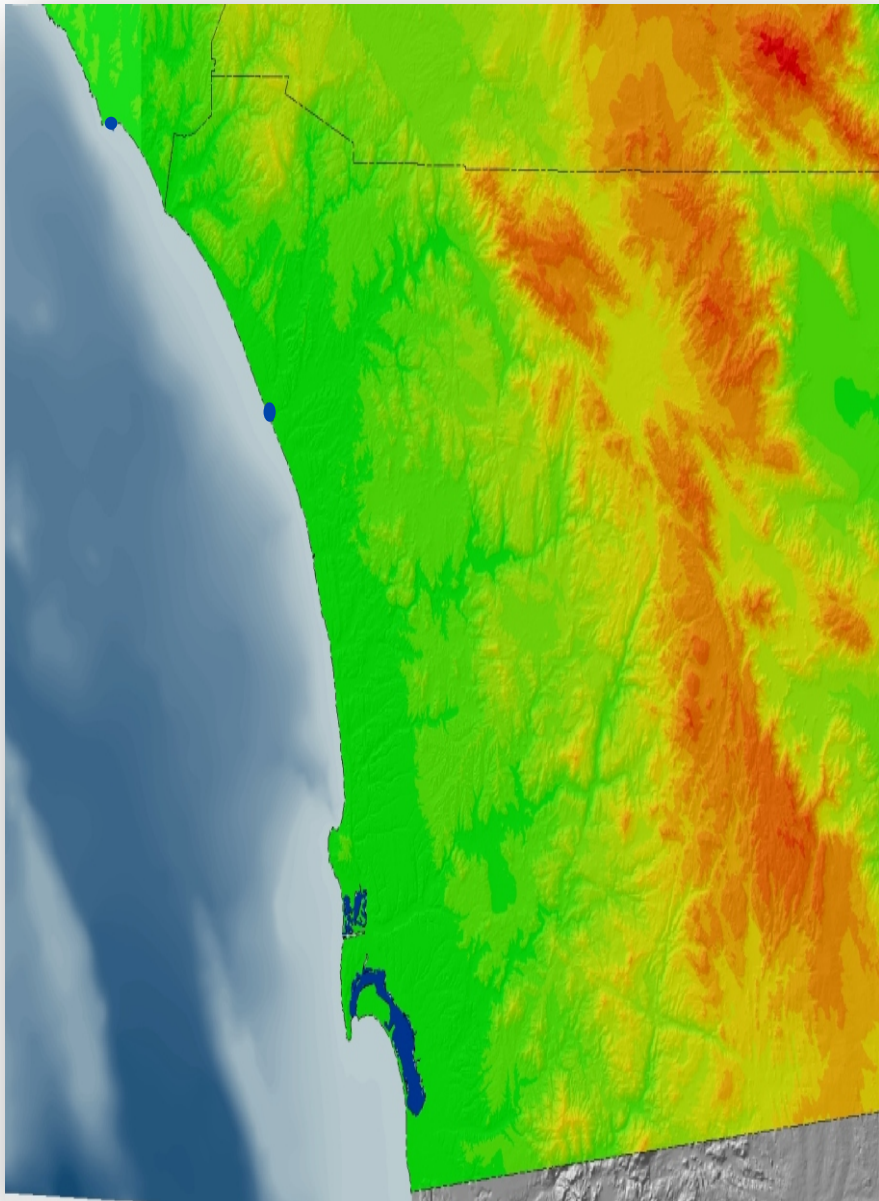


Regional Harbor Monitoring Program Pilot Project 2005-08 Summary Final Report



*Prepared for:
Port of San Diego
City of San Diego
City of Oceanside
County of Orange*

Prepared by:
WESTON
SOLUTIONS

**Regional Harbor Monitoring Program
Pilot Project
2005-08 Summary

Final Report**

Prepared for:

**Port of San Diego
3165 Pacific Highway
San Diego, CA 92101
and
City of San Diego
City of Oceanside
County of Orange**

Prepared By:

**Weston Solutions, Inc.
2433 Impala Drive
Carlsbad, California 92010**

May 2008

TABLE OF CONTENTS

EXECUTIVE SUMMARY	v
1.0 INTRODUCTION	1
2.0 METHODS	4
2.1 Field Sampling	4
2.1.1 Station Selection	4
2.1.2 Water Quality Sampling	11
2.1.3 Sediment Sampling	11
2.2 Laboratory Analysis	12
2.2.1 Chemistry	12
2.2.2 Toxicity	15
2.2.3 Infauna	16
2.2.4 Microbiology	17
2.2.5 Profile Data Processing	17
2.3 Data Analysis	17
2.3.1 Water Column Chemistry	20
2.3.2 Sediment Chemistry	21
2.3.3 Sediment Toxicity	22
2.3.4 Benthic Infauna	23
2.3.5 Statistical Analyses	24
3.0 RESULTS	26
3.1 Water Quality	26
3.1.1 Chemistry	26
3.1.2 Bacteria	33
3.1.3 Water Column Measurements	34
3.2 Sediment Analysis	35
3.2.1 Chemistry	35
3.2.2 Toxicity	42
3.2.3 Benthic Infauna	45
3.2.4 Grain Size and Total Organic Carbon	49
3.3 Analysis of Study Design	50
4.0 DISCUSSION	54
4.1 Assessment of Harbor Conditions	55
4.1.1 Surface Waters	55
4.1.2 Sediments	56
4.2 Comparison of Pilot Project Results to Historical Conditions	57
4.3 Spatial Distribution of Indicators	58
4.3.1 Comparisons between Strata	58
4.3.2 Comparisons among Harbors	58
4.4 Assessment of Trends	59
4.5 Recommended Modifications to the RHMP Study Design	60
4.5.1 Sample Size	60
4.5.2 Bight 2003 Data	61

4.5.3	Sediment Quality Objects	61
4.5.4	Benthic Community Assessment	62
4.6	Future Special Studies.....	64
4.6.1	Spatial Extent of Marina Copper Contamination.....	65
4.6.2	Sediment Copper Flux	65
4.6.3	Bioaccumulation	65
4.6.4	Sediment Toxicity Identification Examination.....	66
5.0	REFERENCES	67

APPENDICES

- A – Locations of Sample Stations
- B – Supplemental Results Tables and Figures
- C – 2007 Water and Sediment Metals and PAHs
- D – 2007 Water Column Physical Indicators
- E – 2007 Toxicology Summary Results
- F – 2007 Benthic Infauna Species List and Abundances

LIST OF TABLES

Table 2-1. RHMP constituents monitored and corresponding analytical methods	13
Table 2-2. Studies used to establish reference ambient values	18
Table 2-3. Summary of threshold values and preset targets	19
Table 2-4. Characterization and BRI ranges for response levels of benthic community conditions.....	23
Table 3-1. Percentage of stations below CTR thresholds for marina and freshwater- influenced surface water metals.....	26
Table 3-2. Percentages of stations at enterococcus concentrations of <10, 10, and 20 MPN/100mL within marina and freshwater-influenced strata	33
Table 3-3. Percentages of stations below ER-L thresholds for marina and freshwater- influenced sediment indicators	35
Table 3-4. Percentage of stations with toxicity below the threshold value	43
Table 3-5. Percentage of stations with benthic infaunal community measures better than reference ambient values.....	45
Table 3-6. Percentages of marina and freshwater-influenced strata classified by the BRI from reference to response level 4	47
Table 3-7. Comparison of sample variability in primary and secondary indicators.....	51
Table 3-8. Observed powers and estimated sample sizes for binomial tests of proportions based on Pilot Project results	52
Table 4-1. Benthic index categorization values for southern California marine bays.....	63
Table 4-2. Reference ranges for IBI metrics in southern California marine bays	64

LIST OF FIGURES

Figure 2-1. Sampling stations in Dana Point Harbor	5
Figure 2-2. Sampling stations in Oceanside Harbor	6
Figure 2-3. Sampling stations in Mission Bay	7
Figure 2-4. Sampling stations in northern San Diego Bay	8
Figure 2-5. Sampling stations in central San Diego Bay	9
Figure 2-6. Sampling stations in southern San Diego Bay	10
Figure 2-7. Example of a distribution curve that can be used as a reference guide.....	20
Figure 2-8. Relationship of copper to iron.....	22
Figure 3-1. Cumulative distribution curves for surface water dissolved and total copper	27
Figure 3-2. Comparison of surface water copper concentrations among years and harbors	29
Figure 3-3. Comparisons of surface water nickel and zinc among years and harbors.....	31
Figure 3-4. Comparison of surface water total detectable PAHs among years and harbors	32
Figure 3-5. Comparisons of surface water dissolved and total organic carbon among years and harbors	33
Figure 3-6. Cumulative distribution curves for sediment mean ER-M quotients.....	36
Figure 3-7. Comparisons of sediment average mean ER-M quotients among years and harbors.....	37
Figure 3-8. Comparisons of sediment copper concentrations among years and harbors.....	39
Figure 3-9. Comparisons of sediment zinc concentrations among years and harbors.....	40
Figure 3-10. Comparison of sediment cadmium, chromium, lead, and nickel concentrations among years and harbors	41
Figure 3-11. Comparisons of sediment total detectable PAHs among years and harbors.....	42
Figure 3-12. Cumulative distribution curves for <i>E. estuarius</i> mortality.....	43
Figure 3-13. Comparisons of toxicity as measured by <i>E. estuarius</i> mortality among years and harbors.....	44
Figure 3-14. Relationship between <i>E. estuarius</i> mortality and percentage of clay in sediments.....	45
Figure 3-15. Cumulative distribution curves for the Benthic Response Index.....	46
Figure 3-16. Comparisons of average Benthic Response Index values among years and harbors.....	47
Figure 3-17. Comparisons of average benthic infaunal community measures among years and harbors.....	49
Figure 3-18. Comparisons of percent fine sediments and total organic carbon among years and harbors	50

EXECUTIVE SUMMARY

The Regional Harbor Monitoring Program (RHMP) was developed by the Port of San Diego, City of San Diego, City of Oceanside, and County of Orange to address questions regarding the general water quality and condition of aquatic life in the four harbors within Region 9 (San Diego) of the State Water Resources Control Board. The RHMP was developed to address the overall condition of the harbors through core monitoring and supplemental focused studies by answering the questions set forth by the San Diego Regional Water Quality Control Board (SDRWQCB) regarding the spatial distribution of pollutants and their impacts, the safety of the waters for human contact, the safety of fish for human consumption, the abilities of the waters and sediments to sustain healthy biota, and the long-term trends in harbor conditions. The core monitoring program assesses the conditions found in the harbors based on comparisons to historical reference values for the four harbors and comparisons of contaminant concentrations to known surface water and sediment thresholds using chemistry, bacterial, toxicology, and benthic infaunal community indicators.

Prior to the initiation of the RHMP in the summer of 2008, a Pilot Project was initiated to assess the effectiveness of the proposed study design, determine the level of sampling effort needed to increase the statistical power of the study, and refine the design as needed. Sampling for the Pilot Project was conducted from 2005 to 2007, once per year in the month of August. The Pilot Project is a scaled-down version of the RHMP that focuses on a limited number of indicator measurements sampled within two of the five identified strata. The strata sampled in the Pilot Project, marinas and freshwater influenced, were selected because the variability within these areas was anticipated to be greater than in the other three strata, and thus would provide a conservative estimate of the amount of sampling effort needed to detect statistically significant differences from historical conditions. Given that only two strata were surveyed, the results of the Pilot Project cannot be used to assess whether conditions have improved or deteriorated from historical conditions, since preset targets were determined using stations located within the shallow, deep, and port/industrial strata in addition to the marina and freshwater-influenced strata. Although the Pilot Project was specifically designed to assess the validity of the design, inferences can be made on whether conditions in the marina and freshwater-influenced strata are better or worse than the historical harbor-wide conditions.

Based on the results of the Pilot Project, the following statements can be made:

- Copper concentrations in marinas exceeded water quality objectives, while concentrations of other metals and polyaromatic hydrocarbons were below water quality objectives.
- All bacterial concentrations were well below AB 411 levels.
- Physical water column measures largely occurred at levels that were suitable to support healthy biota.
- Sediment concentrations of copper in marinas and zinc in marina and freshwater-influenced strata occurred at levels likely to cause adverse biological effects
- Polyaromatic hydrocarbons and all other sediment metals (except arsenic and mercury) primarily occurred at levels that were not likely to result in adverse biological effects.
- The majorities of the marina and freshwater-influenced strata contained sediments that were not toxic.

- Benthic infaunal communities in both strata occurred at intermediate levels of disturbance.
- In the marina stratum, the primary surface water, sediment, and benthic infaunal community indicators occurred at levels that were worse than harbor-wide historical conditions
- Toxicity levels in the marina sediments generally were better than harbor-wide historical conditions.
- In the freshwater-influenced stratum, primary surface water, sediment, and toxicity indicators were better than historical conditions, while only benthic infauna was worse.
- The marina stratum tended to have higher concentrations of surface water and sediment chemistry indicators than the freshwater-influenced stratum.
- Toxicity levels and benthic infaunal communities did not differ between the two strata.
- From 2005-2007, no negative short-term trends were evident for any indicator that would be indicative of a degrading condition.

The results of the Pilot Project validated the effectiveness of the RHMP study design in answering the SDRWQCB questions. The use of a stratified random design that was repeated among years allowed for the assessment of the spatial distributions of pollutants (i.e., differences between strata and among harbors), as well as changes in the levels of pollution through time (i.e., short-term trends from year to year). Additionally, the Pilot Project illustrated that the study design is appropriate for analysis of trends through comparisons of the percentages of stations below threshold values between present-day and historical conditions, as well as comparisons of changes in indicator values from year to year. Evidence for the effectiveness of the approaches is seen in the prevalence of statistically significant results in most cases where it was reasonable to assume that they would occur.

Since the initiation of the Pilot Project additional historical data for the harbors have been released, the Benthic Response Index (BRI) has been modified, and new sediment quality objectives (SQO) for bays and estuaries have been developed. Moreover, successful completion of the Pilot Project has provided valuable insights into the validity of the approach and how it can be enhanced. As a result of recent methodological innovations and analyses of the effectiveness of Pilot Project results, the following modifications to the RHMP study design are recommended:

- Increase the sample size in strata to 15.
- Integrate Bight 2003 data into historical distribution curves.
- Analyze sediment conditions with new SQOs.
- Revise benthic community assessment and BRI calculation.
- Include tributyltin as an analyte.

Based on Pilot Study findings, special studies are recommended to determine the spatial extent of copper pollution surrounding marinas, quantify the level of copper flux from marina sediments, assess the bioavailability of copper using bioaccumulation studies and the biotic ligand model, and identify the causes of high toxicity levels in areas where they occur through sediment toxicity identification evaluations.

1.0 INTRODUCTION

The Regional Harbor Monitoring Program (RHMP) was developed by the Port of San Diego, County of Orange, the City of San Diego, and the City of Oceanside in response to a July 24, 2003 request by the San Diego Regional Water Quality Control Board (SDRWQCB) under §13225 of the California Water Code. The RHMP is a comprehensive effort to survey the general water quality and condition of aquatic life in the harbors and to determine whether beneficial uses are being met for the following four local harbors: San Diego Bay, Mission Bay, Oceanside Harbor, and Dana Point Harbor. The program is composed of a core monitoring program supplemented by potential focused studies to answer specific questions. The core monitoring program was designed to address the following five major questions posed in the SDRWQCB's request:

1. What are the contributions and spatial distributions of inputs of pollutants to harbors in the San Diego Region and how do these inputs vary over time?
2. Are the waters in the harbors safe for body contact activities?
3. Are fish in harbors safe to eat?
4. Do the waters and sediments in the harbors sustain healthy biota?
5. What are the long-term trends in water quality for each harbor?

To answer these questions, the RHMP study design was created through a multistep iterative process that included extensive research of historical information for the four harbors, detailed mapping of the harbors into strata, identification of indicators to be monitored, establishment of reference ambient values (i.e., threshold levels) and preset targets, and development of statistical tests to evaluate findings in a scientifically rigorous manner that is complimentary to the larger Southern California Bight (Bight) regional monitoring program. This program utilized a weight-of-evidence approach to assess the condition of the harbors and compare findings to recent historical conditions to determine whether conditions are improving or deteriorating. Contaminants within surface waters and sediments, toxicity levels, and conditions of biological communities are quantified to determine the health and overall status of the harbors.

Understanding the spatial distribution of pollutants and their impacts (Question 1) requires that indicators be compared among different areas of the harbors (i.e., strata) as well as among the individual harbors. Partitioning the harbors into five strata, classified as freshwater influenced, marinas, port/industrial, deep, and shallow, is essential to better understanding the impacts of specific activities on surface waters, sediments, toxicity, and infaunal communities. The freshwater-influenced stratum includes areas that may be affected by urban runoff. The marina stratum includes areas in close proximity to docks and anchorages that may be impacted by boating and maintenance activities. The port/industrial stratum occurs exclusively within San Diego Bay and includes areas in close proximity to heavy industrial activities. The two remaining strata, shallow and deep, include areas that do not meet the other categories and are classified by depth, using a 12-foot depth cutoff (mean lower low water [MLLW]), as described in detail in Weston 2005a.

To understand whether the waters are safe for human contact (Question 2) and if the waters and sediments sustain healthy biota (Question 4), multiple indicators of harbor condition are

measured at stations, including water and sediment contaminants, bacterial levels, sediment toxicity, and benthic infaunal community condition. Observed indicator levels can then be compared with established thresholds for adverse effects, such as effects range low (ER-L) for sediments, California Toxics Rule (CTR) values for surface waters, AB411 for bacterial levels, and Benthic Response Index (BRI) characterizations for infaunal communities to establish whether conditions are likely to be adverse to humans and biota.

Assessing long-term trends (Question 5) requires an innovative approach whereby present-day conditions within harbors are compared to historical conditions. Historical conditions of the harbors were determined based on reviews of prior studies performed over a 10-year period from 1994 to 2004 within Dana Point Harbor, Oceanside Harbor, Mission Bay, and San Diego Bay. Using the historical dataset, preset targets were established as the percentages of historically-sampled stations at or below threshold levels. By comparing the observed percentage of stations sampled throughout the harbors during the RHMP to the preset targets, one can determine whether conditions in the harbors have improved or have declined. Additionally, proportions of stations below threshold levels within a given harbor or stratum also can be tracked through time as the program progresses to measure changes in the health of the harbors (i.e., trends).

Prior to the full implementation of the RHMP in the summer of 2008, a Pilot Project was initiated to assess the effectiveness of the proposed study design, refine the list of indicators to be monitored, improve threshold levels and preset targets, and determine the level of sampling effort needed to increase the statistical power of the study. Sampling for the Pilot Project was conducted from 2005 to 2007, once per year in the month of August. The Pilot Project is a scaled-down version of the RHMP that focuses on a limited number of indicator measurements sampled within two of the five identified strata. The strata sampled in the Pilot Project, marinas and freshwater influenced, were selected because the variability within these areas was anticipated to be greater than in the other three strata, and thus would provide a conservative estimate of the amount of sampling effort needed to detect statistically significant differences from historical conditions.

Given that only two strata were surveyed, the results of the Pilot Project cannot be used to assess whether conditions have improved or deteriorated from historical conditions, since preset targets were determined using stations located within the shallow, deep, and port/industrial strata in addition to the marina and freshwater-influenced strata. Although the Pilot Project was specifically designed to assess the validity of the design, inferences can be made on whether conditions in the marina and freshwater-influenced strata are better or worse than the historical harbor-wide conditions. Additionally, although the full RHMP will be able to assess long-term trends in individual harbor conditions (Question 5), the Pilot Project did not include enough stations in each harbor in each year to do so effectively.

This report presents the results of the three-year RHMP Pilot Project. Although the conclusions that can be drawn from this study must be limited to the two strata that were assessed, the Pilot Project was able to assess the effectiveness of the study design in answering the following questions:

1. Are conditions in the marina and freshwater-influenced strata better or worse than historical conditions?
2. What is the average measure of relevant indicators in each stratum and harbor?

3. Are conditions in the marina stratum different from those of the freshwater-influenced stratum?
4. Do indicators change significantly from year to year throughout the study, and are trends evident within strata?
5. How do conditions within the marina and freshwater-influenced strata differ among harbors?
6. Is the sampling effort sufficient to detect statistically significant differences and/or changes?

Lastly, the report discusses: how the results of the Pilot Project can be used to enhance the RHMP core study design, modifications to the study design and analyses based on new sediment quality objectives and special focused studies that may augment core monitoring efforts.

2.0 METHODS

2.1 Field Sampling

2.1.1 Station Selection

Station selection in Dana Point Harbor, Oceanside Harbor, Mission Bay, and San Diego Bay was based on a stratified random sampling design similar to ones used in southern California regional monitoring programs. Uniformly sized hexagons were overlaid on maps of each of the bays. The size of the hexagons was determined by the smallest freshwater-influenced area that could be safely sampled. Hexagons were set at 100 feet (ft) (~30 meters) per side with the nominal sampling station at the center of the hexagon. Ten stations were randomly selected in the marina and freshwater-influenced strata annually with the stipulation that at least one station was set in each harbor. In total, 30 stations were sampled per stratum over the entire study period. Because Oceanside Harbor does not contain the freshwater-influenced stratum, only the marina stratum was sampled, while all other harbors contained both strata. Sampling was conducted within a 30-meter (m) radius of the nominal station coordinates, and coordinates of the actual sampling locations were recorded.

The locations of the sampling stations in each of the harbors and the years in which they were sampled are shown in Figure 2-1 to Figure 2-6. Marina stations in Dana Point Harbor and Oceanside Harbor were located near boat slips. Marina stations in Mission Bay were located near Dana Landing, in Quivera Basin, and in Santa Barbara Cove. In San Diego Bay, marina stations were located in the Shelter Island Yacht Basin, America's Cup Harbor, Sunroad Resort Marina, and the Coronado Cays. Most marina stations in San Diego Bay were concentrated in the northern portion of the bay near Shelter Island Yacht Basin and in America's Cup Harbor.

In Dana Point Harbor, freshwater-influenced stations were located adjacent to storm drains. No freshwater-influenced areas were identified in Oceanside Harbor. Freshwater -influenced stations in Mission Bay were located near Rose Creek Inlet and Tecolote Creek. In San Diego Bay, they were located in South Bay near the power plant and at the mouths of Chollas Creek and Sweetwater River.

