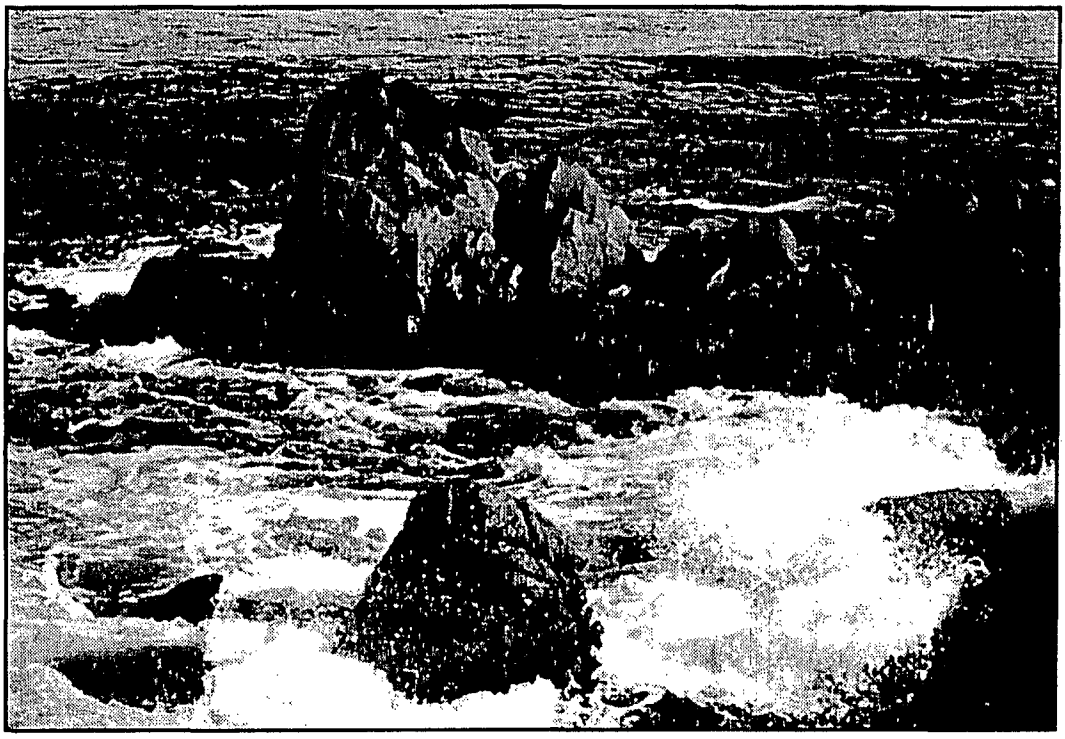


Scientific Planning and Review Committee:

# Recommendations on The Bay Protection and Toxic Cleanup Program Monitoring Activities

January 1997

State Of California
State Water Resources Control Board
Regional Water Quality Control Boards
Department of Fish and Game

**STATE WATER RESOURCES CONTROL BOARD**
**CALIFORNIA ENVIRONMENTAL PROTECTION AGENCY**

STATE OF CALIFORNIA
STATE WATER RESOURCES CONTROL BOARD
REGIONAL WATER QUALITY CONTROL BOARDS
DEPARTMENT OF FISH AND GAME

# SCIENTIFIC PLANNING AND REVIEW COMMITTEE:

# RECOMMENDATIONS ON THE BAY PROTECTION AND TOXIC CLEANUP PROGRAM MONITORING ACTIVITIES

JANUARY 1997

# PREFACE

The Scientific Planning and Review Committee (SPARC) was convened by the State Water Resources Control Board's Bay Protection and Toxic Cleanup Program (BPTCP) to review the scientific aspects of the Program's monitoring activities. SPARC has held two meetings. This report summarizes the SPARC recommendations.

The SPARC recommendations have been used by the BPTCP staff to (1) improve the Statewide monitoring approach and the Program's Quality Assurance Project Plan, (2) develop better ways to effectively identify polluted sites, and (3) train the scientists employed by the Department of Fish and Game, the Regional Water Quality Control Boards and the State Water Resources Control Board to provide more informed assessments of polluted sites.

## SCIENTIFIC PLANNING AND REVIEW COMMITTEE

Dr. Rick Swartz
Environmental Protection Agency

Ms. Rachel Friedman-Thomas
Washington State Department of Ecology

Dr. Bruce Thompson
San Francisco Estuarine Institute

Dr. Mel Suffet
Environmental Science and
Engineering Program
University of California, Los Angeles

Dr. John Knezovich
Health & Ecological Assessment
Lawrence Livermore Laboratory

Dr. Don Stevens
Environmental Statistics
and Aquatic Monitoring
ManTech Environmental
Research Services, Inc.

Dr. Ed Casillas
Environmental Conservation Division
National Marine Fisheries Service

TABLE OF CONTENTS


PREFACE . . . . . i

LIST OF APPENDICES . . . . . iv

EXECUTIVE SUMMARY . . . . . v

INTRODUCTION . . . . . 1

APRIL 1995 RECOMMENDATIONS . . . . . 2
Focus of the April 1995 Workshop . . . . . 2
Recommendations . . . . . 2
Issue 1. Toxicity . . . . . 2
Selection of Reference Sites Within Each Region . . 3
Issue 2. Association of Chemistry and Biological Effects . . . . . 4
Issue 3. Benthic Impacts . . . . . 4
Replication of Benthic Ecological Analysis . . . . 4
Issue 4. What is the most appropriate sampling design . 5
Field Replication . . . . . 5
Issue 5. Toxic Hot spot designation (Screening and Confirmation approach) . . . . . 5
Issue 6. Appropriate Biological Methods . . . . . 6
Biomarkers . . . . . 7
Issue 7. Appropriate Chemical Methods . . . . . 7
Metals . . . . . 7
Organics . . . . . 7
Region-specific Recommendations . . . . . 8
Region 1 . . . . . 8
Region 2 . . . . . 8
Region 5 . . . . . 8

MAY 1996 RECOMMENDATIONS . . . . . 9
Focus of the May Workshop . . . . . 9
Recommendations . . . . . 9
Issue 1: Determination of Significant Toxicity Relative to the Surrounding Water Body . . . . 9
Issue 2: Selection of Reference Sites . . . . . 11
Issue 3: Proposed Tiered Comparison to Determine Significant Toxicity . . . . . 13
Issue 4: Central Valley Monitoring . . . . . 14
Issue 5: Organic Chemistry Issues . . . . . 14
Issue 6: Bioaccumulation . . . . . 16
Issue 7: Benthic Community Analyses . . . . . 16
Issue 8: Weight of Evidence Approach . . . . . 17
Issue 9: Toxicity Identification Evaluations (TIEs) . . . . . 19

MAJOR SUMMARY RECOMMENDATIONS OF THE
SCIENTIFIC PLANNING AND REVIEW COMMITTEE . . . . . 20
Major SPARC Recommendations (from the 1995 meeting) . . 20
Major SPARC Recommendations (from the 1996 meeting) . . 21

CONTRIBUTORS TO THE SUMMARY DOCUMENT . . . . . 23

LIST OF APPENDICES

APPENDIX A: Scientific Planning and Review Committee Briefing Document for the Bay Protection and Toxic Cleanup Program (March 1995)

APPENDIX B: Scientific Planning and Review Committee Briefing Document for Recommendations on the Bay Protection and Toxic Cleanup Program Monitoring Activities (May 1996)

# EXECUTIVE SUMMARY

The Scientific Planning and Review Committee (SPARC) was established by the State Water Resources Control Board in 1994 to review the scientific aspects of the Bay Protection and Toxic Cleanup Program (BPTCP) monitoring activities. The SPARC members are independent experts representing the fields of toxicology, benthic ecology, organic and inorganic chemistry, program implementation and direction, experimental design, and statistics. This report contains the recommendations of the SPARC that were solicited at technical workshops held April 12-13, 1995 and May 15-17, 1996. This report also contains the briefing documents provided to the SPARC prior to the two workshops.

During the two meetings the SPARC made over 100 recommendations on all aspects of BPTCP monitoring. The SPARC discussed approaches for interpreting the toxicity, chemistry, and benthic data collected during the BPTCP monitoring efforts. SPARC also addressed bioaccumulation of contaminants and several Region-specific issues. While differences of opinion are shared among the members, the SPARC reached a strong consensus on the BPTCP monitoring and data interpretation approaches.

There was a strong vote of confidence by SPARC for using a triad of measures (i.e., toxicity testing, sediment chemical measures, and assessments of benthic organisms) to identify the worst toxic hot spots. There was also agreement on the criteria for identifying toxic hot spots using the triad of measures.

Overall, it was clear that the SPARC endorsed the BPTCP's approaches for monitoring and data interpretation. SPARC also encouraged the BPTCP to publish the results of the monitoring efforts in peer-reviewed scientific literature.

# BAY PROTECTION AND TOXIC CLEANUP PROGRAM
# SCIENTIFIC PLANNING AND REVIEW COMMITTEE

## INTRODUCTION

The California Water Code established the Bay Protection and Toxic Cleanup Program (BPTCP) to protect the existing and future beneficial uses of California's bays and estuaries. The BPTCP has provided a new focus on identifying polluted and contaminated locations in California's bays and estuaries. The BPTCP has four major goals: (1) protect beneficial uses of bay and estuarine waters; (2) identify and characterize toxic hot spots; (3) plan for the prevention and control of further pollution at toxic hot spots; and (4) develop plans for remedial action at existing toxic hot spots and prevent the creation of new hot spots. The primary focus of the BPTCP has been on the identification of toxic hot spots.

The SWRCB established the SPARC in 1994. The SPARC brings together independent experts in the fields of toxicology, benthic ecology, organic and inorganic chemistry, program implementation and direction, experimental design, and statistics to review the monitoring approaches taken by the BPTCP. The committee has provided comments on the Program's monitoring approach(es), given input on the scientific merit of the approach(es) taken, and provided suggestions for monitoring improvement.

In 1995 and 1996 the Bay Protection and Toxic Cleanup Program (BPTCP) sponsored two meetings of the Scientific Planning and Review Committee (SPARC). The purpose of this report is to present the recommendations provided by the SPARC.

SCIENTIFIC PLANNING AND REVIEW COMMITTEE

APRIL 1995 RECOMMENDATIONS

## Focus of the April 1995 Workshop

The workshop centered around the following key questions:

1. What is toxic?
2. How should we show association between toxicity, benthic community, etc. and chemical concentrations?
3. What is a benthic impact?
4. Should we use a probability-based sampling design (random sampling) or directed point sampling approach (i.e. based on best professional judgment)?
5. Should we use a screening and confirmation approach?
6. What biological methods should we use?
7. What chemical methods should we use?

Please refer to Appendix A for the issue papers that describe each of these issues.

## Recommendations

The SPARC recommendations from the April 1995 meeting were:

### Issue 1. Toxicity

1. The selection of toxic and reference sites will ultimately be a policy decision based on best available scientific approaches for determining biological response.
2. The reference envelope approach is preferred over simple comparison to laboratory controls, and there is agreement that this is the statistical approach to pursue for determining the level of toxicity suitable for meeting toxic hot spot toxicity criterion.
3. All toxicity data should be normalized to laboratory controls to account for any variation in laboratory factors or test organism condition.
4. Compare test site response to large reference envelope population from a comprehensive data base of reference site results for the protocol used.

5. Compare test site response to reference envelope population from samples collected concurrently with test samples.

6. A site is toxic if it falls below the reference envelope lower bounds for both the reference site data base and concurrent samples.

7. If a site is toxic relative to the large reference envelope population from the comprehensive database, but concurrent reference site results are low, the site should be revisited.

Selection of Reference Sites Within Each Region

Some level of pollution will always be unavoidable. However, reference sites should be selected through the following process:

1. Reference sites should not include those sites where toxicity is observed in association with pollution. Common sense and knowledge of local conditions should be used in order to avoid areas known to be disturbed or polluted.

2. Randomly sample the rest of the water body, conducting analyses of chemistry, benthic community structure, and toxicity.

3. Allow trained benthic ecologists to select the sites that have moderate to high species richness, abundant presence of amphipods or other indicator species, absence of indicators known to be characteristic of polluted sediments, and any other indicator of ecological health that can be argued convincingly.

4. Evaluate the chemistry data and narrow the sites to those that do not exceed more than one upper value of a PEL or ERM for existing chemistry guidelines.

5. Evaluate the toxicity data and eliminate only those sites that have extremely high toxicity, as determined by a qualified toxicologist, not by a priori criteria.

6. Once reference sites are chosen they are sampled along with test sites. Include the new reference site toxicity results in the reference envelope regardless of the magnitude of the toxicity response. The reference envelope toxicity result will fall where it may.

7. Compile a data base of toxicity responses from appropriately selected reference sites, and include past and current reference site data in the reference envelope. Allow the number of data points in the reference envelope to grow as more studies are completed in the area.

## Issue 2. Association of Chemistry and Biological Effects

1. Causal relationships between chemistry and biological effects are desirable to provide evidence of links between pollutant concentrations and biological effects. However, correlation does not necessarily establish causality.

2. Development of spiked bioassay data could be used to unequivocally identify chemicals responsible for observed effects.

3. Simultaneous Extracted Metals and Acid Volatile Sulfides (SEM/AVS) data is essential for understanding metal effects.

4. Measurement of Total and Dissolved Organic Carbon (TOC and DOC) in the pore water is recommended to help understand organic and metal bioavailability.

5. The effect of oxidation state of the environment and of the chemical compounds should be investigated.

6. Pore water toxicity and chemistry are valuable in determining causal relationships.

7. It is recognized that sorbed pollutants may become bioavailable after ingestion and metabolism.

8. Professional judgement and knowledge of local conditions should be used to decide how best to allocate resources to determine causal relationships.

9. The Program should use all available criteria and biological measurements in assessing the relationships between chemistry and biological effects (i.e., use weight of evidence approach).

## Issue 3. Benthic Impacts

No single index is defensible in a regulatory setting. A site should be characterized as "healthy", "intermediate", or "degraded" based on the best professional judgement of a qualified ecologist, using whatever methods are most appropriate to the site.

### Replication of Benthic Ecological Analysis

An analysis of existing data should be conducted to determine benthic replication, keeping in mind the types of analyses that can be done with benthic data, the cost of the analysis and benefits derived. Do not replicate unless there is a clear reason to do so. Broad spatial/temporal coverage of sampling is usually preferable to replication at fewer stations/times.

**Issue 4. What is the most appropriate sampling design**

1. During the screening phase, sampling should incorporate a stratified random design in order to provide an opportunity to find unknown toxic hot spots.

2. Confirmation phase sampling should be based on grids covering the site of concern, with random placements of stations within grid blocks.

3. Grids should be configured to match site characteristics.

4. Temporal variations should be accounted for with repeated sampling at locations at least one meter apart.

5. Spatial and temporal scales should be based on knowledge of the site.

Field Replication

6. Random sampling over suitably sized grids may be preferable to replication. There is no need to replicate unless there is a clear and defensible reason why.

7. It would be best to conduct statistical analysis of past data to determine replication needs for future work.

**Issue 5. Toxic Hot spot designation (Screening and Confirmation approach)**

1. A three tiered data analysis approach should be used. This would include chemical, toxicity, and benthic community analyses. Having hits in all three components of a triad analysis, would classify a site as a worst case toxic hot spot. Hits on fewer than all three would result in classification as a site of concern. All sites could be ranked in this way.

2. Under the BPTCP, the screening phase would consist of using either toxicity or benthic community analysis or chemistry or bioaccumulation data or some combination of all of these. Screening should be flexible, designed to fit the Regional Board's needs. Analysis in this phase should be done only when needed to provide sufficient information to convince the Regional Boards to list or consider the site as a priority site of concern for further action. A hit in any of these analyses would elicit concern, trigger confirmation phase monitoring under the BPTCP and/or perhaps prompt a specific Regional Board to pursue some other type of regulatory review action. It would be very important to involve potential responsible parties as early in the process as possible and coordinate studies and funding.

3. The confirmation phase should consist of toxicity and chemistry and benthic community analyses on a previously visited site of concern or wherever previous evidence indicates a site may be impacted. A confirmatory hit in all three analyses performed during this phase would classify a site as a worst case toxic hot spot. This phase could also include intensive investigations to identify causal relationships, and intensive grid sampling necessary to show gradients and spatial extent.

4. Allow for a mechanism for de-listing sites if intensive studies prove preliminary designation was in error.

5. It is important to focus on the most impacted sites for successful toxic hot spot designation and application of regulatory actions.

**Issue 6. Appropriate Biological Methods**

1. Use the amphipod 10 day solid phase test and the sea urchin 96 hour larval development test in pore water for screening sites.

2. Use the amphipod solid phase test, the sea urchin larval development test in pore water, and the sea urchin larval development test at the sediment water interface (SWI) for confirmation. (A sensitive chronic test, such as the 28 day protocol for Leptocheirus, or tests using resident species may also be useful for confirmation).

3. Centrifuge pore water for bioassay test. Use non-sorbing centrifuge tubes such as stainless steel, glass and/or Teflon. Frozen storage is not acceptable for biological testing.

4. Pore water dilutions are not necessary for screening, but do provide additional information for confirmation.

5. Pore water toxicity coupled with chemical analyses may be useful for establishing correlations between chemistry and biological effects.

6. Use of the Neanthes test should be discontinued because it provides no additional information beyond that provided by the amphipod and sea urchin protocol.

7. Studies should be conducted to investigate whether inhibition of embryo/larval development in pore water and solid phase (SWI) exposures can be correlated, or is associated with ecological perturbation, such as impacts on benthic community structure.

Biomarkers

1. Biomarker analyses are currently difficult to interpret in terms of ecological effects. These types of analyses should not be used for toxic hot spot designation at present.

2. Biomarker analyses may be useful in monitoring cleanup activities to determine if there is continued exposure to pollutants.

Bioaccumulation

Recruit the services of a bioaccumulation expert into SPARC and examine how bioaccumulation can be used in the BPTCP.

**Issue 7. Appropriate Chemical Methods**

Metals

1. Perform SEM/AVS with caution in evaluating potential for metal toxicity. This value may change over time at individual sites due to fluctuations in the concentration of AVS.

2. Use performance-based approach rather than rigid protocols.

3. Do bulk-phase metals in screening.

4. Do pore water metals when deemed necessary. It may help determine causality for confirmation and cleanup planning.

5. Preserve original samples for pore water chemistry.

6. Sediment extracts can be frozen for a year for chemical analysis. The time listed in standard methods for water and waste water should be the maximum holding time (Mel Suffet, personal communication, December 1996).

Organics

The April 1995 meeting ended before the organic chemical methods could be fully discussed. Nevertheless, similar recommendations to metal chemical methods were made. Further examination of this topic is scheduled for the next SPARC meeting.

1. The analyte list should be expanded to include Diazinon and other organophosphate pesticides

2. Use performance-based approach rather than rigid protocols.

3. Do bulk-phase organics and TOC in screening.

4. Do pore water organics to help determine causality for confirmation and cleanup planning.

5. Preserve original samples for pore water chemistry.

6. Sediment extracts can be frozen for a year for chemical analysis.

**Region-specific Recommendations**

Region 1

If local problems can be identified without toxicity screening then proceed to use the available resources as effectively as possible.

Bioaccumulation data may be appropriate to identify problem chemicals, biological exposure and potential sources of pollution in Region 1.

Biological effects measurements (toxicity screening or benthic community analysis) should be considered in cases where unknown toxic hot spots are present.

Region 2

Sampling should be done at a predetermined standard depth in a way to avoid mixing oxic and anoxic sediments. It would be desirable to show the effects of changes in oxidation state on toxicity and toxicity/chemistry relationships.

Use appropriate amphipod species based on knowledge of species tolerance limits to ammonia, salinity, and grain size.

Determine how to include bioaccumulation data into toxic hot spot screening.

Region 5

Pursue monitoring of pesticide degradation products.

Request that the SWRCB, Regional Boards, and Federal agency executive management agree to coordinate monitoring programs and share information from studies in the Bay-Delta. Also that the two Regional Boards pursuing BPTCP work in the Bay-Delta coordinate the planning and monitoring work.

SCIENTIFIC PLANNING AND REVIEW COMMITTEE

MAY 1996 RECOMMENDATIONS

## Focus of the May Workshop

The topics discussed in the May meeting addressed the following topics:

1. Review and incorporation of the SPARC recommendations into the Statewide monitoring approach.

2. Interpretation of toxicity data collected.

3. Interpretation of the benthic community data collected.

4. Setting priorities using a weight-of-evidence approach.

5. Review of the studies of water column toxicity and chemistry in the Central Valley Region.

6. Completion of the discussion on organic chemistry methods.

7. The use of bioaccumulation monitoring techniques.

The briefing document that describes each of these issues is presented in Appendix B.

## Recommendations

The workshop centered around the following key issues:

### Issue 1: Determination of Significant Toxicity Relative to the Surrounding Water Body

1. There is consensus support for the reference envelope concept because it includes all sources of laboratory and field variation affecting toxicity test results.

2. Unexplained toxicity in samples from reference sites should be considered a problem if it occurred in more than 25% of reference samples, and should not be considered a problem if it occurred in less than 10%. There was no SPARC resolution on how to use the reference envelope approach if unexplained toxicity occurred in 10%-25% of reference site samples.

3. Investigation of unexplained toxicity should be focussed on identifying either: (a) pollutants that have not been considered previously, or (b) natural toxicity. Identification of either would be a significant finding consistent with program goals.

