatively low levels of organic carbon (<1.2%) seen in most of these samples from San Francisco Bay suggest that interstitial water concentrations of neutral narcotics, especially of the less lipophilic LPAHs, could be relatively high and constitute a significant route of exposure to the test organisms. Unfortunately, measurements of organics in interstitial water were not available with which to test this hypothesis.

On the other hand, if neutral narcotic compounds associated with the sediment were a significant contribution to the bioaccumulation observed in San Francisco Bay sediments, differences between samples in their grain size distribution and organic carbon content may

⁸ In the interest of brevity, we have left out a complete discussion of all the factors influencing the bioaccumulation of sediment contaminants. A more complete review is provided in Rodgers *et al.* (1987), Adams (1987), and Knezovich *et al.* (1987).

have influenced the relative bioavailability of sediment-bound organic chemicals. For example, in the Oakland Harbor Phase 3A retest study, neutral narcotics in *M. nasuta* tissue were highest at the Alcatraz mound reference station while the highest OC-normalized sediment concentrations were at station I-C12 located in the Oakland Inner Harbor.

Although sediment concentrations of neutral narcotics measured at the Alcatraz station (317 $\mu\text{g}/\text{kgOC}$) were significantly lower than those measured at the inner harbor station (1012 $\mu\text{g}/\text{kgOC}$), the contaminants associated with the finer and organically richer sediments at Alcatraz might have been more bioavailable to deposit feeding organisms.

Some evidence does exist which indicates that ingestion of sediment-bound neutral narcotics can be a significant exposure pathway for infaunal bivalves and polychaetes. Results from several studies have demonstrated that accumulation of PCBs from ingested sediment is at least as important as uptake from interstitial water for infaunal polychaetes (Fowler, *et al.*, 1978; Rubinstein, *et al.*, 1983). Studies evaluating the bioaccumulation of PAHs, on the other hand, have been less conclusive about the relative influence of the sediment-sorbed and water fractions on bioaccumulation. Roesijadi *et al.* (1978) conducted short- and long-term exposures of deposit feeding clams (*Macoma inquinata*) to sediments with radio-labeled PAHs. They found that accumulation of sediment-bound contaminants was lower than that from the water, thereby inferring that sediment interstitial water was the primary source of hydrocarbons accumulated in these organisms. However, the fact that both Roesijadi *et al.* (1978) and Augenfeld *et al.* (1982) observed higher hydrocarbon accumulation in tissue of the deposit feeders *M. inquinata* and the polychaete *Abarenicola pacifica* as compared to invertebrate suspension feeders is an indication that ingestion of sediment-bound contaminants can be an important additional route of exposure for deposit-feeding organisms.

Conclusions

Overall, there was agreement between the predictions of the ΣPAH model and those using the CBR method regarding the influence of non-polar organic compounds on sediment toxicity in this San Francisco Bay data set. Based on the results given by the ΣPAH model

as well as a consideration of the amphipod toxicity data, it appears that PAH contamination alone is unlikely to be an important determinant of the toxicity in the samples from these Harbor studies. Some other contaminant(s) or stress appear to be responsible for the toxicity observed. Tissue concentrations of neutral narcotic chemicals (which include PAHs and PCBs) were very low for all the sediment samples evaluated and were rarely associated with acute toxic responses. Moreover, the highest tissue concentrations observed were more than an order of magnitude lower than those reported in the literature to be associated with acute toxicity. We conclude that the neutral narcotic organic compounds are unlikely to be a source of acute toxicity in sediments from these Harbor studies.

The results of CBR analysis confirm that the bioavailability information provided by bioaccumulation testing cannot be derived indirectly from measurements of bulk sediment chemistry. Although higher tissue concentrations of neutral narcotic chemicals were generally observed from samples having higher sediment concentrations, the highest tissue concentrations in both the bivalve and polychaete test species never occurred in the samples with the highest sediment concentrations.

CBR analysis and the Σ PAH model appear to be useful (and underutilized) alternative approaches for determining potential causes of acute toxicity in San Francisco Bay sediments. Our results, however, can only be seen as preliminary due to the small size of the data set tested. A more thorough test of the utility of both methods for San Francisco Bay sediments would require a larger data base in which a greater range of sediment/tissue concentrations and significant toxicity was observed.

9. LITERATURE CITED

Adams, W.J. 1987. Bioavailability of neutral lipophilic organic chemicals contained on sediments: a review. In Dickson, K.L., A.W. Maki, and W.A. Brungs (Eds.), Fate and Effects of Sediment-Bound Chemicals in Aquatic Systems. New York: Pergamon Press.