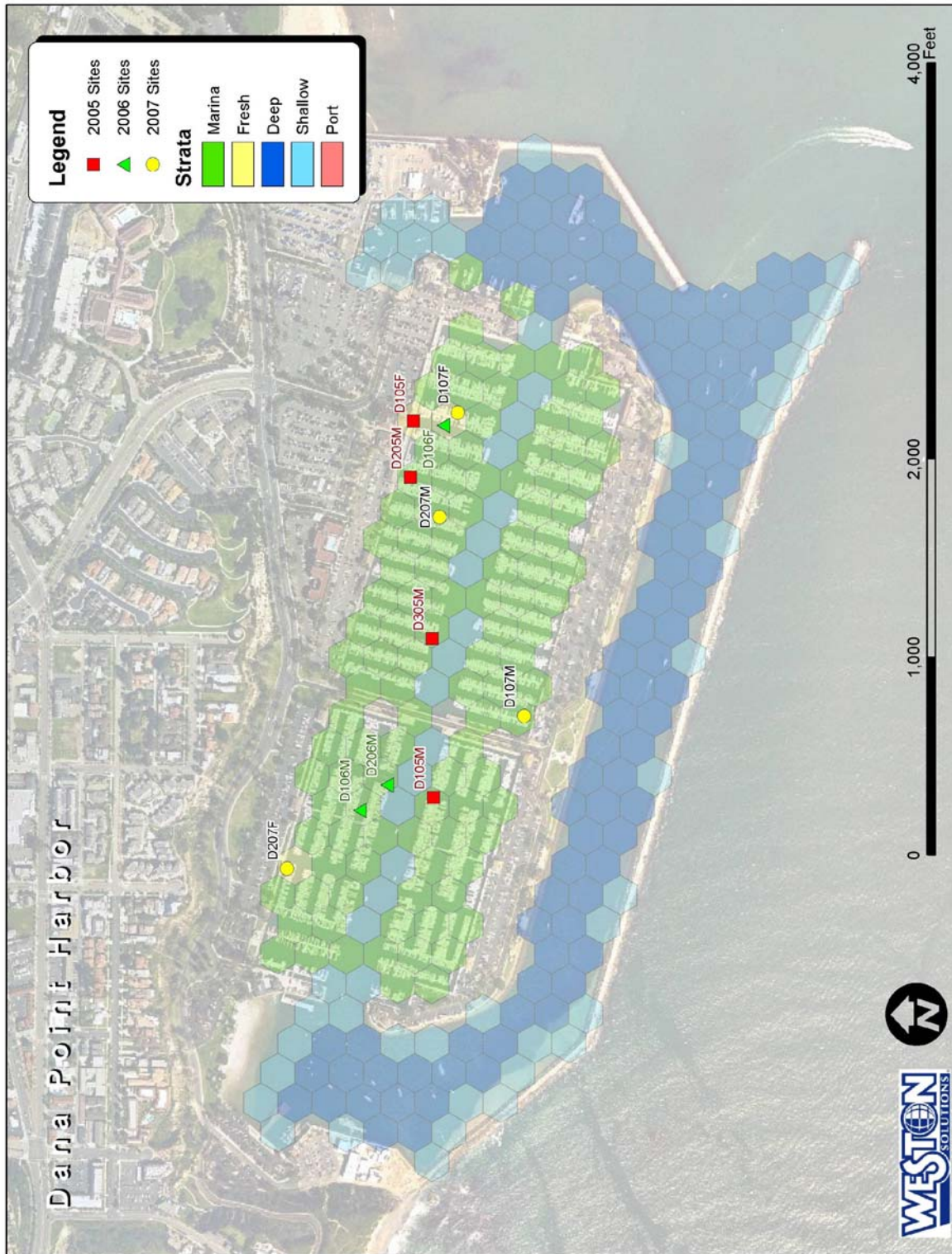


Figure 2-1. Sampling stations in Dana Point Harbor



Figure 2-2. Sampling stations in Oceanside Harbor

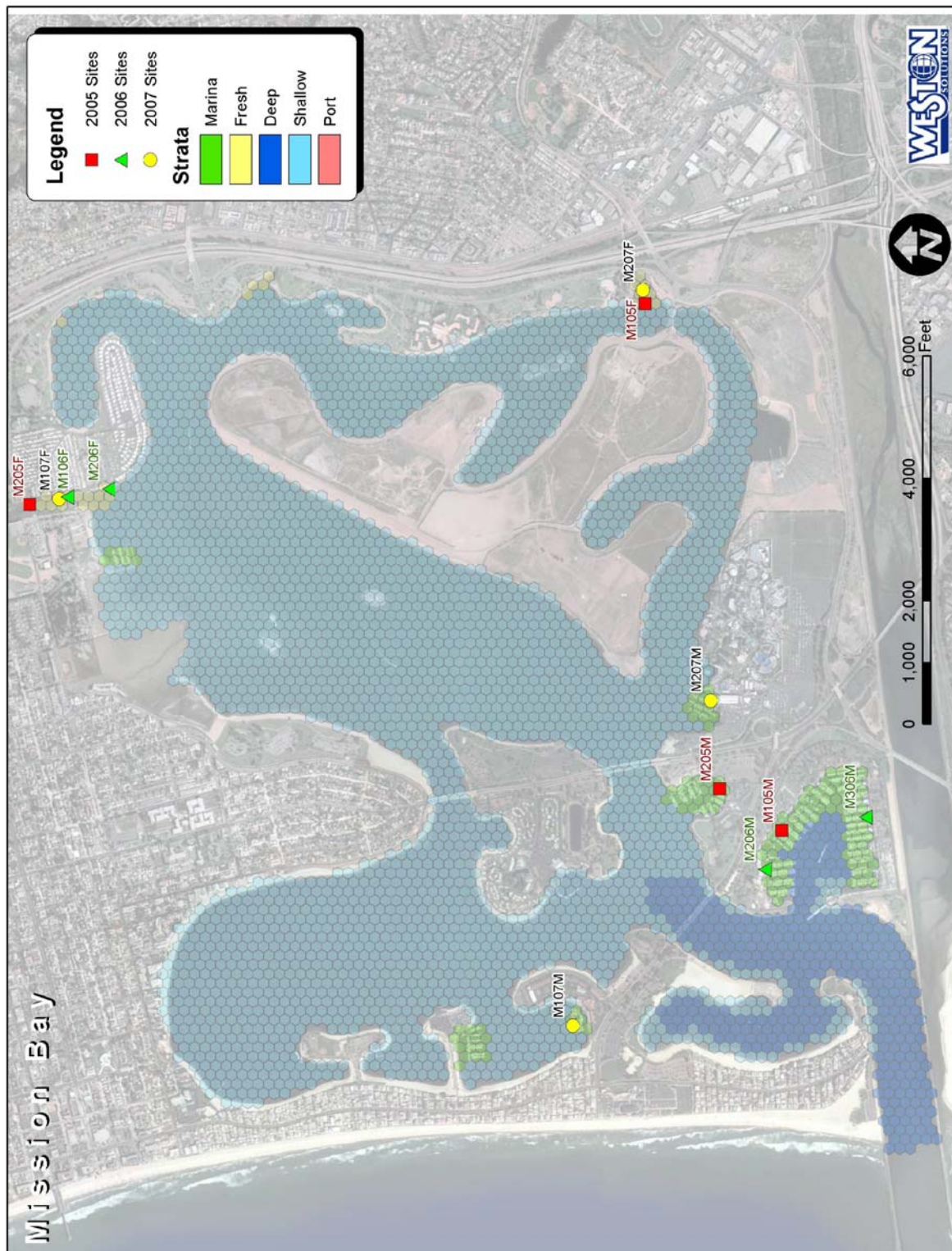


Figure 2-3. Sampling stations in Mission Bay

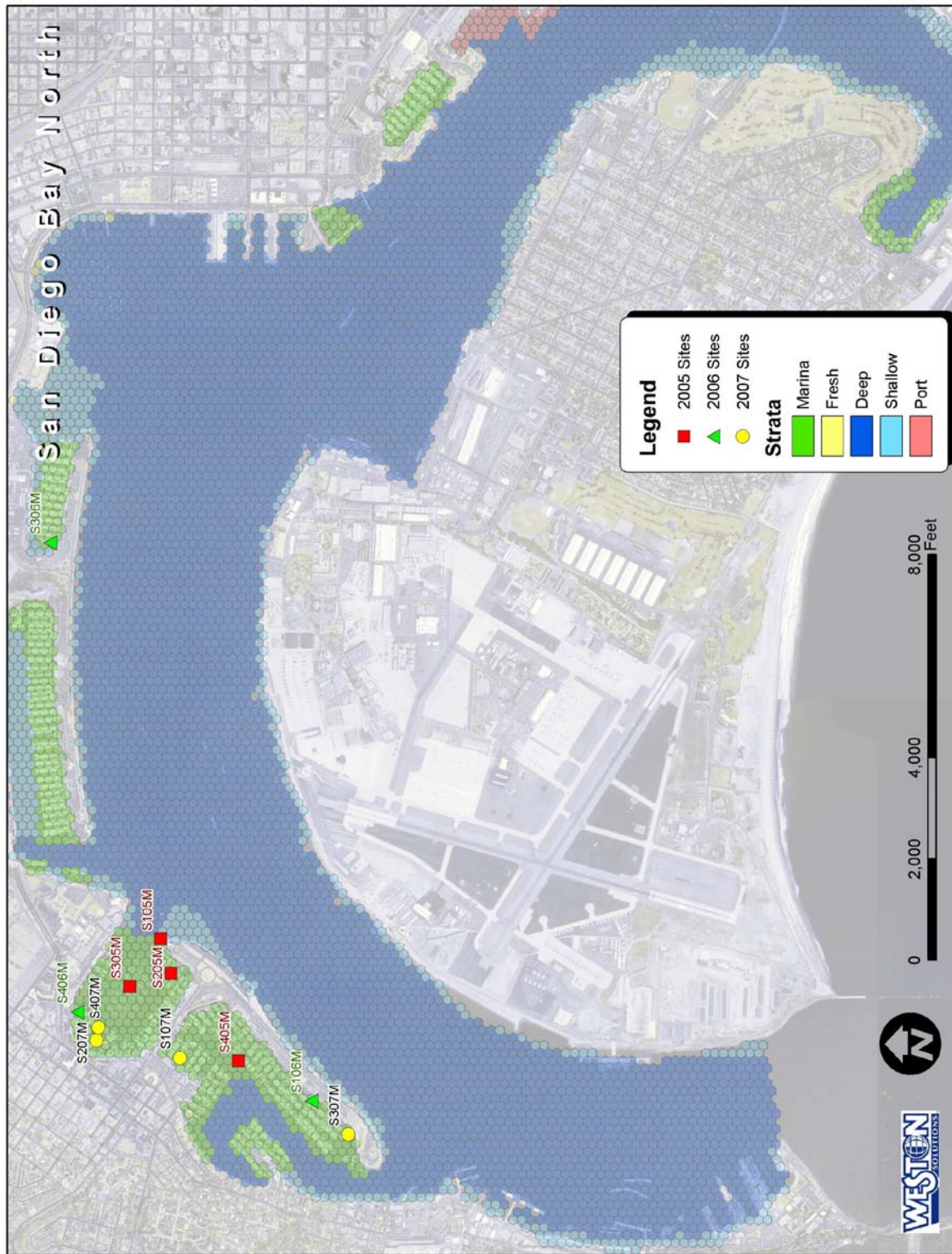


Figure 2-4. Sampling stations in northern San Diego Bay

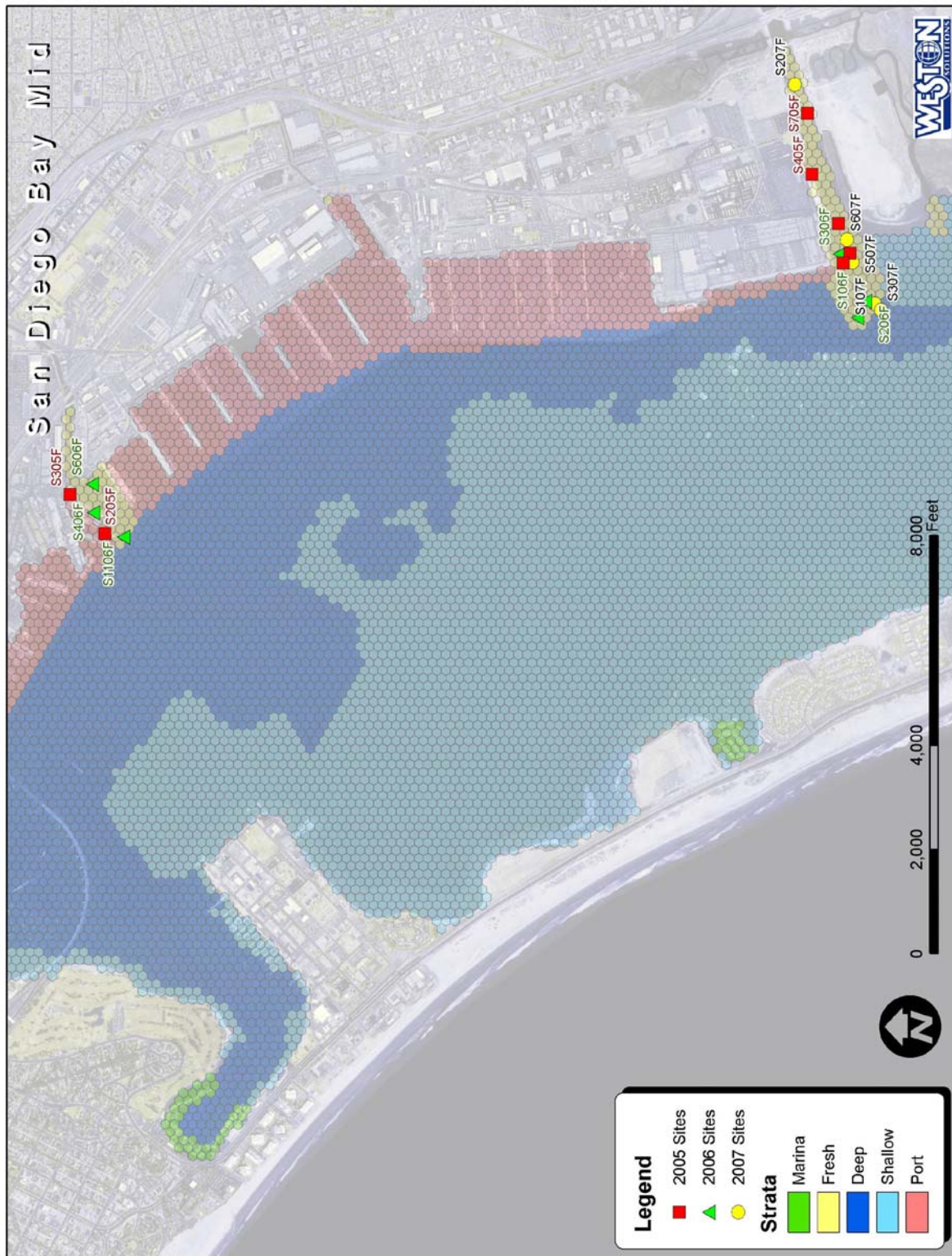


Figure 2-5. Sampling stations in central San Diego Bay

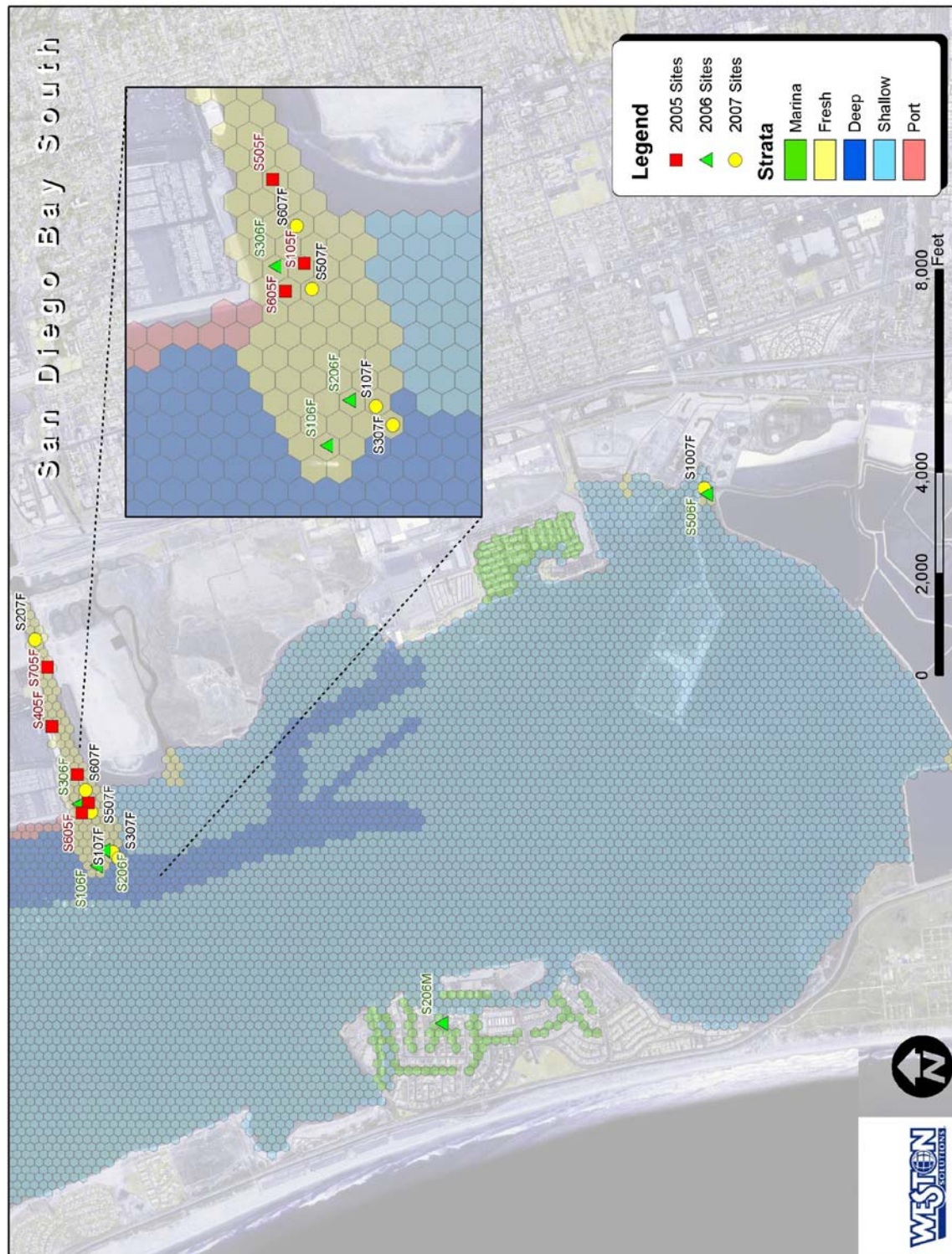


Figure 2-6. Sampling stations in southern San Diego Bay

2.1.2 Water Quality Sampling

Water column sampling was performed by Weston Solutions, Inc. (Weston) in August of 2005, 2006, and 2007. Each year, a total of twenty stations were sampled, ten marina and ten freshwater influenced. Field observations and coordinates of sample locations were recorded on sediment sampling data forms. Station locations and sampling dates are listed in Appendix A Tables A-1 – A-3.

Water column sampling was conducted using a Seabird SBE-25 Sealogger CTD (conductivity-temperature-depth) equipped with sensors that measure temperature, specific conductance, dissolved oxygen (DO), hydrogen ion concentration (pH), and light transmission. Casts were taken at stations located with a differential global positioning system (dGPS). Dissolved oxygen and pH sensors were calibrated prior to the week of monitoring. Transmissivity, conductivity, and temperature are calibrated annually by Sea-Bird Electronics, Inc. Before beginning a cast, a 3-minute equilibration was performed to bring the CTD sensors to thermal equilibration with the ambient seawater and to ensure that all of the pumps had turned on. The CTD was lowered at a speed of 0.25-0.50 m/sec until it was within one meter of the bottom. The instrument operated at a scan rate of 8 scans/sec.

After casts in each harbor were performed, the data were uploaded and saved onto a field computer. Data were checked to ensure the CTD turned on properly, the depth was accurate, and that all water quality measurements were recorded throughout a cast. Data were transferred to a disk upon returning to the laboratory. A post cruise calibration was performed following the week of sampling.

Discrete water samples were collected at each station one meter below the surface using a Niskin bottle. Water samples were transferred to labeled sample containers. Additional data, including weather, wind speed and direction, and water color and odor, were recorded on field data sheets. Samples to be analyzed for total organic carbon (TOC), dissolved organic carbon (DOC), dissolved and total metals, total hardness (measured as Calcium Carbonate [CaCO₃]), and polynuclear aromatic hydrocarbons (PAHs) were sent to CRG Marine Laboratories, Inc. (CRG) for chemical analyses. The CTD profiles and the samples for indicator bacteria, enterococcus, were analyzed by Weston. All of the samples were sent to the designated laboratories under the proper storage conditions and within holding times (six hours).

2.1.3 Sediment Sampling

Sediment sampling was performed in August of 2005, 2006, and 2007 at the same stations as those sampled for water quality using a dGPS to locate the stations. Field observations and coordinates of sample locations were recorded on sediment sampling data forms. Appendix A Tables A-1 – A-3 show the locations of the stations and sampling dates.

Benthic sediments were collected using a stainless steel, 0.1-m² Van Veen grab sampler. A minimum of four sediment grabs per station were collected for the following analyses: benthic infauna, acute toxicity, grain size, and chemistry (including TOC, total metals, and PAHs). A sample was determined to be acceptable if the surface of the grab was even, there was minimal surface disturbance, and there was a penetration depth of at least 5 centimeters (cm). Rejected grabs were discarded and re-sampled.

Samples collected for infaunal analysis were rinsed through a 1.0-mm mesh screen and transferred to a labeled quart jar. A 7% magnesium sulfate (MgSO_4) seawater solution was added for approximately 30 minutes to relax the collected specimens. The samples were then fixed in a 10% buffered formalin solution. Infaunal samples were analyzed by Weston.

Acute toxicity and sediment chemistry samples were collected from the top 2 cm of the grab, avoiding sediment within 1 cm of the sides of the grab. A total of 3 liters (L) of sediment was collected for toxicity and placed in three 1-L jars. Toxicity samples were kept at 4°C on ice in coolers. Sediment for trace metals and organics (PAHs) analysis was placed in one 4-ounce jar, stored at 4°C on ice, and frozen within 24 hours. Approximately 150-200 grams of sediment were collected for TOC and grain size. Each sample was placed in a 1-quart Ziploc™ bag and kept on ice. The TOC samples were frozen within 24 hours of collection, while grain size samples were stored at 4°C. Acute toxicity, grain size, and TOC samples were analyzed by Weston, while trace metals and PAH samples were shipped frozen to CRG within one week of collection for analyses.

2.2 Laboratory Analysis

2.2.1 Chemistry

Chemical analyses were performed on both water and sediment samples; a complete list of chemical analytes with corresponding analytical methods and detection limits is provided in Table 2-1. For water samples, analyses included TOC and DOC, total and dissolved metals, total hardness measured as CaCO_3 , and PAHs. For the sediment samples, TOC, trace metals and PAHs were analyzed. All chemical analyses were conducted to meet or exceed the specifications of the Surface Water Ambient Monitoring Program (SWAMP). Sediment samples were also analyzed for grain size (partitioned into gravel, sand, silt, and clay).

Table 2-1. RHMP constituents monitored and corresponding analytical methods

Analyte	Method	Reporting Limit	Units
Water Samples			
pH	Collected in field	-	-
Specific Conductance	Collected in field	-	µS/cm
Dissolved Oxygen	Collected in field	-	mg/L
Temperature	Collected in field	-	°C
Salinity	Collected in field	-	PSU
Transmissivity	Collected in field	-	%
Total Organic Carbon	EPA 415.1	1	%
Dissolved Organic Carbon	EPA 415.1	0.5	%
Total Hardness as CaCO ₃	SM 2340B	5.00	mg/L
Enterococcus	SM 9223	< 10	MPN/100ml
Dissolved Metals			
Aluminum (Al)	EPA 1640	0.125	µg/L
Antimony (Sb)	EPA 1640	0.015	µg/L
Arsenic (As)	EPA 1640	0.015	µg/L
Beryllium (Be)	EPA 1640	0.01	µg/L
Cadmium (Cd)	EPA 1640	0.01	µg/L
Chromium (Cr)	EPA 1640	0.01	µg/L
Cobalt (Co)	EPA 1640	0.01	µg/L
Copper (Cu)	EPA 1640	0.01	µg/L
Iron (Fe)	EPA 1640	0.025	µg/L
Lead (Pb)	EPA 1640	0.01	µg/L
Manganese (Mn)	EPA 1640	0.01	µg/L
Mercury (Hg)	EPA 245.7	0.02	µg/L
Molybdenum (Mo)	EPA 1640	0.01	µg/L
Nickel (Ni)	EPA 1640	0.01	µg/L
Selenium (Se)	EPA 1640	0.015	µg/L
Silver (Ag)	EPA 1640	0.01	µg/L
Thallium (Tl)	EPA 1640	0.01	µg/L
Tin (Sn)	EPA 1640	0.01	µg/L
Titanium (Ti)	EPA 1640	0.01	µg/L
Vanadium (V)	EPA 1640	0.01	µg/L
Zinc (Zn)	EPA 1640	0.01	µg/L
Total Metals			
Aluminum (Al)	EPA 1640	0.125	µg/L
Antimony (Sb)	EPA 1640	0.015	µg/L
Arsenic (As)	EPA 1640	0.015	µg/L
Beryllium (Be)	EPA 1640	0.01	µg/L
Cadmium (Cd)	EPA 1640	0.01	µg/L
Chromium (Cr)	EPA 1640	0.01	µg/L
Cobalt (Co)	EPA 1640	0.01	µg/L
Copper (Cu)	EPA 1640	0.01	µg/L
Iron (Fe)	EPA 1640	0.025	µg/L
Lead (Pb)	EPA 1640	0.01	µg/L
Manganese (Mn)	EPA 1640	0.01	µg/L
Mercury (Hg)	EPA 245.7	0.02	µg/L
Molybdenum (Mo)	EPA 1640	0.01	µg/L
Nickel (Ni)	EPA 1640	0.01	µg/L

Table 2-1. RHMP constituents monitored and corresponding analytical methods

Analyte	Method	Reporting Limit	Units
Selenium (Se)	EPA 1640	0.015	µg/L
Silver (Ag)	EPA 1640	0.01	µg/L
Thallium (Tl)	EPA 1640	0.01	µg/L
Tin (Sn)	EPA 1640	0.01	µg/L
Titanium (Ti)	EPA 1640	0.01	µg/L
Vanadium (V)	EPA 1640	0.01	µg/L
Zinc (Zn)	EPA 1640	0.01	µg/L
Polynuclear Aromatic Hydrocarbons			
1-Methylnaphthalene	EPA 625	5	ng/L
1-Methylphenanthrene	EPA 625	5	ng/L
2,3,5-Trimethylnaphthalene	EPA 625	5	ng/L
2,6-Dimethylnaphthalene	EPA 625	5	ng/L
2-Methylnaphthalene	EPA 625	5	ng/L
Acenaphthene	EPA 625	5	ng/L
Acenaphthylene	EPA 625	5	ng/L
Anthracene	EPA 625	5	ng/L
Benz[a]anthracene	EPA 625	5	ng/L
Benzo[a]pyrene	EPA 625	5	ng/L
Benzo[b]fluoranthene	EPA 625	5	ng/L
Benzo[e]pyrene	EPA 625	5	ng/L
Benzo[g,h,i]perylene	EPA 625	5	ng/L
Benzo[k]fluoranthene	EPA 625	5	ng/L
Biphenyl	EPA 625	5	ng/L
Chrysene	EPA 625	5	ng/L
Dibenz[a,h]anthracene	EPA 625	5	ng/L
Fluoranthene	EPA 625	5	ng/L
Fluorene	EPA 625	5	ng/L
Indeno[1,2,3-c,d]pyrene	EPA 625	5	ng/L
Naphthalene	EPA 625	5	ng/L
Perylene	EPA 625	5	ng/L
Phenanthrene	EPA 625	5	ng/L
Pyrene	EPA 625	5	ng/L
Sediment Samples			
Total Organic Carbon	EPA 415.1	0.05	%
Grain Size Analysis	Plumb 1981	-	-
Acute Toxicity	EPA/600/R-94/025	-	%
Benthic Infauna	-	-	-
Total Metals			
Aluminum (Al)	EPA 6020	5	mg/kg
Antimony (Sb)	EPA 6020	0.05	mg/kg
Arsenic (As)	EPA 6020	0.05	mg/kg
Barium (Ba)	EPA 6020	0.05	mg/kg
Beryllium (Be)	EPA 6020	0.05	mg/kg
Cadmium (Cd)	EPA 6020	0.05	mg/kg
Chromium (Cr)	EPA 6020	0.05	mg/kg
Cobalt (Co)	EPA 6020	0.05	mg/kg
Copper (Cu)	EPA 6020	0.05	mg/kg
Iron (Fe)	EPA 6020	5	mg/kg
Lead (Pb)	EPA 6020	0.05	mg/kg

Table 2-1. RHMP constituents monitored and corresponding analytical methods

Analyte	Method	Reporting Limit	Units
Manganese (Mn)	EPA 6020	0.05	mg/kg
Mercury (Hg)	EPA 245.7	0.02	mg/kg
Molybdenum (Mo)	EPA 6020	0.05	mg/kg
Nickel (Ni)	EPA 6020	0.05	mg/kg
Selenium (Se)	EPA 6020	0.05	mg/kg
Silver (Ag)	EPA 6020	0.05	mg/kg
Strontium (Sr)	EPA 6020	0.05	mg/kg
Thallium (Tl)	EPA 6020	0.05	mg/kg
Tin (Sn)	EPA 6020	0.05	mg/kg
Titanium (Ti)	EPA 6020	0.05	mg/kg
Vanadium (V)	EPA 6020	0.05	mg/kg
Zinc (Zn)	EPA 6020	0.05	mg/kg
Polynuclear Aromatic Hydrocarbons			
1-Methylnaphthalene	EPA 8270C	5	µg/kg
1-Methylphenanthrene	EPA 8270C	5	µg/kg
2,3,5-Trimethylnaphthalene	EPA 8270C	5	µg/kg
2,6-Dimethylnaphthalene	EPA 8270C	5	µg/kg
2-Methylnaphthalene	EPA 8270C	5	µg/kg
Acenaphthene	EPA 8270C	5	µg/kg
Acenaphthylene	EPA 8270C	5	µg/kg
Anthracene	EPA 8270C	5	µg/kg
Benz[a]anthracene	EPA 8270C	5	µg/kg
Benzo[a]pyrene	EPA 8270C	5	µg/kg
Benzo[b]fluoranthene	EPA 8270C	5	µg/kg
Benzo[e]pyrene	EPA 8270C	5	µg/kg
Benzo[g,h,i]perylene	EPA 8270C	5	µg/kg
Benzo[k]fluoranthene	EPA 8270C	5	µg/kg
Biphenyl	EPA 8270C	5	µg/kg
Chrysene	EPA 8270C	5	µg/kg
Dibenz[a,h]anthracene	EPA 8270C	5	µg/kg
Fluoranthene	EPA 8270C	5	µg/kg
Fluorene	EPA 8270C	5	µg/kg
Indeno[1,2,3-c,d]pyrene	EPA 8270C	5	µg/kg
Naphthalene	EPA 8270C	5	µg/kg
Perylene	EPA 8270C	5	µg/kg
Phenanthrene	EPA 8270C	5	µg/kg
Pyrene	EPA 8270C	5	µg/kg

2.2.2 Toxicity

Solid phase (SP) bioassays were performed to estimate the potential toxicity of the collected sediments to benthic organisms. SP bioassays followed established protocols from the 2003 Regional Monitoring Program for the Southern California Bight (Bight '03). The sediments were tested in a 10-day SP test using the marine amphipod *Eohaustorius estuarius*. On the day before test initiation a 2-cm aliquot of sample sediment was placed in a test chamber followed by prepared seawater. The samples were left overnight to allow establishment of equilibrium between the sediment and overlying water. On day one of the test, 20 amphipods were randomly placed in each of the test chambers. Any amphipods that did not bury in the sediment within 5 to

10 minutes were removed and replaced. Control sediment was used to determine the health of the amphipods; this was done by exposing the amphipods to clean sediment according to the same protocols used for the test sediments. Samples were monitored daily for the emergence of amphipods. At the end of the test, organisms were removed from the test chambers by sieving the sediment through a 0.5-mm mesh screen and the numbers of live and dead amphipods in each test chamber were recorded. The percent survival was calculated for the control and test sediments. The acceptability of the test was determined by evaluating the response of the control organisms. The test was considered acceptable if there was 90% mean control survival.