4. The synergistic effect of mixtures of chemicals found at low concentrations should be considered in any investigation of unexplained bioeffects.

5. The reference envelope should include toxicity data from many different sampling times. Temporal variability should be investigated. If temporal variance exists (i.e., if multiple sites vary concurrently), then the reference envelope equations must be revised to take this factor into account.

6. The reference envelope for toxicity could include reference sites from a broad geographical area (as big as the entire West Coast) or be limited to the local study area, depending on study objectives.

7. Statistical power should be analyzed to determine the minimum number of reference site samples necessary for appropriate use of the reference envelope method. Effects of sample size on data distribution (e.g., normality) should also be examined.

8. To determine statistical significance, study site results should be compared to both:

   a. the tolerance limit derived from a reference envelope that includes previous data, and

   b. results from concurrently collected local reference site sample(s).

9. Regional Boards should set reference envelope "p" values appropriate for their Regions and study objectives. The "p" is the percentile of the reference distribution used to set tolerance limits. There was SPARC consensus that this value is critical in establishing toxicity thresholds, provides an explicit means of selecting the statistical parameters relevant to study objectives, and should be established through policy decisions.

10. Guidelines for selection of "p" values include:

    a. the degree of confidence that reference site samples are indicative of desired ambient water body conditions,

    b. the level of degradation exhibited by reference site samples, and

    c. the political or economic goals associated with designating study sites as toxic.

    Low "p" values would be appropriate for situations where there is high confidence that reference sites are indicative

of desired environmental conditions, and the economic or political costs related to a finding of toxicity are high. Higher "p" values are more appropriate when reference sites are assumed to represent less than optimal conditions, or when policy impacts are less severe.

11. Economic analyses could be used in conjunction with information on reference site quality and regulatory goals to help establish suitable "p" values for reference envelope calculations.

12. There may be greater uncertainty associated with the use of low "p" values. The lower the "p" value, the farther it extends into the tail of the reference population distribution, where deviations from normality are most extreme. This should be investigated as part of an examination of sample size and data distribution.

13. The reference envelope approach is strongly tied to an assumption of normality of the underlying data distribution, and that distribution should be checked as a matter of routine. Any suggestion of strong departure from a bell-shaped or triangular distribution (e.g., skewness, multiple modes, or a flat distribution) should be cause to use the reference envelope approach results with caution. If the reference envelope approach produces tolerance limits that are counter to best professional judgment, the following steps should be taken:

    a. Check the data distribution, transform data if necessary.

    b. Consider switching test protocols (Criteria for protocol rejection should be established).

    c. Check that reference sites were selected appropriately.

    d. Check if the "p" value is appropriate. This may involve re-evaluation of reference sites, program goals, and/or policy considerations.

    e. If unexplained reference site toxicity exists, investigate it. Do not use a statistical test based on reference site data that are poorly understood.

**Issue 2: Selection of Reference Sites**

1. Do not consider nickel in evaluating reference site chemical pollution. However, use common sense in cases with highly elevated nickel concentrations.

2. While evaluation of SEM - AVS (simultaneously extracted metals minus acid volatile sulfide) is useful in evaluating potential for metal toxicity in reference samples, this

value may change over time at individual sites due to fluctuations in the concentration of AVS. In addition, generalizations regarding AVS effects on bioavailability may not apply to all toxic metals. The issue of whether or not AVS - SEM should be used in reference site selection was not resolved by SPARC at this meeting.

3. Effects Range-Median (ERM) and Probable Effects Level (PEL) values are very similar. The lower of the two should be used in screening concentrations of individual chemicals in reference site selection.

4. For reference site selection, a Total DDT concentration of 100 $\mu$g/g TOC was suggested as a cutoff value, based on toxicity studies.

5. For reference site selection, use the sum of ERM quotients that totals less than 5. This value was supported by data from numerous studies described at the meeting by Ed Long. However, all available data and criteria (including EPA EqP and lowest AET) should be evaluated, especially in cases of unexplained toxicity.

6. Benthic community data should not be the sole basis for reference site selection because:

   a. benthic community impacts can be hard to measure and/or interpret,

   b. the community may have adapted to pollutants, and

   c. relatively healthy benthic communities can exist in surface layers above polluted strata.

7. There was no resolution on the use of toxicity data in reference site selection. Contrasting issues of unexplained toxicity and potential for subjective data screening could not be resolved by the entire committee.

8. $H_2S$ and $NH_3$ at reference sites:

   a. Use toxicity test species that can tolerate reference site concentrations.

   b. Use exposure systems that can minimize reference site concentrations (e.g., Sediment Water Interface tests).

   c. $H_2S$ and $NH_3$ are less of an issue with amphipods than with embryos or larvae exposed in pore water tests.

   d. The program should use written guidelines for rejecting reference sample toxicity data when $H_2S$ or $NH_3$ are above threshold values for test species.

**Issue 3: Proposed Tiered Comparison to Determine Significant Toxicity**

Significant toxicity relative to the surrounding water body should be determined by comparing the test sample result to:

1. a tolerance limit calculated from a "universal" reference distribution, and/or a tolerance limit calculated from a "local" reference distribution,

2. results from one or more concurrently collected local reference site sample(s), and

3. 80% of the laboratory control survival.

Significant toxicity would be indicated if the sample result was below the tolerance limit selected for the study (either "universal" or "local" or both, above), and significantly lower than the result from a concurrently collected reference site sample (using a one-tailed t-test), and the sample mean survival was less than 80% of the laboratory control mean. (A "universal" reference distribution refers to one derived from sites from a broad geographical area, such as the entire West Coast of the United States.)

The first comparison [to the reference envelope tolerance limit(s)] accounts for all sources of laboratory and field variation affecting toxicity test results. The second comparison addresses the possibility of a unique toxicity event occurring in the water body at the time of sampling. The third comparison precludes a determination of toxicity when a statistically significant difference is smaller than generally believed to be biologically relevant.

The following should be considered in selecting local versus universal reference populations:

a. The "universal" envelope should be used if local reference site sample results fall within the "universal" reference envelope.

b. In "cleaner" areas or Regions, the local reference envelope should take precedence over the "universal".

c. In areas where local reference samples are more toxic than "universal" reference samples, Regional Board staff should select the reference distribution appropriate to meet study objectives.

**Issue 4: Central Valley Monitoring**

1. Consider measuring selenium.

2. Mercury is likely to become bioavailable in areas where high residence time allows methylation.

3. Mercury source tracking, *Ceriodaphnia* toxicity studies, and TIEs were well done. Suggestions for obtaining additional evidence for pesticide effects:

   a. Benthic communities should be evaluated and linked to toxicity.

   b. Water column community effects should be linked to toxicity.

   c. Investigate effects on Salmonid prey species and larval fish.

   d. Investigate sediment toxicity tests with flow-through site water.

   e. Model hydraulic system inputs and flow to further demonstrate fate.

4. EPA staff working with the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) should be made aware of pesticide data to allow better coordination of management programs.

5. Coordinate Delta toxicity studies with California Endangered Species Act studies.

**Issue 5: Organic Chemistry Issues**

1. The SPARC supports the modification of current BPTCP organic analytical procedures to allow additional analytes to be measured from a single extraction, thereby expanding the analyte list in a cost effective way.

2. Additional analytes of concern that the program should consider measuring include:

   a. Cholinesterase inhibitors, such as the organo-phosphates diazinon and chlorpyrifos, and the carbamate carbofuran. BPTCP currently looks for chlorpyrifos but not the others (e.g., carbamates (methomyl) are used heavily in Elkhorn Slough). Organo-phosphates are important in Regions 5 and 2, and probably elsewhere.

c. Triazines (Atrazine in particular). Both Atrazine and Simazine are used in California. These are highly phytotoxic compounds.

d. Higher molecular weight polynuclear aromatic hydrocarbons (HMW PAHs) may be appropriate to add, though consideration should be given to determining the best HMW PAHs to add.

e. Nonylphenolic surfactants are estrogenic compounds which appear to have synergistic effects at low concentrations, and bioaccumulate. Analytical methods are poorly defined but these compounds may come through our current methods.

f. Alachlor and pthalates.

3. Sample matrix is important. As a guideline, for compounds with a low to moderate Log Octanol/Water Partition Coefficient (Log $K_{ow}$), it would be more useful to analyze for diazinon in water, pore water, and tissue rather than in sediment. Whereas for moderate to high Log $K_{ow}$, it would be best to measure the sediment and tissue rather than the aqueous phase.

4. PAH fingerprinting can be added to BPTCP analyses for minimal cost. All PAH signatures are not created equal. Rather than comparing the sum of 26 compounds in samples with different PAH profiles, the BPTCP should develop an index to describe a sample's PAH signature so that samples can be "typed" prior to statistical comparison.

5. Samples exhibiting bioeffects without concomitant elevated concentrations of measured chemicals (that may be related to unexplained toxicity) should be investigated to identify the source and nature of the toxicological agent in these cases.

6. For analysis of water samples, samples must be filtered using glass fiber filters. Plastic in filters actively binds organics. Total Organic Carbon (TOC), Dissolved Organic Carbon (DOC), and Total Suspended Solids (TSS) measurements should be taken on these samples in order to provide a normalizing index for analytical results. There remains an unresolved argument in the literature about filters vs centrifugation for sample analysis, but Dr. Suffet has found filtration to work well.

7. All chemistry data should continue to be reported in units of dry weight, along with normalizing factors like TOC and AVS, if possible.

**Issue 6: Bioaccumulation**

1. Bioaccumulation data and related health advisories should be used to identify chemicals of concern in a study area. The concentrations of those chemicals in test sediments should be given added consideration in the designation or ranking of sites.

2. A large area (e.g., an entire bay) can be considered an area of concern based on tissue contamination. In such cases, source control would be the preferred cleanup option, as activities such as sediment removal may be impractical.

3. Salmon should be considered for use in bioaccumulation studies.

4. Using models to back-calculate tissue concentrations affecting human and ecosystem health from sediment concentrations can lead to estimates of very low chemical concentrations of concern in sediments. However, the effects of bio-accumulating chemicals, and hot spot designation based on those chemicals, should not be totally dismissed because of low concentrations in sediments.

5. Persistence is not the only issue to consider when evaluating bioaccumulation information. Events of limited duration may still affect ecosystem and human health.

6. Fish (and other organism) tissue burdens in the Sacramento/San Joaquin River Delta should be investigated. The contamination observed in previous studies warrants an evaluation of potential risks to human and ecosystem health.

**Issue 7: Benthic Community Analyses**

1. Choice of indicator species used in BPTCP/EMAP Southern California Coastal Lagoons and San Diego Bay studies was appropriate. There was very little overlap in the presence of positive and negative indicator species.

2. Indicator species selection should be specific to study area. Indicator species should be selected prior to sample analysis, and should include species whose distributions are not limited by natural sediment characteristics likely to be found at study sites (such as grain size, TOC, etc.).

3. The following parameters should be measured (or sampled and preserved) *in situ* to assist with interpretation of benthic community analyses: grain size, salinity and concentrations of dissolved oxygen, ammonia, hydrogen sulfide, and TOC.

4. Numerical scaling of the benthic index should be re-evaluated and discussed with interested SPARC members and program staff.

5. The cutoff point indicating community degradation should not be chosen arbitrarily. Samples ranked between 1 and 2 on the present index should be individually re-evaluated to determine "degraded" status.

**Issue 8: Weight of Evidence Approach**

1. BPTCP should evaluate all three legs of the triad (chemistry, toxicity and benthic community analysis) to most effectively use the Weight of Evidence Approach. In the San Diego study, samples missing one leg of the triad should not be ranked as if there were no effect for that analysis. Missing data should be obtained before ranking all sites together, especially in cases where available data suggests possible degradation.

2. Weight of Evidence could be quantified using an approach similar to Chapman/Long's Ratio to Reference. However, it is informative to present each site with numerical values for each leg of the triad. These values could be either the data values from each analysis (such as percent survival for the toxicity tests), or the rank or percentage relative to other sites studied. These values should not be summed, but each leg should be presented individually. This was suggested in addition to color coding on maps, so that color would indicate hot spot status and numerical values would give a sense of the degree of impact.

3. The legs of triad should be applied independently and should not be expected to agree. Information from one type of analysis should not be disregarded because of different information from another type of analysis. Such cases should be evaluated individually to tease out useful information and supporting evidence.

4. It is not necessary to have two toxicity hits; toxicity, chemistry and benthic ecology should be treated equally.

5. Consider a sampling design that allows samples for all triad analyses to be taken from a single sediment grab. This allows synoptic sampling for all analyses, even if benthic or chemistry samples are archived, and could make sampling more economical.

6. High priority stations are sufficiently confirmed by the BPTCP weight of evidence approach to be considered for the next level of Regional Board or responsible party investigations. Moderate priority stations, and stations for which not all triad data are available, still need additional evidence from BPTCP triad approach prior to

follow-up by Regional Board or responsible party investigations.

7. Adjacent stations should be evaluated together to look for similar chemistry and bioeffects. A number of closely spaced sites exhibiting impacts and pollution from similar chemicals may qualify as an area of concern.

8. Confirmation should include consideration of spatial extent. Sites should be characterized by at least three stations.

9. The following points should be considered in using chemistry data in ranking sites:

   a. Do not use nickel at all (unless concentrations are extremely high) because there is little confidence in the available sediment guidelines.

   b. Use MacDonald's Palos Verdes data for DDT.

   c. Use both single chemical ERM quotients and quotient averages.

   d. Use the average of ERM or PEL quotients in applying the weight of evidence approach, as opposed to the sum of the quotient. This provides a natural cutoff point where averages exceeding 1 indicate elevated chemistry. This number should be used as a guide along with best professional judgment.

   e. Subdivide chemicals into groups likely to have additive effects to better estimate combined effects. For example, low molecular weight PAHs are likely to be additive in their biological effects.

   f. Even though the effects of many different chemicals are not always additive, combinations of chemicals are still likely to produce increased effects. ERMs and PELs do work empirically and should be used.

10. It was suggested that the BPTCP examine Washington State's algorithms for combining data to establish weight of evidence.

11. Weight of evidence assessments should always include graphical evaluation of the data.

12. The reference envelope approach has been applied to benthic community data and chemistry data (by Bob Smith). There was no consensus on whether this approach should be used by the BPTCP.

**Issue 9: Toxicity Identification Evaluations (TIEs)**

1. TIE of sediment pore water should be conducted if it furthers study objectives. TIE is especially important in establishing causal relationships.

2. The TIE approach may provide additional information to guide chemical analysis. There was general agreement that Region 5's investigation of pesticide toxicity supported the power of this approach.

3. For sediments, focus on pore water for TIEs, but realize that removing interstitial water from the sediment matrix may alter the physical availability of analytes. Sorption onto system components may effectively alter the characteristics of the sample and the outcome of the TIE. Removal of pore water from the sediment could be considered one step in the TIE process.

4. A non-filtered pore water treatment should be included in the TIE process. Total suspended solids and dissolved organic carbon are important in determining bioavailablity. These should be measured, although measuring TSS in pore water may be difficult.

5. Chemical analysis should be used as part of the TIE process to verify the compounds identified. Chemicals should be measured at the beginning and end of the TIE toxicity exposures to verify stability.

6. Be aware that there are multiple contaminants everywhere, which may confound the ability to remove toxicity in a TIE. Cumulative effects make it difficult to establish cause/effect relationships.

7. Be aware that TIE procedures may not always provide clear answers, and do not eliminate consideration of a site of concern solely on the basis of the inability of a TIE to identify responsible compounds.

MAJOR SUMMARY RECOMMENDATIONS OF THE
SCIENTIFIC PLANNING AND REVIEW COMMITTEE

**Major SPARC Recommendations (from the 1995 meeting)**

1. Base program decisions on defensible science to provide common ground for all participants and interested parties.

2. Prepare workplans in advance to allow adequate scientific review, efficient allocation of funds, and timely reporting.

3. Use a carefully considered weight-of-evidence approach to accomplish program goals.

4. Include a bioaccumulation expert on the SPARC and examine how bioaccumulation can be used in the BPTCP. Thought should be given to reconciling the two different aspects of toxic hot spot designation: human health risk vs. observed ecological effects.

5. Food web models are not sophisticated enough to allow development of sediment quality criteria based on fish tissue concentrations. The mobility of most fish species limits utility for designation of toxic hot spots on a reasonable scale.

6. Site specific investigations are necessary for toxic hot spot designations. Focus immediately on sites most likely to be successfully designated as a toxic hot spot.

7. Regional Boards must have authority and take responsibility for the planning of work in their respective regions. Local knowledge should be used to focus on the most relevant sites and analyses.

8. In designating toxic hot spots, follow a three-tiered approach: (1) carry out a flexible screening phase using any analysis of the triad or bioaccumulation technique; (2) a confirmation phase using all triad analyses (and); (3)intensive site specific studies demonstrating spatial extent, and causal relationships between pollutants and observed biological effects. It is very important to bring the potential responsible parties into the process as early as possible. Potential responsible parties, and other appropriate entities, should be brought into the process to cooperate in the funding and execution of post-confirmation studies.

9. Confirmation and intensive cleanup studies should use a stratified random sampling design, with grids of suitable size to cover the area of concern. Field replication of all measures (toxicity, chemistry, benthic community structure,

and bioaccumulation) should only be used when there is a clear and valid reason. Bioaccumulation studies should be focussed on contaminants in tissues of fish or other organisms.

10. Statistical significance of toxicity should be determined based on a comparison to a reference envelope.

11. Benthic community degradation should not be based on a single index. A single community index is too easily discredited. Benthic community degradation should be based on convincing evidence determined on a site specific basis by a qualified ecologist.

12. Performance-based chemistry should be used.

13. Pore water toxicity, concurrent chemistry and spiked assays may be useful to determine associations between pollutants and biological effects. Correlations are not nearly as convincing in demonstrating associations. The presence of multiple pollutants may complicate interpretation of toxicity test results. A TIE approach would also provide evidence of cause-effects relationships but should be used judiciously because of cost.

14. SEM/AVS are recommended for all samples.

15. Statewide and site-specific chemical objectives should be pursued.

16. Bioavailability concerns complicate interpretation of solid-phase sediment toxicity testing in evaluating the relationships between pollutant and biological effects.

17. Solid-phase sediment toxicity testing is useful for sediment quality assessment and toxic hot spot designation.

**Major SPARC Recommendations (from the 1996 meeting)**

1. The triad approach now used by the BPTCP is appropriate for identifying the most and least impacted sites, allowing the program to achieve its major goals.

2. BPTCP data collected to date allows for a scientifically defensible ranking of high priority sites. If further study, as part of confirmation or remediation, shows a site to be less of a problem than originally indicated, the site's status can be changed as part of the process. The data is currently sufficient to justify regulatory actions.

3. The State and Regional Boards should be actively cooperating with potential responsible parties to develop funding and study designs for the next level of investigation at sites identified by the BPTCP as sites of concern.

4. Moderately impacted sites should not be disregarded, especially if there are a number of moderately impacted sites in close proximity. Some action, such as source control, may be necessary even if there is not a single high priority station.

5. Sites that have significant toxicity, high chemistry, or a degraded benthic community, but are missing a leg of the triad, should be resampled to complete all three analyses. Information from sites of concern with only two legs of the triad measured should not be compared to sites with all triad components measured until the missing data are collected. Priority should not be downgraded (for sites with two legs of the triad measured) because of missing data.

6. "Other deleterious substances" (ODS), such as hydrogen sulfide, low dissolved oxygen, etc. that are likely to have resulted from human inputs should be considered as chemicals of concern.

7. The BPTCP provides a model for identifying problem sites that other states may wish to follow. SPARC encouraged the program to support publication of objectives, criteria, methods and results in the peer-reviewed literature to make them more widely accepted and available.