Adams, W.J., R.A. Kimerle, and J.W. Barnett, Jr. 1992. Sediment Quality and Aquatic Life Assessment. *Environ. Sci. Technol.*, 10:1864-1875.

Anderson, S.L., J.E. Hose, and K.P. Knezovich. 1994. Genotoxic and developmental effects in sea urchins are sensitive indicators of effects of genotoxic chemicals. *Environ. Toxicol. Chem.*, 7:1033-1041.

Anderson, S., J. Knezovich, J. Jelinski, and D. Steichen. In preparation. Development of site-specific sediment quality objectives for San Francisco Bay.

Augenfeld, J.M., J.W. Anderson, R.G. Riley, and B.L. Thomas. 1982. The fate of polyaromatic hydrocarbons in an intertidal sediment exposure system: bioavailability to *Macoma inquinata* (Mollusca: Pelecypoda) and *Abarenicola pacifica* (Annelida: Polychaeta). *Mar. Environ. Res.*, 7:31-50.

Barrick, R., S. Becker, L. Brown, H. Beller, and R. Pastorok. 1988. Sediment Quality Values Refinement: 1988 Update and Evaluation of Puget Sound AET. Volume 1. Prepared for Tetra Tech Inc. and U.S. Environmental Protection Agency Region 10, Office of Puget Sound by PTI Environmental Services, Bellevue, WA.

Becker, D.S., R.C. Barrick, and L.B. Read. 1990. Evaluation of the AET approach for assessing contamination of marine sediments in California, Report No. 90-3 WQ. Prepared for the State Water Resources Control Board by PTI Environmental Services, Sacramento, California.

Berry, W. June 14, 1994. Personal communication. Senior Biologist, Science Applications International Corp. Narragansett, RI.

BPTC. 1993. Status of the Bay Protection and Toxic Cleanup Program: staff report. Prepared by the Bay Protection and Toxic Cleanup Program for the State Water Resources Control Board and the Regional Water Quality Control Boards, Sacramento, CA.

Burton, G.A., Jr. (Ed.). 1992. Sediment Toxicity Assessment. Chelsea, MI: Lewis Publishers.

Casas, A.M. and E.A. Crecelius. 1994. Relationship between acid volatile sulfide and the toxicity of zinc, lead, and copper in marine sediments. *Environ. Toxicol. Chem.*, 3:529-536.

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. *Marine Ecology Progress Series* 37:75-96.

Chapman, P.M., E.R. Long, R.C. Swartz, T.H. DeWitt, and R. Pastorok. 1991. Sediment toxicity tests, sediment chemistry and benthic ecology do provide new insights into the significance and management of contaminated sediments - a reply to Robert Spies. *Environ. Toxicol. Chem.*, 10:1-4.

Daniel, W.D. 1990. Applied Parametric Statistics (Second Edition). Boston: PWS-Kent Publishing Company.

Davies, T.T., D.G. Davis, and J.P. Elmore. December 21, 1993. EPA Memorandum: Technical Panel Recommendations Concerning Use of Acute Amphipod Tests in Evaluation of Dredged Material.

Davis, J., A. Gunther, B. Richardson, J. O'Conner, R. Spies, E. Wyatt, E. Larson, and E. Meiorin. 1990. San Francisco Estuary Project: Status and trends report on pollutants in the San Francisco estuary. Prepared by the San Francisco Bay-Delta Aquatic Habitat Institute, Richmond, CA.

Davis, J., A. Gunther, B. Richardson, J. O'Conner, R. Spies, E. Wyatt, E. Larson, and E. Meiorin. 1991. San Francisco Estuary Project: Status and trends report on pollutants in the San Francisco estuary. Prepared by the San Francisco Bay-Delta Aquatic Habitat Institute, Richmond, CA.

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.

DeWitt, T.H., R.J. Ozretich, R.C. Swartz, J.O. Lamberson, D.W. Schultis, G.R. Ditsworth, J.K.P. Jones, L. Hoselton, and L.M. Smith. 1992. The influence of organic matter quality on the toxicity and partitioning of sediment-associated fluoranthene. *Environ. Toxicol. Chem.*, 11:197-208.

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

Dillon, T.M., D.W. Moore, and A.B. Gibson. 1993. Development of a chronic sublethal bioassay for evaluating contaminated sediment with the marine polychaete worm *Nereis (Neanthes) arenaceodentata*. *Environ. Toxicol. Chem.*, 12:589-605.