Standard procedure calls for measuring pore water ammonia in the sediments prior to test initiation to determine whether the concentration is within acceptable limits. If concentrations exceed 60 mg/L there is a potential that any observed toxicity may be due to high ammonia rather than some other contaminant. In 2006, the pore water ammonia concentration in sample M2M06 (94.0 mg/L) was found to be significantly elevated above recommended limit of 60 mg/L. Consequently, to reduce ammonia concentrations, sample material was acclimated by performing daily renewals with fresh seawater until pore water ammonia concentrations were reduced to 60 mg/L or below. An acclimated control was set up concurrently with sample M2M-06 to ensure that the acclimation procedure did not contribute to toxicity of *E. estuarius*. An un-acclimated control was also set up the day before test initiation to reflect normal test procedures.

A 96-hour reference toxicity test was also conducted concurrently with the sediment test to establish sensitivity of the test organisms used in the evaluation of the sediments. The reference toxicant test was performed using the reference substance, cadmium chloride, with concentrations of 2.5, 5, 10, 20, and 40 mg Cd^{2+} /L. Ten test organisms were added to each of these concentrations. The concentration that caused 50% mortality of the organisms (the median lethal concentration, or LC_{50}) was calculated from the data. The LC_{50} values were then compared to historical laboratory data for the test species with the reference substance. The results of this test were used in combination with the control mortality to assess the health of the test organisms.

An additional reference toxicant test was also conducted using ammonium chloride with target concentrations of 15.62, 31.25, 62.5, 125, and 250 mg NH_4 /L to evaluate potential influence of ammonia toxicity on the test results of the sediments.

2.2.3 Infauna

The benthic samples were transported from the field to the laboratory and stored in a formalin solution for 7 days. The samples then were transferred from formalin to 70% ethanol for laboratory processing. The organisms were initially sorted using a dissecting microscope into five groups: polychaetes, crustaceans, molluscs, echinoderms, and miscellaneous minor phyla. While sorting, technicians kept a rough count for quality assurance/quality control (QA/QC) purposes, as described in the following paragraph. After initial sorting, qualified taxonomists identified each organism to the lowest possible taxon, and species counts were tabulated.

A QA/QC procedure was performed on each of the sorted samples to ensure a 95% sorting efficiency. A 10% aliquot of a sample was re-sorted by a senior technician trained in the QA/QC procedure. The number of organisms found in the aliquot was divided by 10% and added to the total number found in the sample. The original total was divided by the new total to calculate the

percent sorting efficiency. When the sorting efficiency of the sample was below 95%, the remainder of the sample (90%) was re-sorted.

2.2.4 Microbiology

Water samples were analyzed for the indicator bacteria, enterococcus, using IDEXX Enterolert™ methodology. All results were reported to a most probable number (MPN) value with a minimum reporting limit of <10 MPN/100mL and a maximum reporting limit of 24,196 MPN/100mL. All samples were delivered to the analytical laboratory within the 6-hour holding time requirement. Sample analysis was initiated immediately upon receipt.

2.2.5 Profile Data Processing

Sea-Bird profile scans were processed by Sea-Bird data processing software to convert recorded voltages to parametric values. Scans were averaged into one-meter bins for analysis.

2.3 Data Analysis

In Phase I of this project, historical data were compiled to establish threshold levels and preset targets by which to measure changes in the harbors (Table 2-2). The majority of the data were from the 1998 Regional Monitoring Program for the Southern California Bight (Bight '98) and the Bay Protection and Toxic Cleanup Program (BPTCP). Data that had similar detection limits (chemistry), test species (toxicity), and sampling equipment and screen size (benthic infauna) were used to determine a threshold level (Weston, 2005b).

The selection of which indicators were going to be monitored in the Pilot Project was based on whether there was sufficient historical data to create a threshold level. The threshold levels were established as concentration levels for chemical constituents, toxicity levels for bioassays, and diversity measures and the BRI for infauna (Smith et al., 2003). Preset targets were determined by defining the proportion of historical samples collected in the harbors that were below the established threshold levels. Preset target proportions were defined to be the constant in the binomial model for comparison to Pilot Project data from the harbors and bays. Proportions of stations below the threshold level were compared to the preset target to determine differences between the historical conditions of the harbors/bays and the present conditions of each stratum (marina and freshwater influenced). Since the Pilot Project was designed to assess the effectiveness of the overall approach within specific strata rather than assess the overall conditions of the four harbors and bays, direct comparisons to preset targets, even when statistically significant, do not indicate a state that has improved or worsened from the historical conditions. However, upon implementation of the full RHMP when samples will be collected in all five strata, then a significantly greater proportion of observed samples above the preset targets for all strata combined would indicate that water or sediment quality conditions are improving (Weston, 2005b). In the case of the three infaunal parameters, conditions will be considered to be better than historical levels when the proportions of stations below the BRI score threshold are lower than the historical preset target and when proportions for the Shannon-Wiener Diversity Index and the number of taxa are above the preset target. A summary of the established threshold levels and preset targets is presented in Table 2-3.

Indicators were partitioned into primary and secondary indicators. Primary indicators for the study were selected because they are either major constituents of concern (e.g., copper in water) or they provide information on a suite of measurements (e.g., the mean ER-M quotient). Secondary indicators are used as supporting data to enhance the interpretation of the primary indicators (Weston, 2005b). The selection of individual primary and secondary indicators for water column chemistry, sediment chemistry, sediment toxicity, and benthic infauna is further discussed in Sections 2.3.1 through 2.3.4.

Table 2-2. Studies used to establish reference ambient values

Study Name	Year	Dana Point Harbor	Oceanside Harbor	Mission Bay	San Diego Bay
Sediment Chemistry					
America's Cup Harbor	2001				X
Bight 98	1998	X		X	X
BPTCP	1994, 1996	X	X	X	X
Central SD Bay Nav. Channel Deepening	1998, 2003				X
Chollas Creek	2003				X
10th Avenue Marine Terminal	2002				X
National City Wharf Extension	1999				X
Navy Arco	2000				X
Navy P-326	2000				X
Paleta Creek	2003				X
Reference reconnaissance	2003				X
Sediment sampling	2003	X			
Toxic Hot Spots Sediment	2003				X
Water and Sediment Testing Project	2001-2003			X	
Benthic Infauna					
Ambient Bay and Lagoon Monitoring	2003		X	X	
America's Cup Harbor	2002				X
Bight 98	1998	X		X	X
Reference reconnaissance	2003				X
Switzer Creek	2002				X
Sediment Toxicity					
Bight 98	1998				X
Benthic Infauna Analysis	2003-2004	X			
National City Wharf Extension	1999				X
Water and Sediment Testing Project	2001-2003			X	
Water Column Chemistry					
Baywide Copper	2002				X
Dana Point monitoring	1992-2002	X			
Paco Bay Water measurements	1992-1999				X

Table 2-3. Summary of threshold values and preset targets

Measure	Threshold Value	Preset Target
Primary Indicators		
Dissolved Copper (water)	4.8 µg/L	70%
Total Copper (water)	5.8 µg/L	26%
ER-M Quotient	0.2	48%
BRI	31	37%
<i>E. estuarius</i> mortality	20%	51%
Secondary Indicators		
Dissolved Zinc (water)	90 µg/L	100%
Total Zinc (water)	95 µg/L	97%
Dissolved Nickel (water)	74 µg/L	100%
Total Nickel (water)	75 µg/L	100%
Sediment Cadmium	1.2 mg/kg	90%
Sediment Chromium	81 mg/kg	78%
Sediment Copper	175 mg/kg	68%
Sediment Lead	46.7 mg/kg	74%
Sediment Nickel	20.9 mg/kg	80%
Sediment Zinc	150 mg/kg	45%
Sediment Total PAHs	4022 µg/kg	74%
Shannon-Wiener diversity	2	90%
Number of taxa	24	92%

Each of the indicators measured in the Pilot Project was plotted for visual comparison to the threshold levels and preset targets. Figure 2-7 shows an example of a distribution curve that can be used as a reference guide. Both the historical and current data were plotted as distribution curves with the current data overlying the historical data. The current data is shown as a step plot rather than a smooth curve because there are only 10 samples analyzed annually and 30 samples analyzed cumulatively from each stratum compared to the larger amount of samples used from historical data. The horizontal blue line is the threshold level for each indicator. The vertical green line is the preset target. The orange line represents where the distribution curve for the current data crosses the threshold level.

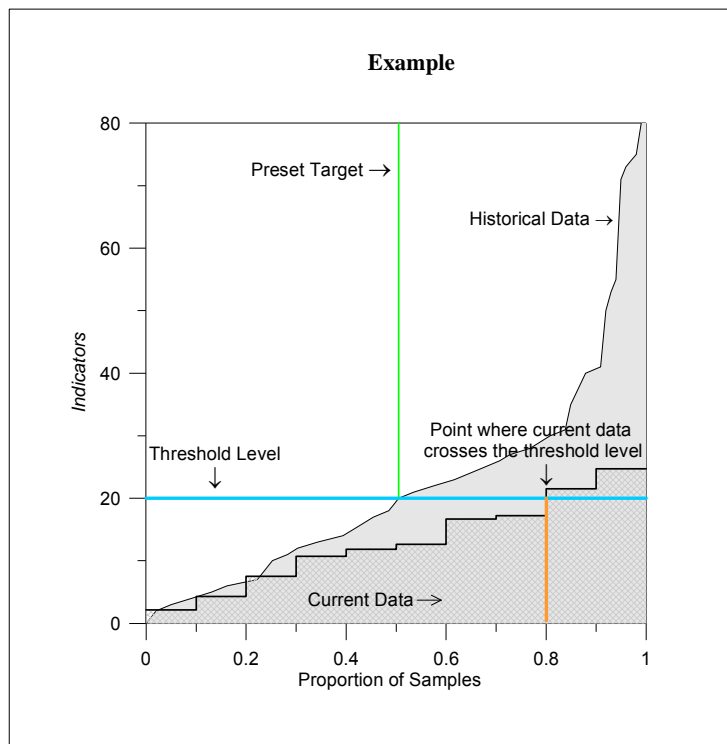


Figure 2-7. Example of a distribution curve that can be used as a reference guide

When the orange line is to the left of the preset target then the proportion of samples that are below the threshold level is lower than the proportion of samples historically observed below this level. Upon implementation of the full RHMP study, this would indicate that water or sediment quality conditions for that particular indicator have degraded from historical conditions throughout the harbors and bays. If the orange line is to the right of the preset target then the proportion of samples below the threshold level is greater than the proportion of samples historically observed below the threshold. This would indicate progress towards improved water or sediment quality in the harbors. The results for each indicator was statistically compared to the preset target to determine if the percent of samples below the threshold level was higher or lower than historical conditions for the four harbors, as detailed in Section 2.3.5.

2.3.1 Water Column Chemistry

Historical observations of water column metal concentration were available for dissolved and total copper, nickel, and zinc (Weston, 2005b). The data along with benchmark values from the California Toxics Rule (CTR) and the California Ocean Plan (COP) were evaluated to establish threshold levels. The CTR was created using both literature and toxicity test data, thus making it the best threshold level to use for aqueous metals (CTR, 2000). Only dissolved and total copper were selected as primary indicators for aqueous metals because of the large numbers of historical observations above the CTR. Dissolved and total zinc and nickel were selected as secondary indicators. If the percent of current samples below the threshold level for a particular stratum was found to be greater than the preset target it would indicate that water quality in the stratum

was better than historically observed across all five strata within the harbors (Weston, 2005b). The threshold levels and preset targets for these metals are listed in Table 2-3.

2.3.2 Sediment Chemistry

For sediment chemistry, the mean ER-M quotient is the primary indicator for comparing results in the monitoring program to preset targets. Briefly, the effects range-low (ER-L) and effects range-median (ER-M) are two effects-based sediment quality values developed to help interpret sediment chemistry measurements and their potential for causing adverse biological effects (Long et al., 1995). These parameters were developed from an extensive database of sediment toxicity bioassays and chemistry measurements. The ER-L was calculated as the lower tenth percentile of the observed effects concentrations and the ER-M as the 50th percentile of observed effects concentrations. Concentrations below the ER-L are not likely to result in biological effects, while concentrations above the ER-M are likely to result in biological effects (Long et al., 1995).

The ER-M quotient, which is the ratio of sample concentration to the ER-M, can be used to evaluate the likelihood of benthic effects based on cumulative sediment chemistry. The quotient is calculated by dividing each measured sediment chemical concentration by its respective ER-M. The mean ER-M quotient calculates an average quotient based on concentrations of all known contaminants relative to the ER-M values. Therefore, the mean ER-M quotient is a method of integrating the effects from multiple contaminants (Wenning et al., 2005). For the Pilot Project, the mean ER-M quotient was calculated using concentrations of arsenic, cadmium, chromium, copper, lead, mercury, nickel, silver, zinc, and total detectable PAHs.

Based on recent projects with the SDRWQCB, the threshold level for the mean ER-M quotient was determined to be 0.2. Samples with mean ER-M quotients above 0.2 are more likely to have adverse benthic effects associated with the sediment chemistry. Based on historical data, the preset target for the mean ER-M quotient was established at 48% across all strata. If the percent of current sediment samples with a mean ER-M quotient below 0.2 is significantly higher than 48%, then it would indicate that the overall conditions of sediment quality for the stratum are better than conditions historically observed within the harbors. If the percent of samples continues to be lower than the preset target over the course of the program then other indicators such as individual chemical constituents can be evaluated in conjunction with the mean ER-M quotient to help determine which contaminants are problematic in the harbors (Weston, 2005b).

Total PAHs and metals, including cadmium, chromium, copper, lead, nickel, and zinc, are used as secondary indicators for the Pilot Project. These measures will be used to help interpret the mean ER-M quotient by showing which of the parameters are predominant or changing in the mean ER-M quotient. For total PAHs and all of the metals except copper, the ER-L was determined to be the best threshold level. The threshold level for copper was based on the level at which anthropogenic origins may be contributing to the overall copper concentrations in the sediment. To determine this concentration, historical data were used to plot copper concentrations against iron concentrations, both of which are common in harbor sediments. Normalization to iron is a common approach to understanding the influence of potential enrichment via anthropogenic inputs since iron is a reliable indicator of “geological background” levels. When trace metals, such as copper, co-vary with iron, they are generally viewed as being within geological background, i.e., they are not attributed to anthropogenic influences (Schiff

and Weisberg, 1999). At lower concentrations of copper within the historical dataset there is a constant linear relationship with iron; however, this relationship changes at a copper concentration of about 175 mg/kg as shown in Figure 2-8. As a consequence, the threshold level for sediment copper was set at 175 mg/kg due to the relatively pronounced shift in the relationship between copper and iron. A higher percent of current samples below the threshold level compared to the preset target would indicate that the measure of sediment quality in the stratum was better than historically observed throughout the harbors (Weston, 2005b). Table 2-3 shows the threshold levels and preset targets.

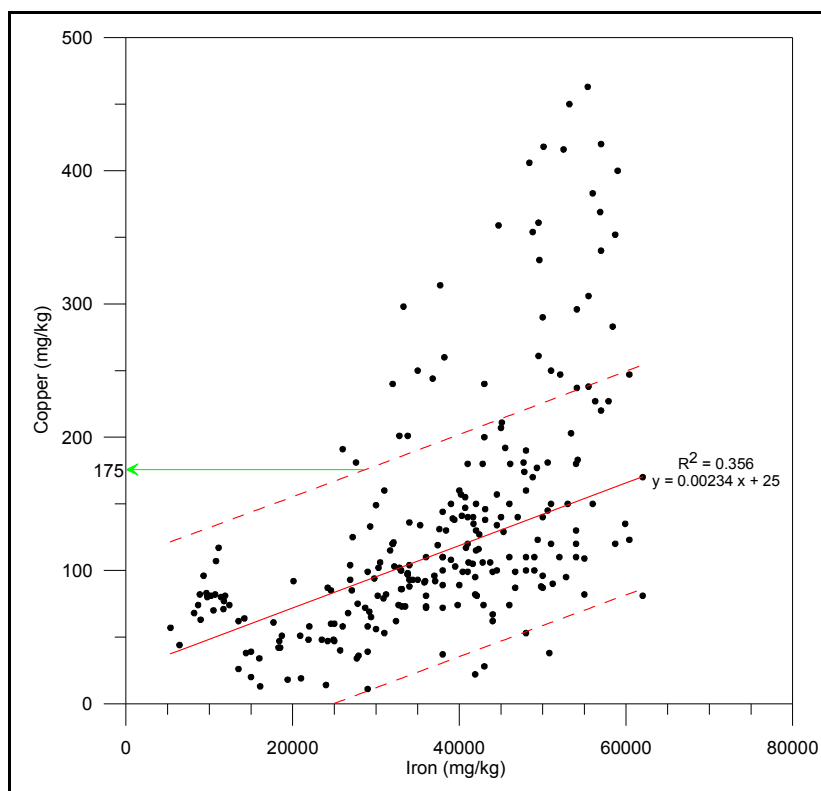


Figure 2-8. Relationship of copper to iron

2.3.3 Sediment Toxicity

Historical toxicity test results for *Eohaustorius estuarius* were used to establish the threshold levels for sediment toxicity. *E. estuarius* was selected as the test species due to its relatively high sensitivity to toxic substances and the availability of data for this species within the study area. Mortality, rather than survival, was analyzed to be consistent with the majority of the other indicators since higher values indicate poorer conditions. Test results were adjusted for control mortality prior to analysis of the data. The threshold level was set at 20% mortality; a value that is typically used as an indicator of non-toxic sediments. Conditions within a stratum were considered to be better than historical conditions if the current percent of stations below the toxicity threshold value was greater than the preset target (i.e., more than 51% of samples show less than 20% mortality) (Weston, 2005b).

2.3.4 Benthic Infauna

Benthic infauna data from each of the harbors was assessed using various indices common to ecological community structure evaluations, including the BRI, Shannon-Wiener diversity index, abundance, and number of taxa. The BRI is the primary indicator for evaluating infaunal assemblages in the harbors. The numerical criterion (i.e. community response levels) for this index is calculated by applying an abundance-weighted-average gradient that is correlated with sediment/habitat quality to the pollution tolerance of infaunal species. A reference threshold and four response levels help to characterize the degrees to which habitat conditions are deviating from reference conditions. Response level 1 is characterized as marginal deviation. Level 1 includes BRI values at which 5% of the reference species were lost. Response Levels 2-4 indicate increasingly disturbed benthic environments. Response level 2 is characterized as a biodiversity loss of 25% of reference species. Response level 3 is when there is a community function loss. BRI values at this level indicate a loss of 50% of reference species. Response level 4 is characterized by defaunation, which indicates a loss of 80% of reference species (Ranasinghe et al., 2003). The range of BRI levels for each of these response levels is shown in Table 2-4.

Table 2-4. Characterization and BRI ranges for response levels of benthic community conditions

BRI Threshold	Level	Characterization	Definition
<31	Reference		
31 - <42	Response Level 1	Marginal deviation	>5% of reference species lost
42 - <53	Response Level 2	Biodiversity loss	>25% of reference species lost
53 - <73	Response Level 3	Community function loss	>50% of reference species lost
>73	Response Level 4	Defaunation	>80% of reference species lost

The BRI threshold level for the Pilot Project was set at 31, which is the currently established value for reference conditions in embayments. After applying this value to historical data, a preset target proportion was determined to be 37%. If more than 37% of the current samples are below the threshold level of 31, then the benthic infaunal community is considered to be impaired relative to what was observed historically. Alternatively, conditions within a stratum were considered to be better than historical conditions if the current percent of stations below the benthic infaunal community threshold value was greater than the preset target.

The Shannon-Wiener diversity and number of taxa are used as secondary indicators. For both of these indicators, higher values indicate healthier benthic communities. The Shannon-Wiener diversity index takes into account the number of species and the evenness of the species, where higher values are indicative of greater diversity and/or evenness. Evenness provides an indication of the equality of different species abundances within a community. Number of taxa also provides a measure of diversity as it is a count of the number of species (or lowest taxonomic units) encountered within a sample. For Shannon-Wiener diversity, the threshold level was determined to be 2 with a preset target proportion of 90%. The threshold for number of taxa was 24 with a preset target of 92%. In contrast to all other indicators, a healthier state than historical harbor conditions is said to occur for the benthic community when the observed

Shannon-Wiener diversity index and number of taxa are GREATER than the threshold levels (Weston, 2005b).

2.3.5 Statistical Analyses

A binomial model was selected to assess differences in benthic infaunal and sediment and water quality parameters between the Pilot Project and established thresholds following the methods of Cohen (1977). Parameters were compared separately for each stratum (marina and freshwater influenced) for all bays combined, including San Diego Bay, Mission Bay, Oceanside Harbor, and Dana Point Harbor. Thus, the states of the strata were assessed without respect to differences among harbors and bays through the use of the binomial analysis. Results for each indicator were statistically compared to the preset target to determine if the proportion of samples below the threshold level was higher or lower. Differences were considered to be significant at $p \leq 0.10$, which indicates a 90 % certainty that the difference that we detect is not due simply to chance. The null hypothesis was that the proportion of current samples below the threshold level was equal to the historical proportion of samples below the threshold level. When the null hypothesis was rejected, it was determined that the current value is significantly different from the preset target. Binomial tests were used to compare the proportion of all stations across all years below the threshold (referred to as the cumulative proportion) to the historical preset target, as well as each individual year to the preset target (i.e., separate tests for 2005, 2006, and 2007). Since the Pilot Project only utilized two strata, differences in the cumulative proportions of samples below thresholds were also compared between the marina and freshwater-influenced strata using the binomial model.

Differences in surface water, sediment, and benthic infaunal parameters also were compared statistically between strata (marina vs. freshwater influenced), among years (2005, 2006, and 2007), and among harbors separately by strata (marina: Dana Point Harbor, Oceanside Harbor, San Diego Bay, and Mission Bay; and freshwater influenced: all except Oceanside Harbor since the freshwater-influenced stratum were absent). To determine whether parametric or nonparametric statistics were required, data were tested for normality and equality of variances. Normality was tested with Kolmogorov-Smirnov tests and equality of variances was tested with Bartlett's tests. If data did not meet the criteria, then data transformations were performed to improve normality and equality of variances. Transformations included arcsine transformations for percentages and square-root and log transformations for the other indicators, following the methods of Zar (1999). If either untransformed or transformed data met requirements, then parametric statistical tests were used (Analysis of Variance [ANOVA]); otherwise, non-parametric tests were performed (Mann-Whitney tests and Kruskal-Wallis tests).

Differences between the marina and freshwater-influenced strata were compared by ANOVA when data met the parametric criteria and by Mann-Whitney tests when the data did not. Additionally, ANOVAs and Kruskal-Wallis tests were used to assess differences in indicator measures or values among years and among harbors, performed separately for each stratum. Rather than comparing proportions, these tests directly compared differences in the indicators (for example dissolved copper concentrations) between the strata, among sample years, and among harbors. Differences were considered to be significant at $p \leq 0.05$, which indicates a 95% certainty that the difference that we detect is not due simply to chance. When significant differences were detected by ANOVAs or Kruskal-Wallis tests, follow-up pair-wise tests were performed to test for differences between any two given years or harbors.

Regression analysis was performed to test the relationship between toxicity (i.e., *E. estuarius* mortality) and grain size (percent clay), since *E. estuarius* has the potential to be negatively impacted by grain size, experiencing higher levels of mortality when sediments are fine grained (Weston, unpublished data).

Coefficients of variation (CV) were calculated to compare the level of variability among indicators, since this measure provides a relative measure of variability that is normalized by the average (i.e., $CV = 100\% \times \text{standard deviation}/\text{mean}$). Typically, measures with higher averages also have larger levels of variation, making it inappropriate to compare different measures without normalization.

Power analysis was used to test the probability that a statistical test will yield a significant result (i.e., reject the null hypothesis). An analysis was performed to test the ability of the binomial test to detect a significant difference between observed proportions and preset targets for the primary and secondary indicators. Based on the observed level of variability among stations, the level of difference between observed proportions and preset targets, and a power of 80%, estimated sample sizes to detect a significant difference were calculated (Cohen, 1977)..

3.0 RESULTS

3.1 Water Quality

3.1.1 Chemistry

Surface water samples collected from marina and freshwater-influenced stations were analyzed for total and dissolved metals, hardness, DOC, TOC, and PAHs. Of the indicator metals (copper, nickel, and zinc), only the primary indicator (copper) exceeded CTR thresholds for dissolved and total concentrations, and none of the metals exceeded COP standards.