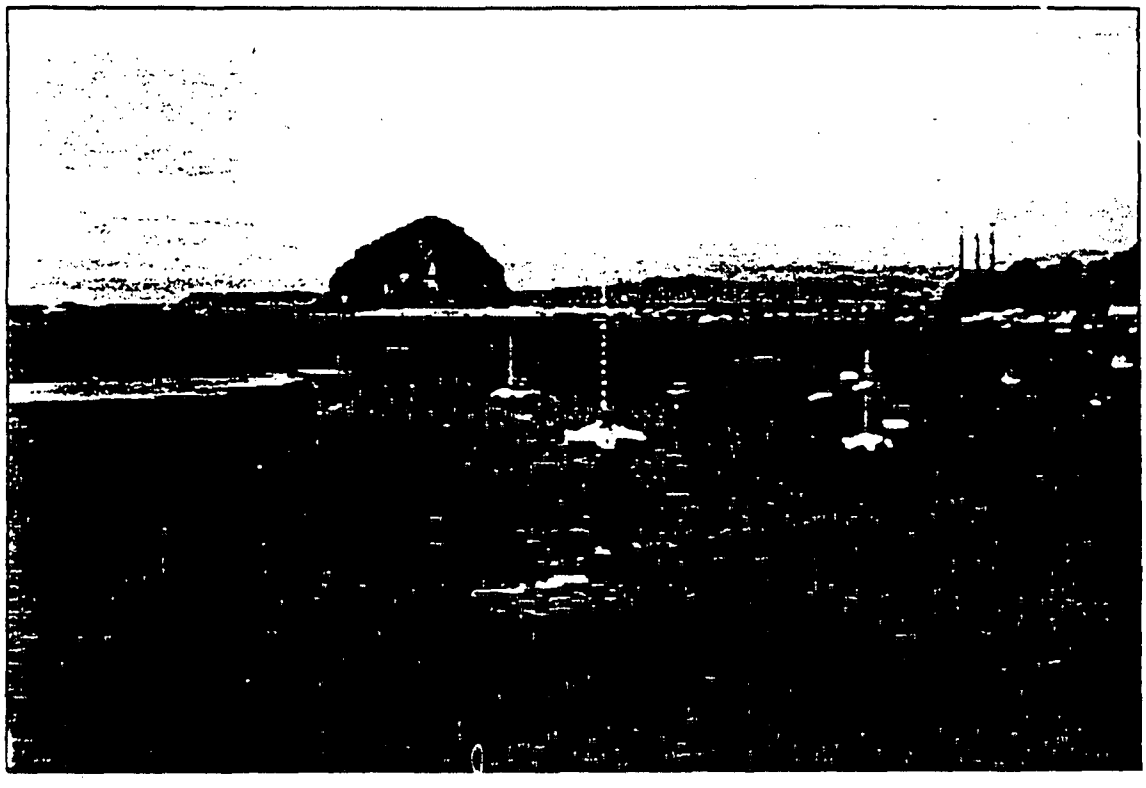
## CONTRIBUTORS TO THE SUMMARY DOCUMENT

| | |
|---|---|
| John Hunt | University of California, Santa Cruz |
| Gita Kapahi | State Water Resources Control Board |
| Brian Anderson | University of California, Santa Cruz |
| Rusty Fairey | San Jose State University |
| John Newman | University of California, Santa Cruz |
| Fred LaCaro | State Water Resources Control Board |
| Max Puckett | Department of Fish and Game |
| Mark Stephenson | Department of Fish and Game |
| Craig J. Wilson | State Water Resources Control Board |

# A P P E N D I X  A

Scientific Planning and Review Committee
Briefing Document for the
Bay Protection and Toxic Cleanup Program

March 1995



# Scientific Planning and Review Committee Briefing Document for the Bay Protection and Toxic Cleanup Program

Department of Fish and Game
Regional Water Quality Control Boards
State Water Resources Control Board

STATE OF CALIFORNIA
STATE WATER RESOURCES CONTROL BOARD
REGIONAL WATER QUALITY CONTROL BOARDS
DEPARTMENT OF FISH AND GAME

# SCIENTIFIC PLANNING AND REVIEW COMMITTEE:

# BRIEFING DOCUMENT FOR THE THE BAY PROTECTION AND TOXIC CLEANUP PROGRAM

MARCH 1995

# PREFACE

This briefing document was developed to assist the Scientific Planning and Review Committee (SPARC) in preparing for a technical workshop to review the monitoring programs of the State of California's Bay Protection and Toxic Cleanup Program (BPTCP). The purpose of the workshop is to solicit comments from the SPARC on the BPTCP monitoring approach(es), to give input on the scientific merit of the approach(es) taken, and to provide suggestions for monitoring improvement in the future.

The document is organized to focus SPARC on the most fundamental questions and concerns about the BPTCP monitoring approaches. The document presents the workshop agenda, a brief summary of the BPTCP, the overall monitoring approach to identify toxic hot spots, and issue papers describing the fundamental questions posed for SPARC including the approach used by the BPTCP. The issue papers are followed by regional summaries that generally contain specific monitoring objectives, overview of water bodies in the Region, studies completed to date or in progress, and regional questions for SPARC. The last chapter of the briefing document contains a complete list of the questions for SPARC developed by the Department of Fish and Game.

## MEMBERS OF THE SCIENTIFIC PLANNING AND REVIEW COMMITTEE

Dr. Rick Swartz
Environmental Protection Agency
2111 S.E. Marine Science Drive
Newport, OR 97365-5260

Ms. Rachel Friedman-Thomas
Washington State Department of Ecology
P.O. Box 47600
Olympia, WA 98504-7600

Dr. Bruce Thompson
San Francisco Estuarine Institute
1301 South 46th Street
180 Richmond Field Station
Richmond, CA 94804

Dr. Mel Suffet
Environmental Science and
Engineering Program
University of California, Los Angeles
10833 Le Conte Avenue
Los Angeles, CA 90024-1772

Dr. John Knezovich
Health & Ecological Assessment
Division L-453
Lawrence Livermore Laboratory
Livermore, CA 94550

Dr. Don Stevens
Environmental Statistics
and Aquatic Monitoring
ManTech Environmental
Research Services, Inc.
200 S.W. 35th Street
Corvallis, OR 97333

TABLE OF CONTENTS

PREFACE . . . . . ii

MEMBERS OF THE SCIENTIFIC PLANNING AND REVIEW COMMITTEE . . . iii

TABLE OF CONTENTS . . . . . iv

BAY PROTECTION AND TOXIC CLEANUP PROGRAM
SCIENTIFIC PLANNING AND REVIEW COMMITTEE (SPARC)
TECHNICAL WORKSHOP
Focus of the Workshop . . . . . 1
Agenda . . . . . 3

PROGRAM SUMMARY . . . . . 5
Program Activities . . . . . 5
Toxic Hot Spot Identification . . . . . 6
Ranking Criteria . . . . . 7
Sediment Quality Objectives . . . . . 7
Toxic Hot Spot Cleanup Plans . . . . . 7

THE DESIGN OF THE BAY PROTECTION AND TOXIC CLEANUP PROGRAM
MONITORING PROGRAM . . . . . 10
Legislative Mandate . . . . . 10
Specific Definition of a Toxic Hot Spot . . . . . 10
Proposed Specific Definition . . . . . 11
Potential Toxic Hot Spot . . . . . 11
Candidate Toxic Hot Spot . . . . . 12
Known Toxic Hot Spot . . . . . 15
Monitoring Program Objectives . . . . . 15
Review of Preliminary Studies and Research . . . . . 15
Biological Monitoring Methods . . . . . 15
Chemical Methods . . . . . 16
Sampling Strategy . . . . . 16
Screening Sites and Confirming Toxic Hot Spots . . 16
A Battery of Screening Tests . . . . . 18
Site Selection . . . . . 21
Toxicity Screening . . . . . 22
Confirmation (i.e., Qualification as Candidate Toxic Hot Spots) . . . . . 22
Quality Assurance . . . . . 23
References . . . . . 24

ISSUE PAPERS . . . . . 27
What is Toxic? . . . . . 28
How do we show association between toxicity, benthic community measurements, etc. and chemical concentration? . . . . . 32
What is a benthic impact? . . . . . 36
Should we use a probability-based sampling design (random sampling) or directed point estimates (based on best professional judgement)? . . . . . 39

Should we use a screening and confirmation approach? . . 41
Are the toxicity testing methods the most appropriate for meeting program objectives? . . . . . 43
Are BPTCP Analytical Chemistry Methods Scientifically Sound and Appropriate? . . . . . 45

REGIONAL SUMMARIES . . . . . 48

REGIONAL WATER QUALITY CONTROL BOARD
NORTH COAST REGION (REGION 1) . . . . . 49
Monitoring Goals and Objectives . . . . . 49
Arcata Bay and Humboldt Bay Segments . . . . . 49
Bodega Harbor . . . . . 50
Eel River Estuary . . . . . 51
Klamath River Estuary . . . . . 51
Mad River Estuary . . . . . 51
Noyo River Estuary . . . . . 51
Russian River Estuary . . . . . 51
Smith River Estuary . . . . . 51

REGIONAL WATER QUALITY CONTROL BOARD
SAN FRANCISCO BAY REGION (REGION 2) . . . . . 52
Physical Description of the Region . . . . . 52
Philosophy of Monitoring in the Region . . . . . 53
Bay Protection and Toxic Cleanup Studies . . . . . 53
Projects to Collect Information for the Cleanup of the South Bay . . . . . 59
Questions and Issues Particular to this Region . . . . . 62
Additional Data . . . . . 63

REGIONAL WATER QUALITY CONTROL BOARD
LOS ANGELES REGION
(REGION 4) . . . . . 64
Results of Previous Studies (State Mussel Watch/Toxic Substance Monitoring/Regional Board Sediment Sampling) . . . . . 64
Sampling Goals . . . . . 65
BPTCP-related Goals and Objectives . . . . . 65
Issues/Questions Generated by This Work . . . . . 66
General Issues/Questions . . . . . 66

REGIONAL WATER QUALITY CONTROL BOARD
CENTRAL VALLEY REGION (REGION 5) . . . . . 68
Physical Features . . . . . 68
Goals and Objectives . . . . . 68
Summary of Studies . . . . . 68
Additional Data Available . . . . . 70
General Issues and/or Questions . . . . . 70

REGIONAL WATER QUALITY CONTROL BOARD
SANTA ANA REGION (REGION 8) . . . . . 72
Overview . . . . . 72

Anaheim Bay/Huntington Harbor Complex . . . . . 72
Newport Bay . . . . . 72
Regional Monitoring Goals and Objectives . . . . . 73
Regional Questions . . . . . 73
Data Available . . . . . 73

REGIONAL WATER QUALITY CONTROL BOARD
SAN DIEGO REGION (REGION 9) . . . . . 75
San Diego Region Bays and Estuaries . . . . . 75
San Diego Bay . . . . . 75
Mission Bay . . . . . 76
Dana Point Harbor, Oceanside Harbor, and Del Mar Boat Basin at Camp Pendleton . . . . . 76
Coastal Lagoons . . . . . 77
San Diego Region Monitoring Goals and Objectives . . . . 77
San Diego Region Results . . . . . 77
San Diego Bay Questions . . . . . 78

CALIFORNIA BAY PROTECTION AND TOXIC CLEANUP PROGRAM
TECHNICAL QUESTIONS FOR THE
SCIENTIFIC PLANNING AND REVIEW COMMITTEE . . . . . 81
General Questions . . . . . 81
Toxicity Testing . . . . . 81
Benthic Community Analyses . . . . . 83
Trace Organics Chemistry . . . . . 84
Trace Metals Chemistry . . . . . 84
Biomarkers . . . . . 85
Natural Toxins and Unknowns . . . . . 86
Statistics . . . . . 86

CONTRIBUTORS TO THE BRIEFING DOCUMENT . . . . . 89

BAY PROTECTION AND TOXIC CLEANUP PROGRAM
SCIENTIFIC PLANNING AND REVIEW COMMITTEE (SPARC)
TECHNICAL WORKSHOP

April 12 and 13, 1995
Monterey, California

The Bay Protection and Toxic Cleanup Program (BPTCP) is sponsoring this technical workshop for the Program to bring together experts in the fields of toxicology, benthic ecology, organic and inorganic chemistry, program implementation and direction, experimental design, and statistics. The purpose of the workshop is to solicit comments from the Scientific Planning and Review Committee on the Program's monitoring approach(es), to give input on the scientific merit of the approach(es) taken, and to provide suggestions for monitoring improvement in the future.

The BPTCP is a Statewide Program legislatively mandated to identify toxic hot spots, to develop Toxic Hot Spot Cleanup Plans for each of the seven coastal Regional Water Quality Control Boards, and to prepare a consolidated Statewide Toxic Hot Spot Cleanup Plan.

**Focus of the Workshop**

The workshop will center around a discussion of the following key questions that have been identified by the State and Regional Boards and the Department of Fish and Game:

1. What is toxic?

2. How should we show association between toxicity, benthic community, etc. and chemical concentrations?

3. What is a benthic impact?

4. Should we use a probability-based sampling design (random sampling) or directed point sampling approach (i.e. based on best professional judgment)?

5. Should we use a screening and confirmation approach?

6. What biological methods should we use?

7. What chemical methods should we use?

For each of these questions, a brief issue paper outlining the options that have been evaluated is presented.

Each of the fundamental questions posed to the SPARC could take several days of discussion to fully evaluate and assess each facet of the question. It is the intent for this first workshop that SPARC hear the approaches being pursued by the program and comment on their appropriateness and usefulness. The SPARC is charged with determining if the approaches the Program is taking are scientifically credible and, if not, what approaches the Program should evaluate for use.

The BPTCP has two critical short-term needs: (1) to report monitoring data collected is San Diego Bay and (2) to plan for new monitoring scheduled for FY 1995-1996 (which begins July 1, 1995). To complete these tasks, the BPTCP needs to develop interim solutions on how to (1) evaluate the toxicity information collected and (2) associate biological effects with observed chemistry measurements.

It is anticipated that the Workshop discussion will lead to further questions for SPARC. The Program plans to convene another meeting of the group by the end of June, 1995 to continue the discussion on the BPTCP.

BAY PROTECTION AND TOXIC CLEANUP PROGRAM
SCIENTIFIC PLANNING AND REVIEW COMMITTEE (SPARC)
TECHNICAL WORKSHOP

April 12 and 13, 1995
Doubletree Hotel, Monterey CA

A G E N D A

**Day 1: April 12, 1995**

| | |
|---|---|
| 8:00 to 8:30 | Welcome |
| 8:30 to 8:45 | Introductions<br>Max Puckett |
| 8:45 to 9:00 | Committee Goals and Anticipated Products<br>Max Puckett |
| 9:00 to 9:45 | Program Overview<br>Craig J. Wilson |
| 9:45 to 10:00 | Coffee Break |
| 10:00 to 11:00 | Regional Board Presentations |
| 11:00 to 11:30 | The Fundamental Questions<br>Max Puckett |

What is toxic?
What measure of association between chemistry and biological effects?
What is a benthic impact?
Deterministic or probability-based sample collection?
Screening and confirmation approach?
What biological methods should be used?
What chemical methods should be used?

| | |
|---|---|
| 11:30 to 1:00 | Lunch |
| 1:00 to 3:00 | Toxicity Endpoint<br>John Hunt and Brian Anderson |
| 3:00 to 3:15 | Coffee Break |
| 3:15 to 5:00 | Association with Toxic Pollutants<br>Craig Wilson and Max Puckett |

**Day 2: April 13, 1995**

| | |
|---|---|
| 8:00 to 8:30 | Coffee |
| 8:30 to 10:00 | Benthic Impacts<br>Carrie Bretz |
| 10:00 to 10:15 | Coffee |
| 10:15 to 11:30 | Random sampling vs. directed point sampling<br>Craig Wilson and Rusty Fairey |
| 11:30 to 12:30 | Lunch |
| 12:30 to 1:45 | Screening and confirmation approach<br>Craig Wilson and Rusty Fairey |
| 1:45 to 2:00 | Coffee Break |
| 2:00 to 4:00 | Biological and Chemical Methods<br>John Hunt, Brian Anderson and Mark Stephenson |
| 4:00 to 5:00 | Wrap-Up |

BAY PROTECTION AND TOXIC CLEANUP PROGRAM

PROGRAM SUMMARY

California Water Code, Division 7, Chapter 5.6 established a comprehensive program within the State Water Resources Control Board (State Water Board) to protect the existing and future beneficial uses of California's bays and estuaries. The Bay Protection and Toxic Cleanup Program (BPTCP) provides new focus on the State Water Board and the California Regional Water Quality Control Boards' (Regional Water Boards) efforts to control pollution of the State's bays and estuaries and to establish a program to identify toxic hot spots and plan for their cleanup. SB 475 (1989), SB 1845 (1990), and AB 41 (1989) added Chapter 5.6 Bay Protection and Toxic Cleanup (Water Code Sections 13390-13396.5) to Division 7 of the Water Code. Recent legislation (SB 1084 (1993)) extended program funding through 1998, the deadline for the regional toxic hot spot cleanup plans to 1998, and the Statewide cleanup plan until 1999.

**Program Activities**

The BPTCP has four major goals: (1) protect existing and future beneficial uses of bay and estuarine waters; (2) identify and characterize toxic hot spots; (3) plan for the prevention of further pollution and the remediation of existing hot spots; and (4) develop prevention and control strategies for toxic pollutants that will prevent creation of new hot spots or perpetuation of existing hot spots.

The BPTCP is a comprehensive effort by the State and Regional Water Boards to programmatically link standards development, environmental monitoring, water quality control planning, and site cleanup planning. The primary program activities are:

1. Development and amendment of the California Enclosed Bays and Estuaries Plan. This plan will contain the State's water quality objectives for enclosed bays and estuaries and contain the implementation measures for the objectives.

2. Development and implementation of regional monitoring programs designed to identify toxic hot spots. These monitoring programs includes analysis for a variety of chemicals, the completion of a variety of toxicity tests, measurements of biological communities, and various special studies to support the program.

3. Development of a consolidated database that contains information pertinent to describing and managing toxic hot spots.

4. Development of narrative and numeric sediment quality objectives for the protection of California enclosed bays and estuaries.

5. Preparation of criteria to rank toxic hot spots that are based on the severity of water and sediment quality impacts.

6. Development of regional and statewide toxic hot spot cleanup plans that include identification and priority ranking of toxic hot spots, strategies for preventing formation of new toxic hot spots, and cost estimates for remedial action recommendations.

7. Implementation of a fee system to support all BPTCP activities.

**Toxic Hot Spot Identification**

The Water Code defines toxic hot spots as locations in enclosed bays, estuaries, or the ocean where pollutants have accumulated in the water or sediment to levels which (1) may pose a hazard to aquatic life, wildlife, fisheries, or human health, or (2) may impact beneficial uses or (3) exceed State Water Board or Regional Water Board adopted water quality or sediment quality objectives.

To identify toxic hot spots, water bodies of interest have been assessed both on a regional and site-specific basis. Regional assessments require evaluating whether water quality objectives are attained and beneficial uses are supported throughout the waterbody. Existing data on enclosed bays and estuaries are relatively limited for the purposes of determining impacts on beneficial uses.

Where sites are not well characterized, regional monitoring programs have been implemented. This monitoring activity has been performed by the California Department of Fish and Game under contract with the State Water Board.

The consolidated statewide database required by legislation will eventually include all data generated by the regional monitoring programs. The statewide database will be updated regularly to serve as the information source for making toxic hot spot determinations. It will contain information on pollutant concentrations in water, sediment, and tissue and the impacts on water bodies. The database will also include geographic information system (GIS) capabilities to allow mapping and accurate site identification.

## Ranking Criteria

The Water Code (Section 13393.5) requires the State Water Board to develop criteria for ranking toxic hot spots. The ranking criteria must consider the pertinent factors relating to public health and environmental quality. These factors include: (1) potential hazards to public health, (2) toxic hazards to fish, shellfish, and wildlife, and (3) the extent to which the deferral of a remedial action will result or is likely to result in a significant increase in environmental damage, health risks, or cleanup costs.

## Sediment Quality Objectives

State law defines sediment quality objectives as "that level of a constituent in sediment which is established with an adequate margin of safety, for the reasonable protection of beneficial uses of water or prevention of nuisances" (Water Code Section 13391.5). Water Code Section 13393 further defines sediment quality objectives as: "...objectives...based on scientific information, including but not limited to chemical monitoring, bioassays or established modeling procedures." The Water Code requires adequate protection for the most sensitive aquatic organisms." Sediment quality objectives can be either numerical values based on scientifically defensible methods or narrative descriptions implemented through toxicity testing or other methods.

## Toxic Hot Spot Cleanup Plans

The Water Code requires that each Regional Water Board must complete a toxic hot spot cleanup plan and the State Water Board must prepare a consolidated toxic hot spot cleanup plan. The State Water Board will develop a water quality control policy with guidance to the Regional Water Boards for consistent implementation of the BPTCP.

Each cleanup plan must include: (1) a priority listing of all known toxic hot spots covered by the plan; (2) a description of each toxic hot spot including a characterization of the pollutants present at the site; (3) an assessment of the most likely source or sources of pollutants; (4) an estimate of the total costs to implement the cleanup plan; (5) an estimate of the costs that can be recovered from parties responsible for the discharge of pollutants that have accumulated in sediments; (6) a preliminary assessment of the actions required to remedy or restore a toxic hot spot; and (7) a two-year expenditure schedule identifying State funds needed to implement the plan.