DiToro, D.M., J.D. Mahony, D.J. Hansen, K.J. Scott, A.R. Carlson, and G.T. Ankley. 1992. Acid volatile sulfide predicts the acute toxicity of cadmium and nickel in sediments. *Environ. Sci. Technol.*, 1:96-101.

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

DiToro, D., C. Zarba, D. Hansen, W. Berry, R. Swartz, C. Cowan, S. Pavlou, H. Allen, N. Thomas, and P. Paquin. 1991. Technical basis for establishing sediment quality criteria for nonionic organic chemicals using equilibrium partitioning. *Environ. Toxicol. Chem.*, 10:1541-1583.

Donkin, P., J. Widdows, S. Evans, C. Worrall, and M. Carr. 1989. Quantitative structure-activity relationships for the effect of hydrophobic organic chemicals on rate of feeding by mussels (*Mytilus edulis*). *Aquat. Toxicol.*, 14:277-294.

Fowler, S., G. Polikarpov, D. Elder, P. Parsi, and J. Villeneuve. 1978. Polychlorinated biphenyls: accumulation from contaminated sediment and water by the polychaete *Nereis diversicolor*. *Mar. Biol.*, 48:303-309.

Ginn, T. March 13, 1993. Meeting Comments for February 19, 1993 Technical Review Panel for the Long-Term Management Strategy, PTI Project C421-01-01.

Ginn, T. and R. Pastorok. 1992. Assessment and management of contaminated sediments in Puget Sound. In G.A. Burton Jr. (Ed.), Sediment Toxicity Assessment. Chelsea, MI: Lewis Publishers, Inc. p. 391-397.

Hermens, J., H. Könemann, P. Leeuwangh, and A. Musch. 1985. Quantitative structure-activity relationships in aquatic toxicity studies of chemicals and complex mixtures of chemicals. *Environ. Toxicol. Chem.*, 4:273-279.

Hornberger, M. February 29, 1994. Personal communication. Physical Scientist, U.S. Geological Survey, Menlo Park, CA.

Hornberger, M.I., S.N. Luoma, A. van Geen, C. Fuller, and R. Anima. In review. Historical trends in metal contamination in cores collected in San Francisco Bay, CA.

Kemp, P.F. and R.C. Swartz. 1988. Acute toxicity of interstitial and particle-bound cadmium to a marine infaunal amphipod. *Mar. Environ. Res.*, 26:135-153.

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

Kohn, N.P., H.L. Mayhew, J.A. Ward, N.A. Cadoret, J.Q. Word. 1992a. Ecological evaluation of proposed dredged material from Oakland Harbor berthing areas. Prepared for the Port of Oakland under contract number 18157 by Battelle/Pacific Northwest Laboratories, Sequim, WA. Vols. I and II.

Kohn, N.P., J.A. Ward, H.L. Mayhew, J.Q. Word, E.S. Barrows, S.M. Barrows, S.M. Goodwin, L.F. Lefkovitz. 1992b. Ecological evaluation of proposed discharge of dredged material from Oakland Harbor into ocean waters (Phase IIIB of -42-foot project). Prepared for the U.S. Army Corps of Engineers under contract number DE-AC06-76RLO 1830 by Battelle/Pacific Northwest Laboratories, Sequim, WA. Vols. I and II.

Kohn, N.P., J.Q. Word, D.K. Niyogi, L.T. Ross, T. Dillon, and D.W. Moore. 1994. Acute toxicity of ammonia to four species of marine amphipod. *Mar. Environ. Res.*, 38:1-15.

Krauskopf, K. 1967. Introduction to Geochemistry. New York: McGraw-Hill.

Lamberson, J.O., T.H. DeWitt, and R.C. Swartz. 1992. Assessment of sediment toxicity to marine benthos. In G.A. Burton Jr. (Ed.), Sediment Toxicity Assessment. Chelsea, MI: Lewis Publishers.

Landrum, P., B. Eadie, and W. Faust. 1991. Toxicokinetics and toxicity of a mixture of sediment-associated polycyclic aromatic hydrocarbons to the amphipod *Diporeia* sp. *Environ. Toxicol. Chem.*, 10:35-46.

Lee, R., E. Furlong, and S. Singer. 1977. Metabolism of hydrocarbons in marine invertebrates: aryl hydrocarbon hydroxylase from the tissues of the blue crab, *Callinectes sapidus* and the polychaete worm, *Nereis* sp. In C. Giam (Ed.), Pollutant Effects on Marine Organisms. Lexington, MA.: Lexington Books.

Long, E.R. and M.F. Buchman. 1989. An evaluation of candidate measures of biological effects for the National Status and Trends Program. NOAA Tech. Memo NOS OMA 45. Prepared for the National Oceanic and Atmospheric Administration. Seattle, Washington.