Surface water chemistry results for primary and secondary indicators at all stations surveyed throughout the Pilot Project are reported in Tables B-1 – B-3 and cumulative distribution curves used to determine percentages of area that did not exceed threshold levels for secondary indicators are presented in Figures B-1 – B-2 in Appendix B. Additionally, surface water chemistry results for all indicators assessed in 2007 are provided in Appendix C and those of previous years are reported in 2005-2006 and 2006-2007 RHMP Pilot Project Reports (Weston, 2006 and Weston, 2007). Percentages of stations that did not exceed copper CTR thresholds were significantly different from the historically-defined preset targets, with the marina stratum having a higher percentage of area exceeding CTR thresholds and the freshwater-influenced stratum having a lower percentage of area with exceedances (Table 3-1, Figure 3-1). Although exceedances were only observed for copper, there were often pronounced differences in metal concentrations between strata and among harbors, but concentrations rarely changed significantly with time over the duration of the Pilot Project (significance values reported in Appendix B Tables B-10 –B-11). For DOC, TOC, and total detectable PAHs, CTR and COP thresholds have not been established; however differences in concentrations were apparent between strata (DOC and TOC), among years (DOC and TOC), and among harbors (total detectable PAHs) as described in the following sections.

Table 3-1. Percentage of stations below CTR thresholds for marina and freshwater-influenced surface water metals

Metal	Preset Target	Marina				Freshwater Influenced			
		2005	2006	2007	Cumulative	2005	2006	2007	Cumulative
dissolved copper	70	40*	30*	40*	37*	90*	80	80	83*
total copper	26	10	30	20	20	60*	90*	80*	77*
dissolved nickel	100	100	100	100	100	100	100	100	100
total nickel	100	100	100	100	100	100	100	100	100
dissolved zinc	100	NA	100	100	100	NA	100	100	100
total zinc	97	100	100	100	100*	100	100	100	100*

* Indicates result that is significantly different from preset target ($p \leq 0.1$); green indicates a higher percentage and yellow a lower percentage than preset target.

NA = Not Available

Primary Indicators

Dissolved and Total Copper

Historically, 70% of the sites sampled did not exceed the CTR threshold for dissolved copper and 26% did not exceed the threshold for total copper when assessed across all areas of the

harbors. Throughout the three-year Pilot Project, 37% of the marina and 83% of the freshwater-influenced strata did not exceed the dissolved copper CTR threshold (4.8 $\mu\text{g/L}$), and 20% of the marina and 77% of the freshwater-influenced strata did not exceed the total copper CTR threshold (5.8 $\mu\text{g/L}$) (Figure 3-1). Within the marina stratum, the percentage of stations that did not exceed the dissolved CTR threshold was significantly lower than the preset target, while the difference for total copper was not significant (Table 3-1). Within the freshwater-influenced stratum, the percentage of stations that did not exceed the total copper CTR threshold was three times higher than the preset target, resulting in a significant difference. Based on the primary indicator copper, conditions in the marina surface waters were worse than historical conditions, while conditions in the freshwater-influenced stratum were better.

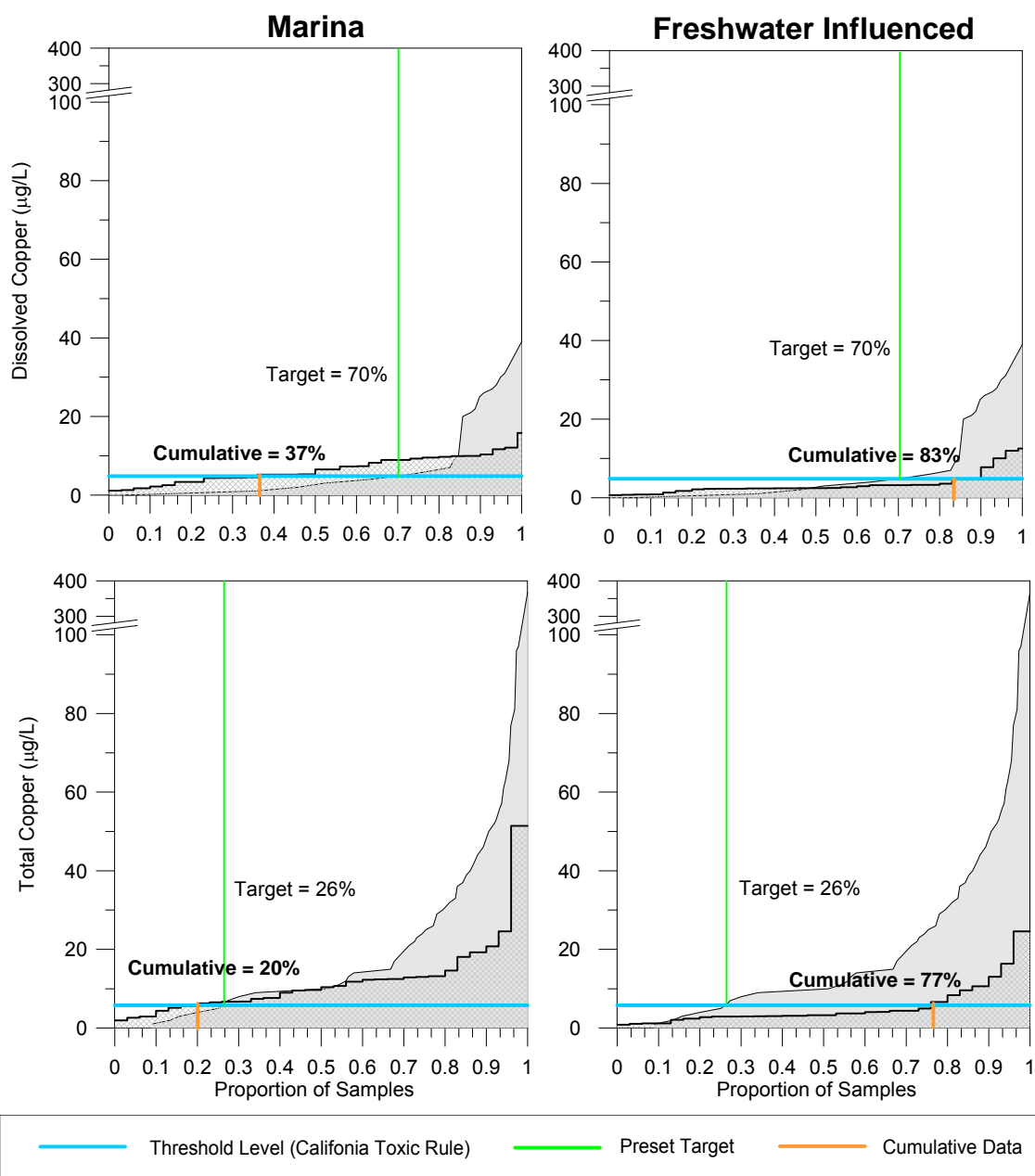


Figure 3-1. Cumulative distribution curves for surface water dissolved and total copper

Percentages of area that did not exceed dissolved and total copper CTR thresholds were approximately two times higher in the freshwater-influenced stratum than in the marina stratum, resulting in a significant difference between the strata. Within the marina stratum, dissolved copper concentrations ranged from 1.10-15.80 µg/L and total copper ranged from 1.33-51.40 µg/L. In the freshwater-influenced stratum, concentrations ranged from 0.70-12.51 µg/L for dissolved copper and 0.81-24.60 µg/L for total copper. Accordingly, average (reported as mean \pm standard error) copper concentrations were approximately two times higher in the marinas (6.56 ± 0.66 µg/L) than the freshwater-influenced stratum (3.54 ± 0.56 µg/L), resulting in a significant difference.

Throughout the duration of the Pilot Project, the percentages of area that did not exceed dissolved copper threshold levels were relatively consistent from year to year, ranging from 30-40% in the marinas and 80-90% in the freshwater-influenced stratum (Table 3-1). For total copper, interannual differences in percentages were of a greater magnitude, increasing from 10% in 2005 to 60% in 2006 before declining to 50% in 2007 in marinas. The freshwater-influenced stratum had a similar pattern with percentages ranging from 60-90%. Annual comparisons of copper concentrations showed that neither dissolved nor total copper concentrations changed significantly from year to year in either stratum (Figure 3-2). Mean dissolved copper concentrations within marinas consistently exceeded the CTR threshold in all years, reaching a maximum of 8.12 ± 1.24 µg/L in 2006, while those of the freshwater-influenced stratum consistently did not exceed the CTR threshold. Within the marinas, mean total copper levels declined by approximately 50% from 2005 (16.38 ± 4.45 µg/L) to 2007 (8.0 ± 1.21 µg/L), although average values were still in exceedance of the total copper CTR threshold for all years. In freshwater-influenced areas, declines in total copper concentrations were less pronounced; however, the average of 2005 (7.23 ± 2.07 µg/L) exceeded the CTR threshold, while averages of subsequent years were not in exceedance.

Large differences in the percent of area with copper exceedances were observed among harbors (Figure 3-2). All 11 stations in Dana Point Harbor exceeded dissolved and total copper CTR thresholds. At Oceanside Harbor, one station (20%) did not exceed the thresholds. In San Diego Bay, 78% of stations (25 of 32) did not exceed the dissolved copper threshold and 59% of stations (19 of 32) did not exceed the total copper CTR thresholds. In Mission Bay, 83% of stations (10 of 12) did not exceed thresholds. Additionally, dissolved and total copper concentrations were found to be significantly different among harbors (Figure 3-2). Average copper concentrations in the marina stratum exceeded the dissolved and total copper thresholds for Dana Point Harbor, Oceanside Harbor, and Mission Bay, while those of San Diego Bay were not in exceedance. Within the freshwater-influenced stratum, average copper concentrations exceeded CTR thresholds only for Dana Point Harbor, while San Diego Bay and Mission Bay were not in exceedance. Differences in copper concentrations between strata were most evident for San Diego and Mission Bays, while copper concentrations in the marina and freshwater-influenced waters of Dana Point Harbor were much more similar, since freshwater-influenced areas of Dana Point Harbor occurred immediately adjacent to the marinas.

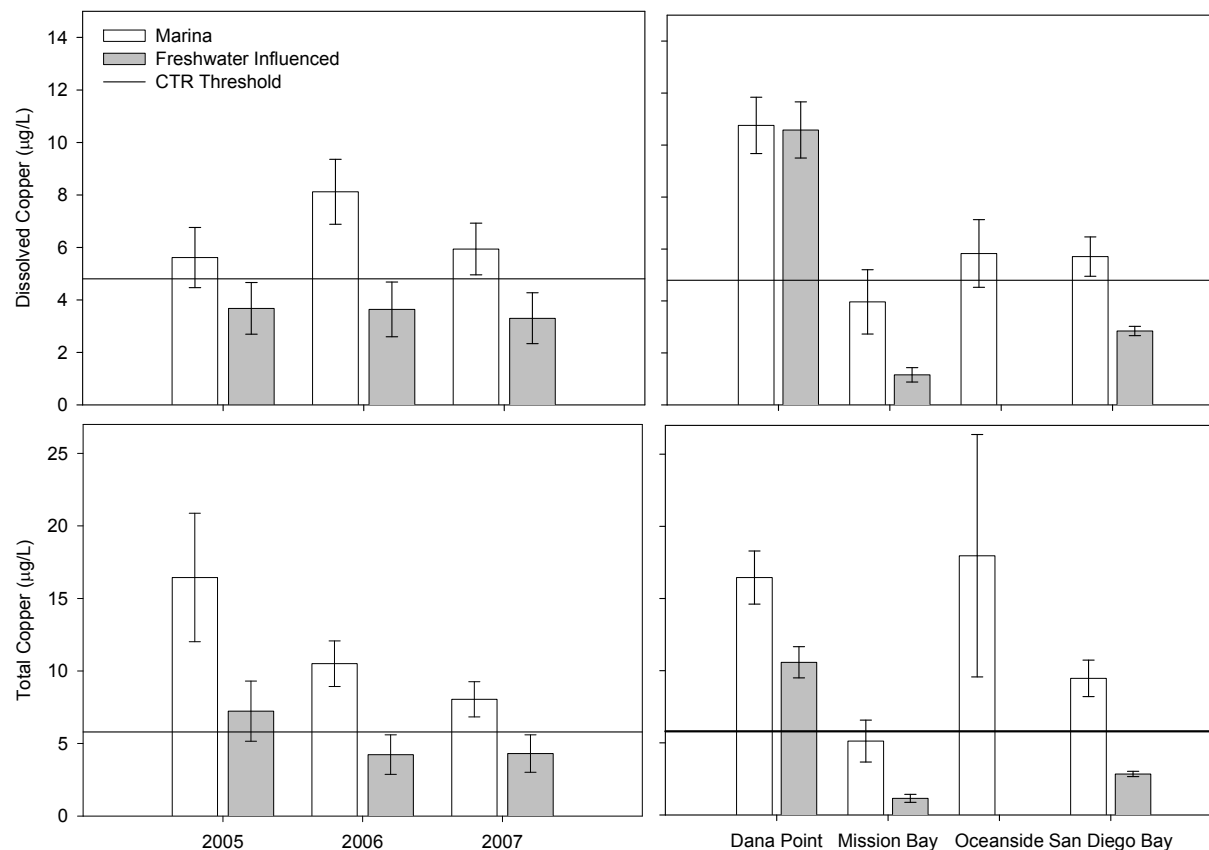


Figure 3-2. Comparison of surface water copper concentrations among years and harbors

Secondary Indicators

Nickel & Zinc

All stations located in both marina and freshwater-influenced strata had concentrations of dissolved and total nickel and dissolved and total zinc that did not exceed CTR and COP thresholds throughout the duration of the Pilot Project, which is largely consistent with historical conditions. Preset target percentages were determined to be 100% for dissolved and total nickel and dissolved zinc (i.e., no historical stations exceeded CTR thresholds) and 97% for total zinc. In accordance, there were no significant differences in the observed Pilot Project percentages for dissolved and total nickel and dissolved zinc from historical harbor-wide conditions; however, the percentage of stations that did not exceed the CTR threshold for total zinc was significantly greater than the preset target, although the difference was only 3% (Table 3-1). Since percentages remained constant at 100%, comparisons of percentages among years or among harbors were unnecessary, although comparisons of concentrations were still informative.

Mean concentrations of dissolved and total nickel were far below the CTR thresholds across all years (Figure 3-3). The mean dissolved nickel concentration within the marina stratum from 2005-2007 was 0.36 ± 0.03 µg/L, while that of the freshwater-influenced stratum was 0.60 ± 0.06 µg/L. In addition, mean total nickel concentrations were 0.41 ± 0.03 µg/L for marina stations and 0.66 ± 0.04 µg/L for freshwater-influenced stations. Dissolved and total nickel concentrations were significantly greater in the freshwater-influenced stratum than the marinas,

although concentrations were only a fraction of their respective CTR threshold values. Furthermore, nickel was the only metal that showed a significant change in concentration with time as total nickel concentrations declined from 2005 to 2007 for both marina and freshwater-influenced stations, while dissolved nickel reached its apex level in 2006 before again declining in 2007. Lastly, dissolved and total nickel concentrations did not differ significantly among harbors.

Mean dissolved and total zinc concentrations were approximately twice as high in the marinas as in the freshwater-influenced areas, resulting in significant differences between strata. The mean dissolved zinc concentration within the marina stratum was 20.12 ± 2.17 $\mu\text{g/L}$, while that of the freshwater-influenced stratum was 7.43 ± 2.45 $\mu\text{g/L}$. In addition, mean total zinc concentrations were 21.40 ± 2.06 $\mu\text{g/L}$ for marina stations and 10.01 ± 2.32 $\mu\text{g/L}$ for freshwater-influenced stations.

As was the case for copper, dissolved and total zinc concentrations were nearly identical within marina and freshwater-influenced strata of Dana Point Harbor, while inter-strata differences were most evident for Mission Bay and San Diego Bay. Dissolved and total zinc concentrations significantly differed among harbors both in the marina and freshwater-influenced strata, since mean zinc concentrations within Dana Point Harbor were at least 1.8 times higher than any other harbor (Figure 3-3). This pattern was even more pronounced within the freshwater-influenced stratum, since mean zinc concentrations in Dana Point Harbor were over six times higher than San Diego Bay and Mission Bay. Even so, the maximum zinc concentration at any Dana Point Harbor station was approximately half the CTR threshold.

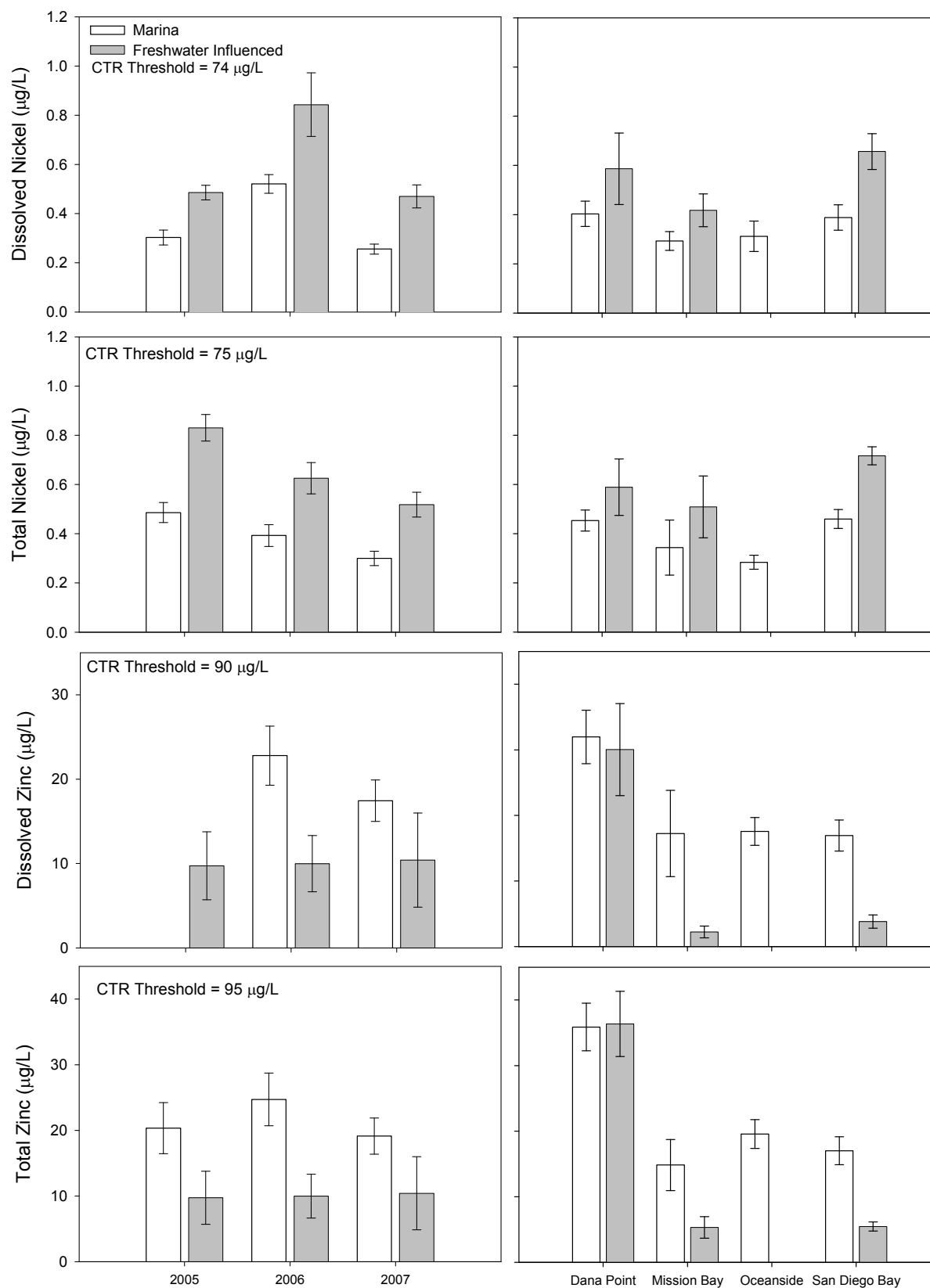


Figure 3-3. Comparisons of surface water nickel and zinc among years and harbors

Total Detectable PAHs

Concentrations of total detectable PAHs were extremely variable, ranging from zero (non-detectable) to 8225.10 ng/L. The single highest concentration was recorded in 2007 at a marina station in Oceanside Harbor, with the next highest measurement of 308.60 ng/L recorded at a San Diego Bay freshwater-influenced station in 2007. The highest value was determined to be an outlier and was excluded from statistical analyses since it would have had an undue influence on the results. Concentrations of total detectable PAHs were not significantly different between strata, since the mean concentration was 35.24 ± 9.74 ng/L in the marina and 32.52 ± 10.11 ng/L in the freshwater-influenced stratum. Additionally, concentrations did not change significantly from year to year. There were significant differences among harbors in both strata (Figure 3-4). Within the marina stratum, mean concentrations of total detectable PAHs in San Diego Bay were over five times greater than Dana Point Harbor and Mission Bay. Within the freshwater-influenced stratum, San Diego Bay also had significantly higher levels than both Dana Point Harbor and Mission Bay.

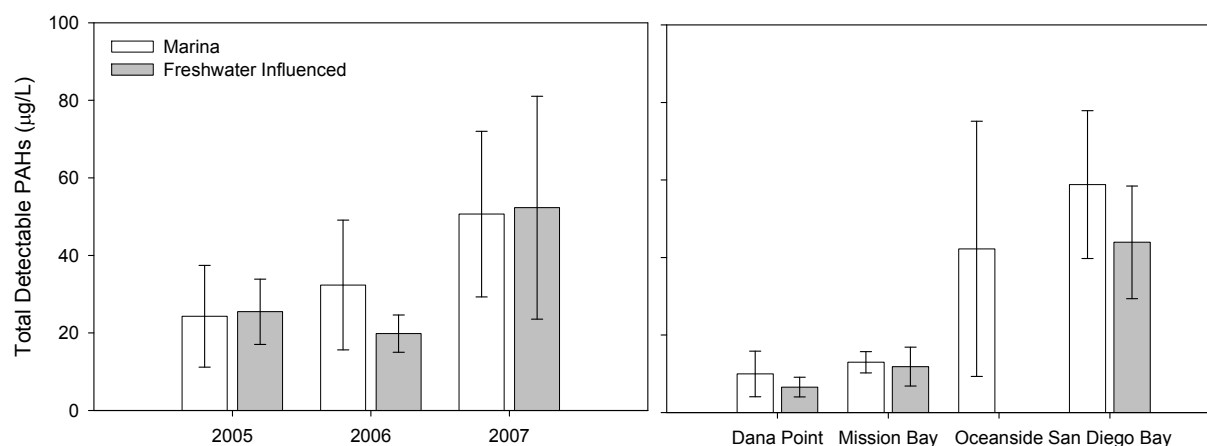


Figure 3-4. Comparison of surface water total detectable PAHs among years and harbors

Dissolved and Total Organic Carbon

Mean concentrations of DOC and TOC for 2006 and 2007 combined were significantly higher in the freshwater-influenced stratum than in the marinas. Mean DOC concentrations were 0.54 ± 0.02 mg/L in the marina stratum and 0.70 ± 0.04 mg/L in the freshwater-influenced stratum, and mean TOC concentrations were 0.64 ± 0.03 mg/L in the marina stratum and 0.76 ± 0.04 mg/L in the freshwater-influenced stratum. Averages and statistical analyses excluded data from 2005 since chemical analyses were performed by a different laboratory than in subsequent years and method detection limits and reporting limits were much higher in 2005 than in 2006 and 2007 (Appendix B Tables B-1 – B-3). Although 2005 DOC and TOC concentrations were not statistically analyzed, results are still reported in figures (Figure 3-5). Concentrations of DOC increased significantly from 2006 to 2007 in the marina and freshwater-influenced strata, and TOC concentrations increased significantly over the same period in the marina stratum (Figure 3-5). Lastly, TOC and DOC concentrations did not differ significantly among harbors in either stratum.

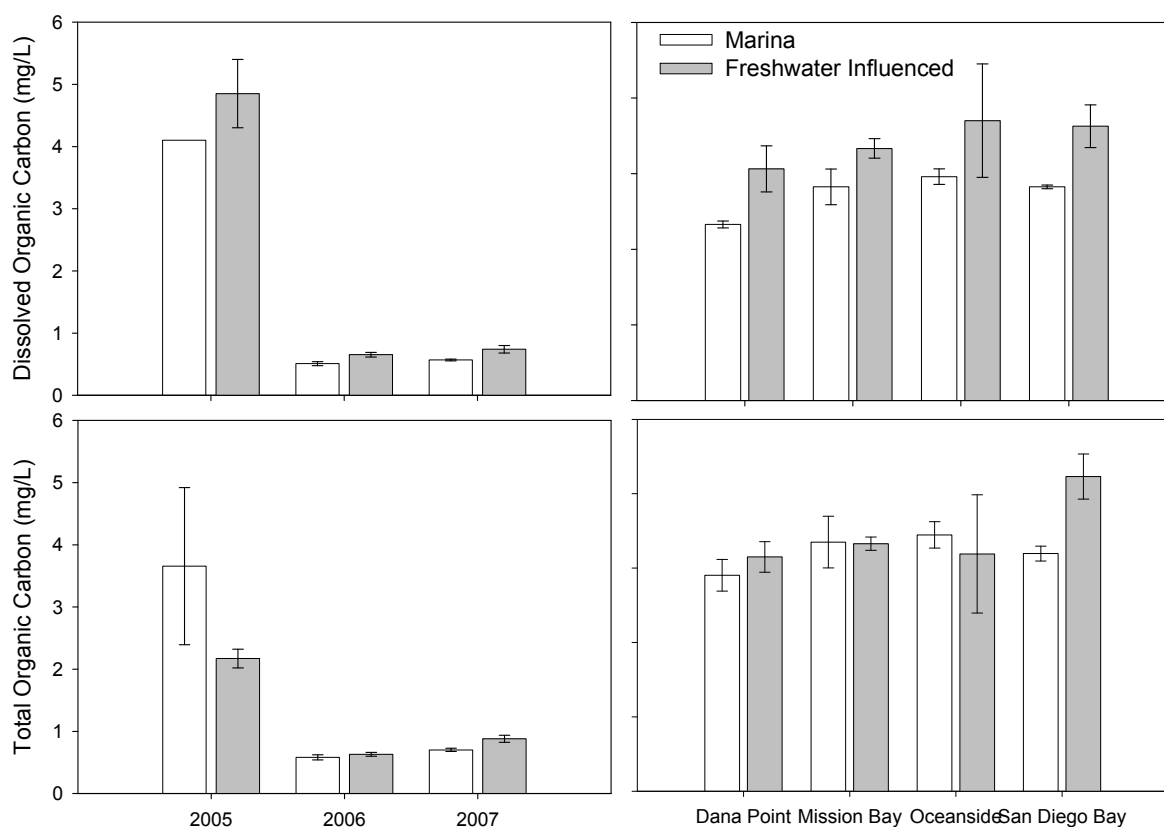


Figure 3-5. Comparisons of surface water dissolved and total organic carbon among years and harbors

3.1.2 Bacteria

Bacterial levels never exceeded the AB411 threshold of 104 MPN/100mL throughout the duration of the Pilot Project. From 2005 to 2007, only two stations had enterococcus counts of 20 MPN/100 mL, nine had counts of 10 MPN/100 mL, and the majority was below the reporting limit of 10 MPN/100 mL (Appendix B, Tables B-1 - B-3). The two stations with the highest enterococcus concentrations occurred within the freshwater-influenced stratum at stations sampled in 2006 in San Diego Bay and Dana Point Harbor, while those stations with enterococcus concentrations of 10 MPN/100 mL were more prevalent in the marina stratum, occurring across all harbors in all years (Table 3-2).