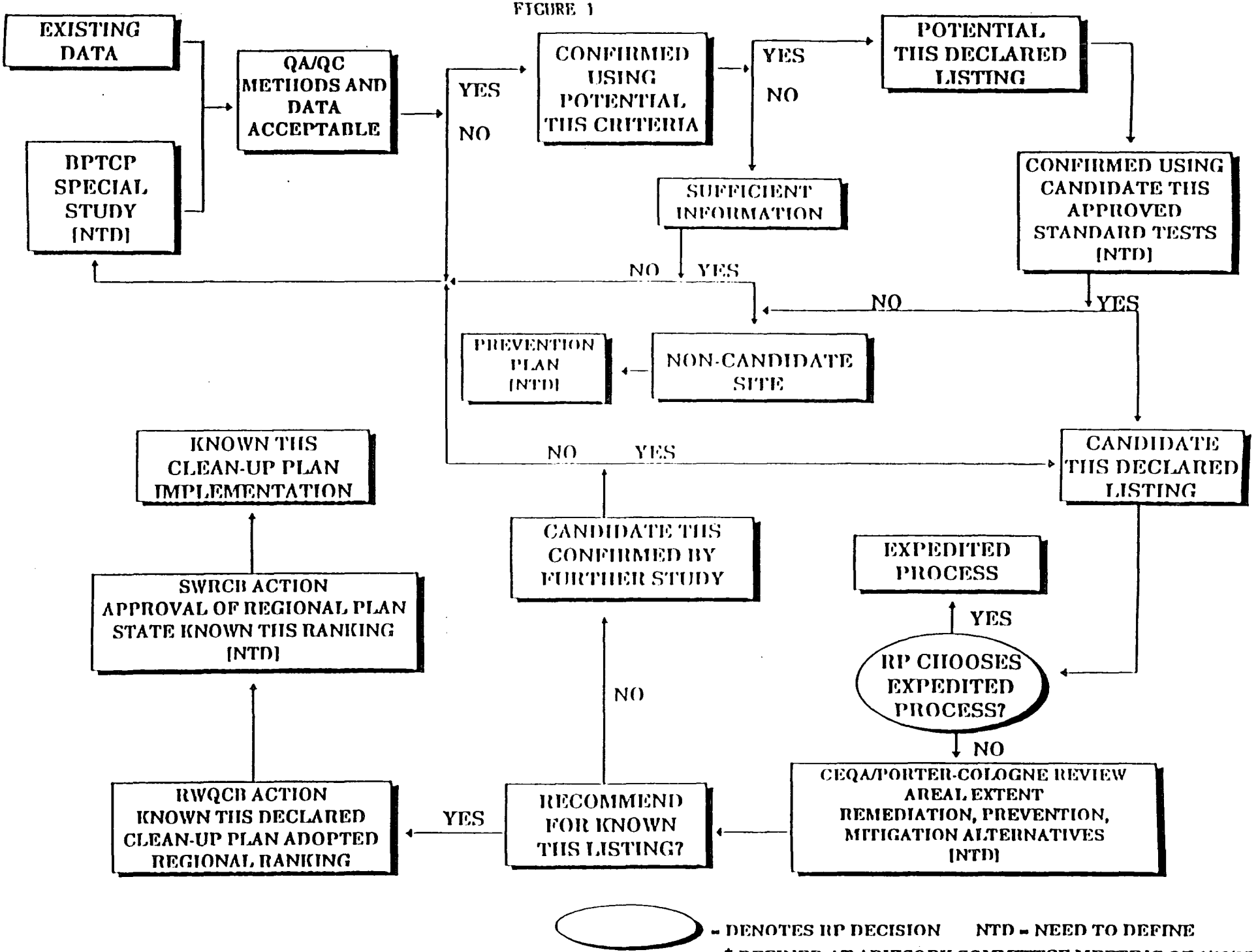
Within 120 days from the ranking of a toxic hot spot in a Regional cleanup plan, each Regional Water Board is required to begin reevaluating waste discharge requirements for dischargers

who have contributed any or all or part of the pollutants which have caused the toxic hot spot. These reevaluations shall be used to revise water quality control plans and water quality control plan amendments wherever necessary; reevaluations shall be initiated according to the priority ranking established in cleanup plans.

Figure 1 is a flow chart that presents the relationships between the various program activities.

*PROPOSED PROCESS TO DEVELOP REGIONAL TOXIC HOT SPOTS CLEAN-UP AND PREVENTION PLAN-FINAL* : 1/26*

FIGURE 1




= DENOTES RP DECISION  NTD = NEED TO DEFINE

* DEFINED AT ADVISORY COMMITTEE MEETING OF 1/10/05

# THE DESIGN OF THE BAY PROTECTION AND TOXIC CLEANUP PROGRAM MONITORING PROGRAM

The Bay Protection and Toxic Cleanup Program (BPTCP) was initiated by the State Water Resources Control Board (SWRCB) in April 1990. As part of the legislated requirements of the program, the BPTCP has begun implementation of regional monitoring programs, development of a consolidated database, identification of toxic hot spots, and begun planning for the cleanup and prevention of toxic hot spots.

Section 13392.5 requires, in part, that the State develop monitoring programs that are composed of at least the following components:

1. Guidelines to promote standardized analytical methodologies and consistency in data reporting; and

2. Additional monitoring and analyses that are needed to develop a complete toxic hot spot assessment for each enclosed bay and estuary.

This briefing document is to present the approach used to identify toxic hot spots in California enclosed bays and estuaries. The Scientific Review Committee is requested to review the approach, to give input on the scientific merit of the approaches taken, and to provide suggestions for monitoring improvement in the future.

## Legislative Mandate

Section 13391.5 of the Water Code defines toxic hot spots as "...locations in enclosed bays, estuaries, or adjacent waters in the 'contiguous zone' or the 'ocean' as defined in Section 502 of the Clean Water Act (33. U.S.C. Section 1362), the pollution or contamination of which affects the interests of the State, and where hazardous substances have accumulated in the water or sediment to levels which (1) may pose a substantial present or potential hazard to aquatic life, wildlife, fisheries, or human health, or (2) may adversely affect the beneficial uses of the bay, estuary, or ocean waters as defined in the water quality control plans, or (3) exceeds adopted water quality or sediment quality objectives."

## Specific Definition of a Toxic Hot Spot

One of the most critical steps in the development of toxic hot spot cleanup plans is the identification of hot spots. Once they are identified the parties responsible for the sites could be liable for the cleanup of the site or further prevention of the

discharges or activities that caused the hot spot. Because the cost of cleanup or added prevention could be very high, the SWRCB is considering categorizing toxic hot spots to distinguish between sites with little information (potential toxic hot spots) and areas with significantly more information (candidate toxic hot spots).

**Proposed Specific Definition**

Although the Water Code provides some direction in defining a toxic hot spot, the definition presented in Section 13391.5 is broad and somewhat ambiguous regarding the specific attributes of a toxic hot spot. The following specific definition provides the RWQCBs with a specific working definition and a mechanism for identifying and distinguishing between "potential," "candidate" and "known" toxic hot spots. A Candidate Toxic Hot Spot is considered to have enough information to designate a site as a Known Toxic Hot Spot except that the candidate hot spot has not been approved by the appropriate Regional Water Quality Control Board. Once a candidate toxic hot spot has been adopted into a toxic hot spot cleanup plan then the site shall be considered a known toxic hot spot and all the requirements of the Water Code shall apply to that site.

a. Potential Toxic Hot Spot

The Water Code requires the identification of suspected or "potential" toxic hot spots (Water Code Section 13392.5). Sites with existing information indicating possible impairment, but without sufficient information to be classified further as a "candidate" or "known" toxic hot spot are classified as "potential" toxic hot spots. Four conditions sufficient to identify a "potential" toxic hot spot are defined below. If any one of the following conditions is satisfied, a site can be designated a "potential" toxic hot spot:

1. Concentrations of toxic pollutants are elevated above background levels, but insufficient data are available on the impacts associated with such pollutant levels to determine the existence of a known toxic hot spot;

2. Water or sediments which exhibit toxicity in screening tests or test other than those specified by the State or Regional Boards;

3. Toxic pollutant levels in the tissue of resident or test species are elevated, but do not meet criteria for determination of the site as a known toxic hot spot, tissue toxic pollutant levels exceed maximum tissue residue levels (MTRLs) derived from water quality objectives contained in appropriate water quality

control plans, or a health advisory for migratory fish that applies to the whole water body has been issued for the site by OEHHA, DHS, or a local public health agency, the waterbody will be considered a potential toxic hot spot. Further monitoring is warranted to determine if health warnings are necessary at specific locations in the waterbody.

4. The level of pollutant at a site exceeds Clean Water Act Section 304(a) criterion, or sediment quality guidelines or EPA sediment toxicity criteria for toxic pollutants.

b. Candidate Toxic Hot Spot:

A site meeting any one or more of the following conditions is considered to be a "candidate" toxic hot spot.

1. The site exceeds water or sediment quality objectives for toxic pollutants that are contained in appropriate water quality control plans or exceeds water quality criteria promulgated by the U.S. Environmental Protection Agency.

   This finding requires chemical measurement of water or sediment, or measurement of toxicity using tests and objectives stipulated in water quality control plans. Determination of a toxic hot spot using this finding should rely on recurrent measures over time (at least two separate sampling dates). Suitable time intervals between measurements must be determined.

2. The water or sediment exhibits toxicity associated with toxic pollutants, based on toxicity tests acceptable to the State Water Resource Control Board or the Regional Water Quality Control Boards.

   To determine whether toxicity exists, recurrent measurements (at least two separate sampling dates) should demonstrate an effect. Appropriate reference and control measures must be included in the toxicity testing. The methods acceptable to and used by the BPTCP may include some toxicity test protocols not referenced in water quality control plans (e.g., the Bay Protection and Toxic Cleanup Program Quality Assurance Project Plan). Toxic pollutants should be present in the media at concentrations sufficient to cause or contribute to toxic responses in order to satisfy this condition.

3. The tissue toxic pollutant levels of organisms collected from the site exceed levels established by

the United States Food and Drug Administration (FDA) for the protection of human health, or the National Academy of Sciences (NAS) for the protection of human health or wildlife. When a health advisory against the consumption of edible resident non-migratory organisms has been issued by OEHHA or DHS, on a site or waterbody, the site or waterbody is automatically classified a "candidate" toxic hot spot if the chemical contaminant is associated with sediment or water at the site or water body.

Acceptable tissue concentrations are measured either as muscle tissue (preferred) or whole body residues. Residues in liver tissue alone are not considered a suitable measure for known toxic hot spot designation. Animals can either be deployed (if a resident species) or collected from resident populations. Recurrent measurements in tissue are required. Residue levels established for one species for the protection of human health can be applied to any other consumable species.

Shellfish: Except for existing information, each sampling episode should include a minimum of three replicates. The value of interest is the average value of the three replicates. Each replicate should be comprised of at least 15 individuals. For existing State Mussel Watch information related to organic pollutants, a single composite sample (20-100 individuals), may be used instead of the replicate measures. When recurrent measurements exceed one of the levels referred to above, the site is considered a known toxic hot spot.

Fin-fish: A minimum of three replicates is necessary. The number of individuals needed will depend on the size and availability of the animals collected; although a minimum of five animals per replicate is recommended. The value of interest is the average of the three replicates. Animals of similar age and reproductive stage should be used.

4. Impairment measured in the environment is associated with toxic pollutants found in resident individuals.

Impairment means reduction in growth, reduction in reproductive capacity, abnormal development, histopathological abnormalities, or identification of adverse effects using biomarkers. Each of these measures must be made in comparison to a reference condition where the endpoint is measured in the same species and tissue is collected from an unpolluted

reference site. Each of the test shall be acceptable to the SWRCB or the RWQCBs.

Growth Measures: Reductions in growth can be addressed using suitable bioassays acceptable to the State or Regional Boards or through measurements of field populations.

Reproductive Measures: Reproductive measures must clearly indicate reductions in viability of eggs or offspring, or reductions in fecundity. Suitable measures include: pollutant concentrations in tissue, sediment, or water which have been demonstrated in laboratory tests to cause reproductive impairment, or significant differences in viability or development of eggs between reference and test sites.

Abnormal Development: Abnormal development can be determined using measures of physical or behavioral disorders or aberrations. Evidence that the disorder can be caused by toxic pollutants, in whole or in part, must be available.

Histopathology: Abnormalities representing distinct adverse effects, such as carcinomas or tissue necrosis, must be evident. Evidence that toxic pollutants are capable of causing or contributing to the disease condition must also be available.

Biomarkers: Direct measures of physiological disruption or biochemical measures representing adverse effects, such as significant DNA strand breakage or perturbation of hormonal balance, must be evident. Biochemical measures of exposure to pollutants, such as induction of stress enzymes, are not by themselves suitable for determination of "candidate" toxic hot spots. Evidence that a toxic pollutant causes or contributes to the adverse effect are needed.

5. Significant degradation in biological populations and/or communities associated with the presence of elevated levels of toxic pollutants.

This condition requires that the diminished numbers of species of individuals of a single species (when compared to a reference site) are associated with concentrations of toxic pollutants. The analysis should rely on measurements from multiple stations. Care should be taken to ensure that at least one site is not degraded so that a suitable comparison can be made.

In summary, sites are designated as "candidate" hot spots after generating information which satisfies any one of the five conditions constituting the definition.

c. Known Toxic Hot Spot:

A site meeting any one or more of the conditions necessary for the designation of a "candidate" toxic hot spot and has gone through a full State or Regional board hearing process, is considered to be a "known" toxic hot spot. A site will be considered a "candidate" toxic hot spot until approved as a known toxic hot spot in a Regional Toxic Hot Spot Cleanup Plan by the Regional Water Quality Control Board and approved by the State Water Resources Control Board.

## Monitoring Program Objectives

The four objectives of BPTCP regional monitoring are:

1. Identify locations in enclosed bays, estuaries, or the ocean that are toxic hot spots;

2. Determine the extent of biological impacts in portions of enclosed bays and estuaries not previously sampled (areas of unknown condition);

3. Confirm the extent of biological impacts in enclosed bays and estuaries that have been previously sampled; and

4. Assess the relationship between toxic pollutants and biological effects.

## Review of Preliminary Studies and Research

Each of the seven RWCQBs participating in the program has assembled information that was used to develop a preliminary list of potential and candidate toxic hot spots (SWRCB, 1993).

## Biological Monitoring Methods

The tests listed in Table 1 are acceptable to measure water and sediment toxicity. Other tests may be added to the list as deemed appropriate by the State or Regional Water Boards provided the tests have a detailed written description of the test method; Interlaboratory comparisons of the method; Adequate testing with water, wastewater, or sediments; and measurement of an effect

that is clearly adverse and interpretable in terms of beneficial use impact.

## Chemical Methods

The BPTCP measures a variety of organic and inorganic pollutants in estuarine sediments (Stephenson et al. 1994). The BPTCP requires its laboratories to demonstrate comparability continuously through strict adherence to common Quality Assurance/Quality Control (QAQC) procedures, routine analysis of certified reference materials, and regular participation in an on-going series of interlaboratory comparison exercises (round-robins). This is a "performance-based" approach of quality assurance.

The method used by the BPTCP are those used in the NOAA National Status and Trends Program (Lauenstein et al. 1993) and the methods documented in the DFG QAQC Manual (DFG, 1992). Under the BPTCP performance-based chemistry QA program, laboratories are not required to use a single, standard analytical method for each type of analysis, but rather are free to choose the best or most feasible method within the constraints of cost and equipment.

## Sampling Strategy

### Screening Sites and Confirming Toxic Hot Spots

In order to identify known toxic hot spots a two-tier process was used. The first tier was a screening step where at least two toxicity tests were used at a site (Tables 2 and 3). Sediment grain size, total organic carbon (TOC) and $H_2S$ concentration were measured to differentiate pollutant effects found in screening tests from natural factors. Chemical analyses (metals and organics) were performed on a subset of the screening samples.

If effects were found at sites by these screening steps, some sites were retested (depending on available funding) to confirm the effects. In the confirmation step measurements were replicated and compared to reference sites or conditions. Chemical measurements (metals, organics, TOC, $H_2S$) and other factors (e.g., sediment grain size) were measured. Measurements of benthic community structure and, if needed, bioaccumulation were also made.

Table 1
Water and Sediment Toxicity Tests That Meet
the Criteria For Acceptability

| Type of Toxicity Test | Organism Used | | Reference |
|---|---|---|---|
| | Common Name | Scientific Name | |
| Solid Phase | Amphipod | Rhepoxinius | ASTM, 1993 |
| Sediment | Amphipod | Eohaustorius | ASTM, 1993 |
| | Amphipod | Ampelisca | ASTM, 1993 |
| | Amphipod | Hyalella | ASTM, 1993 |
| | Polychaete | Neanthes | Johns et. al., 1990 |
| Sediment Pore Water* | Bivalve larvae | Crassostrea | ASTM, 1993 |
| | | Mytilus | ASTM, 1993 |
| | Abalone larvae | Haliotis | Anderson et al., 1990 |
| | Echinoderm fertilization | Strongy-locentrotus | Dinnel et al., 1987; with modification by EPA, 1992 |
| | Giant kelp | Macrocystis | Anderson et al., 1990 |
| | Red alga | Champia | Weber et al., 1988 |
| | Fish embryos | Atherinops | Anderson et al., 1990 |
| | | Menidia | Middaugh et al., 1988 |
| | | Pimephales | Spehar et al., 1982 |
| | Cladoceran | Daphnia | Nebecker et al., 1984 |
| | | Cereodaphnia | Horning and Weber, 1985 |
| Ambient Water | Bivalve larvae | Crassostrea | ASTM, 1993 |
| | | Mytilus | ASTM, 1993 |
| | Abalone larvae | Haliotis | Anderson et al., 1990 |
| | Echinoderm fertilization | Strongylocen-trotus | Dinnel et al., 1987; with modifications by EPA, 1992 |
| | Giant kelp | Macrocystis | Anderson et al., 1991 |
| | Red alga | Champia | Weber et al., 1988 |
| | Mysid | Holmesimysis | Hunt et al., 1992 |
| | Fish embryos | Atherinops | Anderson et al., 1990 |
| | | Menidia | Middaugh et al., 1988 |
| | | Pimephales | Spehar et al., 1982 |
| | Fish larvae | Atherinops | Anderson et al., 1990 |
| | | Menidia | Peltier and Weber, 1985<br>Weber et al., 1988 |
| | | Pimephales | Peltier and Weber, 1985<br>Weber et al., 1988 |
| | Cladocerans | Daphnia | Nebecker et al., 1984 |
| | | Cereodaphnia | Horning and Weber, 1985 |

*Pore water tests (other than amphipods) alone can not be used to designate a candidate toxic hot spot.

Table 2

Screening Tests for
Toxic Hot Spot Identification

| Test Organism | Type | End Point |
|---|---|---|
| Rhepoxynius, Eohaustorius (Amphipod) | Bedded sediment | Survival |
| Haliotus, Mytilus, Crassostrea | Overlying water | Shell development |
| Strongylocentrotus (Sea urchin) | Sediment pore water | Fertilization, development, and/or anaphase aberration |
| Neanthes (Polychaete worm) | Bedded sediment | Survival and growth |

**A Battery of Screening Tests**

Selecting a battery of toxicity screening tests (Table 2) can improve cost-effectiveness by expanding the range of potential impacts to be evaluated. Although recurrent toxicity must be demonstrated to qualify a site as a "candidate" toxic hot spot, the degree of certainty for each of the measurements does not necessarily have to be equivalent. The cost of confirming toxicity at a site can be prohibitively high, especially if it includes a large number of field replicates and extensive reference site testing. The screening tests should allow for a relatively rapid lower cost assessment of the site.

Even though the list of acceptable tests is long (see Table 1), the State and Regional Water Boards have used between two and four tests to screen sites (Table 2). For all screening, at least one amphipod test was performed. Other tests were performed as needed depending on funding availability, the needs of collaborators (such as the National Oceanic and Atmospheric Administration or the EPA Environmental Monitoring and Assessment Program), test organisms sensitivity to the

Table 3

Types of Data Collected in Regional Monitoring Programs for the Identification of Toxic Hot Spots

| Type of Data | Screening | Confirmation |
|---|---|---|
| Toxicity testing | Suite of 4 tests (see Table 5) | Repeat of positive results |
| Field replicates | None | Three (if needed) |
| Lab replicates | Five | Five |
| Reference sites | None | Several |
| Physical analysis | Grain size | Grain size |
| Chemical analyses | Ammonia, hydrogen sulfide, TOC, pesticides, PCB, PAH, TBT, metals | Ammonia, hydrogen sulfide, TOC, pesticides, PCB, PAH, TBT, metals |
| Benthic community analysis | None | Five replicates |
| Bioaccumulation | None | Occasionally (sites with no pre-existing bioaccumulation data) |

Table 4

Sequence of Tasks for Designating Toxic Hot Spots

| | |
|---|---|
| 1. | Select toxicity screening sites. |
| 2. | Sample screening sites. |
| 3. | Conduct battery of four toxicity screening tests; analyze for hydrogen sulfide, ammonia, TOC, and grain size. |
| 4. | Determine whether quality assurance requirements have been met. |
| 5. | Report on Items 3 and 4. |
| 6. | Select and match hits and potential reference sites for ammonia, hydrogen sulfide, and grain size. |
| 7. | Conduct metals and organic chemical analysis on subset of screening sites from Item 6. |
| 8. | Determine whether quality assurance requirements have been met. |
| 9. | Report on Items 7 and 8. |
| 10. | Select sites and toxicity tests for confirmation and reference sites. |
| 11. | Sample confirmation and reference sites. |
| 12. | Conduct subset of the battery of toxicity tests which were screening hits; analyze for hydrogen sulfide, TOC, and conduct benthic community analysis. |
| 13. | Conduct metals and organic chemical analyses. |
| 14. | Determine whether quality assurance requirements have been met. |
| 15. | Report on Items 12 through 15. |
| 16. | Conduct statistical and other analyses to determine whether sites qualify as toxic hot spots. |

pollutants expected to be present, and the media (bedded sediment or pore water) thought to be contaminated.

## Site Selection

Two somewhat different approaches were used in BPTCP monitoring. Six of the coastal RWQCBs have used a design that combines toxicity testing, chemical analysis, and benthic community analysis in a two-phased screening-confirmation framework (Tables 3 and 4).

The Central Valley RWQCB, with jurisdiction over the Sacramento-San Joaquin Delta, has designed its program to respond to Delta conditions and to the water quality problems characteristic of that area. Fresh water toxicity testing combined with water chemistry (metals and pesticides) constitutes the main program components. Sediment toxicity testing could be added to the monitoring design at a later stage.