Long, E., D. MacDonald, M. Matta, K. VanNess, M. Buchman, and H. Harris. 1988. Status and trends in concentrations of contaminants and measures of biological stress in San Francisco Bay. NOAA Tech Memo NOS OMA 41. Prepared for National Oceanic and Atmospheric Administration. Seattle, WA.

Long, E., D. MacDonald, S. Smith, and F. Calder. In press. Incidence of adverse biological effects within ranges of chemical concentrations in marine and estuarine sediments. *Environ. Management*.

Long, E.R. and R. Markel. 1992. An evaluation of the extent and magnitude of biological effects associated with chemical contaminants in San Francisco Bay, California NOAA Tech Memo NOS ORCA 64. Prepared for the National Oceanic and Atmospheric Administration, Seattle, Washington.

Long, E.R. and L.G. Morgan. 1990. The potential for biological effects of sediment-sorbed contaminants tested in the National Status and Trends Program, NOAA Tech Memo NOS OMA 52. Prepared for the National Oceanic and Atmospheric Administration, Seattle, Washington.

Luoma, S.N. and J.L. Carter. 1993. Editorial: Understanding the toxicity of contaminants in sediments: beyond the bioassay-based paradigm. *Environ. Toxicol. Chem.*, 12:793-796.

MacDonald, D. 1992. Development of an approach to the assessment of sediment quality in Florida coastal waters. Prepared for Florida Department of Environmental Regulation, Ladysmith, B.C.

Mackay, D. and A. Hughes. 1984. Three-parameter equation describing the uptake of organic compounds by fish. *Environ. Sci. Technol.*, 18:439-444.

McCarty, L. 1986. The relationship between aquatic toxicity QSARS and bioconcentration for some organic chemicals. *Environ. Toxicol. Chem.*, 5:1071-1080.

McCarty, L.S., P.V. Hodson, G.R. Craig, and K.L.E. Kaiser. 1985. The use of quantitative structure-activity relationships to predict the acute and chronic toxicity of organic chemicals to fish. *Environ. Toxicol. Chem.*, 4:595-606.

McCarty, L.S. and D. Mackay. 1993. Enhancing ecotoxicological modeling and assessment. *Environ. Sci. Technol.*, 9:1719-1728.

McCarty, L.S., D. Mackay, A.D. Smith, G.W. Ozburn, and D.G. Dixon. 1992. Residue-based interpretation of toxicity and bioconcentration QSARs from aquatic bioassays: neutral narcotic organics. *Environ. Toxicol. Chem.*, 11:917-930.

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

Oakden, J.M. 1984. Feeding and substrate preference in five species of phoxocephalid amphipods from Central California. *J. Crust. Biol.*, 2:233-247.

Payne, J. 1977. Mixed function oxidase in marine organisms in relation to petroleum hydrocarbon metabolism and detection. *Mar. Pollut. Bull.*, 8:112-116.

Pinza, M.R., J.A. Ward, H.L. Mayhew, J.Q. Word, D.K. Niyogi, N.P. Kohn. 1992. Ecological evaluation of proposed dredged material from Richmond Harbor. Prepared for the U.S. Army Corps of Engineers under contract number DE-AC06-76RLO 1830 by Battelle/Pacific Northwest Laboratories, Sequim, WA. Vols. I and II.

Risebrough, B. October 13, 1993. Letter to Thomas Wakeman, US Army Corps of Engineers.

Risebrough, R.W. 1994. Contaminants in San Francisco Bay sediments: relationships with toxicity studies undertaken by the National Oceanic and Atmospheric Administration, 1990. 205j project, Prepared for the San Francisco Bay Regional Water Quality Control Board, The State Water Resources Control Board, and the U.S. Environmental Protection Agency.

Roesijadi, G., J.W. Anderson, and J.W. Blaylock. 1978. Uptake of hydrocarbons from marine sediments contaminated with Prudhoe Bay crude oil: influence of feeding type of test species and availability of polycyclic aromatic hydrocarbons. *J. Fish. Res. Board. Can.*, 35:608-614.

Rogers, J., Jr., K.L. Dickson, F.Y. Saleh, and C.A. Staples. 1987. Bioavailability of sediment-bound chemicals to aquatic organisms: some theory, evidence and research needs. In K.L. Dickson, A.W. Maki, and W.A. Brungs (Eds.), Fate and Effects of Sediment-Bound Chemicals in Aquatic Systems. New York: Pergamon Press. p. 245-266.