Table 3-2. Percentages of stations at enterococcus concentrations of <10, 10, and 20 MPN/100mL within marina and freshwater-influenced strata

Enterococcus Concentration	Marina				Freshwater Influenced			
	2005	2006	2007	Cumulative	2005	2006	2007	Cumulative
< 10 MPN/ 100mL	70	80	80	77	90	70	100	87
10 MPN/ 100mL	20	20	23	23	0	20	0	7
20 MPN/ 100mL	0	0	0	0	0	10	0	7

3.1.3 Water Column Measurements

Surface water measurements for the 60 stations sampled from 2005-2007 are presented in Appendix B Tables B-1 – B-3, and depth profiles summary data of physical water column measurements are presented in Appendix D. Measurements include temperature, salinity, pH, dissolved oxygen, and transmissivity. These measures, while not being compared to threshold levels, are useful in providing information about water quality that can help explain biological results and determine if the harbor waters can sustain a healthy biota.

Temperature

During August surveys, temperatures did not change greatly with depth since differences between surface and bottom temperatures for individual stations were less than 4 degrees Celsius ($^{\circ}\text{C}$). When surface temperatures were compared between strata across all years, surface water temperatures in the freshwater-influenced stratum ($25.2 \pm 0.5^{\circ}\text{C}$) were significantly higher than in the marina stratum ($21.8 \pm 0.4^{\circ}\text{C}$). In addition, there was also a significant increase in surface water temperatures from 2005 to 2007 as the average temperatures increased by approximately 4°C over this period. Surface water temperatures declined moving from southerly to northerly harbors on average by approximately 3°C , leading to significant differences between San Diego Bay and Dana Point Harbor.

Salinity

Salinity tended to decrease slightly with depth, with a maximum difference of approximately 6 psu. The lowest salinities were encountered in bottom waters, occurring around 28-29 psu in both strata alike. Mean salinity was 33.7 ± 0.1 psu in the marina and 34.5 ± 0.2 psu in the freshwater-influenced stratum. Although there was less than a 0.8-psu difference, salinity within surface waters in the freshwater-influenced stratum was significantly greater than that of the marina stratum. However, this difference was not likely to be biologically important. Similar to temperature, salinity also changed along a gradient from northerly to southerly harbors, with significantly higher salinities in the warmer San Diego Bay than in the cooler Dana Point and Oceanside Harbors.

pH

Hydrogen ion concentrations (pH) did not change more than by 0.1 between surface and bottom waters and were slightly basic in both strata. The average pH was only minimally different between marinas (7.7 ± 0.1) and freshwater-influenced areas (7.9 ± 0.0). Additionally, pH levels remained at relatively consistent levels from year to year and did not change markedly among harbors.

Dissolved Oxygen

Dissolved oxygen (DO) concentrations tended to decrease slightly with depth for most stations, although in a specific case concentrations ranged from 6.4 to 1.8 mg/L. Dissolved oxygen concentrations in the marinas (6.2 ± 0.2 mg/L) were on average nearly identical to those of the freshwater-influenced stratum (6.4 ± 0.2 mg/L) when compared across all years. However, there were instances where DO occurred at extremely low levels around 0.1 mg/L during a 2005 red tide event in Oceanside Harbor.

Transmissivity

Average values of surface water light transmittance were nearly indistinguishable between the marina and freshwater-influenced strata, with means equal to $65.1 \pm 2.5\%$ and $67.8 \pm 2.2\%$, respectively. However, in the marina stratum, transmissivity commonly decreased with depth, with declines ranging from 9% to 39%, while in the freshwater-influenced stratum, transmissivity remained relatively constant with increasing depth and in several instances it increased; the only exceptions being the freshwater-influenced stations of Dana Point, which showed declines in transmissivity similar to marina stations.

3.2 Sediment Analysis

3.2.1 Chemistry

Sediment samples collected from marina and freshwater-influenced stations were chemically analyzed to determine concentrations of metals and total detectable PAHs, from which mean ER-M quotients were calculated. In addition, sediment samples were analyzed for TOC and grain size.

Sediment chemistry results for primary and secondary indicators at all stations surveyed throughout the Pilot Project are reported in Tables B-4 – B-6, cumulative distribution curves used to determine percentages of stations that did not exceed threshold levels for secondary indicators are presented in Figures B-3 – B-4, and statistical significance values are provided in Tables B-12 – B-13 in Appendix B. Additionally, sediment chemistry results for 2007 indicators are provided in Appendix C, and those of previous years are reported in 2005-2006 and 2006-2007 RHMP Pilot Project Reports (Weston, 2006 and Weston, 2007). All indicator metals exceeded their ER-L values at least at one station, although only copper and zinc exceeded their respective ER-M values. Mean ER-M quotients also exceeded the 0.2 threshold, although the percentage of stations with ER-M quotients that did not exceed the threshold was significantly lower in the marina stratum than the freshwater-influenced stratum (Table 3-3). As was the case for surface water indicators, there were significant differences in both primary and secondary indicators between strata (mean ER-M quotient and copper), among years (cadmium, chromium, and nickel), and among harbors (mean ER-M quotient, chromium, copper, nickel, and TOC), as discussed in detail in the following sections.

Table 3-3. Percentages of stations below ER-L thresholds for marina and freshwater-influenced sediment indicators

Indicator	Preset Target	Marina				Freshwater Influenced			
		2005	2006	2007	Cumulative	2005	2006	2007	Cumulative
mean ER-M quotient	48	20*	40	40	33*	40	60	70*	60*
cadmium	90	100*	100*	100*	100*	90	100*	100*	97*
chromium	78	100*	90	100*	93*	100*	100*	100*	100*
copper ¹	68	40	50	50	47*	80	80	90	83*
lead	74	90*	90*	70	83	80	90*	100*	90*
nickel	80	80	100*	90	87*	90	100*	100*	97*
zinc	45	40	40	40	40	40	40	40	40
total detectable PAHs	74	100*	100*	100*	100*	100*	100*	100*	100*

¹ Reference ambient value for copper was not based on the ER-L, as described in Section 2.3.2.

* Indicates result that is significantly different from preset target ($p \leq 0.1$); green indicates a higher percentage and yellow a lower percentage.

Primary Indicator

Mean ER-M Quotient

The mean ER-M quotient is one of the primary indicators of sediment quality for the Pilot Project. Samples with ER-M quotients above 0.2 (threshold) were more likely to have adverse benthic effects. Historically, 48 % of stations did not exceed the mean ER-M quotient threshold of 0.2. Within the freshwater-influenced stratum, 57% of the area did not exceed the mean ER-M quotient threshold, while only 33% of the marina stratum did not exceed the threshold (Figure 3-6). Accordingly, there were significantly fewer exceedances in the freshwater-influenced stratum than historically encountered throughout the harbors, while a significantly higher percentage of the marina stratum exceeded the threshold (Table 3-3), indicating that sediment chemistry conditions in the freshwater-influenced stratum appeared to be better than the historical conditions, while those of the marina stratum were worse.

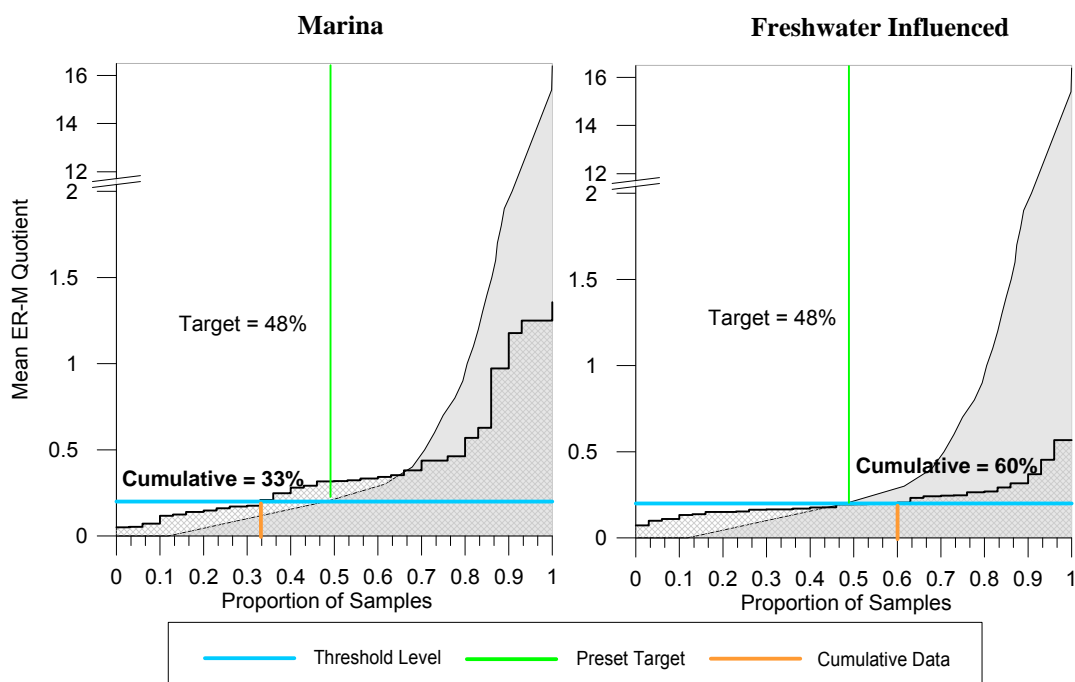


Figure 3-6. Cumulative distribution curves for sediment mean ER-M quotients

The percentage of the freshwater-influenced stratum that did not exceed the mean ER-M quotient threshold was 1.7 times higher than that of the marina stratum. Additionally, the three-year average mean ER-M quotient within the freshwater-influenced stratum was 0.22 ± 0.02 , and the average for the marina stratum was 0.40 ± 0.06 . Therefore, the mean ER-M quotient of the freshwater-influenced stratum was on average very close to the threshold for adverse effects, while the average for the marina stratum was approximately twice the threshold. In accordance, mean ER-M quotient values were significantly higher in the marina than the freshwater-influenced stratum.

For both strata there was a general trend in which the percentage of stations that did not exceed the threshold increased with time. Percentages doubled from 20% in 2005 to 40% in 2006 and

2007 in the marina stratum, and increased from 40% in 2005 to 60% in 2006 and 70% in 2007 for the freshwater-influenced stratum (Table 3-3). In the marina stratum, mean ER-M quotients became more variable as they increased from 2005 into 2006 and 2007, while in the freshwater-influenced areas, average values decreased from 0.29 ± 0.04 in 2005 to 0.18 ± 0.02 in 2007 (Figure 3-7); however, interannual differences did not meet the criteria for significance in either strata.

The percentages of strata below the threshold differed greatly among harbors, with 27% of Dana Point, 40% of Oceanside Harbor, 53% of San Diego Bay, and 75% of Mission Bay stations below. Within the marina stratum, average mean ER-M quotients were typically at or above the threshold of 0.2 for each harbor, and did not differ significantly among harbors (Figure 3-7). In the freshwater-influenced stratum, mean ER-M quotients significantly differed among harbors due to differences between Dana Point Harbor, where the average was above the threshold, and Mission Bay, where the average was below the threshold.

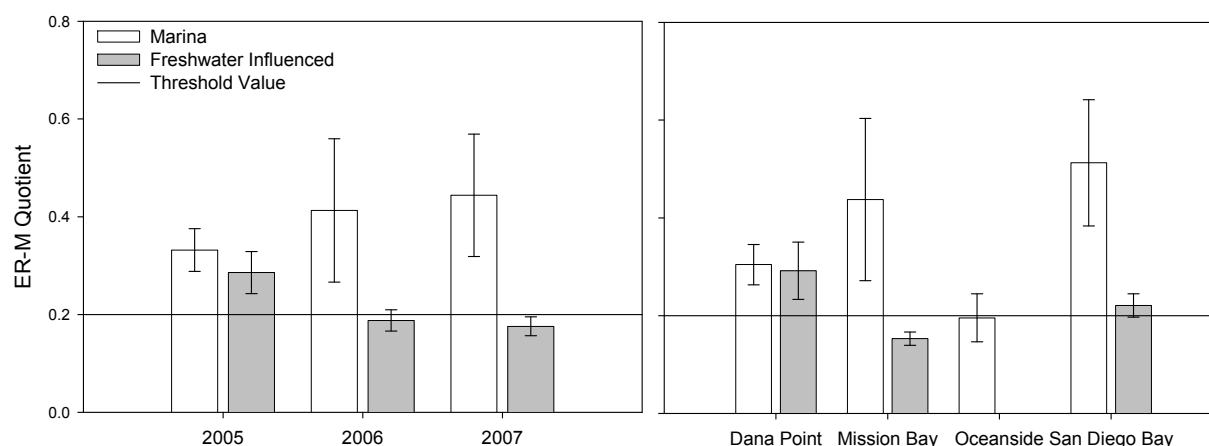


Figure 3-7. Comparisons of sediment average mean ER-M quotients among years and harbors

Secondary Indicators

Six metals (cadmium, chromium, copper, lead, nickel, and zinc) and total detectable PAHs were identified as secondary indicators of sediment chemistry conditions since concentrations of these analytes were used to calculate mean ER-M quotients in addition to three other metals (arsenic, mercury, and silver). Of the 60 stations sampled in the three-year Pilot Project, ER-M exceedances were observed only for copper (12 stations) and zinc (2). Further analyses comparing differences from historical conditions (i.e., preset targets), between strata, among years, and among harbors are provided below, with particular attention paid to copper and zinc due to the higher numbers of exceedances.

Copper

Concentrations of copper within marina and freshwater-influenced sediments rarely did not exceed the ER-L value of 34 mg/kg and occasionally exceeded the ER-M value of 270 mg/kg in all four harbors. Only one marina station (i.e., 3% of the stratum) and four freshwater-influenced stations (13%) had copper concentrations that did not exceed the ER-L. In addition, 30% of

marina stations exceeded the ER-M, while only 10% of freshwater-influenced stations did. In Dana Point Harbor, all stations exceeded the ER-L value and 5 stations exceeded the ER-M. Additionally, all stations in Oceanside Harbor had sediment copper concentrations that exceeded the ER-L and 2 stations exceeded the ER-M. In Mission Bay, three stations (25%) did not exceed the ER-L and only 1 exceeded the ER-M. Lastly, two San Diego Bay stations (6%) did not exceed the ER-L, while four stations exceeded the ER-M. Although copper concentrations commonly exceeded the ER-L across all harbors and in all years, elevated copper levels are partly due to high natural levels rather than entirely being due to anthropogenic influences. As a consequence, the threshold was set at 175 mg/kg, as described in Section 2.3.2.

Historically 68% (preset target) of samples did not exceed the threshold of 175 mg/kg. Throughout the three-year Pilot Project, 47% of the marinas and 83% of the freshwater-influenced stratum did not exceed the threshold. In accordance, the percentage of marina stations that exceeded threshold levels was significantly lower than the preset target, indicating that copper levels are worse in the marina stratum than historic conditions. Alternatively, the percentage of freshwater-influenced stations below the threshold was significantly greater than the preset target, indicating that conditions in this stratum were better than historic conditions (Table 3-3).

The percent of stations that did not exceed the copper threshold in the freshwater-influenced stratum was 1.76 times that of the marina stratum, resulting in a significant difference between the two strata (Table 3-3). Similarly, mean copper concentrations were 2.4 times higher in the marina stratum (285.8 ± 87.7 mg/kg) than the freshwater-influenced stratum (119.1 ± 18.7 mg/kg) when averaged over the three-year study, also resulting in a significant difference.

Percentages of stations below the copper reference ambient value changed minimally in both strata from year to year (Table 3-3). In the marina stratum, observed percentages were always below the preset target, ranging from 40% in 2005 to 50% in 2006 and 2007. In the freshwater-influenced stratum, percentages were always above the threshold, ranging from 80% in 2005 and 2006 to 90% in 2007. Although the mean copper concentration in 2006 in the marina stratum was approximately twice that of the other years, differences were not significant due to high levels of variability among stations (Figure 3-8).

The percentages of strata below the copper threshold differed greatly among harbors, with 18% of Dana Point Harbor, 20% of Oceanside Harbor, 75% of San Diego Bay, and 83% of Mission Bay below 175 mg/kg threshold when analyzed for both the marina and freshwater-influenced strata combined. Within the marina stratum, the average sediment copper concentration in Dana Point Harbor, Oceanside Harbor, and San Diego Bay exceeded the threshold, while Mission Bay did not exceed the threshold; however, differences were not significant. For the freshwater-influenced stratum, average sediment copper concentrations exceeded the threshold in Dana Point Harbor, but were below the threshold in Mission Bay and San Diego Bay. Mean copper concentrations at Dana Point were over twice that of San Diego Bay and concentrations at San Diego Bay were over three times that of Mission Bay, resulting in significant differences among harbors.

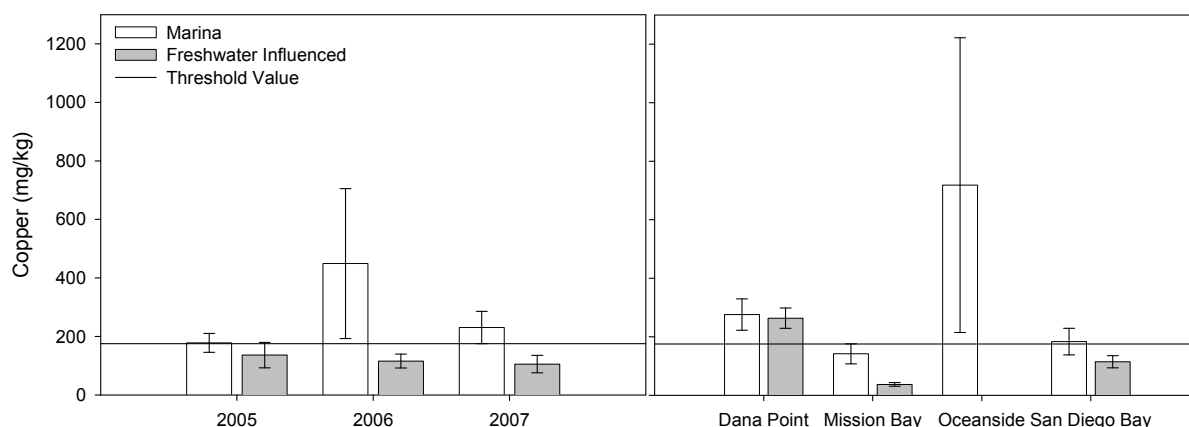


Figure 3-8. Comparisons of sediment copper concentrations among years and harbors

Zinc

Concentrations of zinc within marina and freshwater-influenced sediments exceeded the ER-L value of 150 mg/kg in all four harbors and exceeded the ER-M value of 410 mg/kg at only one station in Dana Point Harbor. Historically, 45% (preset target) of samples did not exceed the Zinc ER-L threshold. Throughout the duration of the Pilot Project, 40% of marina and 40% of freshwater-influenced strata had concentrations that did not exceed the Zinc ER-L. Although both strata had lower percentages of stations that did not exceed the threshold than the preset target, differences were not significant (Table 3-3). Consequently, Zinc levels in the two strata appear to be consistent with historically-observed levels.

Zinc concentrations ranged from 31.0-410.1 mg/kg in the marinas and from 55.4-555.0 mg/kg in the freshwater-influenced stratum. In both strata, 40% of stations had concentrations of zinc that did not exceed the ER-L. The average marina concentration (202.5 ± 19.4 mg/kg) was higher than the average freshwater-influenced concentration (187.0 ± 17.0 mg/kg); however, due to relatively high variability, the difference in concentrations was not significant.

The percentage of stations that did not exceed the ER-L did not change at all from year to year in either stratum, remaining constant at 40%. Although mean zinc concentrations were always in exceedance of the ER-L, concentrations in both strata changed more than the constant percentages would indicate (Figure 3-9). Zinc concentrations in marina sediments increased from 2005 to 2007, while those of the freshwater-influenced sediments tended to decrease; however, differences were not significant due to the high degree of variability within strata.

Eighteen percent of Dana Point Harbor, 20% of Oceanside Harbor, 44% of San Diego Bay, and 68% of Mission Bay marina and freshwater-influenced strata did not exceed the ER-L. Mean zinc concentrations at least slightly exceeded the ER-L in both strata across all harbors (Figure 3-9), and although mean concentrations ranged from 163.9 ± 24.1 mg/kg in Mission Bay to 248.2 ± 44.7 mg/kg in Dana Point Harbor, the differences were not significant.

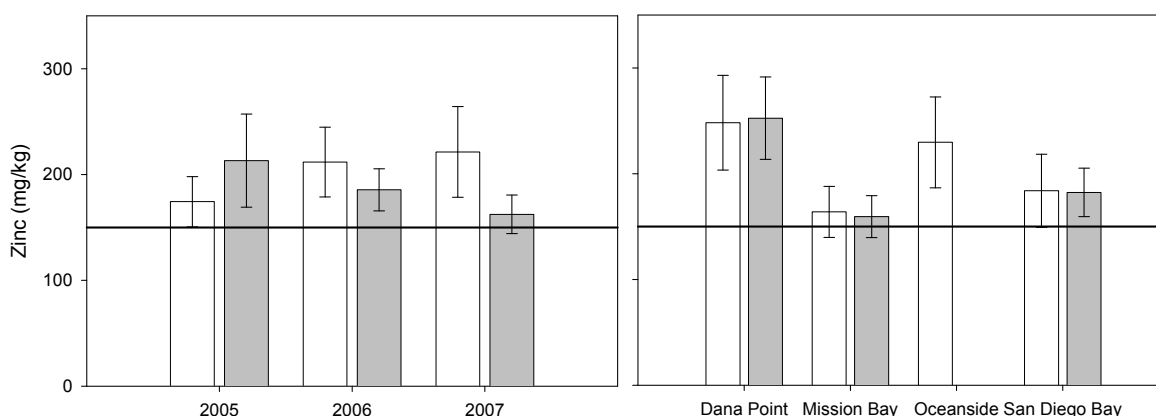


Figure 3-9. Comparisons of sediment zinc concentrations among years and harbors

Cadmium, Chromium, Lead, and Nickel

For the four secondary indicators (cadmium, chromium, lead, and nickel), the majority of stations had concentrations that did not exceed their respective ER-Ls throughout the duration of the Pilot Project. Additionally, percentages of stations that did not exceed ER-Ls were significantly greater than historically-derived preset targets for cadmium, chromium, and nickel in both marina and freshwater-influenced strata and for lead in the freshwater-influenced stratum (Table 3-3).

There were no pronounced differences between the marina and freshwater-influenced strata in the percentage of stations that did not exceed ER-Ls for any of the four metals (Table 3-3). Additionally, concentrations did not significantly differ between strata either (Figure 3-10).

Changes in the percentages of stations that did not exceed ER-Ls from year to year were generally small (i.e., less than 20 %) in both strata. Percentages of stations that did not exceed ER-Ls increased slightly for cadmium, nickel, and lead in the freshwater-influenced stratum. It was only in the marina stratum that percentage of stations that did not exceed the lead ER-L declined from 90% to 70% (Table 3-3). In no case was there a significant increase in the concentrations of these metals from 2005 to 2007 (Figure 3-10). In contrast, concentrations of cadmium, chromium, and nickel all significantly declined in the freshwater-influenced stratum over this period.

Differences in the percentage of stations that did not exceed ER-Ls among harbors were most evident for lead and nickel, respectively. Lead ER-L exceedances only occurred within San Diego Bay, although 67% of the marina and 80% of the freshwater-influenced stations had concentrations that did not exceed the ER-L. Nickel exceedances occurred predominantly in Dana Point Harbor, although the majority of marina (71%) and freshwater-influenced (75%) stations did not exceed the ER-L. There was also another nickel exceedance in Oceanside Harbor, as well as one cadmium exceedance in Dana Point Harbor and one chromium exceedance in San Diego Bay throughout the three-year study. Although exceedances of lead only occurred in San Diego Bay, there were no significant differences in lead concentrations among harbors. However, there were significant differences in chromium and nickel concentrations among harbors. Chromium concentrations were significantly higher in San Diego

Bay than Mission Bay for the freshwater-influenced stratum, and nickel concentrations were significantly higher in Dana Point Harbor than San Diego Bay and Mission Bay for the marina stratum.