Four different categories of sites have been identified for sampling in the BPTCP monitoring program: (1) potential toxic hot spots base on existing information, (2) high risk sites based on existing information, (3) stratified random sites, and (4) reference sites. Potential toxic hot spots are the highest priority sites because some indication already exists that these sites have a pollution-related problem. These data are typically sites with information available on chemical contamination of mussel tissue, data documenting water and sediment toxicity, measurements of metals or organic chemicals in sediments, and, occasionally, biological impairment. These sampling efforts are typically point estimates.

There are many other sites that are considered "high risk" even though we have no monitoring information to support this contention. High risk sites are locations where a nearby activity (such as marinas, storm drains, and industrial facilities) are thought to be associated with a certain risk of toxicity. The measurements at high risk sites are either point estimates or selected probabilistically.

When little is known about the quality of a waterbody segment, the monitoring efforts should use a stratified, random sampling approach. These random sites are useful in determining the quality of larger areas in the State's enclosed bays and estuaries. This probabilistic approach will allow for the State and Regional Water Boards to make better estimates of area (percentage) of water bodies that is impacted. The State and Regional Water Boards have used the techniques used by the U.S. Environmental Protection Agency's Environmental Monitoring and Assessment Program (SWRCB et al. 1994).

Locating reference sites requires identification and testing of a variety of potential reference sites encompassing the expected range of grain size, TOC, and other characteristics. Existing data sets that describe chemical contamination, grain size, and TOC at marine and estuarine sites are reviewed. Since these sources yield an insufficient number of sites, fine-grained areas presumed to be relatively free of contamination are also examined. These sites may likewise prove to be rare, so sites with chemicals present, but experiencing low energy tidal flushing, will also be sampled. Sites with previous indication of no pollution, and those lacking sediment toxicity measurements will also be sampled. Finally, random selection of sites (as described above) may prove useful in locating reference sites.

**Toxicity Screening**

All tests included controls which were conducted in media known to exert minimal stress on test organisms. Both positive (toxicant present) and/or negative (toxicant absent) controls were used to ensure that test organisms are responding within expected limits (Table 3).

The screening step began with the collection of a single field sample from each site (Table 4, Steps 1 and 2). Five laboratory replicates were required to accommodate statistical comparison with the control. Although the lack of field replicates restricts statistical comparisons with other sites, this approach allowed the BPTCP to test more locations for toxicity within the allocated funding. Ammonia and hydrogen sulfide analyses are then performed on the media of all tests (Table 4, Step 3) to determine their relative contribution to any observed toxic affects. Grain size and TOC values were determined on all sediment samples to evaluate the presence of naturally occurring toxicity.

All these data, along with an assessment of quality assurance performance, were then reviewed. Toxicity hits and potential reference sites were selected and matched for ammonia, hydrogen sulfide, grain size, and TOC. A subset of the sites is selected for analysis of metals and organics after conducting confirmation testing (Table 4, Steps 4-9). Some of these sites were revisited for confirmation.

**Confirmation (i.e., Qualification as Candidate Toxic Hot Spots)**

Some of the screening sites (Table 4, Steps 10 and 11) with at least one positive test result were revisited to evaluate both the recurrent nature of the toxicity and impacts on the benthic community. This required repeat testing of potential toxic hot spots to ensure that toxicity was present or absent. Confirmation testing was more intensive because of (1) addition

of field replicates (three to a site); (2) comparison to reference sites (unless water toxicity is the focus); and (3) benthic community analysis (Table 3).

For each positive toxicity test at a screening site, confirmation was performed for the same test. Generally, benthic analysis was also performed and added to an ever-enlarging nearshore benthic community database which will be periodically evaluated to determine whether impacted and non-impacted sites can be distinguished (Table 4, Step 12). When either recurrent toxicity was demonstrated with a positive confirmation test or benthic impacts were suspected, chemical analysis were also performed (Table 4, Step 13). Careful review of all quality assurance procedures was conducted and, upon approval, will be followed by statistical analysis of the data. Compared to screening, this analysis will be more comprehensive and will include measures of field variability in toxicity, benthic data, and reference site conditions.

Once both toxicity and benthic impacts have been confirmed through comparison with an appropriate reference site and appear to be due to human-causes the site will be declared a candidate toxic hot spot. When toxicity is present but benthic impacts are lacking, careful analysis will be performed to determine whether the two results are in conflict. Similarly, when toxicity is not demonstrated but benthic impacts are observed, careful review will be conducted to determine whether the same explanation prevails or whether some factor other than toxicants may be responsible. Further characterization of the site (such as areal extent, range of effects, and source determination) will be described in the cleanup plan and is not intended (unless samples are collected using a random or stratified random design) under this phase of the program.

## Quality Assurance

The BPTCP Quality Assurance Project Plan (Stephenson et al. 1994) presents a systematic approach that has been implemented within each major data acquisition and data management component of the program. Basic requirements specified in the QAPP are designed to: (1) ensure that collection and measurement procedures are standardized among all participants; (2) monitor the performance of the various measurement systems being used in the program to maintain statistical control and to provide rapid feedback so that corrective measures can be taken before data quality is compromised; (3) assess the performance of these measurement systems and their components periodically; and, (4) verify that reported data are sufficiently complete, comparable, representative, unbiased, and precise.

## References

Anderson, B.S., J.W. Hunt, S.L. Turpen, A.R. Coulon, M. Martin, D.L. McKeown and F.H. Palmer. 1990. Procedures manual for conducting toxicity tests developed by the marine Bioassay Project. California State Water Resources Control Board 90-10WQ. Sacramento, California. 113 p.

ASTM. 1987. Standard practice for conducting static acute tests with larvae of four species of bivalve molluscs. Procedure E724-80. Annual book of ASTM standards; water and environmental technology. Vol 11.4: 382-388. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1993. Designation E 1367: Standard guide for conducting 10-day static sediment toxicity tests with marine and estuarine amphipods. Volume 11.04: Pesticides; resource recovery; hazardous substances and oil spill responses; waste disposal; biological effects. Annual book of standards; water and environmental technology. American Society for Testing and Materials, Philadelphia, PA.

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

Department of Fish and Game. 1992. Laboratory Quality Assurance Program Plan. Environmental Services Division. Sacramento, CA.

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

Dinnell, P.J., J. Link, and Q. Stober. 1987. Improved methodology for sea urchin sperm cell bioassay for marine waters. Arch. Envir. Cont. and Toxicol. 16: 23-32.

Di Toro, D.M., J.D. Mahony, D.J. Hansen, K.J. Scott, M.B. Hicks, S.M. Mayr, and M.S. Redmond. 1990. Toxicity of Cadmium in Sediments: The Role of Acid Volatile Sulfide. Environmental Toxicology and Chemistry, Vol. 9, pp. 1487-1502. 1990.

Environmental Protection Agency (EPA). 1992. Sea urchin (Strongylocentrotus purpuratus) fertilization test method. Final Draft. Gary A. Chapman, U.S. Environmental Protection Agency. ERL - Pacific Ecosystems Branch, Newport, Oregon.

Horning, W.B. II and C.I. Weber (eds.). 1985. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms. Environmental Monitoring and Research Laboratory - Cincinnati Office of Research and

Development. U.S. Environmental Protection Agency, Cincinnati, Ohio. EPA/600/4-85/014.

Hunt, J.W., B.S. Anderson, S.L. Turpen, H.R. Barber, D.L. Johns, D.M., T.C. Ginn and D.J. Reish. 1990. The juvenile neanthes sediment bioassay. Puget Sound Notes, No. 24, U.S. EPA, Seattle, WA.

Lauenstein, G.L., A.Y. Cantillo, and S. Dolvin (eds.) 1993. A compendium of methods used in the NOAA National Status and Trends Program. National Ocean Service, Office of Ocean Resources Conservation and Assessment, Silver Springs, MD.

Long, E.R. and L. Morgan 1990. The potential for Biological Effects of Sediment-sorbed Contaminants Tested in the National Status and Trends Program. NOAA Technical Memorandum NOS OMA 52.

McKeown D., F.H. Palmer and M. Martin. 1991. Marine Bioassay Project Sixth Report: Interlaboratory comparisons and protocol development with four marine species. Report #91-21-WQ. State Water Resources Control Board, California.

Middaugh, D.P., M.J. Hemmer, and E.M. Lores. 1988. Teratological effects of 2,3-dinitrophenol, produced water, and naphthalene on embryos of the inland silverside Menidia beryllina. Dis. Aquat. Org., 4,53 - 65.

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

Peltier, W.H., and C.I. Weber. 1985. Methods for measuring the acute toxicity of effluents to freshwater and marine organisms. Environmental Monitoring and Support Laboratory - Cincinnati Office of Research and Development. U.S. Environmental Protection agency, Cincinnati, Ohio. EPA/600/4-85/013.

Spehar, R.L., D.K. Tanner and J.H. Gibson. 1982. Effects of kelthane and pydrin on early life stages of fathead minnows (Pimephales promelas) and amphipods (Hyalella azteca). In J.G. Peatson, R.B. Foster and W.E. Bishop, eds., Aquatic Toxicity and Hazard Assessment: Fifth Conference. STP 766. American Society for Testing and Materials, Philadelphia, PA. pp. 234-244.

State Water Resources Control Board. 1993. Staff Report: The Status of the Bay Protection and Toxic Cleanup Program. Sacramento, CA. 231 pp + 5 Appendices.

State Water Resources Control Board, National Oceanic and Atmospheric Administration, and U.S. Environmental Protection Agency Environmental Monitoring and Assessment Program. 1994. Technical proposal for a cooperative agreement between the State Water Resources Control Board, and the U.S. Environmental

Protection Agency Environmental Monitoring and Assessment Program. Sacramento, CA. 26 pp + 3 appendices.

Stephenson, M., M. Puckett, N. Morgan, and M. Reid. 1994. Bay Protection and Toxic Cleanup Program Quality Assurance Project Plan. 12 Sections and 1 appendix.

## ISSUE PAPERS

1. What is toxic?

2. What measure of association between chemistry and biological effects?

3. What is a benthic impact?

4. Deterministic or probability-based sample collection?

5. Screening and confirmation approach?

6. What biological methods should be used?

7. What chemical methods should be used?

## ISSUE 1

**What is Toxic?**

or, more specifically:

What level of response in a sediment toxicity test demonstrates that the sample is toxic, and what statistical tests should be used to make that determination?

While very low survival of test organisms is clearly indicative of toxicity, many test results are in the intermediate range (50% to 80% survival or normal development). For hot spot identification, the program must state exactly where to draw the line between responses that do or do not indicate significant toxicity. A number of statistical methods have been suggested and employed, but we need to reach agreement on which method is the most appropriate and defensible for splitting the hair in a regulatory setting.

We have considered two main approaches. The first is to simply compare each sample against the negative control (such as home sediment or dilution water). If a statistical comparison shows a significant difference, then it can be assumed that the observed effect was caused by something inherent in the sample, and not by laboratory conditions or organism handling. However, no assumption can be made about the specific sample characteristic responsible for the observed effect (i.e. we have no experimental basis to assume the effect was caused by anthropogenic contaminants as opposed to grain size or other factors). In order to use this approach for hot spot identification, a fairly strong association would have to be established between toxicity and chemistry to independently determine that contamination was the probable cause of the observed biological effect.

In the second approach, each sample could be compared against one or more reference sites. If multiple reference sites are sampled, covering a range of sediment grain size and other characteristics, it is possible to account for a large portion of the natural variation between sites (i.e. the variation occurring in the absence of contaminant effects). Any test sample that had significantly lower survival or normal development relative to the population of reference sites could be considered significantly toxic, and it would be reasonable to assume that the toxicity was due to anthropogenic contamination. This approach attempts to consider the cause and effect in a single analysis. While this second approach is more directly defensible for hot spot designation, it has the disadvantage that reference site characteristics are hard to define, and reference sediments are difficult to locate in the field. It is not uncommon to observe low rates of survival or normal development in samples with low concentrations of measured contaminants. In these cases, the observed effect could be due to natural toxins, in which case the site might still be

considered as a reference site. If, however, unmeasured anthropogenic contaminants are the cause of toxicity, and the site is used for reference, then the results of statistical analyses may be misinterpreted.

A variety of statistical methods could be used for either of the two main approaches. Statistical methods employed or considered for the first approach include the following:

1. t-tests have been used to compare each test sediment to the laboratory negative control. This method assumes that each comparison is a complete experiment and is not affected by other comparisons with other sites. Separate-variance t-tests have been used to adjust the degrees of freedom for unequal variances, which are commonly observed.

2. Analysis of Variance and Dunnett's tests have been used to compare all test sediments to the laboratory negative control, as above. Sample variances would have to be homogeneous.

3. We have also used a detectable difference approach (as suggested by Glen Thursby), where the Minimum Significant Difference (MSD) is calculated for a large number of individual comparisons, and the difference detectable in 90% of the cases is then used to determine significant difference from the control for all samples. For example, our data with the Rhepoxynius test indicate that the test can detect an 18% difference from the control 90% of the time. Therefore, if Rhepoxynius survival was 95% in the control, a sample with mean Rhepoxynius survival of $\leq$ 77% would be considered significantly less than the control. This approach is similar to t-tests and ANOVA, but depends on general trends in between-replicate variability, rather than on the variability found in a single comparison. The method tends to eliminate "skinny hits", small differences detected because of low between-replicate variability.

4. Equivalency tests could be used to compare the mean response from a test sediment to some standard toxic level. If, for example, we could state with confidence that 60% survival indicated toxicity, an equivalency test could use the between-replicate variability from the sediment toxicity test to determine whether that sediment was toxic (i.e. the mean result from that sample was significantly equal to or lower than the level considered toxic).

5. A standard cutoff line could be established based on previous data. For example, 80% of the control could be given as the cutoff, and anything less would be considered toxic. Schimmel et al. (1991) (EMAP), use this level to indicate toxicity, if the sample was also significantly different from the control in a t-test. Their objective,

however, was to discern general trends rather than identify hot spots for cleanup.

Statistical methods employed or considered for the second approach, in which test sites are compared to reference sites, include the following:

1. Any of the above methods could be used by substituting a reference site for the control.

2. A "reference envelope" analysis could be employed if results were available from multiple reference sites. This approach has been investigated by Bob Smith of EcoAnalysis, both in studies using benthic community data and in analyses of BPTCP data sets. In its simplest form, the method defines the mean and lower confidence limit of the reference site population, and any test site with a mean that is below the lower confidence limit is considered significantly toxic.

3. Outlier identifier methods, such as a Hampel Outlier Identifier, could be used to determine which sites were not part of the population of reference sites. This approach requires data from a relatively large number of reference sites.

Any method dependent on comparisons with reference sites must be preceded by adoption of reference site criteria and location of sites that consistently meet those criteria. A number of questions have arisen regarding reference sites: Must samples from reference sites be uncontaminated (using what analyte list and concentration limits)? Must they be non-toxic? Must they be both uncontaminated and non-toxic? What range of grain size, TOC, salinity, etc. must be included in the reference site population? What are the geographical constraints (i.e. same water body, same state)? Can one fine-grained reference site suffice for all tests?

If toxicity tests are not evaluated in the context of reference site or background conditions, will the results have sufficient credibility for hot spot designations?

A final issue for consideration: What level of field replication is necessary for hot spot designation? A single replicate allows us to say that the sample (not the site) is toxic. Disregarding concerns about the spatial extent of toxicity, how many field replicates are sufficient to indicate that a site is toxic? How should field replication be considered in the statistical approach to determining sediment toxicity?

Reference

Schimmel, S.C., B.D. Melzian, D.E. Campbell, C.J. Strobel, S.J. Benyi, J.S. Rosen, H.W. Buffum, N.I. Rubenstein. 1991. Statistical Summary EMAP-Estuaries - Virginian Province. EPA/620/R-94/005. US EPA, Office of Research and Development, Washington, DC.

## ISSUE 2

**How do we show association between toxicity, benthic community measurements, etc. and chemical concentration?**

The definition of a toxic hot spot requires that a determination of association of biological effect be associated with the response. There are several approaches available that allow a determination of chemical concentration in sediments can potentially contribute to the observed benthic or toxic effect.

Options Evaluated

1. Environmental Protection Agency (EPA) Sediment Quality Criteria (SQC)--Equilibrium Partitioning

   The EqP approach assumes that pollutants are generally in a state of thermodynamic equilibrium and that the relative concentration of a pollutant in any particular environmental compartment (sediment, pore water, ambient water, etc.) can be predicated using measured partitioning coefficients for specific substances in equilibrium equations. The EqP approach is currently limited to nonpolar, nonionic compounds although methods for metals are under development. The protection of sediment ingesting organisms is not addressed in this approach. Also the assumptions stated above have not been adequately tested. EPA has recently published (EPA, 1993a; 1993b; 1993c; and 1993d) draft SQC that could be used for this purpose.

2. Effects Range Low (ERL), Effects Range Median (ERM), Probable Effects Level (PEL), Threshold Effects Level (TEL)

   Two related efforts have been completed that provide an alternative approach for evaluating the quality of marine and estuarine sediments. These are the National Oceanic Atmospheric Administration (NOAA) (Long et al. 1995) and the sediment weight-of-evidence guidelines developed for the Florida Coastal Management Program (1993) and MacDonald, in press).

   Long et al. (1995) assembled data from throughout the country for which chemical concentrations had been correlated with effects. These data included spiked bioassay results and field data of matched biological effects and chemistry. The product of the analysis is the identification of two concentrations for each substance evaluated. One level, the Effects Range-Low (ER-L) was set at the 10th percentile of the ranked data and was taken to represent the point below which adverse effects are not expected to occur. The second level, the Effects Range-Median (ER-M), was set at the 50th percentile and

interpreted as the point above which adverse effects are expected. A direct cause and effect linkage in the field data was not a requirement for inclusion in the analysis. Therefore, adverse biological effects recorded from a site could be attributed to both a high concentration of one substance and a low concentration of another substance if both substances were measured at the site. The adverse effect in field data could be caused by either one, or both, or neither of the two substances of concern.

The State of Florida efforts (1993, in press) revised and expanded the Long and Morgan (1990) data set and then identified two levels of concern for each substance: the "TEL" or threshold effects level, and the "PEL" or probable effect level. Some aspects of this work represent improvements in the original Long and Morgan analysis. First, the data was restricted to marine and estuarine sites, thereby removing the ambiguities associated with the inclusion of freshwater sites. Second, a small portion of the original Long and Morgan (1990) database was excluded, while a considerable increase in the total data was realized due to inclusion of new information. The basic criteria for data acceptance and for classifying the information within the database were essentially the same as used by Long and Morgan (1990).

The development of the TEL and PEL differ from Long and Morgan's development of ER-L and ER-M in that data showing no effects were incorporated into the analysis. In the weight-of-evidence approach recommended for the State of Florida, two databases were assembled; a "no-effects" database and an "effects" database. The PEL was generated by taking the geometric mean of the 50th percentile value in the effects database and the 85th percentile value of the no-effects database. The TEL was generated by taking the geometric mean of the 15th percentile value in the effects database and the 50th percentile value of the no-effects database. By including the no effect data in the analysis, a clearer picture of the chemical concentrations associated with the three ranges of concern; no-effects, possible effects, and probable effects, can be established.

3. Apparent Effects Thresholds (AET) and scatterplots

The AET approach is an empirical method applying the triad of chemical, toxicological, and benthic community field survey measures to determine a concentration in sediments above which adverse effects are always expected (statistically significant different of adverse effects are predicted at $p<0.05$) (EPA 1989). Each suite of measures consists of chemical and toxicological measures taken from subsamples of a single sample and benthic analysis conducted on separate samples collected at the same time and place. A large suite of chemical measures and a large number of sites

are required before an AET value can be estimated. The method assumes a single toxicant is responsible for effects measured at a given site. In addition, the value generated is by design, an effect level rather than a protective level. While above the AET one can expect adverse effects, the method does not recognize that below the AET adverse effects may be attributed to the substance of concern. A major limitation of the method is that the observed relationships between effects and chemical concentrations are based on correlations only (the relationship does not demonstrate cause and effect).

4. Correlations

Correlations between toxicity or benthic community effects and chemical concentration can be used to show the relationship between these factors. Correlation analysis is most useful in assessing which chemicals study-wide (or throughout a specific dataset) may contribute to toxicity or benthic effects.

5. Multivariate Analysis

Patterns of occurrence of pollutants can be identified using multivariate techniques (cf. Anderson et al. 1988). Procedures such as Principal Components Analysis can be used to reduce a dataset from a large number of individual measurements which are often correlated with each other to a small number of uncorrelated factors, each group representing a group of pollutants that have a similar pattern distribution. These groups can be used in scatterplots, correlation calculations or subsequent multivariate analysis.

6. Sediment Toxicity Identification Evaluation

Sediment toxicity identification evaluation (TIE) methods can be used to make a better estimate of the cause-and-effect relationship between chemicals and toxicity. TIEs provides strong scientific evidence that a chemical or chemical is causing toxicity. When a specific discharger is identified and the chemical of concern is known, a study can be performed to link the observed effects with the chemical on a site-by-site basis. Standard procedures for TIEs are unavailable.

7. Weight of Evidence

Use any available sediment guidelines outline in Alternatives 1 through 4. This approach relies on a preponderance of evidence with all available chemical screening levels to indicate when effects produced by specific pollutants are likely to occur.

The program has used individual measures such as the PEL or ERM as the values to make determinations of association between chemicals and toxicity.

References

Anderson, J.W., S.M. Bay, and B.E. Thompson. 1988. Characteristics and effects of contaminated sediments from southern California. Southern California Coastal Water Research Project. Final Report. SCCWRP Contribution No. C-297.