Ross, B. July 14, 1994. Personal communication. Sediment and Dredging Management Team, U.S. Environmental Protection Agency Region 9, San Francisco, CA.

Rubinstein, N., E. Lores, and N. Gregory. 1983. Accumulation of PCB, mercury, and cadmium by *Nereis virens*, *Mercenaria mercenaria*, and *Palaemonetes pugio* from contaminated harbor sediments. *Aquat. Toxicol.*, 3:249-260.

Samoiloff, M.R., J. Bell, D.A. Birkholz, G.R.B. Webster, E.G. Arnott, R. Pulak, and A. Madrid. 1983. Combined bioassay-chemical fractionation scheme for the determination and ranking of toxic chemicals in sediments. *Environ. Sci. Technol.*, 6:329-334.

SFBRWQCB. 1994. The San Francisco Estuary Pilot Regional Monitoring Program: sediment studies final report. Submitted to the San Francisco Bay Regional Water Quality Control Board and the State Water Resources Control Board.

Spies, R.B. 1989. Editorial. Sediment bioassays, chemical contaminants and benthic ecology: new insights or just muddy water? *Mar. Environ. Res.*, 27:73-75.

Spies, R.B., D.W. Rice Jr., and J.W. Felton. 1988. The effects of organic contaminants on reproduction of starry flounder, *Platichthys stellatus* (Pallas), in San Francisco Bay. I. Hepatic contamination and mixed-function oxidase (MFO) activity during the reproductive season. *Mar. Biol.*, 98:191-200.

Swartz, R. March 15, 1994. Personal communication. Marine Science Center, U.S. Environmental Protection Agency. Newport, Oregon.

Swartz, R.C., F.A. Cole, J.O. Lamberson, S.P. Ferraro, D.W. Schults, W.A. DeBen, H. Lee II., and R.J. Ozretich. 1994. Sediment toxicity, contamination and amphipod abundance at a DDT- and dieldrin-contaminated site in San Francisco Bay. *Environ. Toxicol. Chem.*, 6:949-962.

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

Swartz, R.C., P.F. Kemp, D.W. Schults, G.R. Ditsworth, and R.J. Ozretich. 1989. Acute toxicity of sediment from Eagle Harbor, Washington, to the infaunal amphipod *Rhepoxynius abronius*. *Environ. Toxicol. Chem.*, 8:215-22.

Swartz, R.C., D.W. Schults, R.J. Ozretich, J.O. Lamberson, F.A. Cole, S.P. Ferraro, T.H. DeWitt, and M.S. Redmond. In review. Summed PAH: A model to predict the toxicity of field-collected marine sediment contaminated by polynuclear aromatic hydrocarbons.

Taberski, K. September 5, 1994. Personal communication. Regional Monitoring Program, San Francisco Bay Regional Water Board. Oakland, California.

US ACOE. 1992. Sediment budget study for San Francisco Bay. Prepared by Ogden Beeman & Associates and Ray B. Krone & Associates for the U.S. Army Corps of Engineers under contract number DACW07-89-D-0029. San Francisco, CA.

US ACOE and US EPA. February 1, 1993. Public Notice 93-2 Testing Guidelines for Dredged Material Disposal at San Francisco Bay Sites. San Francisco, CA.

US EPA. 1993. Technical basis for deriving sediment quality criteria for nonionic organic contaminants for the protection of benthic organisms by using equilibrium partitioning. EPA-822-R-93-011. Prepared by Office of Water, Washington, DC.

Vonburg, R. and D. Liu. 1993. Chromium and hexavalent chromium. *J. Appl. Toxicol.*, 3:225-230.

Ward, J.A., J.Q. Word, M.R. Pinza, H.L. Mayhew, E.S. Barrows, and L.F. Lefkovitz. 1992. Ecological evaluation of proposed discharge of dredged material from Oakland Harbor into ocean waters (Phase IIIA of -42-foot project), Prepared for the U.S. Army Corps of Engineers under contract number DE-AC06-76RLO 1830 by Battelle/Pacific Northwest Laboratories, Sequim, WA. Vols. I and II.

Wolfenden, J. and M. Carlin. 1992. Draft sediment screening criteria and testing requirements for wetland creation and upland beneficial reuse. Prepared for the California Environmental Protection Agency and San Francisco Regional Water Quality Control Board, Oakland, CA.

Word, J. May 18, 1994. Personal communication. Biological Oceanography Group, Battelle/Pacific Northwest Laboratories, Sequim, Washington.

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
TECHNICAL AND ELECTRONIC
INFORMATION DEPARTMENT
BERKELEY, CALIFORNIA 94720