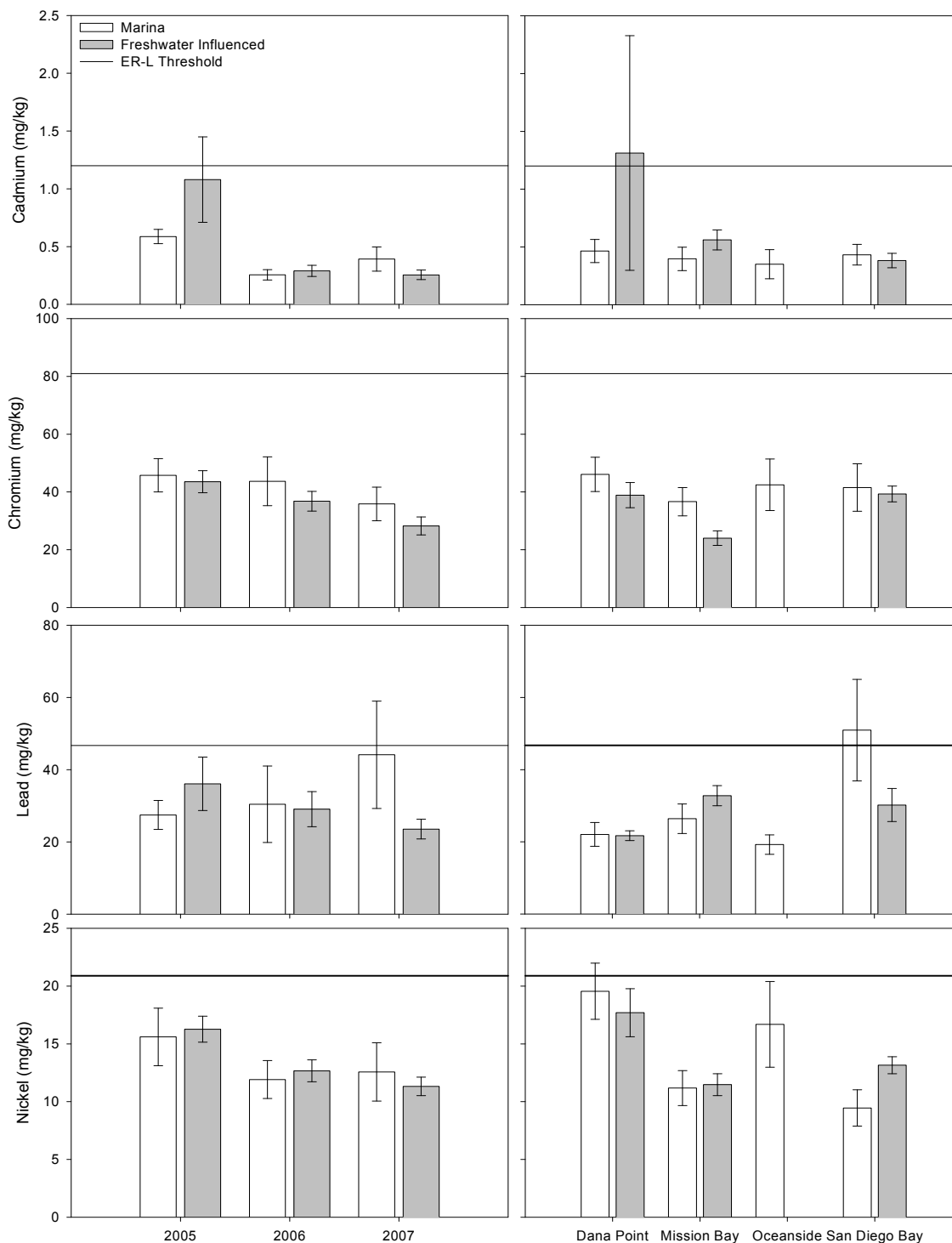


Figure 3-10. Comparison of sediment cadmium, chromium, lead, and nickel concentrations among years and harbors

Total Detectable PAHs

Concentrations of total detectable PAHs in sediments ranged from 43 to 3975 $\mu\text{g/kg}$, with no stations exceeding the ER-L of 4022 $\mu\text{g/kg}$ throughout the entire Pilot Project. Historically, 74% of samples had total detectable PAH levels that did not exceed the ER-L, while 100 % of the marina and freshwater-influenced stations did not exceed the ER-L. Accordingly, percentages for both strata were significantly greater than the preset target, indicating that conditions in the two strata are better than historical conditions. Since proportions did not change between strata, among years, or among harbors, no further analysis for proportions are provided.

The mean concentration of total detectable PAHs was 550 ± 154 $\mu\text{g/kg}$ within the marina stratum and 708 ± 141 $\mu\text{g/kg}$ in the freshwater-influenced stratum; however, the difference between the two strata was not significant. Furthermore, total detectable PAHs did not differ significantly among years or among harbors for either stratum due to the high degree of variability in concentrations among stations. Variability in total detectable PAHs increased from 2005 into subsequent years (Figure 3-11), with concentrations ranging from 55-1246 $\mu\text{g/kg}$ in 2005, 90-3975 $\mu\text{g/kg}$ in 2006, and 43-3489 $\mu\text{g/kg}$ in 2007. A similar pattern was observed when PAHs were compared among the harbors, since harbors with the highest average levels also had the widest range of PAHs.

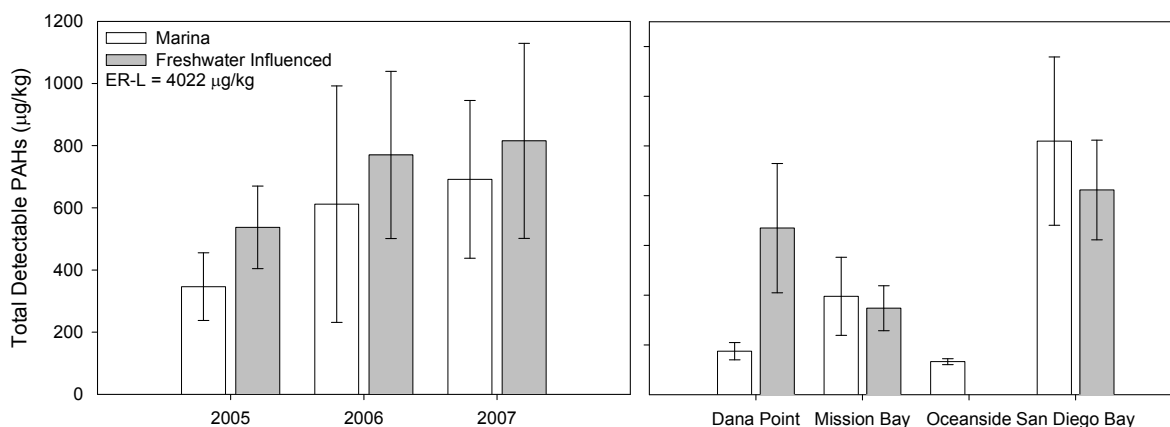


Figure 3-11. Comparisons of sediment total detectable PAHs among years and harbors

3.2.2 Toxicity

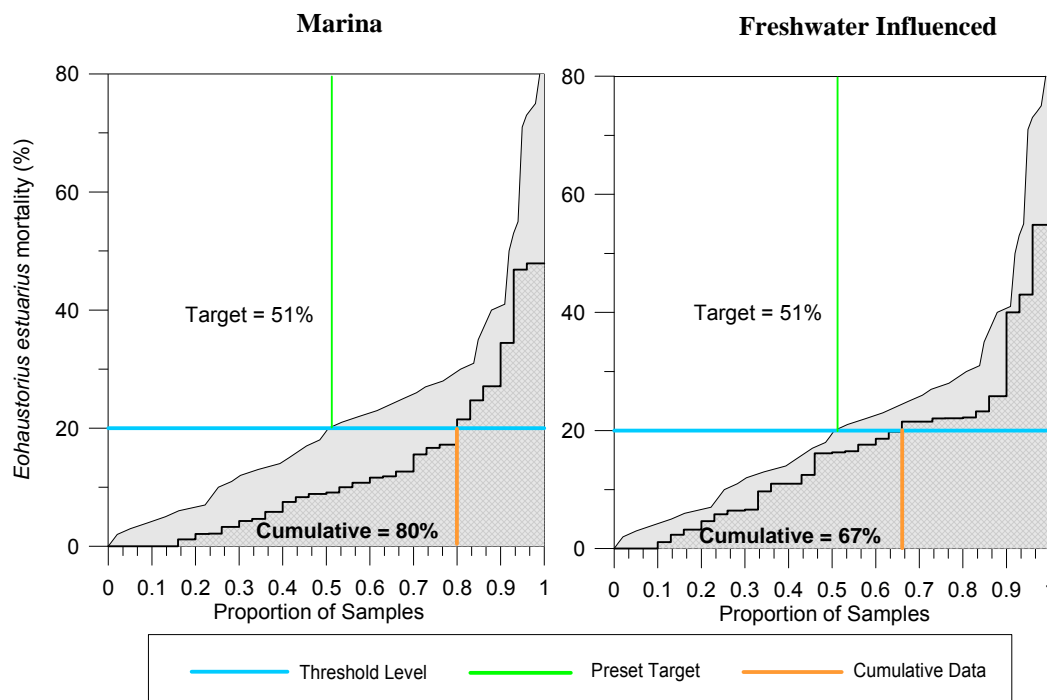
Toxicity was measured as the control-adjusted percent mortality for *E. estuarius* conducted with marina and freshwater-influenced sediments. The results of the sediment toxicity tests for all stations are presented in Table B-4 – B-6 and significance values of statistical tests Tables B-12 – B-13 in Appendix B. Cumulative distribution curves from which percentages below reference ambient values were determined are shown in Figure 3-12. Additional supporting data are provided in Appendix E.

Table 3-4. Percentage of stations with toxicity below the threshold value

Indicator	Preset Target	Marina				Freshwater Influenced			
		2005	2006	2007	Cumulative	2005	2006	2007	Cumulative
<i>E. estuarius</i> mortality	51	80*	100*	60	80*	40	80*	80*	67*

* Indicates result that is significantly different from preset target ($p \leq 0.1$); green indicates a higher percentage and yellow a lower percentage than preset target.

Historically, 51 % of stations had toxicity levels that did not exceed the 20% *E. estuarius* mortality threshold. Toxicity ranged from 0% to 55% in the marina sediments and from 0% to 47.9% in the freshwater-influenced sediments throughout the Pilot Project. Within the marina stratum, 80% of stations had percent mortality levels that did not exceed the threshold value and 67% of freshwater-influenced stations were not in exceedance (Figure 3-12). Both strata had significantly higher percentages of stations that did not exceed the threshold than the preset target of 51% (Table 3-4), indicating that toxicity levels in the two strata were lower than historically observed throughout the harbors and bays.

**Figure 3-12. Cumulative distribution curves for *E. estuarius* mortality**

There were significantly more marina stations with toxicity levels that did not exceed the threshold than freshwater-influenced stations. The mean level of toxicity within the marina stratum was $12.2 \pm 2.37\%$, while that of the freshwater-influenced stratum was $15.88 \pm 2.41\%$ over the three-year study. Although mortality was on average slightly higher in the freshwater-influenced stratum, the difference was not significant.

The percentage of marina stations that did not exceed the 20% mortality threshold ranged from 60% in 2007 to 100% in 2006, with 2005 being intermediate between the two at 80% for the

marina stratum. Within the freshwater-influenced stratum, percentages that did not exceed the threshold ranged from 40% in 2005 to 80% in 2006 and 2007. On average, the lowest levels of toxicity were observed in 2006 for both strata (Figure 3-13). As a result, toxicity was significantly lower in 2006 than 2007 in the marina stratum and significantly lower in 2006 than 2005 and 2007 in the freshwater-influenced stratum.

The majority of all harbors had toxicity levels that did not exceed the threshold, with 60% of Oceanside Harbor, 64% of Dana Point Harbor, 75% of Mission Bay, and 83% of San Diego Bay stations below the reference ambient value. Since toxicity levels for the harbors and bays were highly variable, often ranging from 0 to >20%, differences in toxicity levels among harbors were not significant for either stratum. Higher mean levels of toxicity were accompanied by higher levels of variability, as seen in the freshwater-influenced stratum of Dana Point Harbor and Mission Bay and the marina stratum of San Diego Bay (Figure 3-13).

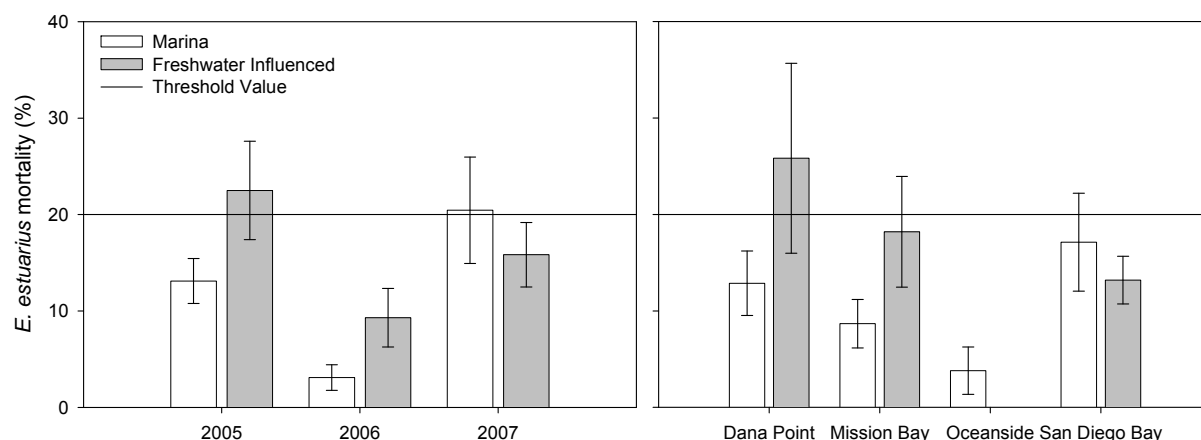


Figure 3-13. Comparisons of toxicity as measured by *E. estuarius* mortality among years and harbors

One of the factors that may be contributing to *E. estuarius* mortality, and by inference toxicity, is sediment grain size. At smaller grain sizes (i.e., higher percentages of clays) mortality of *E. estuarius* was higher. Consequently, there was a significant relationship between toxicity and percent clay (Figure 3-14).

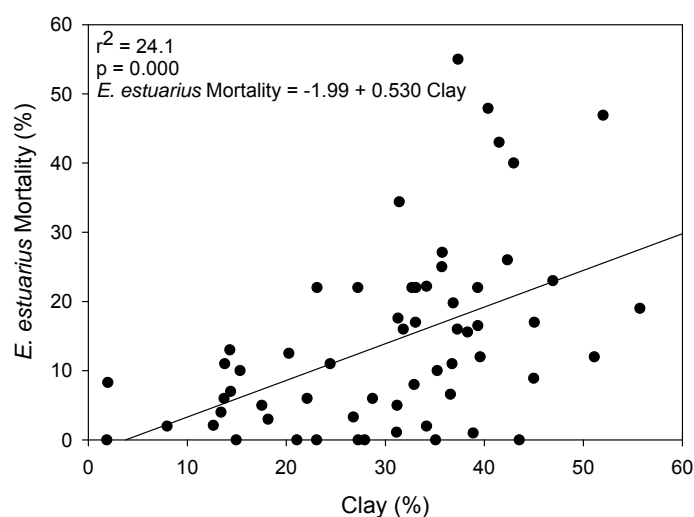


Figure 3-14. Relationship between *E. estuarius* mortality and percentage of clay in sediments

3.2.3 Benthic Infauna

Benthic infaunal samples were collected and analyzed to determine the relative health of the benthic community. The primary indicator of benthic community status was the BRI, while the Shannon-Wiener diversity index and number of taxa were used as secondary indicators. Primary and secondary indicator values for all stations assessed over the three-year Pilot Project are provided in Appendix B Tables B-7 – B-9, as are the secondary indicator cumulative distribution curves depicting percentages of stations with biological conditions better than the thresholds (Figure B-6). For the BRI, lower values are indicative of a less disturbed benthic community, while for the secondary indicators (Shannon-Wiener diversity and number of taxa) lower values are associated with more disturbed benthic communities. Additionally, species names and abundances for each taxon encountered in 2007 surveys are provided in Appendix F, and those of prior surveys are reported in the 2005-2006 and 2006-2007 RHMP Pilot Project Reports (Weston, 2006 and Weston, 2007).

Table 3-5. Percentage of stations with benthic infaunal community measures better than reference ambient values

Indicator	Preset Target	Marina				Freshwater Influenced			
		2005	2006	2007	Cumulative	2005	2006	2007	Cumulative
BRI	37	20	30	10*	20*	10	0	10*	7*
Shannon-Wiener ¹	90	60*	80	50*	63*	80	50*	60*	63*
Number of Taxa ¹	92	50*	70*	50*	60*	60*	70*	80	70*

* Indicates result that is significantly different from preset target ($p \leq 0.1$). Statistical comparisons were made for cumulative values only; green indicates a higher percentage and yellow a lower percentage than preset target.

¹ Reported as percentage of stations ABOVE the reference ambient value.

Primary Indicator

Benthic Response Index

Historically, 37% of stations were categorized as having reference benthic infaunal assemblages (i.e., 37% of stations did not exceed the threshold) (Figure 3-15). Throughout the Pilot Project, 20% of marina and 7% of freshwater-influenced stations had BRI scores within the reference level. BRI values ranged from 17.9-64.6 in the marina stratum and from 29.1-61.0 in the freshwater-influenced stratum, with characterization values of both marina and freshwater-influenced stations ranging from the reference condition (response level 0) to community function loss (response level 3). Consequently, both strata had significantly lower percentages of stations within the reference condition (i.e., did not exceed the threshold) than the preset target (Table 3-5), indicating that the benthic communities in the marina and freshwater-influenced strata were more disturbed than historically-observed conditions throughout the harbors.

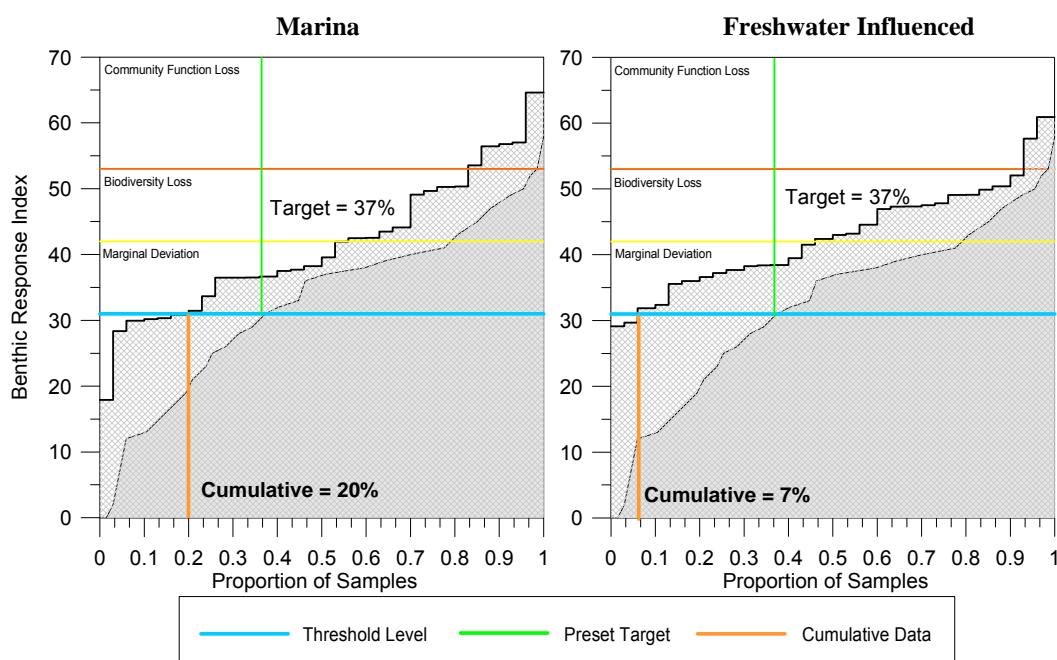


Figure 3-15. Cumulative distribution curves for the Benthic Response Index

Stations with BRI values indicative of a reference condition were 2.8 times more prevalent in the marina than in the freshwater-influenced stratum, resulting in a significant difference. However, the three-year average BRI value in the marina (41.1 ± 1.9) was not significantly different from that of the freshwater-influenced stratum (42.7 ± 1.4). In both strata the average BRI values were consistent with a benthic community characterized as intermediate between response level 1 (marginal deviation) and level 2 (biodiversity loss); however, the marina had a higher percentage of area (53%) below the threshold of biodiversity loss (response level 2) than did the freshwater-influenced stratum (43%) (Table 3-6).

Table 3-6. Percentages of marina and freshwater-influenced strata classified by the BRI from reference to response level 4

Harbor	Reference (%)	Level 1 Marginal Deviation (%)	Level 2 Biodiversity Loss (%)	Level 3 Community Function Loss (%)	Level 4 Defaunation (%)
Marina	23.3	30	30	16.7	0
Freshwater Influenced	6.7	36.7	50	6.7	0

Percentages of stations with BRI values below the reference threshold ranged from 10-30% in the marina and from 0-10% in the freshwater stratum from 2005-2007, showing no clear trend in benthic community status changes over time. Additionally, BRI values did not change significantly from year to year over the Pilot Project, since mean scores generally vacillated around 40 (i.e., marginal deviation) in both strata (Figure 3-16).

In all harbors, the majority of stations exceeded the reference threshold of 31. In Dana Point Harbor, all stations were categorized as having biodiversity loss or community function loss (i.e., response levels 2 and 3). Oceanside Harbor stations were primarily classified as having marginal deviation, as well as biodiversity loss. In Mission Bay and San Diego Bay, benthic infaunal community assemblages at stations ranged from a reference condition to community function loss. In accordance, BRI values significantly differed among harbors, with Dana Point having significantly higher BRI values than San Diego Bay (Figure 3-16).

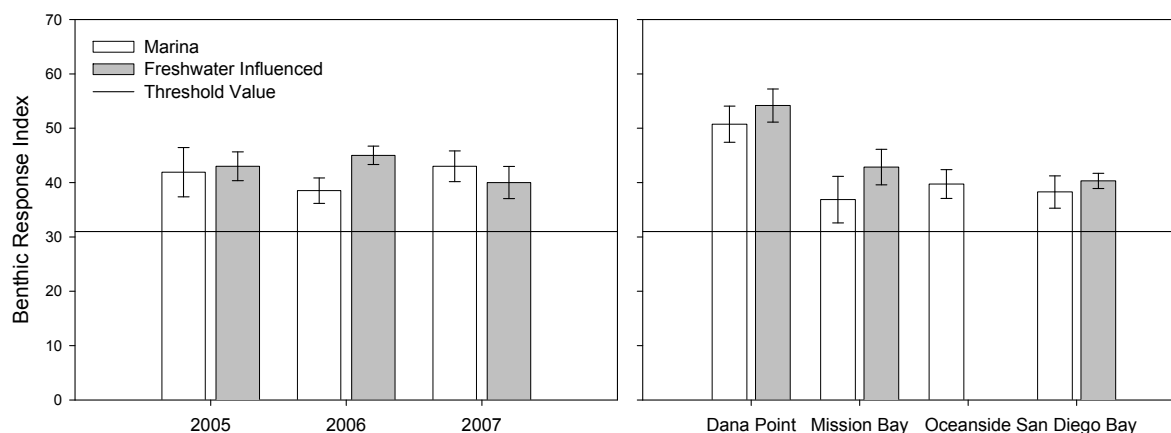


Figure 3-16. Comparisons of average Benthic Response Index values among years and harbors

Secondary Indicators

Shannon-Wiener Diversity Index and Number of Taxa

The Shannon-Wiener diversity index and number of taxa were used as secondary indicators of benthic infauna community condition. Since both indicators are a measure of diversity, higher values are indicative of healthier benthic infaunal communities; therefore, percentages of stations above thresholds of 2 for Shannon-Wiener diversity and 24 for number of taxa are reported rather than the percent below as is the case for all other indicators. Historically, at least 90% of

stations had Shannon-Wiener diversity values and numbers of taxa above respective thresholds. During the Pilot Project, 63% of stations had Shannon-Wiener diversity values above the threshold in both strata, and 60% of marina and 70% of freshwater-influenced stations had greater than 24 taxa. In both strata, the percentages of stations in exceedance of thresholds for the two secondary indicators were significantly lower than the preset targets (Table 3-5), indicating that benthic community conditions within the marina and freshwater-influenced strata were less diverse and more disturbed than communities historically encountered throughout the harbors and bays.

Within the marina stratum, Shannon-Wiener diversity values ranged from 1.21-3.25 and numbers of taxa ranged from 4-63, while in the freshwater-influenced stratum values ranged from 0.37-3.09 and 4-63, respectively. Percentages of stations above thresholds for the two secondary indicators did not differ significantly between strata. Additionally, mean Shannon-Wiener diversity (2.24 ± 0.10 for the marina and 2.14 ± 0.11 for the freshwater-influenced strata) and mean numbers of taxa (29.0 ± 3.2 for the marina and 30.9 ± 2.2 for freshwater-influenced strata) were nearly equivalent for both strata throughout the three-year study (Figure 3-17).

Neither measure of diversity experienced pronounced changes from year to year in the Pilot Project. Within the marina stratum, percentages of stations above the Shannon-Wiener and number of taxa above thresholds increased in synchrony from 2005 to 2008 before declining again in 2007. In the freshwater-influenced stratum, percentages of stations above the Shannon-Wiener diversity threshold were lower in 2006 and 2007 than 2005, while the percentages of stations above the threshold for number of taxa increased every year (Table 3-5). Changes in percentages were the result of minor fluctuations in diversity values about the threshold, since means for both Shannon-Wiener diversity and numbers of taxa remained relatively constant from year to year, hovering just above the thresholds (Figure 3-17). As a consequence, there were no significant differences among years for the secondary indicators in either stratum.

There were substantial differences in the percentages of stations above thresholds among the harbors. For Shannon-Wiener diversity, 18% of Dana Point Harbor, 60% of Oceanside Harbor, 73% of Mission Bay, and 78% of San Diego Bay stations were above the threshold. Similarly, 36% of Dana Point Harbor, 60% of Oceanside Harbor, 78% San Diego Bay, and 92% of Mission Bay stations had at least 24 taxa. Mean Shannon-Wiener diversity was lower than the threshold for Dana Point Harbor, while the mean measures of diversity of all other harbors were above the threshold (Figure 3-17). A similar pattern was observed when mean numbers of taxa were compared, with both Dana Point and Oceanside Harbors having mean numbers below the threshold, while those of Mission and San Diego Bays were above. Both secondary indicators of community condition significantly differed among harbors due to greater Shannon-Wiener diversity in Mission Bay than San Diego Bay marina stations, higher numbers of taxa in marina stations of Mission Bay than any other harbor, and higher numbers of taxa in freshwater-influenced stations of Mission and San Diego Bays than those of Dana Point Harbor.