Long, E.R. and L. Morgan 1990. The potential for Biological Effects of Sediment-sorbed Contaminants Tested in the National Status and Trends Program. NOAA Technical Memorandum NOS OMA 52.

Long, E.R. D.D. MacDonald, S.L. Smith, and F.D. Calder. 1995. Incidence of adverse biological effects within ranges of chemical concentrations in marine and estuarine sediments. Environmental Management 19: 81-97.

MacDonald, D.D., K. Brydges, and M.L. Haines. 1993. Development of an approach to the assessment of sediment quality in Florida coastal waters. Prepared for the Florida Coastal Management Program, Florida Department of Environmental Regulation.

U.S. Environmental Protection Agency. 1989. Report of the Sediment Criteria Subcommittee: Evaluation of the Apparent Effects Threshold (AET) Approach for Assessing Sediment Quality. Office of the Administrator. Science Advisory Board. Washington, D.C.

U.S. Environmental Protection Agency. 1993a. Technical Basis for Deriving Sediment Quality Criteria for Nonionic Organic Contaminants for the Protection of Benthic Organisms by Using Equilibrium Partitioning. Office of Water. Washington, D.C.

U.S. Environmental Protection Agency. 1993b. Sediment Quality Criteria for the Protection of Benthic Organisms: Phenanthrene. Office of Water, Office of Research and Development, and Office of Science and Technology. Washington, D.C.

U.S. Environmental Protection Agency. 1993c. Sediment Quality Criteria for the Protection of Benthic Organisms: Fluoranthene. Office of Water, Office of Research and Development, and Office of Science and Technology. Washington, D.C.

U.S. Environmental Protection Agency. 1993d. Sediment Quality Criteria for the Protection of Benthic Organisms: Acenaphthene. Office of Water, Office of Research and Development, and Office of Science and Technology. Washington, D.C.

## What is a benthic impact?

Analytical Procedures For Assessing Benthic Community Degradation

The following issue paper summarizes the multistep procedures used previously for the analyses of benthic community data for the Bay Protection and Toxic Cleanup Program.

### Sampling Design and Collection

Samples for benthic community analyses were collected from the same sediment grab as samples for grain size, total organic carbons, biotoxicity and chemical tests. However, unlike for the other analyses, samples for benthic communities are collected at only one time period per site. Therefore, spatial distribution, including replication, is the major concern in a sampling design for benthic data collection. To date, the design for the collection of benthic community samples has been evolving with each successive project. We used previous data to determine the appropriate number of replicates -- a standardized method is needed.

### Sorting and Identifying

Individual benthic samples were processed and preserved immediately following collection. Laboratory processing of the benthic cores consisted of both rough and then fine sorting. Initial sorting separated animals into large taxonomic groups such as polychaetes, crustaceans, mollusks, and others (e.g. phoronids). These grouped animals were placed in separate vials. The vials contained pre-printed duplicate labels identifying the project, IDORG number, date collected, site/station and sample replicate number. Vials were bundled together according to station and placed in a specific area designated for the project. Sample residues were placed back into the original, internally-labeled jars for later re-examination by the QA Officer. Species identification and enumeration was conducted by highly experienced taxonomists. On occasion, specimens were sent to specialized expert taxonomists for species verification.

### Data Analyses and Interpretation

The identification of degraded and undegraded habitat (as determined by benthic community structure) was conducted using several common and well-documented methods. The following tests have been used to assess benthic community data only- no formal integration with results of laboratory exposures or chemical analyses were made. Results from benthic analyses alone often warranted further examination of certain sites.

Post species identification analyses included initial statistical tests that defined individual stations by mean species abundance (using replicates), standard deviation, standard error, confidence limits, etc. Following these statistical summaries, several analyses were performed to identify relationships between community structure at each site including, diversity/evenness

indices, habitat -- species composition analyses, dissimilarity matrices, assessment of indicator species and development of a benthic index, classification (cluster) and ordination (multidimensional scaling) analyses. Initially, a correlation matrix was produced from species density data from each site. From this matrix we ran several tests for association of variables.

Cluster analysis is a multivariate procedure for detecting natural groupings in data, and, for our purposes, data were grouped by average similarities in total composition and species abundance. We have used the average-linkage method which uses the average similarity of all species at each site. From this information we looked at site-related patterns such as which species dominated the community. Grouped stations were typically clustered at a conservative distance limit of 50-60% similarity- however, this level is purely arbitrary. At this juncture, physical parameters, typically grain size, were evaluated to determine if station clusters were influenced solely by habitat type. Since classification analyses have the tendency to force data into artificially distinct groups, another method, involving statistical rigor, was required to confirm the validity of group clusters. We chose multidimensional scaling.

Multidimensional scaling (MDS) has been used extensively in the analyses of benthic communities, particularly in estuarine and marine pollution studies. MDS is a procedure for fitting a set of points in space such that the distance between points correspond as closely as possible to a given set of dissimilarities. We chose multidimensional scaling over principal co-ordinate analyses because MDS is more flexible in terms of handling the large number of zero counts generally characteristic of species-samples matrices. It is important to note that, as with cluster analyses, MDS results are not definitive and must be used in conjunction with addition ecological information.

After classification and ordination patterns were determined, the raw data was reevaluated for species differences to determine which one(s) may have been responsible for influencing the observed patterns. Often, the presence of specific species indicated non-contaminated areas or sometimes sites of environmental recovery. Indicator species were selected on the basis of literature review (to determine distribution, life history strategies and habitat preference), and discussions with experienced benthic taxonomists (to address the benefits and limitations of using certain species as environmental stress indicators). Objective techniques from published literature have also been used.

Although there are problems with trying to simplify complex biological communities, we needed to develop a quantitative method that created a partition between degraded and undegraded areas. We previously realized that we could not conclusively

identify "hot spots" using only results from benthic community analyses- but that benthic analyses could justly describe "environmentally stressed" areas. The benthic index was based on species (indicators) and group (general taxa) information-mainly community parameters such as species richness, abundance and presence of pollution indicators- that identify the "extremes" of the community characteristics. Sites were ranked according to these extremes and were represented by a single value. In general, decreasing numbers of species, increasing numbers of individuals, and decreasing diversity values are common responses observed near polluted areas. These trends were incorporated into the index. One of the important restrictions with the existing method is that it evaluates only a very limited data set in dividing groups and subsequent ranking. Sites identified as degraded (or undegraded) are derived from a combination of test documentation- indicator species, benthic index, diversity analyses. Data has been presented mainly as figures and summary tables.

Data Integration

Analyses of patterns associated with biological, chemical and biotoxic variables were conducted separately so as to not confound results by creating circular arguments from data interpretation. The final strategy of analyses would be to relate biological patterns with environmental data, both chemical and toxicological, to see if assumptions of site degradation are valid.

## ISSUE 4

**Should we use a probability-based sampling design (random sampling) or directed point estimates (based on best professional judgement)?**

The major objective of the BPTCP is to find toxic hot spots. Once these hot spots are identified the program needs to determine the areal extent of the toxic hot spots identified. The BPTCP has used both non-random and random sampling designs. The approaches used by the Environmental Protection Agency's Environmental Monitoring and Assessment Program has been used during the screening steps of BPTCP monitoring.

Options Evaluated

1. Use a worst-case sampling design for site selection (i.e. point estimates of pollution).

   This approach is based on previous knowledge about the presence and distribution of potential sources of sediment pollution in the water body or previously known pollutants or biological effects in the water body. This sample design is useful as an initial survey to determine the potential for pollution-related problems, followed by a more complete sampling later (if needed). This approach is most useful when there is adequate information available from previous studies on biological effects present, measurements of chemicals present, sources and other information.

   A limitation of this approach is that the data collected from this type of survey can only be evaluated in terms of the sampling stations that are sampled. The areal extent of the pollution or biological effects can not be determined.

2. Use a random or stratified random sampling design for site selection.

   This design is most useful when little is known about the likely distribution of pollutants or biological effects in a water body. To use this design a grid is established and stations are randomly selected with each location having an equal probability of being sampled. The number of samples can be selected statistically based on the requirements of the survey (i.e., the objectives of the study) and acceptability of error rates. A stratified random design is distinguished from a purely random design by the selection of zones (based on available information) that exhibit similar levels of pollution, similar source type, or other characteristics. Samples are randomly collected in the various zones that are selected.

Using these approaches provides a statistical basis for determining the areal extent of the identified pollution or biological effects.

3. Use a combination of Options 1 and 2.

The BPTCP has used Alternative 3. Most of the screening and confirmation sampling stations have been selected using available information or the likelihood of effects being present at a site (some human activity that raised concern). Random or stratified random sampling designs have been used to support screening of water bodies (e.g., San Diego Bay, Newport Bay and several coastal lagoons in Southern California).

## ISSUE 5

**Should we use a screening and confirmation approach?**

Options Evaluated

1. Sample sites a single time.

   Under this option sites would be sampled one time and repeated sampling would not be required. This approach would only work with the definition of a toxic hot spot if information were available form other studies conducted prior to any new sampling because of the need for repeated measurements of effect.

   With this approach the samples collected may be collected with different equipment and tests may be performed with different test species.

2. Sample sites at least two times before toxic hot spots can be designated.

   In order to identify known toxic hot spots a two-tier process was used. The first tier was a screening step where a suite of toxicity tests is used at a site (one amphipod test and at least one other toxicity test (pore water, bedded sediment or overlying water test)). Sediment grain size, total organic carbon (TOC) and $H_2S$ concentration are measured to differentiate pollutant effects found in screening tests from natural factors. Chemical analyses (metals and organics) were performed on a subset of the screening samples.

   If effects were found at sites by these screening steps, the highest priority sites were retested to confirm the effects. In the confirmation step measurements were replicated. Chemical measurements (metals, organics, TOC, $H_2S$) and other factors (e.g., sediment grain size) were also measured. Measurements of benthic community structure were also be made.

   With this approach, the program measurements will be affected by temporal variability of the sites (between year variation if sampled in same season in following year or seasonal variation if sampled in different season).

3. Continue to sample at worst sites until well characterized (more than two samples).

   This option would repeat the monitoring identified in Option 2 until a few sites are very well characterized. Under this option uncertainty about a few sites would be decreased. New toxic hot spots would not be identified because effort

would be focussed on characterizing sites already identified.

The program has implemented Option 2. While at least one amphipod test is performed at each site, the additional test(s) have not been consistently performed.

## ISSUE 6

**Are the toxicity testing methods the most appropriate for meeting program objectives?**

Toxicity tests, using a suite of organisms and protocols, have been the primary tool used to screen potential hot spots and reference sites, and have also been part of the "confirmation" phase of the program. If significant toxicity ("associated with toxic pollutants") is observed at least twice in samples from a given site, then that site can be considered a hot spot under the BPTCP hot spot criteria. Toxicity testing methods are described in the BPTCP QAPP.

Toxicity tests used by the BPTCP to date include:

Solid-phase tests:

Amphipod 10-d survival test (Rhepoxynius, Eohaustorius, and Ampelisca)
Polychaete 20-d growth and survival (Neanthes)
Sea urchin 96-h embryo/larval development test at the sediment/water
interface

Pore water tests:

Sea urchin 1-h fertilization test
Sea urchin 96-h embryo/larval development test
Abalone 48-h embryo/larval development
Bivalve 48-h embryo/larval development
Amphipod (Eohaustorius) 96-h survival test
(Pore water was extracted initially by piston squeezing and currently by centrifugation.)

Specific methods for each test are included in laboratory SOPs based on ASTM protocols (amphipods, bivalves), draft ASTM protocols (Neanthes, sea urchin larval development), or draft EPA protocols (abalone, sea urchin fertilization). The methods for the sea urchin larval development test at the sediment water interface are currently in peer-review, and are similar to methods described for bivalves in the Puget Sound Protocols, except that a screen is used to allow for more complete recovery of test larvae. To date, these methods have met test acceptability criteria a high percentage of the time, and have shown a broad range of sensitivity to test sediments, from highly sensitive (pore water tests) to highly tolerant (Neanthes test).

**Biomarkers**

Bioaccumulation data seems to be useful to the BPTCP because it can indicate a direct association between contaminants and organisms. Mussel watch has pinpointed many hot spots throughout

the state, and the recent effort on bioaccumulation in fish from San Francisco Bay has indicated that most of the fish collected exceeded the EPA screening levels for PCBs and other contaminants. This has promulgated health risk warnings from the State and would appear to be a fairly useful method worthy of further consideration for classifying areas as hot spots. The major drawback to this approach is that fish are extremely mobile, and to use them to pinpoint a specific hot spot site is difficult, unless perhaps one can also show a link between sediment contaminants at the site with tissue contaminants in fish caught at the same site. One solution that has been suggested is that mussels and fish be used in concert. The mussels could be used to pinpoint hot spots, and the fish could be used to trigger health warnings. Is this mussel and fish approach worthwhile? Is bioaccumulation data of cost-effective and interpretable value to the program?

## ISSUE 7

**Are BPTCP Analytical Chemistry Methods Scientifically Sound and Appropriate?**

Analytical Methods, Analyte Lists, Detection Limits Currently Used in the BPTCP:

- Please see list of BPTCP organic and metal analytes and detection limits in QAPP
- Please see methods employed for organic and metal analyses in QAPP

What chemical methods should we be using?

Should we use EPA standard methods or use performance-based techniques? Many of the BPTCP fee-payers use EPA standard methods, due to permit requirements of the EPA, SWRCB, and US Army Corps of Engineers. Most of the national monitoring programs such as the NOAA Status and Trends program and EPA's EMAP program use a performance-based system, in which the participating laboratories must qualify to do the analysis by participating in the NOAA Status and Trends Program's Intercalibration Exercise.

The benefits of using EPA's methodology are:

1. They are well defined
2. There are many data sets that are available for comparison that are EPA methodology-based.

The disadvantages of using the EPA techniques are:

1. They can give inaccurate numbers.
2. The detection limits are almost invariably much higher than other techniques.
3. The techniques were developed 10 or 20 years ago for different equipment that was not as sophisticated as today's equipment (i.e. bench top GC/MS).
4. Two different laboratories can obtain very different sets of numbers using the same EPA technique, thus not insuring data comparability.
5. The laboratories using EPA techniques invariably state that the techniques have been modified, which further adds to doubts of comparability.

The pros of using a performance-based technique are:

1. They give accurate numbers and the detection limits are usually very low.
2. They are customized to take into consideration the latest in development of equipment or extraction techniques, thus leading to constant improvements.

3. The data is compatible and comparable with other programs participating in the NOAA program.
4. All our data to date has been collected by this technique, and if we changed there would be an unknown amount of incompatibility and incomparability.

The cons of the performance-based technique are:

1. The data may differ somewhat from that produced by the EPA technique.
2. The fee-payers, most of whom are required to utilize EPA standard techniques do not seem to understand the benefits/strengths of using performance-based techniques.

Other issues/questions regarding chemical methods

Should chemical analyses be performed upon pore water?

Trace organic and trace element compounds have been measured in bulk sediment exclusively, with the exception of 21 pore water samples which we performed a limited trace element analysis upon. If toxicity evaluations of pore water are to be incorporated into the final assessment of sediment quality, then it would seem that trace organics and elements should be measured in pore water.

It should be realized that the levels of organic compounds in pore water will be a function of the bulk concentration in the sediments, the water solubility of the compounds, and the organic content of the water (should we be measuring DOC in porewater, and not just TOC in sediments?). Preliminary toxicity tests could be performed to indicate the necessary detection limits to assess significant correlations with the chemistries.

Should the number/type of organic compounds currently analyzed for be increased/changed?

Please see the current analyte list in the BPTCP QAPP. Our thoughts on this question are that we should re-examine the list of compounds and make some changes/additions. Due to the nature of the toxicity tests being performed, there may be a higher tendency to indicate toxicity resulting from more water soluble compounds than those presently being determined by the BPTCP program.

Therefore, since correlations between chemistries and toxicity have been weak, it would seem desirable to expand the analyses into new classes of chemicals, such as aliphatics, pthalates, additional PCB's, etc. In order to expand this analysis in a coherent and cost-effective fashion, these expanded analyses might only be performed once a site has had fairly clear weight of evidence of being a hot spot, and a TIE approach would then seem to be very useful.

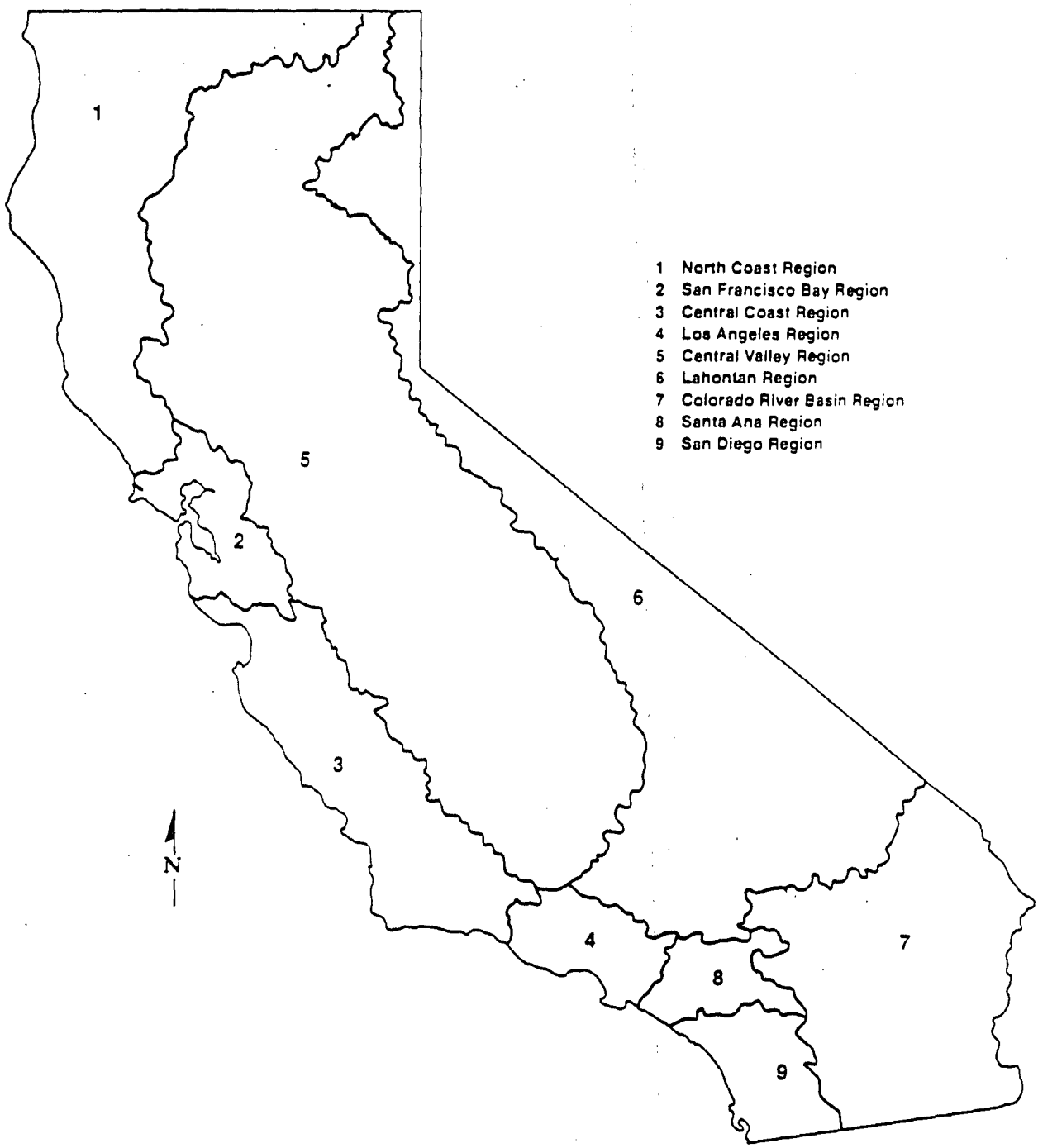
Should effort be directed toward identification of unknown peaks?

Almost invariably, numerous well-defined peaks appear on the chromatographs for compounds/classes of organic chemicals that we are not authorized to analyze. Perhaps by allocating as little as 5% of the resources dedicated toward organic analyses to attempting to at least index and quantitate these unknown peaks, we may find extremely useful information. The TIE approach may be a better approach, as analyte intensity does not necessarily correlate with toxicological impact.

Should chemistry be performed on screening samples, or just on confirmation samples?

To date, we have performed limited chemistry on screening samples, and on all confirmation samples taken. The hot spot criteria require an association with some level of anthropogenic chemical contaminants, necessitating chemistry to be performed on confirmation samples. It does not appear that there is any overwhelming need to perform chemistry on screening samples, unless perhaps we are trying to determine that a particular site is CLEAN and free of particular contaminants, for possible use as a field reference site (if indeed lack of contaminants is a prerequisite for field reference sites). A discussion of the rationale and timing for authorizing chemistry to be performed would be useful and helpful.

# REGIONAL SUMMARIES



REGIONAL WATER QUALITY CONTROL BOARD
NORTH COAST REGION
(REGION 1)

REGIONAL SUMMARY

## Monitoring Goals and Objectives

The overall goal is to collect data on sediment, water column and soil pore water quality in order to:

1. Identify types and distribution of toxic pollutants in North Coastal bays and estuaries, including spatial and temporal variations, for the purpose of identifying the location, extent and degree of toxicity of toxic hot spots.