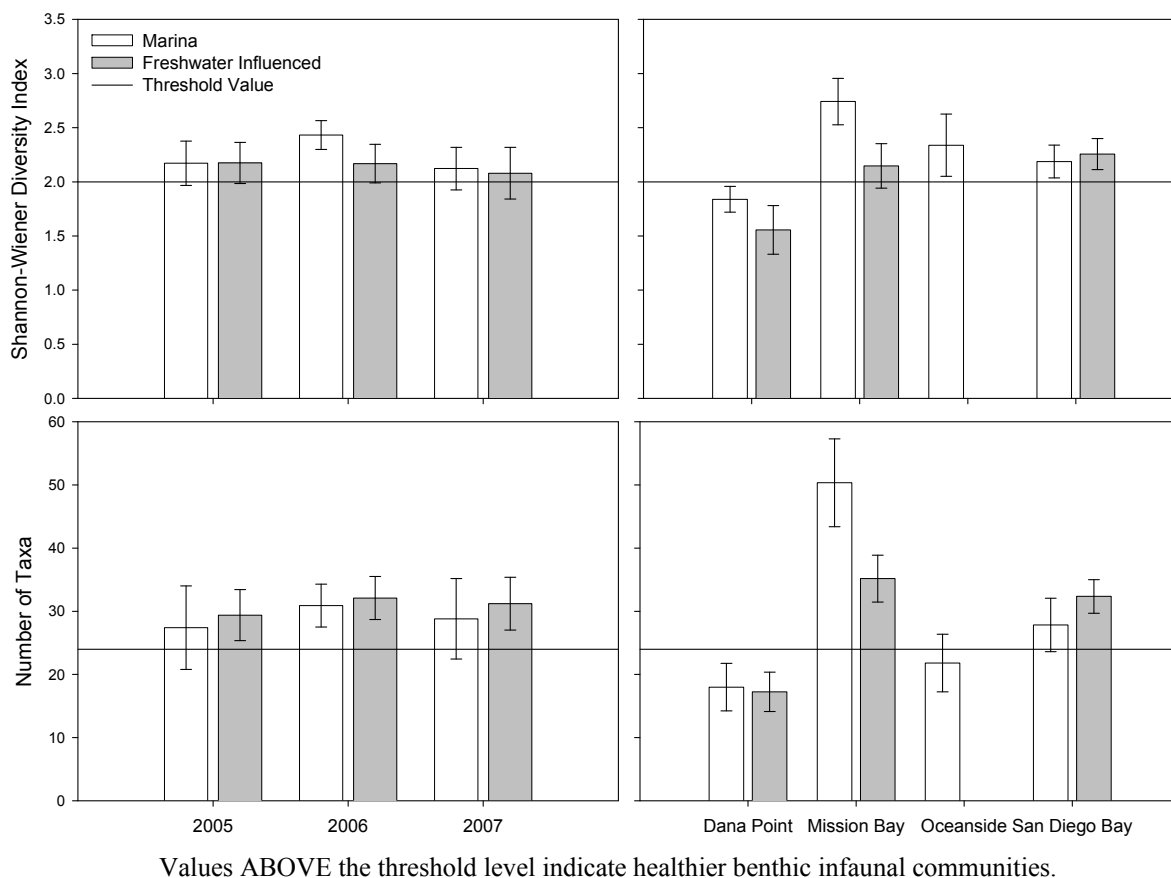


Figure 3-17. Comparisons of average benthic infaunal community measures among years and harbors

3.2.4 Grain Size and Total Organic Carbon

The results of sediment grain size and TOC analyses for the 60 stations sampled in Pilot Project are summarized in Appendix B Table B-4 – B-7. These measurements have no threshold levels for comparison; however, they can be used to help interpret biological responses, as well as understand the distribution of contaminants within sediments.

The median sizes of sediments were highly variable, ranging from 3.5-198.7 microns in the marina stratum and from 2.26-215.72 microns in the freshwater-influenced stratum, with average values being slightly higher in the marina (55.2 ± 11.9 microns) than the freshwater-influenced stratum (40.61 ± 8.50 microns). Additionally, fine sediments (i.e., silt and clay) comprised a slightly larger percent of freshwater-influenced sediments (mean = $63.77 \pm 3.09\%$) than the marina sediments ($61.62 \pm 5.06\%$). Sediment sizes changed little from year to year in the Pilot Project and were of such high variability among stations within harbors that differences among harbors were difficult to detect (Figure 3-18). As a result, neither median grain size nor percent fine sediments differed significantly between strata, among years, or among harbors.

Differences in TOC, in contrast, were more apparent between strata and among harbors than grain size. Average TOC was significantly lower in the marina stratum ($1.33 \pm 0.12\%$) than in

the freshwater-influenced stratum ($1.65 \pm 0.15\%$), and TOC differed significantly among harbors both in the marina and freshwater-influenced strata. Within the marina stratum, Mission Bay had a significantly higher percent of TOC than San Diego Bay, and in the freshwater-influenced stratum both Mission Bay and Dana Point Harbor had significantly higher levels than San Diego Bay. As was the case for sediment grain size, TOC did not change markedly from year to year (Figure 3-18).

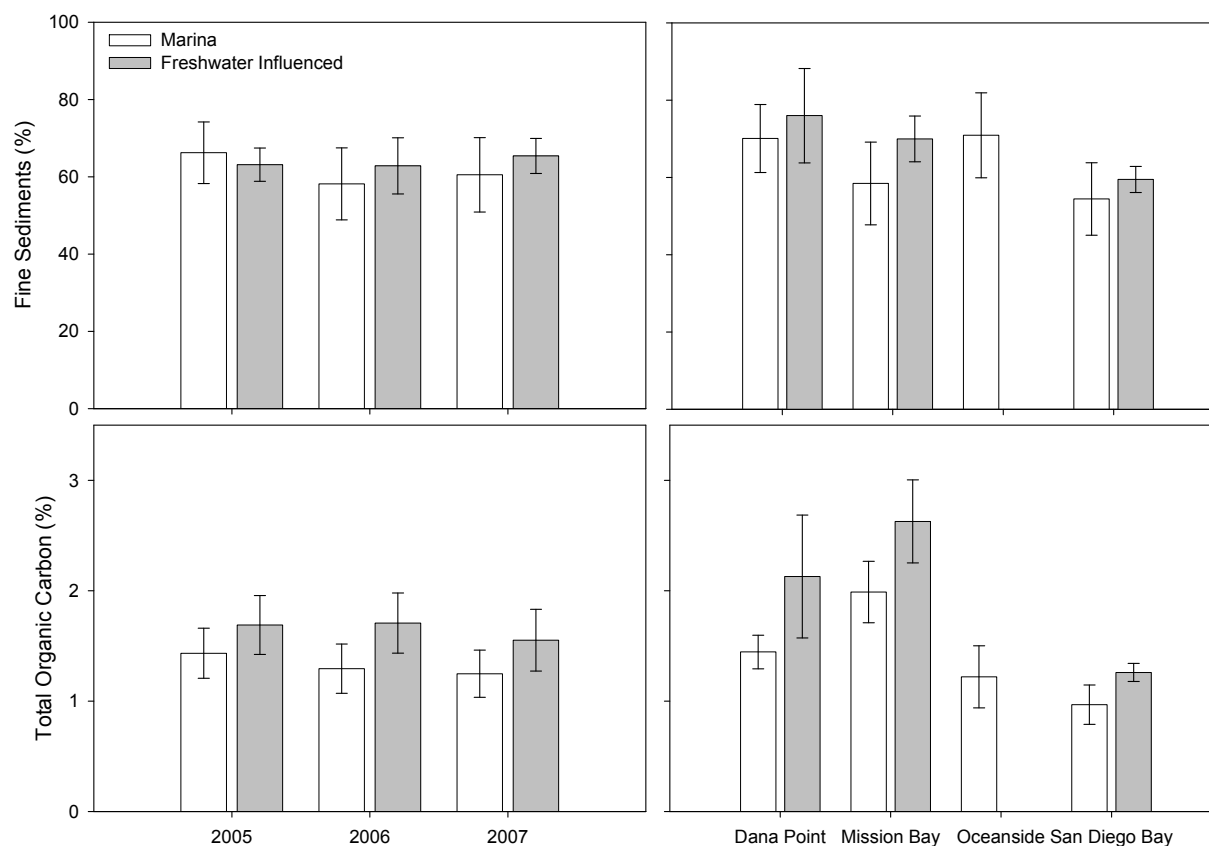


Figure 3-18. Comparisons of percent fine sediments and total organic carbon among years and harbors

3.3 Analysis of Study Design

The marina and freshwater-influenced strata were selected for the Pilot Project because they were expected to have more variability in the results (i.e., higher coefficients of variation (CV) values) than the open water strata, and would, therefore, provide the most conservative test of the study design's ability to detect significant differences. The effectiveness of the study design in detecting significant differences was assessed in three ways. First, CV were compared between strata and among indicators to determine the level of variability that may have limited the ability to detect significant differences. Second, a power analysis was used to test the probability that the binomial test will yield a significant result (i.e., reject the null hypothesis) when observed proportions are compared to preset targets. Third, statistical test results were assessed to determine the percentage of tests that resulted in significant findings.

Coefficient of Variation

The CV (i.e., ratio of standard deviation to mean expressed as percent) provides a relative measure of variability that can be compared directly among different indicators. Coefficients of variation ranged from 36-79% in marina surface waters and from 33-118% in freshwater-influenced surface waters (Table 3-7). The freshwater-influenced surface water indicators on average were more variable (avg. CV = 84%) than the marina (avg. CV = 52%), with the highest level of variability encountered for zinc in the freshwater-influenced stratum. Although the freshwater-influenced strata covered less area than the marina, differing types and levels of freshwater inputs may have increased the variability of surface waters.

For all sediment indicators (i.e., chemistry, toxicity, and infauna) CV ranged from 25-168% in the marinas and from 18-141% in the freshwater-influenced stratum. In contrast to surface waters, variability in the sediments was higher in the marina (avg. CV = 81%) than the freshwater-influenced stratum (avg. CV = 67%), with copper having the single highest CV. Higher levels of variability in the marina stratum appeared to be driven by copper primarily, since it had nearly four times higher CV than any other indicator.

The ability to detect significant differences for indicator values (i.e., concentrations, quotients, and index scores) is predicted to be lowest for those indicators with the highest CV. Therefore, using a CV of 100% as a threshold, significant differences for surface water dissolved and total zinc, sediment zinc, sediment cadmium, sediment copper, total detectable PAHs, and *E. estuarius* mortality would be more difficult to detect, assuming equivalent differences between and among factors (e.g., time, strata, and harbors).

Table 3-7. Comparison of sample variability in primary and secondary indicators

Constituent		Marina			Freshwater Influenced		
		Mean	St Dev	CV (%)	Mean	St Dev	CV (%)
Water	Dissolved Metals						
	Copper (Cu)	6.558	3.617	55.16	3.538	3.056	86.37
	Nickel (Ni)	0.3599	0.1493	41.49	0.5995	0.3033	50.6
	Zinc (Zn)	20.12	9.7	48.2	9.9	11.68	118.01
	Total Metals						
	Copper (Cu)	11.66	9.26	79.41	5.257	5.133	97.63
	Nickel (Ni)	0.393	0.1409	35.86	0.6581	0.2168	32.95
	Zinc (Zn)	21.41	11.25	52.56	10.01	12.04	120.29
Sediment	Mean ER-M Quotient	0.3963	0.351	88.58	0.2169	0.1039	47.92
	Cadmium (Cd)	0.4179	0.2583	61.82	0.541	0.766	141.47
	Chromium (Cr)	41.77	21.16	50.65	36.19	12.29	33.97
	Copper (Cu)	285.8	480.4	168.08	119.1	102.4	86.02
	Lead (Pb)	34	33.73	99.19	29.56	17.13	57.96
	Nickel (Ni)	13.36	7.08	52.97	13.422	3.633	27.06
	Zinc (Zn)	202.5	106.2	52.44	187	93.3	49.89
	Total Detectable PAHs	550	842	153.05	708	774	109.44
Toxicity	<i>E. estuarius</i> mortality	12.22	13	106.41	15.88	13.19	83.08
Infauna	BRI Score	41.13	10.48	25.49	42.66	7.88	18.47
	Number of Taxa	29.03	17.29	59.56	30.9	11.91	38.56
	Shannon-Wiener Diversity Index	2.242	0.569	25.38	2.14	0.624	29.18

Power Analysis

Based on the level of variability and observed differences between measures, the power analysis was used to predict the sample size needed to detect a difference at a certain power (i.e., 80%). Since the levels of variability within the two strata were expected to be the highest of any strata, the power analysis likely provided a conservative estimate of the power of the full RHMP study design.

The power to detect significant differences between observed proportions and preset targets was relatively high for those indicators where observed proportions and preset targets were not equal, with approximately half of the indicators having a power of 60% or greater (Table 3-8). When differences with preset targets were small, as in the case of total copper in marinas and for sediment zinc, the power to detect a significant difference was low, and, consequently, the level of sampling effort needed to detect a difference would be impracticably high (e.g., greater than 300 stations). However, for the majority of indicators the results indicate that there will be sufficient power to detect significant changes of the four RHMP harbors from preset targets for the full RHMP study since a total of 75 stations will be sampled (i.e., 15 stations per stratum by 5 strata). Using a sample size of 75 would meet or exceed the estimated sample size to determine significance for 67% of the indicators.

Table 3-8. Observed powers and estimated sample sizes for binomial tests of proportions based on Pilot Project results

	Constituent	n ¹	Preset Target %	Marina			Freshwater Influenced		
				Area %	Observed Power %	Estimated n for Significance ²	Area %	Observed Power %	Estimated n for Significance ²
Water	Dissolved Copper	30	70	37	98	14	83	52	64
	Total Copper	30	26	20	20	302	77	>99	5
	Dissolved Zinc	20	100	100	N/A	N/A	100	N/A	N/A
	Total Zinc	30	97	100	60	51	100	60	51
	Dissolved Nickel	30	100	100	N/A	N/A	100	N/A	N/A
	Total Nickel	30	100	100	N/A	N/A	100	N/A	N/A
Sediment	ER-M Quotient	30	48	33	51	66	0.57	27	191
	Cadmium	30	90	100	99	15	97	49	71
	Chromium	30	78	97	97	16	100	>99	6
	Copper	30	68	47	76	34	83	61	50
	Lead	30	74	83	34	127	90	76	34
	Nickel	30	80	90	46	77	97	94	18
	Zinc	30	45	40	16	594	40	16	594
	Total Detectable PAHs	30	74	100	>99	5	100	>99	5
Toxicity	<i>E. estuarius</i> mortality	30	51	80	96	16	67	55	58
Infauna	BRI	30	37	20	66	43	7	>99	10
	Shannon-Wiener diversity	30	90	63	98	14	63	98	14
	Number of taxa	30	92	60	>99	10	70	94	18

N/A = not applicable due to no difference between proportions

¹ Pilot Project sample size

² Estimated sample size needed to detect a significant difference based on an alpha of 0.1 and power of 80% for the observed delta (i.e., the difference between the observed proportion and preset target)

Statistical Significance Results

Comparisons of observed percentages of areas below threshold levels to preset targets with the binomial test resulted in significant results in 82% of the surface water cumulative tests (i.e., combined test for years 2005-2007 with a sample size of 30) where differences were greater than zero. For indicators of sediment condition, cumulative binomial tests were significant 87.5% of the time. Based upon the statistical results, the binomial test appears to provide sufficient power to detect significant differences.

Parametric and nonparametric tests, including ANOVA, Mann-Whitney, and Kruskal-Wallis, were used to test for differences in indicator values between strata, among years, and among harbors. Differences between strata were found to be significant in 89% of tests for surface water indicators and 18% for sediment indicators. For assessments of temporal changes, differences were significant in 39% of surface water and 23% of sediment tests. Lastly, interharbor differences were found to be significant in 61% of surface waters and 42% of sediment tests. The much higher level of significant findings in the surface waters than the sediments appeared to be attributable to greater levels of difference (i.e., delta values) and lower levels of variation for surface water indicators. Therefore, the use of ANOVAs and nonparametric tests appeared to be effective in detecting significant differences between strata, among years, and among harbors when it was reasonable to assume that significant differences would exist.

4.0 DISCUSSION

The results of the Pilot Project validated the effectiveness of the RHMP study design in answering four of the five core monitoring questions set forth by the SDRWQCB:

1. What are the contributions and spatial distributions of inputs of pollutants to harbors in the San Diego Region and how do these inputs vary over time?
2. Are the waters in the harbors safe for body contact activities?
3. Are the fish safe to eat?
4. Do the waters and sediments in the harbors sustain healthy biota?
5. What are the long-term trends in water quality for each harbor?

The Pilot Project was not designed to assess the third question regarding the safety of the fish for human consumption; however, this question will be studied upon full implementation of the 2008 RHMP study conducted in coordination with the Bight '08 regional monitoring study.

The use of a stratified random design that was repeated among years allowed for the assessment of the spatial distributions of pollutants (i.e., differences between strata and among harbors), as well as changes in the levels of pollution through time (i.e., short-term trends over the 3-year Pilot Project). Additionally, the Pilot Project illustrated that the study design lends itself to analysis of long-term trends. First, the binomial approach can be used to assess whether present-day proportions of stations within strata are better or worse than historical preset target proportions. Secondly, the design allows for statistical comparisons of changes in conditions (e.g., metal concentrations and BRI values) from year to year using ANOVAs. Thirdly, Chi-squared tests of trends among proportions can be used to analyze trends in the percentages of areas below threshold values. Evidence for the effectiveness of the first two approaches is seen in the prevalence of statistically significant results in most cases where it was reasonable to assume that they would occur. Even though statistical differences were commonplace, high levels of variability among stations for several of the indicators (e.g., total detectable PAHs) still limited the ability to detect significant differences and will likely continue to do so even at extremely high sample sizes based upon the results of power analyses.

Discussion of the Pilot Project results is provided from two basic perspectives. In the first, the major study findings are discussed in relation to the health of the harbors and the likelihood that indicators will have adverse environmental effects. The second evaluates the effectiveness of the study approach and offers suggestions for improvement. Both perspectives address the primary monitoring questions by assessing differences from historical conditions, spatial distributions of indicators (i.e., differences between strata and among harbors) and changes in indicators over time (i.e., trends). Lastly, based on the results of the Pilot Project special focused studies are recommended that may be used to augment core monitoring efforts.

4.1 Assessment of Harbor Conditions

4.1.1 Surface Waters

Chemical Indicators

All chemical indicators in surface waters occurred at concentrations below CTR thresholds, with the exception of copper in marinas, indicating that only copper occurred at concentrations that may result in toxic effects. Copper concentrations exceeded the CTR thresholds for dissolved and total copper throughout the majority of the marina stratum, and copper concentrations occurred at approximately twice the threshold concentration on average for all years, indicating that marina surface waters may have toxic effects due to elevated copper concentrations. This finding is consistent with previous studies that have documented copper as a contaminant of concern in San Diego Bay marinas (McPherson and Peters, 1995; SDRWQCB, 2005) and the larger the San Diego region (Schiff et al., 2006). However, not all harbors experienced the same level of copper exceedances, since the average concentrations of Mission Bay were below threshold values, while the other three harbors were above. Since surface water copper levels generally were not elevated in freshwater-influenced areas, except those areas of Dana Point Harbor that were surrounded by the marina stratum, it appears that boating-related activities have a more detectable and persistent effect on copper concentrations in the harbors than does urban runoff. It is estimated that passive leaching from boat antifouling paints contributes a mass load of 2000 kg/yr of dissolved copper to the Shelter Island Yacht Basin (SDRWQCB, 2001 and Brown and Schottle, 2006). Thus, elevated copper concentrations appear to be most strongly related to the use of antifouling paints on boats, although contributions from urban runoff and aerial deposition should not be overlooked and may still be important to other areas of the harbors.

Bacterial Indicators

Measurements of the bacterial indicator enterococcus were consistently well below AB411 standards, with the vast majority of the harbor areas having bacterial levels that were below detection limits. The highest recorded measurements, 20 MPN/100mL, were less than one fifth the AB411 standard. Therefore, the consistently low enterococcus levels observed in both freshwater-influenced and marina strata alike in all years of the Pilot Project may indicate that this bacterium is not likely to occur at elevated levels during summer months when rain events are extremely rare.

Physical Indicators

Physical water column measures, including temperature, salinity, DO, pH, and transmissivity were suitable to support marine organisms in both strata. Although there were slight and sometimes statistically significant differences between the strata, the changes in physical conditions generally were not of sufficient magnitude to alter the suitability of the waters or underlying benthos to support marine life. However, there were a few potential exceptions. First, transmissivity (i.e., levels of light penetration) in the marine stratum declined much more rapidly with depth than in the freshwater-influenced stratum. Reductions in light have the potential to limit the abundance of primary producers, such as eelgrass and algae, as well as animals that depend on these resources for food and habitat. Secondly, temperature increases from northern to southern harbors, which were on average approximately 3°C, and from 2005-2007 by 4°C have the potential to affect certain marine species. Lastly, during a 2005 red tide in Oceanside Harbor, DO dropped to nearly anoxic levels, around 0.1 mg/L. At such low oxygen

levels, impacts to less mobile benthic species as well as some water column species may have occurred, although benthic infaunal community condition within Oceanside Harbor was not significantly different from other harbors in 2005 or subsequent years.

4.1.2 Sediments

Chemical Indicators

Sediment contamination by pollutants was largely limited to copper, zinc, arsenic, and mercury, since all other contaminants examined on average occurred at concentrations below ER-L thresholds. As a consequence, the other metals and PAHs are unlikely to cause adverse biological impacts (Long et al., 1995). Furthermore, copper, zinc, and mercury reached concentrations that exceeded the ER-M, indicating that pollutant levels have the potential to result in adverse biological effects. Similar to surface waters, elevated copper levels occurred primarily within the marinas, while zinc exceedances of the ER-L were equally likely in both strata.

Mean ER-M quotients above the 0.2 threshold were primarily due to elevated copper, zinc, arsenic, and mercury concentrations. Copper, zinc, arsenic, and mercury concentrations were on average elevated (i.e., consistently occurred at levels above ER-Ls) in the marina stratum, while only zinc, arsenic, and mercury were in the freshwater-influenced stratum. The 0.2 mean ER-M quotient was used as a conservative threshold for biological effects; stations with mean ER-M quotients above 0.2 are more likely to have adverse benthic effects due to toxicity (Weston, 2005b). Based on the 0.2 threshold level, 67% of the marina and 43% of the freshwater-influenced stratum may have adverse benthic effects due to exceedances. Mean ER-M quotients for the freshwater-influenced stratum were only slightly greater than the threshold on average (0.22) while those of the marina were nearly twice as high (0.4). Therefore, it would be reasonable to expect that both toxicity and benthic community condition in the marina stratum also would be worse than the freshwater-influenced stratum; however, both toxicity and infaunal measures were equivalent between the strata. One explanation for the lack of toxicity and infaunal differences between strata could be attributed to the conservativeness of the threshold, since it is more generally accepted that mean ER-M quotients below 0.6 are believed to be nontoxic (Weston, 2005b). In Bight studies, mean ER-M quotients less than 0.5 are considered to have low-to-moderate risks of adverse biological impacts (Schiff et al., 2006).

Toxicity

Evidence of toxicity was less apparent in the marina than the freshwater strata, contrary to predictions based on mean ER-M quotients. Eighty percent of the marina stratum and 67% of the freshwater-influenced stratum contained sediment that was not toxic (i.e., *E. estuarius* mortality was less than 20%). Moreover, the average level of *E. estuarius* mortality was approximately 12%, while that of the freshwater-influenced stratum was 16%, also contrary to predictions. The projected level of participation of the RHMP in Bight 2008 monitoring will greatly improve the estimate of the spatial extent of toxicity in San Diego harbors.

Benthic Infaunal Communities

According to mean ER-M quotient values, benthic infaunal communities would be expected to show greater evidence of disturbance in the marina than the freshwater-influenced stratum, while based on toxicity results the communities would be expected to be equivalent or slightly more disturbed in the freshwater-influenced stratum. In reality, the benthic communities of the two

strata better reflected the toxicity results, since they were nearly equivalent between strata as assessed by the BRI and the measures of diversity (i.e., Shannon-Wiener index and number of taxa). Average BRI values for both strata indicated that the benthic communities were intermediate between response level 2 (marginal deviation from reference) and 3 (biodiversity loss of reference species). Stations in both strata had communities that ranged from reference to community function loss; however, a higher percentage of the marinas were classified in response levels 1 and 2 (63%) than the freshwater-influenced areas (43%). Although previous studies have shown that sites with degraded benthic communities in San Diego Bay were most closely associated with marinas and shipyards (Faurey et al., 1996), the Pilot Project showed that freshwater-influenced infaunal communities are equally or more degraded than the marina communities. Both areas are thought to receive high levels of pollutants, including antifouling paints and fuels for marinas and urban runoff for freshwater-influenced areas (Ranasinghe et al., 2007). In addition to the influence of pollutants, diversity and BRI values are sensitive to other forces that are unrelated to the measured indicators, including the physical environment and disturbance events (Smith et al., 2001). For example, lower benthic community measures in the freshwater-influenced stratum than other areas of the harbors may result from heavy rains during winter storms that serve as episodic and seasonal disturbance events that temporarily alter the physical environment, limiting the abundance and number of infaunal species. Further studies are needed to assess the factors that are impacting benthic infaunal communities within these two strata. Upon implementation of the 2008 RHMP, it will be possible to determine whether benthic communities are in a poorer condition than historically encountered throughout all areas of the harbors or just within these two strata.

4.2 Comparison of Pilot Project Results to Historical Conditions

The Pilot Project demonstrated that the study design was effective in assessing differences from historical conditions using a binomial statistical approach. This approach will provide a first step in examining the long-term trends in conditions for each stratum, as well as in each harbor. Surface water, sediment, and benthic infaunal community measures assessed from 2005-2007 commonly differed significantly from historical conditions. In the marina stratum, the primary indicators for surface waters (dissolved and total copper), sediments (mean ER-M quotient), and benthic infauna (BRI) occurred at levels that were worse than historical conditions, while toxicity levels were better. In contrast, the majority of secondary surface water and sediment indicators were equivalent to or better than historical conditions, while secondary benthic infaunal indicators were poorer. For the freshwater-influenced stratum, surface water copper concentrations, sediment mean ER-M quotients, and toxicity levels were better than historical conditions, while only benthic infauna was worse. Moreover, secondary indicators within the freshwater-influenced stratum consistently corroborated primary indicator results. Thus based on a weight-of-evidence approach, conditions in the marina stratum were generally construed as being worse than historical conditions, while those of the freshwater-influenced stratum were better.

Although the Pilot Project was not implemented to determine changes in the marina and freshwater-influenced strata from historical conditions, the results indicate that the study design should be able to detect statistically significant changes when all strata of the harbors are compared to historical conditions upon implementation of the full RHMP study. The prevalence of statistically significant results for dissolved copper, total copper, total zinc in surface waters and nearly all indicators in sediments except zinc strongly speaks to the adequacy of the design

to detect statistical changes when they would be expected to exist. Given the level of differences between observed percentages and preset targets, the observed power of the tests was relatively high. However, when differences between observed percentages and preset targets were small, the power was low (e.g., comparisons of total copper in marinas) and the sample size that would be needed to detect such a small difference was impracticably high. Additionally, for the surface water metals where differences were absent between observed percentages and preset targets (e.g., dissolved and total nickel and dissolved zinc), application of a statistical test is unneeded. For these indicators, changes can be assessed by comparing the values directly rather than the proportions. Since the Pilot Project intentionally targeted the strata that were believed to have the highest levels of variability, the ability to detect significant differences with reasonably high power is a positive indication that the design is sufficiently robust.

4.3 Spatial Distribution of Indicators

4.3.1 Comparisons between Strata

The RHMP study design provides two alternatives for assessing differences between strata. First, the percent of area below thresholds can be compared between strata with a two-sample binomial test, and, secondly, ANOVAs can be used to test for differences in indicator values between strata. In subsequent years when more than two strata are surveyed, Chi-squared tests can be used to test for differences in the percent of area below thresholds among strata.