2. Identify the condition of the resource/ecosystem and the effects of toxic hot spots on various species of the aquatic community. This will include characterization of background (reference) conditions of the resource/ecosystem.

3. Establish a database to measure future trends in the condition of toxic hot spots and their effects on the aquatic community.

Sampling and analysis techniques will be standardized techniques used statewide as recommended by the SWRCB study group [BPTCP Monitoring and Surveillance Task Force]. Specific water bodies and monitoring goals and objectives are as follows:

## Arcata Bay and Humboldt Bay Segments

1. Determine if dioxins or furans from airborne fallout, rainfall runoff or other form of discharge from two pulp mills have accumulated in sediments of Arcata Bay and all Humboldt Bay segments.

2. Determine if pentachlorophenol, tetrachlorophenol or dioxins/furans discharged near Mad River Slough, McDaniel Slough or south of Eureka as a result of discharges from wood treatment activities at lumber and plywood mills have accumulated in sediments of Arcata Bay and all Humboldt Bay segments.

3. Determine if polynuclear aromatic hydrocarbons or chlorinated solvents from various industrial complexes have accumulated in sediments along the Arcata or Eureka waterfronts.

4. Determine if pesticides in rainfall runoff from lily bulb growing activities in Arcata bottoms have accumulated in sediments of Arcata Bay and North Humboldt Bay.

5. Determine if polynuclear aromatic hydrocarbons from past activities of coal and oil gasification by historic gas utility companies have accumulated in sediments of Arcata Bay or North Humboldt Bay.

6. Determine if petroleum hydrocarbons or heavy metals from petroleum fuel storage/usage along the bay shoreline have accumulated in sediments.

7. Determine if bacteria contained in rainfall runoff from dairies and urban storm drains are adversely affecting commercial or sport shellfish harvesting. Identify climatic effects on bacterial concentrations.

8. Determine if constituents in toxic leachate from extensive redwood bark fill are accumulating in sediments of Humboldt Bay.

9. Determine if solvents, petroleum hydrocarbons, tributyltin or other metals from boats and boat servicing activities have accumulated in sediments around boat basins (Fields Landing, Woodley Island, Eureka small boat basin).

10. Determine if petroleum hydrocarbons or metals contained in urban runoff have accumulated in sediments in the vicinity of storm drain outlets or other portions of the bay where sediments are deposited.

11. Sample animal-sediment pairs at several locations to determine if toxic constituents accumulate to a higher degree in the tissue of the test animal or in the sediment. Identify natural (background) sediment conditions (sulfide and/or physical factors) which favor/disfavor animal recruitment.

12. Compare quality of Arcata and Humboldt Bay sediments with that of other, similar bays in California (Tomales, San Pablo ?). This would necessitate analyzing sediments from other, similar bays for the same constituents such as dioxins, furans, tetrachlorophenol, pentachlorophenol and polynuclear aromatic hydrocarbons.

13. Identify and diagnose stressed biological communities. Distinguish between sulfide and nonsulfide causes, favorable and nonfavorable factors.

14. Characterize sediment types and locations in all bay segments for physical factors related to sediment deposition (grain size, stratigraphy) and pollutant affinity.

**Bodega Harbor**

1. Determine if solvents, petroleum hydrocarbons, tributyltin or other metals from boats and boat servicing activities

have accumulated in sediments around boat basins (Tides wharf area, Mason's Marina, Spud point Marina).

2. Determine if bacteria contained in rainfall runoff from dairies are adversely affecting sport shellfish harvesting. Identify climatic effects on bacterial concentrations.

**Eel River Estuary**

1. Determine if petroleum hydrocarbons or metals contained in urban runoff have accumulated in sediments.

**Klamath River Estuary**

None at this time.

**Mad River Estuary**

None at this time.

**Noyo River Estuary**

1. Determine if solvents, petroleum hydrocarbons, tributyltin or other metals from boats or boat servicing activities have accumulated in sediments.

**Russian River Estuary**

1. Determine if petroleum hydrocarbons or metals contained in urban runoff have accumulated in sediments.

2. Determine if pesticides in rainfall runoff from extensive wine grape vineyards throughout the watershed have accumulated in the sediments.

**Smith River Estuary**

1. Determine if pesticides from lily bulb growing activities have accumulated in sediments.

2. Determine if pentachlorophenol, tetrachlorophenol or other toxic compounds from wood treating activities at a lumber mill have accumulated in sediments.

REGIONAL WATER QUALITY CONTROL BOARD
SAN FRANCISCO BAY REGION
(REGION 2)

REGIONAL SUMMARY

## Physical Description of the Region

The San Francisco Bay/ Delta Estuary, the largest estuary on the west coast of North and South America, is the main waterbody in this Region included in the Bay Protection and Toxic Cleanup Program. The San Francisco Estuary receives runoff from 14 watersheds having a total area of over 5 million acres. The San Francisco Bay Regional Water Quality Control Board has jurisdiction over the area from the vicinity of Antioch at the confluence of the Sacramento and San Joaquin Rivers west to include Suisun Bay, San Pablo Bay, Central San Francisco Bay and South San Francisco Bay. The Central Valley Regional Water Quality Control Board, Region 5, has jurisdiction over the area east of Antioch that makes up the Sacramento-San Joaquin Delta. Like all estuaries, the San Francisco Estuary is a trap for suspended particulate matter. It is estimated that the total annual amount of sediment deposited throughout the Bay is 4.38 metric tonnes. Because the Bay is so shallow, 40% is less than 2 m deep and 70% less than 5 m deep, sediment resuspension and redistribution is very high compared to other estuarine systems (i.e, Chesapeake Bay, Hudson River and Puget Sound). Tidal action, currents and wind play a large role in the resuspension and transport of sediments especially in the large, shallow embayments of Suisun, San Pablo and the South Bay.

Suisun Bay is a shallow embayment surrounded by Suisun Marsh, the largest brackish water marsh in the United States. The narrow Carquinez Strait joins Suisun Bay with San Pablo Bay. San Pablo Bay is a large, shallow, open bay that is largely influenced by outflow from the Delta. It is the deposition site for many of the fine-grained sediments carried out of the Delta by high winter flows. The Central Bay is the deepest part of San Francisco Bay and has the most oceanic influence. South San Francisco Bay receives much smaller amounts of freshwater inflow from the surrounding watershed and, as a result, is more like a shallow tidal lagoon. Tributaries to the San Francisco Estuary, as well as, several coastal embayments and lagoons, which have a relatively low level of anthropogenic impact, are also included in this program.

Sources of contaminants to San Francisco Bay include over 200 permitted discharges, including 50 POTWs with a combined design flow of 829 MGD, urban runoff, many boatyards and marinas, dredging activities and historical dumping. In addition, historical mining activities and agricultural runoff are sources

of metals and pesticides from higher up in the watershed. All of these sources of contaminants plus the fill of wetlands and water diversions have been the major impacts to the health of the Estuary.

## Philosophy of Monitoring in the Region

The main philosophy in the Region towards monitoring is the measurement of ambient trends in the watershed through the Regional Monitoring Program and comparison of those trends and measurements to monitoring programs being conducted near points of discharge. In this Region many dischargers have conducted their own monitoring programs (Local Effects Monitoring Programs) or special studies, many of which have included sediment studies. In 1993, we instituted a Regional Monitoring Program, managed by the San Francisco Estuary Institute, in which water column chemistry and toxicity, sediment chemistry and toxicity, bioaccumulation and benthic community analysis are analyzed several times a year throughout the Estuary (from the South Bay to the Sacramento-San Joaquin Rivers). Since one of the major long term goals of this strategy is to gain a clear understanding of ambient conditions and spatial and temporal trends in the watershed, and compare them to areas where there are current or historical discharges, the identification and characterization of reference sites has been very important, especially for sediment toxicity tests.

## Bay Protection and Toxic Cleanup Studies

These studies are described according to discipline and emphasis in the program (sediment studies, bioaccumulation studies and transport studies) and not necessarily in chronological order.

1. San Francisco Estuary Pilot Regional Monitoring Program: Sediment Studies

   The main objectives of this study were to: 1) screen critical habitats (marshes and mudflats) near potential sources of contamination to identify potential toxic hot spots, 2) develop a baywide sediment monitoring program that would act as a pilot program to define ambient conditions and 3) evaluate the use of various sampling and testing methods to use in monitoring programs. To achieve the first objective, sediment chemistry and toxicity were measured at 32 stations in critical habitats throughout the estuary. To achieve the second objective, sediment was collected at 15 stations that were thought to reflect ambient conditions. These samples were collected during wet and dry seasons and were geographically distributed throughout the Estuary. Sediment chemistry and toxicity were measured. In both the critical habitat study and the baywide study three toxicity tests were used: the solid phase 10 day amphipod test using

Eohaustorius, the bivalve larvae development test using an elutriate and the Menidia growth and survival test using an elutriate. In the second baywide run the Menidia test was eliminated due to lack of sensitivity. For all of these samples the depositional layer was sampled which was characterized by brown, loose sediment lacking the smell of hydrogen sulfide. This layer varied between 2 to over 20 cm. A reference site in Tomales Bay was used to compare sediment chemistry and toxicity with test sites.

To evaluate various sampling and testing methods a study was conducted on a sediment gradient that had been contaminated by a oil refinery. The main purposes of the gradient study were to: 1) determine which toxicity tests or phases (solid phase, elutriate, or pore water) could best distinguish between highly contaminated, moderately contaminated and relatively uncontaminated sites, 2) evaluate the degree to which field replication increases the ability to distinguish between sites, 3) determine the effects of sample depth, 4) determine the relationship between toxicity and factors that may effect toxicity including the levels of chemical contaminants, total organic carbon, grain size, ammonia and sulfides and 5) determine the relationship between toxicity test results and benthic community analysis. Five field replicates were collected at each of four stations on the gradient. Samples of the depositional layer were collected, as well as, samples one foot deep for each of the field replicates. Tests included solid phase and pore water chemistry, the 10 day solid phase amphipod test using Eohaustorius, the bivalve development test using an elutriate and pore water and benthic community analysis. On a subset of samples biomarker measurements (exposing speckled sandabs to sediment in a lab and analyzing for P450, EROD activity, stress proteins and histopathology), as well as, pore water tests that included sea urchin fertilization, development, cytologic and cytogenic effects, nematode broodsize and mutagenic effect, amphipod tests using intact cores and bacterial mutagenicity were conducted.

Using data from the baywide and critical habitat studies, areas were identified that had high levels of contaminants and/or toxicity. These areas are included in the potential toxic hot spot list. The results of the baywide and critical habitat studies showed that nickel exceeded the ERM in all samples and seemed to be the result of geologic deposits. The Tomales Bay reference site, although removed from sources of contamination, was toxic approximately half of the time when compared to controls. Other stations along the coast that were evaluated to be used for reference sites because of the lack of contaminant sources also proved to be toxic in toxicity tests. The Menidia growth and survival

test seemed to be the least sensitive of the three tests conducted in these studies. The baywide studies have formed the basis of the sediment portion of the Regional Monitoring Program.

In the gradient study, contaminants measured in the solid phase significantly correlated with each other and with related variables such as organic carbon and nitrogen. Concentrations of metals, extracted with aqua regia, were poor predictors of pore water metal concentrations. The amphipod test was significantly correlated with all of the contaminant and related variables and had low field variability. Toxicity was higher in the deeper cores where chemical concentrations were higher. For the bivalve larvae tests, pore water tests were more toxic than elutriate tests, field variability was greater than laboratory variability, and toxicity was also greater in the deeper cores. Benthic community analysis could not detect differences between stations along the gradient. Sea urchin development had a strong relation to bivalve larvae development but a poor relation to sea urchin fertilization. In the pore water tests neither ammonia or sulfides seemed high enough to cause toxicity. The PAH content of the sediment was significantly correlated with P-4501A content of the gills, hepatic EROD activity and gill histopathology. Although these were the major findings of the gradient study, analysis of this data is continuing through another Regional Board contract.

In addition to these results, this study provided the groundwork for a data management system currently being used by the Bay Protection and Toxic Cleanup Program and the San Francisco Estuary Regional Monitoring Program.

2. Reference Site Study

The main purposes of this study are to: 1) identify sediment reference sites in San Francisco Bay to use in toxicity tests, 2) recommend sediment toxicity test protocols to use in monitoring sediment toxicity in San Francisco Bay, 3) develop sediment Toxicity Identification Evaluation (TIE) protocols that can be used in San Francisco Bay and 4) identify the cause of toxicity at contaminated and previously identified reference sites. This study is currently in progress but nearing completion. For this study five potential sediment reference sites were chosen. Two sites were in San Pablo Bay, one site was in the Central Bay and two sites were in the South Bay. Chemical analysis has been or will be conducted at all sites that do not show toxicity. Sediment samples from Tomales Bay and several contaminated sites were also collected for comparison. All potential reference sites had three field

replicates. In addition, all potential reference sites, except those in the South Bay, were sampled three times during the year during different hydrologic conditions. Since the most likely locations to find reference sites were in San Pablo and the Central Bay, those sites were chosen first. Since these sites seemed to be good reference sites based on results from the first two sampling events, additional sites were chosen in the South Bay. Between seven to nine toxicity tests were performed on each sample. These tests were: 1) the 10 day solid phase amphipod test using Eohaustorius, 2) the 10 day solid phase amphipod test using Ampelisca, 3) the 10 day amphipod test using Eohaustorius in undisturbed cores, 4) the 10 day amphipod test using Eohaustorius in pore water, 5) the bivalve larvae development test in pore water, 6) the urchin larvae development test in pore water, 7) the urchin larvae development test using a sediment/water interface exposure, 8) the Neanthes growth and survival test and 9) a 10 day solid phase test using Nubelia. Toxicity tests were dropped out of the study based on their level of control survival, performance at reference sites and sensitivity to contaminated sites.

The first step in this project was to develop Sediment TIE protocols for the 10 day amphipod test, the bivalve larvae development test and the urchin larvae development test using pore water. When all laboratory tests were completed including pore water extraction experiments, testing the sensitivity of the various organisms to TIE manipulations and spiking experiments, the field portion of the study began. Samples were collected at the reference sites with enough field replication to try to determine field variability and during different hydrologic cycles to try to determine seasonal variability. By collecting the samples in this way, we hope to identify reference sites, determine the variability at those sites for statistical purposes, and identify sediment toxicity tests that perform well at reference sites but are sensitive to contaminated sites. Once this study is completed and reference sites are identified, testing of these sites will continue and data will be added to develop a "reference envelop" for these sites. In addition, we performed the amphipod test with undisturbed cores and the urchin test using a sediment/water interface to evaluate the environmental relevance of the standard amphipod and urchin tests. These tests could possibly be used in confirming toxic hot spots.

When samples were found to be toxic, a TIE was performed using the pore water test that exhibited toxicity. The first two field TIEs were performed on sediment from Islais Creek, where the City of San Francisco has had their main outfall for decades, and on the Tomales Bay sediment. After

removing ammonia and hydrogen sulfide from the Islais Creek sample, toxicity remained. After running TIEs on both samples results seemed to indicate that in both samples toxicity was being caused by a polar organic degradation product. Additional work has been performed to try to extract and identify the cause of this toxicity. Draft reports for this study are due July 1995.

3. Screening for Sediment Toxicity in San Francisco Bay

In this study, 49 sites will be screened for toxicity using the 10 day solid phase amphipod test using Eohaustorius and the urchin development test using pore water. Preliminary results from the reference site study seem to indicate that these are the two most reliable standard tests. Sediments from reference sites identified in the reference site study will be sampled concurrently. Test results from reference sites will be compared to test site results. This study has just begun.

4. Contaminant Levels in Fish Tissue in San Francisco Bay

Since one of the working definitions of a toxic hot spot involves the suitability of fish for human consumption, we conducted this study to measure contaminant levels in fish caught and consumed by anglers in San Francisco Bay. The main objectives of the study were to identify, to the maximum extent possible, the chemicals, species and geographical areas of concern in San Francisco Bay. This study was designed in a coordinated effort between state agencies, environmental groups and anglers. Thirteen fishing piers were sampled for fish with a small habitat range. Other regions of the Bay were sampled for fish that had a larger habitat range. The species of fish that were collected were white croaker (which was the highest priority fish based on its feeding behavior and lipid content), shiner surfperch, walleye surfperch, leopard sharks, brown smoothhound sharks, striped bass, sturgeon and halibut. EPA Screening Values based on the consumption rate of 30 grams per day were used to screen the data for potential chemicals of concern.

Results showed that: 1) The EPA guidance document, Guidance For Assessing Chemical Contaminant Data For Use In Fish Advisories- Volume 1- Fish Sampling And Analysis (EPA 823-R-93-002, 1993), was an effective tool for designing the pilot study and analyzing data collected from the San Francisco Bay study. 2) Based on EPA screening values six chemicals or chemical groups were identified as potential chemicals of concern in San Francisco Bay. They were PCBs, mercury, dieldrin, total DDT, total chlordane and the dioxin/furans. 3) High levels of the pesticides dieldrin,

total DDT and total chlordane were most often found in fish from the North Bay. 4) Levels of PCBs, mercury and the dioxin/furans were found at concentrations exceeding EPA screening values throughout the Bay. 5) Fish with high lipid content (croaker and shiner surfperch) in their muscle tissue generally exhibited higher organic contaminant levels. Fish with low lipid levels (halibut and shark) generally exhibited lower organic contaminant levels. 6) Of the Bay fish collected, white croaker consistently exhibited the highest tissue lipid concentrations. Lipophilic PCBs and pesticides concentrated to the highest levels in the muscle tissue of these fish. 7) Mercury levels were found to be the highest in the two shark species collected; the leopard shark and the brown smoothhound shark. Both the sharks and white croaker exhibit increasing mercury concentration with increasing fish size indicating bioaccumulation of this metal in Bay area fish. 8) Vallejo-Mare Island was the sampling location from which fish most often exhibited high levels of chemical contaminants. Oakland Inner Harbor also exhibited a high incidence of tissue contamination. As a result of this study, the Office of Environmental Health Hazard Assessment (OEHHA) has issued an interim health advisory for consuming fish caught in San Francisco Bay. OEHHA is currently in the process of using this data to conduct a thorough health risk assessment of consuming fish in San Francisco Bay. This study was designed partially due to the great interest on the part of the public in this issue. The results of this study have produced more public interest than any other of our Bay Protection studies.

5. Bioaccumulation of Trace Metals and Organics in Bivalves in San Francisco Bay

The California Mussel Watch Program, which has been measuring contaminant levels in bivalves throughout the state for the past 16 years, has proven to be a valuable tool for identifying areas with high levels of contaminants and for tracking trends in contaminants. This study was designed to test some of the assumptions inherent in the program and to determine if the program could be better designed for monitoring contaminants that bioaccumulate in San Francisco Bay. The main objectives of the study were to: 1) describe the distribution of trace metals and organics in organisms in San Francisco Bay; 2) determine the difference in contaminants collected during wet and dry seasons; 3) determine the differences between mussels transplanted high in the water column and down by the sediment at the same station; 4) determine the effect of depurating sediment from the guts of organisms on the contaminant levels in the whole body; 5) determine the optimum length of exposure for transplant organisms and

6) determine the differences in species uptake at selected stations. Eight bivalve transplant stations were chosen that were geographically spread from the South Bay to the Sacramento and San Joaquin Rivers. Three species of bivalves were transplanted depending on their salinity tolerances. These were the mussel Mytilus californianus, the fresh water clam Corbicula and the oyster Crassostrea gigas. Multiple species were transplanted at several stations. Bivalves were transplanted for 30, 60, 90 and 120 days. Bivalves were deployed during wet and dry seasons. At selected stations mussels were transplanted high in the water column and down by the sediment. Some mussels were depurated while some were not.

Results showed that most of the stations within San Francisco Bay accumulated contaminant levels that were significantly higher than controls collected at sites in more pristine locations. Stations in the South Bay, especially Coyote Creek, were significantly higher than the Central or North Bay stations for DDT, PCBs, chlordane, and PAHs. This was the first indication that organic contaminants may be a problem in the South Bay. Previously, Regional Board efforts were focusing on metals concentrations. Silver was significantly higher in the South and Central Bay than in the North Bay. There were no significant differences in contaminant levels between wet and dry seasons (this was a dry year) or between surface or bottom deployed mussels. A small number of metals was significantly different between depurated and undepurated mussels. An equilibrium appeared to be attained during the 90 to 120 day transplants for copper, mercury, lead, selenium and chlordane. No equilibrium was obtained in mussels for silver, PCBs and possibly DDT after 120 days. Oysters and mussels exhibited similar concentrations of chlordane, DDT and PCBs. However, PAHs and all metals were different between the two species. Recommendations are made in the report for deploying bivalves in San Francisco Bay based on these results.

**Projects to Collect Information for the Cleanup of the South Bay**

The purpose of the following projects is to develop the information necessary to use a watershed management and wasteload allocation approach to attain water quality objectives in the South Bay. This information will be used to develop cleanup plans for the South Bay based on wasteload allocations, sediment dynamics and hydrodynamic modeling. The South Bay was identified as an impaired water body through the Clean Water Act 304(l) listing process and was designated a candidate toxic hot spot under the Bay Protection Program because of repeated exceedences of water quality objectives. The pollutants of concern identified at that time were heavy metals, and the sources were

three POTWs and storm water. Advanced treatment is already in place at the treatment plants, and the effluent quality is quite high, so that additional treatment was expected to be very costly. In addition, it was unclear to what extent remobilization of sediment-bound pollutants (as opposed to ongoing discharges) was responsible for receiving water conditions. In order to determine what level of pollutant reduction was necessary to clean up the hot spot, additional information was needed to determine what level of discharge would result in the attainment of water quality standards. In addition to conducting these studies, stormwater is being monitored through another Regional Board program.