During the Pilot Project, differences between the marina and freshwater-influenced strata were evident for surface water and sediment chemistry indicators, while differences in toxicity and benthic infauna were not. Mean measures of surface water indicators differed significantly between the marina and freshwater-influenced stratum. Copper and zinc concentrations were consistently higher in the marina than the freshwater-influenced stratum, while nickel concentrations were higher in the freshwater-influenced stratum. Although, there were statistically significant differences for nickel and zinc, mean concentrations for nickel were less than a hundredth of the CTR threshold value and mean zinc concentrations were less than one half the CTR in both strata. Consequently, impacts of these metal concentrations are not anticipated. Copper, on the other hand, regularly exceeded the CTR values in the marina stratum, while exceedances in the freshwater-influenced stratum were largely limited to 2005 occurrences in Dana Point Harbor. Therefore, if copper were to have a toxic effect, it is much more likely to occur in the marina stratum than the freshwater-influenced stratum as previously discussed.

4.3.2 Comparisons among Harbors

By assessing differences among harbors, special studies and management actions can be tailored to specific harbors and bays, consistent with the purpose for implementing stratification. Comparing differences between strata and among harbors increases the ability to determine the spatial distribution of pollutants in the harbors. Average water column, sediment, and benthic infaunal indicators as well as the percentages of stations below target thresholds showed pronounced and often significant differences among harbors, while toxicity levels did not differ significantly among harbors and were consistently better than historical conditions. Larger bays, such as Mission Bay and San Diego Bay, tended to have fewer exceedances of threshold levels than did smaller bays, which were often predominated by the marina stratum. For example,

greater percentages of stations had exceedances of water column copper and zinc concentrations and sediment mean ER-M quotients and copper concentrations in Dana Point and Oceanside Harbor than in San Diego and Mission Bay. However, lead was an exception, since lead exceedances only occurred within San Diego Bay. Although, several indicators within the marina and freshwater-influenced strata differed among the harbors, differences in the overall conditions of the harbors cannot be inferred from these two strata alone. All harbors contain other strata that may be more or less impacted; therefore, comparisons of the overall conditions of the harbors, if desired, may only be made upon completion of the 2008 monitoring study.

The level of difference between strata also varied among harbors. Freshwater-influenced and marina areas of Dana Point Harbor were much more similar to each other than were freshwater-influenced and marina areas of Mission Bay and San Diego Bay. Freshwater-influenced areas within Dana Point Harbor were limited to small areas that were nearly completely surrounded by the larger marina stratum, while freshwater-influenced areas of the larger bays were not in close proximity to the marina stratum and generally encompassed larger areas. Upon implementation of the RHMP in 2008, comparisons of the freshwater-influenced stations to surrounding stations located in the shallow, deep, port/industrial, and marina strata will help determine whether freshwater inputs are having a discernable and persistent impact on areas of the harbor.

The ability to detect significant differences for indicator values, particularly concentrations, was limited by the level of variability among stations and the magnitude of the differences. The use of a three-way ANOVA, with factors strata, harbor, and survey year, enhanced the ability to test for differences between the two strata by assessing the variability attributable to each factor and removing it from the model. The same approach can be used to test for differences among the five strata as well, provided that the data meet the criteria for parametric statistics. However, in cases where the data do not meet the criteria or cannot be transformed to meet the criteria, nonparametric statistics will have to be used in place of ANOVAs. Since significant differences were readily apparent between strata (sample size = 30), among years (sample size = 10), and among harbors (sample sizes varied by harbors) when data were compared over the 3-year Pilot Project, it is expected that similar differences will also be statistically detectable upon implementation of the full RHMP study provided a sample size of 15 stations per stratum is used.

4.4 Assessment of Trends

Over the three-year Pilot Project, no negative trends were evident for any indicator that would be indicative of a degrading condition since there were no consistent increases in percentage of stations above threshold levels or significant increases in concentrations of indicators across all three years of the study. Some indicators had significant increases for one year of the study only to decline in the subsequent year (e.g., surface water dissolved nickel concentrations), while most significant changes from year to year were actually declines. For example, surface water total nickel concentrations, although far below the CTR threshold, declined from 2005 to 2007, as did chromium and nickel concentrations in sediments.

Trends indicating increases in the percentage of stations below threshold levels (i.e., decreases in exceedances at stations) were rarely observed throughout the three-year study. No trends were apparent for surface water indicators; however, sediment conditions within the freshwater-

influenced stratum appeared to be improving, since percentages of stations below mean ER-M quotient and toxicity thresholds increased from 2005 to 2007.

Given that the majority of indicators did not experience pronounced changes in average levels or percentages of stations below threshold levels over the three-year study, performing the study at annual intervals does not appear to be essential to assess long-term trends. Rather, conducting the RHMP at a longer interval, such as a five-year cycle conducted in synchrony with the Bight regional monitoring program, will concentrate sampling effort at a specific time to get a better picture of the health of the harbor, while also allowing for direct comparisons of the San Diego harbors to other harbors of the Bight. Furthermore, comparison of the RHMP harbors as a whole to historical conditions will allow us to determine if conditions are improving or deteriorating within the harbors.

4.5 Recommended Modifications to the RHMP Study Design

Since the initiation of the Pilot Project additional historical data for the harbors have been released, the BRI has been modified, and new sediment quality objectives (SQO) for bays and estuaries have been enacted. Moreover, successful completion of the Pilot Project has provided valuable insights into the validity of the approach and how it can be enhanced. As a result of recent methodological innovations and analyses of the effectiveness of Pilot Project results, the following modifications to the RHMP study design are recommended:

- Increase the sample size in strata to 15.
- Integrate Bight 2003 data into historical distribution curves.
- Analyze sediment conditions with new SQOs.
- Revise benthic community assessment and BRI calculation.
- Include tributyltin (TBT) as an analyte.

4.5.1 Sample Size

Increases in sample sizes are the most efficacious way to enhance the power of the RHMP study design to detect statistically significant changes from historical conditions, differences in the spatial distribution of pollutants, and long-term trends in harbor conditions. Additionally, by monitoring more stations, the RHMP will provide a better estimate of the overall conditions of the harbors, as was clearly illustrated by the comparison of Pilot Project toxicity results to those of the Bight '03 regional monitoring program (Section 4.1). The Pilot Project demonstrated that the ability to detect significant differences of observed percentages from preset targets, as well as between strata, among years, and among harbors was greatly enhanced when three year's of data were analyzed (sample size of 30) rather than just one (sample size of 10). For example, differences between the historical preset target and observed percentages below the sediment copper threshold were only significant when the data from the entire Pilot Project were assessed. At a sample size of 30, the power to detect a 15% difference was 61% for sediment copper. By increasing the sample size to 15 stations per stratum for a total of 75 stations, there should be sufficient power to detect a 20% change from historical conditions for the majority of the indicators given their relative measures of variability.

4.5.2 Bight 2003 Data

In order to enhance the 10-year period historical dataset from which preset targets are determined, it is recommended that the San Diego harbors data from the Bight '03 regional monitoring survey be included. In Bight '03, stations were surveyed in three of the four RHMP harbors, San Diego Bay, Mission Bay, and Dana Point Harbor. Inclusion of the data will result in shifts in the historical cumulative frequency distributions that may alter the preset targets.

In the Bight '03 regional monitoring study, the average copper concentration of RHMP harbor marinas, including Dana Point Harbor, Mission Bay, and San Diego Bay, was 144 mg/kg (Schiff et al., 2008), while for the Pilot Project the average for marinas was much higher at 286 mg/kg. Additionally, the average zinc concentration of the Bight '03 RHMP marinas (153 mg/kg) was also lower than the Pilot Project average of 202 mg/kg, although the difference was less pronounced. Although the Pilot Project detected higher copper and zinc concentrations in marina sediments than did the Bight '03, it is difficult to determine if conditions are actually worsening since far fewer marina stations were surveyed in Bight '03 (12 stations) than were surveyed in the Pilot Project (30 stations). Additionally for copper, the difference is most likely due to the influence Oceanside Harbor copper concentrations, since it was not surveyed during Bight '03 and it had the highest average sediment copper concentrations of all RHMP harbors throughout the Pilot Project. If Oceanside Harbor copper concentrations are removed, then the Pilot Project average is only 199 mg/kg, which while still higher is much closer to the Bight '03 average.

The relatively low levels of toxicity observed throughout the three years of the Pilot Project strongly differ from Bight 2003 findings of approximately 50% of marina areas being toxic (Bay et al. 2005). Additionally, a closer examination of the Bight 2003 data indicated that 47% of San Diego Bay and 50% of Mission Bay were toxic, while none of Dana Point Harbor was toxic; however, this conclusion was based on assessments of 9, 2, and 1 station(s), respectively. In the Pilot Project, 17% of San Diego Bay, 25% of Mission Bay, 36% of Dana Point Harbor, and 40% of Oceanside Harbor were found to be toxic. Although a slightly different criterion was used in the Bight to ascribe toxicity (i.e., 17% mortality relative to controls), the large discrepancy between the two studies is more likely due to the much smaller sampling effort in the San Diego harbors by the Bight 2003 study. The projected level of participation of the RHMP in Bight 2008 monitoring will greatly improve the estimate of the spatial extent of toxicity in San Diego harbors.

4.5.3 Sediment Quality Objects

As of February 2008, new SQOs have been enacted by the State Water Resources Control Board (SWRCB) to assess ambient sediment quality in bays and estuaries. The goals of the SQOs are to determine if pollutants in sediments are present in quantities that are toxic to benthic organisms and/or will bioaccumulate in marine organisms to levels that may be harmful to humans. It is recommended that sediment quality from RHMP harbors be assessed using California's SQOs as described in the *Draft Staff Report, Water Quality Control Plan for Enclosed Bays and Estuaries* (State Water Resources Control Board – California Environmental Protection Agency, 2007), since Bight 2008 results for embayments, including RHMP harbors, will be analyzed accordingly. Evaluating RHMP sediment condition via SQOs will involve the inclusion of a sublethal toxicity test using *Mytilus galloprovincialis* in addition to *E. estuarius* survival tests, analyzing benthic infaunal data with four indices, and ensuring that all chemical

analytes required for the sediment chemistry line of evidence (LOE) are included in the RHMP analytes list.

The SQOs are based on a multiple-lines-of-evidence (MLOE) approach in which sediment toxicity, sediment chemistry, and benthic community condition are the lines of evidence (LOE). The MLOE approach evaluates the severity of biological effects and the potential for chemically-mediated effects to provide a final station level assessment. An overview of the methods is provided as follows.

Sediment Toxicity

Sediment toxicity is assessed using two tests: a 10-day *E. estuarius* survival test and a sublethal test using the mussel *M. galloprovincialis*. Sediment toxicity test results from each station are statistically compared to control test results, normalized to the control survival, and categorized as nontoxic, low, moderate, and high toxicity. The average of the test responses then is calculated to determine the final toxicity LOE category.

Sediment Chemistry

Concentrations of chemicals detected in sediments are compared to the California Logistic Regression Model (CA LRM) and the Chemical Score Index (CSI). The CA LRM is a maximum probability model (P_{MAX}) that uses logistic regression to predict the probability of sediment toxicity. The CSI is a predictive index that relates sediment chemical concentration to benthic community disturbance. Sediment chemistry results according to CA LRM and CSI are categorized as having minimal, low, moderate, and high exposure to pollutants, and the average of the exposure categories are used as the chemistry LOE.

Benthic Community Condition

Benthic community condition is assessed using a combination of four benthic indices: the Benthic Response Index (BRI), Relative Benthic Index (RBI), Index of Biotic Integrity (IBI), and a predictive model based on the River Invertebrate Prediction and Classification System (RIVPACS), following the January 21, 2008 guidance provided by the SCCWRP entitled *Determining Benthic Invertebrate Community Condition in Embayments* for southern California marine bays. Each benthic index result is categorized according to four levels of disturbance, including reference, low, moderate, and high disturbance. The results of the four indices are averaged to determine the benthic community LOE.

Integration of Multiple Lines of Evidence

The station level assessment provides an indication of whether the aquatic life SQOs are being met at each station of interest. The station level assessment is based on the severity of biological effects (i.e., integration of toxicity LOE and benthic condition LOE categories) and the potential for chemically-mediated effects (i.e., integration of toxicity LOE and chemistry LOE categories), using decision matrices.

4.5.4 Benthic Community Assessment

As part of the new SQOs, the BRI calculation and thresholds were modified to increase the level of accuracy in categorizing levels of disturbance experienced by benthic infaunal communities. As previously indicated, not only was the BRI modified, but three other indices were used to categorize the community state as well. It is recommended that the recent updates to the BRI be

included in the RHMP analyses and that the benthic communities be assessed consistent with the SQO benthic community assessment, since it has been demonstrated that the four-index approach provides estimates of disturbance levels of communities that are more consistent with expert opinion than the BRI or any other index alone (Ranasinghe et al. 2007). Thus, it is recommended that in addition to the BRI, that the RBI, IBI, and RIVPACS indices be used to assess benthic community condition. Calculation of the indices is described following the January 21, 2008 guidance provided by the SCCWRP entitled *Determining Benthic Invertebrate Community Condition in Embayments* for southern California marine bays.

Each benthic index result is categorized according to four levels of disturbance, with conditions ranging from a reference condition to high disturbance.

- Reference: Equivalent to a least affected or unaffected site
- Low Disturbance: Some indication of stress is present, but is within measurement error of unaffected condition
- Moderate Disturbance: Clear evidence of physical, chemical, natural, or anthropogenic stress
- High Disturbance: High magnitude of stress

Specific categorization values tailored to southern California marine bays are assigned for each index (Table 4-1). The final step in determining the benthic community condition is the integration of the four indices into a single category. In doing so, the median of the four benthic index response categories is computed. If the median falls between two categories, the value is rounded to the next higher category to provide the most conservative estimate of benthic community condition.

Table 4-1. Benthic index categorization values for southern California marine bays

Benthic Community Guideline				Index
BRI	IBI	RBI	RIVPACS	
< 39.96	0	> 0.27	> 0.90 to < 1.10	Reference
39.96 to 49.14	1	0.17 to 0.27	0.75 to 0.90 or 1.10 to 1.25	Low Disturbance
49.15 to 73.26	2	0.09 to 0.16	0.33 to 0.74 or > 1.25	Moderate Disturbance
> 73.26	3 or 4	< 0.09	< 0.33	High Disturbance

A description of the methods used to calculate the four indices is provided as follows.

Benthic Response Index

The BRI is the 'abundance-weighted pollution tolerance score' of infaunal species, with scores increasing from 0 to 100 with greater levels of disturbance (Smith et al., 2001 and 2003). The BRI scores are calculated using the abundances of species and their respective pollution-tolerance values (P) as shown in the following formula:

$$BRI = \frac{\sum (\sqrt[4]{Abundance}) \times P}{\sum \sqrt[4]{Abundance}}$$

The BRI scores then are compared to categorization values to determine the community condition category of the sample, as shown in Table 4-1.

Relative Benthic Index

The RBI is calculated as the weighted sum of (a) four community parameters (total number of taxa, number of crustacean taxa, number of molluscan taxa, and number of crustacean individuals), (b) three positive indicator organisms, and (c) two negative indicator taxa. Positive indicator taxa included an amphipod (*Monocorpius insidiosus*), a bivalve (*Asthenothaerus diegensis*), and a polychaete (*Goniada littorea*), and negative indicator taxa included *Oligochaeta* and *Capitella capitata* complex. The RBI values are scaled from 0 to 1.0, with lower values indicative of higher levels of disturbance. Scores are compared to categorization values to determine the community condition category of the sample (Table 4-1).

Index of Biotic Integrity

Determination of the IBI involves comparisons of four community measures (total number of taxa, number of molluscan taxa, abundance of *Notomastus* sp., percentage of sensitive taxa) to reference conditions for southern California marina bays (Table 4-2). For every metric that exceeds a reference condition, the IBI value is increased by a score of one; therefore, IBI values potentially range from 0 to 4, with lower values indicative of lower levels of disturbance (Table 4-1).

Table 4-2. Reference ranges for IBI metrics in southern California marine bays

Metric	Reference
Total Number of Taxa	13 to 99
Number of Mollusc Taxa	2 to 25
Abundance of <i>Notomastus</i> sp.	0 to 59
Percentage of Sensitive Species	19 to 47.1

River Invertebrate Prediction and Classification System Index

The RIVPACS index is used to compare the sample assemblages (Observed) to reference species compositions (Expected) from a similar habitat. Calculation of the RIVPACS score involves three steps. (1) The probability of the test sample belonging to the 12 southern California marine bays reference sample groups is calculated. (2) The identity and expected number of reference species are determined based on the probabilities of group membership. (3) The observed number of reference species in the sample is totaled, and then the Observed/Expected RIVPACS score is calculated for comparisons to benthic community categorization values (Table 4-1).

4.6 Future Special Studies

The findings of the Pilot Project confirm that copper is a contaminant of concern within marinas and zinc also may be a contaminant in both marina and freshwater-influenced strata, since both metals occurred at concentrations that have the potential to adversely affect marine organisms (i.e., above ER-Ms). Based on these findings, special studies are recommended to (1) quantify the spatial extent of copper pollution surrounding marinas; (2) quantify the level of copper flux from marina sediments, (3) assess the bioavailability of copper, zinc, and other contaminants of

concern using bioaccumulation studies and the biotic ligand model (BLM); and (4) identify the causes of high toxicity levels in areas where they occur through sediment toxicity identification evaluations (TIEs).

4.6.1 Spatial Extent of Marina Copper Contamination

To determine the spatial extent of copper contamination within and adjacent to marinas, it is recommended that additional sediment and surface water sampling be performed in the four harbors as part of a gradient design. This study calls for the collection of water samples and sediment cores to determine copper concentrations at increasing distances from marinas. Copper within sediment samples will be analyzed using fingerprinting and microscopy to determine the source of particulates where practicable. It is hypothesized that copper concentrations will decline with distance from marinas since boating activities appear to be an important source of copper to harbor waters and sediments. This study will help assess the first component of SDRWQCB Question 1: What are the contributions and spatial distributions of inputs of pollutants to harbors in the San Diego Region?

4.6.2 Sediment Copper Flux

Given the long history of boat-maintenance activities within the marinas, including hull cleaning and sloughing of paint from boats, the build up of copper within sediments can cause a feedback loop where the sediments become a source of dissolved copper back into the water column. To quantify whether marina sediments are serving as a sink or source of copper, a laboratory experiment is recommended. Sediment cores collected within marinas will be transferred to the laboratory, placed in chambers, and exposed to seawater. Diffusive gel technology (DGT) disks will be placed just above the sediment-water interface to measure the release of copper from sediments. A control will also be included where diffusive gel disks are exposed to seawater without sediments to quantify the amounts of copper that may accumulate in disks from seawater alone. Dissolved and total copper concentrations will be quantified before and after experiments to provide an additional measure of copper release from sediments. If the amount of copper trapped in disks is greater in the presence of sediments than in their absence, it can be concluded that marina sediments have sufficiently high copper concentrations to be serving as a source of copper to the water column. This study will help assess the first component of SDRWQCB Question 1.

4.6.3 Bioaccumulation

Mussel Watch

The bioavailability of contaminants within the water column, sediments, and biota can be assessed by several methods. First, bioaccumulation studies can be used to quantify the uptake of pollutants in the water column by measuring tissue concentrations of contaminants within filter-feeding test organisms. The Mussel Watch Program conducted by the National Oceanographic and Atmospheric Administration (NOAA) is already being performed in San Diego Bay. Extension of the program to the other harbors can be used to assess contamination issues in all RHMP harbors, providing a link for direct comparisons to a long-term dataset. Therefore, a mussel bioaccumulation study following the protocols of NOAA's Mussel Watch can help address SDRWQCB Questions 4 and 5: Do the waters and sediments in the harbors sustain healthy biota, and what are the long-term trends in water quality for each harbor?

Biotic Ligand Model

The Biotic Ligand Model (BLM) can be used to calculate metal speciation and predict metal toxicity in marine and aquatic systems. Currently, the BLM is being developed for a variety of metals, including copper, silver, cadmium, zinc, nickel, and lead. The model takes into account water chemistry factors such as hardness, salinity, specific ion levels, hydrogen ion concentration (pH), alkalinity, and dissolved organic carbon (DOC) to determine the projected level of toxicity for a particular metal as measured by the metal's binding affinity to a biotic ligand (for example, the gills of an aquatic organism) (Niyogi and Wood, 2004). Therefore, the copper BLM can be used to assess the potential toxicity of copper in the harbors based on metal concentrations as well as the aforementioned physical parameters. The use of the BLM will help answer Question 4.

4.6.4 Sediment Toxicity Identification Examination

At stations where toxicity levels are determined to be high (i.e., > 50% *E. estuarius* mortality), sediment TIEs can be used to experimentally examine the constituents likely to cause toxic effects. Sediment TIEs can be performed using the amphipod *E. estuarius* to assess lethal effects and the mussel *M. galloprovincialis* to assess sublethal effects. Based on the levels of toxicity observed in the Pilot Project, only station D105F exceeded the *E. estuarius* mortality trigger for TIEs. This station was located in the freshwater-influenced stratum of Dana Point Harbor and was surveyed in 2005. Inclusion of TIEs when triggered will help answer Question 4, since they will identify the indicator(s) that are exerting a toxic effect on biota.

5.0 REFERENCES

- Bay, S.M., T. Mikel, K. Schiff, S. Mathison, B. Hester, D. Young, and D. Greenstein. 2005. Southern California Bight 2003 Regional Monitoring Program: I. Sediment Toxicity. Southern California Coastal Water Research Project, Westminster, CA.
- Brown, P. and R. Schottle. 2006. Copper Loading Assessment from In-water Hull Cleaning Following Natural Fouling Shelter Island Yacht Basin San Diego Bay San Diego, California.
- Cohen, J. 1977. Statistical Power Analysis for the Behavioral Sciences, Revised Edition. Academic Press, New York. 474 p.
- California Toxics Rule (CTR). 2000. Water Quality Standards; Establishment of Numeric Criteria for Priority Toxic Pollutants for the State of California. Federal Register, Vol. 65, No. 97, pp 31682-31719.
- Long, E.R., G.I. Scott, J. Kucklick, M. Fulton, B. Thompson, R.S. Carr, K.J. Scott, G.B. Thursby, G.T. Chandler, J.W. Anderson, and G.M. Sloane. 1995. Final Report. Magnitude and extent of sediment toxicity in selected estuaries of South Carolina and Georgia. NOAA Technical Memorandum NOS ORCA: 178p.
- McPherson, T.N and G.B. Peters. 1995. The effects of copper-based antifouling paints on water quality in recreational boat marinas in San Diego and Mission Bays. California Regional Water Quality control Board San Diego Region.
- Niyogi, S. and C.M. Wood. 1995. Biotic Ligand Model, a Flexible Tool for Developing Site-Specific Water Quality Guidelines for Metals. Environmental Science and Technology. Vol. 38. No. 23. pp 6177-6192.
- Parsons, T.R. and M. Takahashi. 1973. Biological Oceanographic Processes. Pergammon Press. New York. 186 p.
- Ranasinghe, J.A., D.E. Montagne, R.W. Smith, T.K. Mikel, S.B. Weisberg, D. Cadien, R. Velarde, and A. Dalkey. 2003. Southern California Bight 1998 Regional Monitoring Program: VII. Benthic Macrofauna. Southern California Coastal Water Research Project. Westminster, CA. 91 p + 9 Appendices.
- Ranasinghe, J.A., A.M. Barnett, K. Schiff, D.E. Montagne, C. Brantley, C. Beegan, D.B. Cadien, C. Cash, G.B. Deets, D.R. Diener, T.K. Mikel, R.W. Smith, R.G. Belarde, and S.D. Watts, S.B. Weisberg. Southern California Bight 2003 Regional Monitoring Program: III. Benthic Macrofauna. Southern California Coastal Water Research Project. Westminster, CA.
- San Diego Regional Water Quality Control Board (SDRWQCB). 2001. Results of Sampling Survey at Shelter Island Yacht Basin for Dissolved Copper Concentrations, February 15, 2001.

San Diego Regional Water Quality Control Board (SDRWQCB). 2005. Total maximum daily load for dissolved copper in Shelter Island Yacht Basin, San Diego Bay. Technical report. Resolution No. R9-2005-0019.

Schiff, K. and S. Weisberg. 1999. Iron as a reference element for determining trace element enrichment in Southern California coastal shelf sediments. *Marine Environmental Research* 48:161-176

Schiff, K.C., J. Brown, and D. Diehl. 2006. Extent and magnitude of copper contamination in marinas of the San Diego region, CA USA. Southern California Coastal Water Research Project, Westminster, CA.

Schiff, K., K. Maruya, and K. Christenson. 2006. Southern California Bight 2003 Regional Monitoring Program: II. Sediment Chemistry. . Southern California Coastal Water Research Project, Westminster, CA.

Smith, R.W., J.A. Ranasinghe, S.B. Weisberg, D.E. Montagne, D.B. Cadien, T.K. Mikel, R.G. Velarde, and A. Dalkey. 2003. Extending the southern California Benthic Response Index to assess benthic conditions in bays. Technical Report No. 410. Southern California Coastal Water Research Project. Westminster, CA. 36p. plus appendices.

State Waters Resources Control Board – California Environmental Protection Agency. 2007. Draft Staff Report, Water Quality Control Plan for Enclosed Bays and Estuaries. September 27.

Wenning, R.J., G.E. Batley, C.G. Ingersoll, and D.W. Moore, editors. 2005. Use of sediment quality guidelines and related tools for the assessment of contaminated sediments. Pensacola (FL): Society of Environmental Toxicology and Chemistry (SETAC). 815p.

Weston Solutions, Inc. 2005a. Harbor Monitoring Program for San Diego Region. Identification of Indicators to be Sampled and Mapping of Strata. Prepared for Port of San Diego, City of San Diego, City of Oceanside, and County of Orange. Prepared by Weston Solutions, Inc. August 2005.

Weston Solutions, Inc. 2005b. Progress Update. Harbor Monitoring Program for San Diego Region. Establishment of Preliminary Reference Ambient Values and Preset Target Percentages. Prepared for Port of San Diego, City of San Diego, City of Oceanside, and County of Orange. Prepared by Weston Solutions, Inc. March 2005.

Weston Solutions, Inc. 2006. Regional Harbor Monitoring Program Pilot Project 2005-06. Prepared for Port of San Diego, City of San Diego, City of Oceanside, and County of Orange. June 2006.

Weston Solutions, Inc. 2007. Regional Harbor Monitoring Program Pilot Project 2006-07. Prepared for Port of San Diego, City of San Diego, City of Oceanside, and County of Orange. June 2007.

Zar, J.H. 1999. Biostatistical Analysis. 4th Edition. Prentice-Hall, Inc., Upper Saddle River, NJ. 931 pp.