1. Wasteload Allocation Modeling

   The purpose of this project was to use existing EPA models and available data to determine the allowable level of loading of copper, nickel and lead to the South Bay. CEAM used the WASP4 model for this purpose. The model incorporated hydrodynamics, sediment transport and sediment-water partitioning of metals. They concluded that significant reductions in loading were needed to attain water quality standards. However, based on comparisons of model results with current ambient water quality, Regional Board staff concluded that the model was not accurate enough to form the basis of regulatory decisions.

2. 2-D Hydrodynamic Modeling and Sediment Dynamics

   This is the largest project in this category, and it has two components. After the experience with the CEAM model, we decided that existing models and data did not allow accurate modeling of pollutant fate and transport, but that modeling the physical processes could provide valuable information for estimating pollutant residence time. The goals of hydrodynamic modeling were to calculate the dry weather hydrodynamic residence time of the extreme south Bay, and estimate the dry weather sediment residence time of the South Bay, and use these two values as a range for pollutant (metals) residence time. The estimate of sediment residence time will be based on a the idealized approach of tracking a particle (in the model) that deposits and goes into suspension at the appropriate water velocities (determined from the suspended sediment monitoring). Modeling was (and is being) performed using TRIM2D (depth averaged), developed by Cheng and Casulli. Because of the lack of data describing sediment movement in the estuary, the sediment dynamics aspect of the project focuses on data collection rather than modeling. Time series of suspended sediment concentrations are being collected at 15 minute intervals at three locations in the South Bay (2 depths each) using optical backscatter (OBS) sensors. The data are analyzed to

determine the influence of tides, wind, and freshwater inflows on suspended sediment concentrations. In addition to the South Bay stations, there are similar stations in Central Bay funded by the US Army Corps of Engineers and the Regional Monitoring program, and in Suisun Bay, funded by the USGS. Therefore, forcing factors for sediment resuspension can be compared for different parts of the Estuary. In addition to the long term stations, there have been several 30 day deployments of OBS sensors in shallow water, both in North and South Bay. This component of the project has produced three to four (depending on the station) years of suspended sediment data. Data analysis to date has determined that in the South Bay, the spring-neap tidal cycle is the most important factor in determining suspended sediment concentrations. Both data collection and data analysis continue.

3. South Bay Bathymetry

Hydrodynamic modelers have concluded that models (and TRIM2D in particular) are very sensitive to bathymetry. Much of the extreme South Bay is mudflats, for which depths are not included in NOAA maps. Therefore, the purpose of this project was to produce accurate bathymetry of the South Bay, south of Dumbarton Bridge, for use in hydrodynamic models. Aerial photos were taken over the course of a tidal cycle, so that the water level could be used as isobaths. Water levels were adjusted to 1929 NGVD elevations after surveying the benchmark using global positioning system. In addition, it was determined that MLLW is 1.25 below the NGVD datum, so that depths in the channel were corrected as well as in the flats. Products were a bathymetric map and a computerized bathymetric grid with resolution of 0.1 m. Volumes of the South Bay at different tidal elevations were calculated as well. Modelers at USGS district office and at Stanford are now using the new grid.

4. 3-D Hydrodynamic Modeling

The 2-D model described above is depth averaged, and it is unclear whether it can adequately characterize depth dependent phenomena such as stratification in wet weather and sediment transport. That's why the project described above will only estimate residence times for dry weather. The purpose of this project is to apply TRIM3D to the South Bay to estimate residence times. In addition to providing very high quality characterization of the hydrodynamics of the region, the project is a test case to determine whether (or under what conditions) the additional effort involved in 3D modeling is merited. This contract was executed last month, and has not yet produced results.

## Questions and Issues Particular to this Region

1. Reference Sites - This Region has placed a great deal of emphasis on identifying and characterizing sediment reference sites. We believe that this type of characterization is necessary in order to identify toxic hot spots.

2. Sediment Sample Depth - This Region has been sampling sediment for chemistry and toxicity first at the depositional layer and then for consistency at 5 cm. Other Regions have sampled at a depth of 2 cm. We believe that the dynamic nature of this Estuary requires deeper sampling. Results from the Pilot Regional Monitoring Program and USGS indicate that the top 2 centimeters is very mobile due to resuspension and transport. Sediments could be eroded away at a particular site or buried very quickly. Monitoring the top 2 cm in an ongoing monitoring program would make some sense, but sampling the top 2 cm to determine if there is a toxic hot spot, we believe, is not a sufficient characterization. Determining whether an area is depositional or erosional would come in to play when evaluating, during the cleanup plan process, whether an area is being capped or eroded.

3. Benthic Community Analysis - In the San Francisco Bay Estuary fluctuating salinity, water movement and grain size play a major role in determining benthic communities. In addition, exotic species are introduced frequently that play a major role in the makeup of the benthos. Although there has been a considerable amount of work to date on the benthos of the San Francisco Estuary, the effect of contaminants on the benthic community is still too unclear to take a sample and determine the cause of different species assemblages or biomass. In addition, it is very difficult to find appropriate reference sites. Even when sampling a contaminated gradient, the impact was unclear. Should we sample for benthic community analysis? It seems that it is a waste of funds until we know how to interpret the data. On the other hand, it is the most realistic evidence of impact. Any suggestions?

4. Designation of Hot Spots Based on Exceedences of Water Quality Objectives or Elevated Contaminant Levels in Tissues - In this Region we have data on the levels of metals and organics in the water column. We also have health advisories that have been issued for fish that have a fairly wide habitat range. Since a hot spot designation can be triggered by water quality objective exceedences that are contained in our Basin Plan or by Health Advisories, we would like some guidance on how to delineate this type of hot spot. Our main thought, at this time, is to address

both of these types of hot spots by developing watershed management plans and conducting ongoing monitoring programs.

5. Bioaccumulation in Screening - In the Bay Protection Program we are screening sites by measuring toxicity at a station. However, in this Region we believe that bioaccumulation from the sediments into higher trophic levels has led to Public Health Advisories for the consumption of fish and may be contributing to the decline of different populations. Currently, if there is no toxicity at a station that station is not revisited. Are we "missing the boat" by not screening for bioaccumulation?

6. The Use of Sediment Toxicity Identification Evaluations (TIEs) - For the Reference Site Study conducted in this Region we have developed methods for conducting TIEs in pore water with estuarine species. We believe that this is a very useful tool in determining if ammonia, hydrogen sulfide, anthropogenic contaminants or other natural factors are causing the toxicity seen in toxicity tests. Currently, if a station has ammonia or hydrogen sulfide levels that could impact a particular test that station is eliminated as a potential toxic hot spot. Yet, something else could be causing the toxicity. We believe that abbreviated TIEs could be used to determine if toxicity is actually being caused by ammonia or hydrogen sulfide and full TIEs could be used to identify the cause of toxicity either to designate a candidate toxic hot spot or to determine cleanup options for known toxic hot spots.

**Additional Data**

The San Francisco Estuary Regional Monitoring Program continually collects data on water column chemistry and toxicity, sediment chemistry and toxicity, and bioaccumulation. Dischargers conduct their own Local Effects Monitoring Programs. In addition, the Department of Defense and dredgers have conducted many investigations for base closures and dredging operations. In our preliminary toxic hot spot list 110 of these studies are listed. This list has been expanded to include 122 studies and is continuously being updated.

REGIONAL WATER QUALITY CONTROL BOARD
LOS ANGELES REGION
(REGION 4)

REGIONAL SUMMARY

**Physical Description**: The region contains two large deepwater harbors and one smaller harbor. There are small craft marinas within the harbors as well as tank farms, naval facilities, fish processing plants, boatyards, and container terminals. A number of separate small craft marinas occur along the coast; these contain boatyards, other small businesses, and dense residential development.

Several large concrete-lined rivers lead to unlined tidal prisms which are for the most part marine-influenced. Salinity may be greatly reduced following rains since these rivers drain large urban areas composed of mostly impermeable surfaces. Some of these tidal prisms receive a considerable amount of freshwater throughout the year from POTWs discharging tertiary-treated effluent. Lagoons are located at the mouths of other rivers draining relatively undeveloped areas with some degree of agricultural activity (Mugu Lagoon, and lagoons at the mouths of the Ventura and Santa Clara Rivers). There are also a few isolated coastal brackish water bodies receiving runoff from agricultural or residential areas.

**Results of Previous Studies (State Mussel Watch/Toxic Substance Monitoring/Regional Board Sediment Sampling)**: Previous work in deepwater harbors has revealed decreasing, but in some cases, still relatively high levels of DDT and its isomers in tissue and sediment. More recent SMW data for LA Harbor indicates that considerable water transport of DDT may be occurring in some areas since tissue samples rather than sediment are exhibiting high DDT concentrations. PCBs are also on the decrease but still show up in high concentrations in sediment and tissue near "problem sites." Other pesticides, except for TBT, are usually not a problem. Copper, zinc, and sometimes chromium tend to be elevated in sediment and tissue. PAHs are also a problem in inner harbor areas where liver lesions associated with the chemicals have been found in fish. The innermost part of LA Harbor (mouth of Dominguez Channel/Consolidated Slip) continues to show a degraded benthic community. Port Hueneme, the smaller deepwater harbor in Ventura County, is also contaminated with PCBs, DDT, and metals (sediment and tissue). Tissue and sediment samples from small craft marinas are generally moderately high to very high in copper, chromium, and zinc. Some small areas within the marinas are also high in PCBs, DDT, and chlordane.

Most of the tidal prisms of concrete-lined rivers have not been as thoroughly investigated; limited sampling of fish tissue and

sediment indicate some metals and pesticides contamination. Some of the lagoons had not been investigated prior to the BPTCP. Of those previously sampled, some are virtually uncontaminated while others are very contaminated. Malibu Lagoon is located at the mouth of Malibu Creek which drains a large part of the Santa Monica Mountains. Development is mostly residential with some commercial. However, sediment turnover in the lagoon is frequent and contaminants do not reside long enough to bioaccumulate or be found in the sediment. On the other hand, Mugu Lagoon has been occupied by the Navy for many years and its presence appears to have contributed to high sediment metals concentrations in some areas; however, pesticides found in the lagoon seem to be originating from the extensive agricultural land in the area. Very high concentrations of banned chemicals such as DDT and toxaphene still persist in the drains leading to the lagoon. The effects from these persistent chemicals include reduced reproduction of the endangered light-footed clapper rail. The miscellaneous isolated brackish water bodies have been largely uninvestigated but merit attention due to their support of large numbers of migrating and overwintering birds.

**Sampling Goals**: The goal has always been to identify "hot spots", pursue identification of the problem's source, eliminate the source (permits, enforcement orders, etc.), and then go back and monitor for recovery of the hot spot. This is consistent with the goals of the BPTCP but on a much smaller scale.

**BPTCP-related Goals and Objectives:** Because of the results of previous monitoring, certain water bodies were designated candidate toxic hot spots right from the beginning (parts of LA/LB Harbors and Mugu Lagoon). The program goal for these sites was confirmation of candidate toxic hot spot status. The rest of the water bodies were to be screened for sediment toxicity with higher priority given to those water bodies designated as potential toxic hot spots from previous studies. My objectives for LA/LB Harbors were to target the candidate and potential hot spots preferentially in order to resolve whether sediment contamination resulted in an effect other than bioaccumulation (toxicity or benthic impacts). Unfortunately, my goals and those of NOAA, which supplied a large amount of additional money for more generic monitoring, were not compatible and many suspected hot spots (plus one candidate site) were not sampled. It's been suggested that these data (which includes some confirmation work) may be used for screening purposes instead. I would like to be able to do that and see no reason why it can't be done. There was also some concern about the timing of the sampling phases and the possibility that a lot of changes due to storm events had occurred. I don't think that's a problem, at least in the deepwater harbors. Previous sampling seems to indicate that sediment changes occur slowly over the years in these water bodies. At this point I would like to do confirmation work at the suspected hot spots and move on.

As for the rest of the region's water bodies, some time ago we were all requested to formulate monitoring plans for our water bodies. I planned on targeting sites with known and highly suspected problems (near storm drains, confluence of problem areas, etc.). Unfortunately, those plans were tossed when money ran out and instead screening of potential hot spots was accomplished with one sample per water body for the most part. I would prefer to concentrate more sites in higher priority water bodies and completely leave out water bodies I feel previous data tell me are of only moderate concern and extremely unlikely to gain attention to the point where a "cleanup" is conducted. These water bodies will still need remediation plans, but I think source control and prevention programs will be the answer.

**Issues/Questions Generated by This Work**: 1) Do we always need field replicates (for screening or confirmation) considering the extra costs involved and in what situations can we get away with not collecting them? 2) Is collecting AVS and SEM data worth the extra cost? If so, is it recommended this be done on a regular basis or only under certain circumstances? 3) Does porewater toxicity by itself tell us anything or is chemistry always needed? Is porewater toxicity and chemistry giving useful information or just more information? 4) Which would be better: utilizing several acute toxicity tests or having a mix of acute and chronic tests? 5) Should we be gathering chemistry data on nontoxic sites also? We aren't right now.

**General Issues/Questions**: With regards to the toxic hot spot definition, when "the water or sediment exhibits toxicity associated with toxic pollutants" the site is considered to be a toxic hot spot. While "what is toxic?" is certainly one question that immediately arises, that is already being dealt with in a number of ways, especially at DOD sites. The other question that arises is, "how strong an association do we need to have?" There has been a tendency thus far in the program to consider the "conventional (co-occurrence) approach" to be completely unacceptable.

The conventional approach appears to be comparison of test sites with a biased group (nontoxic, low pollutant-level) of reference sites. The argument against doing this is that there are probably sites out there that are nontoxic and relatively high in pollutants that are just not bioavailable. A RP might just argue about cleaning up a site exhibiting toxicity with high contaminants when high contaminants elsewhere don't cause a problem. This argument makes a lot of sense but I think the approach can be changed somewhat and still be useful. Why not compare the test sites to a nonbiased group of reference sites (not pollutant-level dependent).

The recommended approach thus far has been what's called the "internal AET approach." This is very conservative and requires

a very rigid sampling scheme. It seems to be showing more a cause and effect of specific chemicals (probably just one step short of a TIE) rather than just an association. One of my suggestions has been to develop an approach midway between the two. Instead of clearly demonstrating an association between toxicity and manmade pollutants, why not demonstrate lack of association with natural pollutants?

REGIONAL WATER QUALITY CONTROL BOARD
CENTRAL VALLEY REGION
(REGION 5)

REGIONAL SUMMARY

## Physical Features

The Sacramento-San Joaquin Delta estuary is of ecological, aesthetic, and economic significance to California. Total area of the delta encompasses 4,950 square miles, including 90 square miles of water area. The delta provides drainage for one forth of the total area of the State. Major estuarine and tidally-influenced rivers of the delta include the Sacramento River, Mokelumne River, Consumnes River, Old River, Middle River and the San Joaquin River. The delta has major State and federal water project facilities including the Clifton Court Forebay, and the Delta-Mendota and California Aqueducts. Delta facilities provide approximately 40 percent of California's drinking water. Two thirds of the water consumed in California comes from the delta. One half of California's anadromous fishery passes through or lives on the estuary. The Port of Sacramento and the Port of Stockton are on the north and south ends of the delta. Within the delta lies 70 leveed islands, and 550,000 acres of agriculture.

## Goals and Objectives

Regional goals include: implementation of regional surveillance monitoring program to identify hot spots and focus monitoring to define extent of hot spots; use monitoring results to assess and rank hot spots for cleanup; formulate cleanup plans; and adopt or revise waste discharge requirements to bring about cleanup. Additional needs include development of freshwater sediment and water column aquatic life criteria that can be used to further define hot spots in the freshwater and saline portions of the delta.

## Summary of Studies

Originally, the Central Valley Region monitoring plan included 7 fixed station water column sites for metals and 24 for EPA three species water column toxicity located throughout the delta. This was done to define the extent of metal objective exceedences and toxicity throughout the delta. This work would also be used to determine metal loading patterns to the delta during normal and high flow (storm) events. During the same period additional monitoring was performed to assess water column toxicity from urban and agricultural discharges in main channels and back sloughs. In addition to the above monitoring efforts, three special studies were designed to assess the impact of metals and

related toxicity from the Northern Sacramento Valley and Sacramento Urban Storm Run-off. During toxicity tests a study of dissolved metals bioavailability was made.

Results of these projects have shown: significant toxicity from pesticide applications and discharges during peak runoff seasons (late winter, spring) in the form of short to mid-term pulse like movements into the delta from agricultural applications outside the delta; potentially significant mercury loads to the delta from coastal range streams during high flow events; significant toxicity in delta back sloughs due to pesticides in urban runoff from the Stockton area; and pesticides toxicity in urban stormwater sumps in Sacramento.

Current Monitoring program(s) - Because monitoring funds have been cut and problem areas have been identified, the regional fixed station monitoring approach was modified. A scaled back regional program is in place which is weighted to areas of the delta which have shown toxicity or potential problems in the past. Toxicity identification evaluations (TIEs) are being run on water indicating toxicity to determine responsible chemicals. Special projects are being used to determine the temporal and spatial extent and sources of toxicity from metals and pesticides.

- The 94/95 winter storms have provided extreme flow events which have indicated potentially significant and previously unknown sources of mercury to the delta and San Francisco Bay.

- Past bioassays have not indicated the presence of metal toxicity in Delta waters. These results were confirmed this winter.

- During the 94/95 winter the Board implemented a volunteer urban monitoring network to determine pesticide impacts on local creeks in the delta area. The purpose of the network is to sample Sacramento and Stockton area creeks, rainfall and atmospheric deposition and assist the Board in detecting pesticides. Diazinon and Chlorpyrifos have been found in California streams at levels that cause toxicity to bioassay organisms. Comparisons with literature values suggest that sensitive local organisms should also be affected. This study is designed to help determine how pesticides are moving into the urban creeks. The primarily results indicate that both pesticides are coming from orchard spraying and urban uses via runoff and atmospheric scavenging due to rainfall. The urban creeks and orchard drainage basins in the project area discharge to the delta.

**Additional Data Available**

1. Sacramento County and Sacramento City have implemented a semi-regional monitoring program to assessment ambient water quality conditions in the lower Sacramento River watershed primarily just upstream and downstream of the county urban area and including urban and industrial discharges. This program has been operating over the past 3 years with sampling events occurring every two weeks.

2. Deepwater ship channel maintenance projects have been performing sediment sampling and assessment prior to dredging activities during the past few years. This information is being gathered now by the Department of Water Resources and the Army Corp of Engineering for submittal to the Board for review and consideration of revised sediment assessment activities. This information may be limited in nature due to high detection limits and undocumented QA/QC.

3. USGS has been assisting the Board in identifying pesticide pulse movements and their fate in the Delta. They have also conducted a semi daily pesticide monitoring program at Tower Bridge in the City of Sacramento and at Vernalis on the San Joaquin River. These two sites are the legal upstream boundaries of the delta. The USGS has identified several new pesticides that may be of concern.

4. The Department of Pesticide Regulation has a pesticide monitoring program in the Sacramento River and San Joaquin River Watersheds. By in large their monitoring has confirmed Regional Board conclusions about pesticide concentrations in the two rivers.

5. The Department of Fish and Game has developed and continue to work on draft hazard assessment documents for agricultural pesticide commonly observed in the Sacramento and San Joaquin Rivers and delta at concentrations known to be toxic to sensitive aquatic life. Draft reports are out on Molinate, Thiobencarb, Methyl Parathion, Carbofuran, Diazion and Chlorpyrifos. No water quality objectives are available for these compounds. The hazard assessment reports may be helpful in prioritizing hot spot cleanups.

**General Issues and/or Questions**

1. Should we pursue freshwater sediment criteria given the budget constraints of the program when we are finding significant water column toxicity in the delta due to pesticides from surface water discharges (urban, agriculture) from within and outside the delta?

2. Should the program pay or support monitoring and assessment up the watershed (outside of the delta boundary) to provide information needed to write cleanup plans for sources (abandoned mercury mines, orchard runoff) of toxic hot spots?

3. Should the Board consider the entire Delta a Hot Spot for mercury based on the fish advisory or should the Board attempt to define specific areas or reaches of the Delta as a hot spot based on fish tissue exceeding human health protection values?

REGIONAL WATER QUALITY CONTROL BOARD
SANTA ANA REGION
(REGION 8)

REGIONAL SUMMARY

## Overview

### Anaheim Bay/Huntington Harbor Complex

Complex is approximately 5 miles long and one-half mile wide with one ocean inlet and three main freshwater sources (stormwater channels). Watershed is approximately 75 sq. miles, highly urbanized with heavy industrial and commercial activity.

1. Anaheim Marsh, Seal Beach National Wildlife Refuge, Seal Beach Navel Weapons Station, and Bolsa Chica Ecological Reserve - Remnants of larger coastal marshlands complex. Shallow, good tidal mixing in most of marsh, poor tidal mixing in Bolsa Chica.

   Problems: Copper, lead, chromium, zinc, DDT, DDE

2. Huntington Harbor - Heavily developed marina/urban setting. Moderate depth, periodically dredged.

   Problems: Copper, aldrin, chlordane, lead, zinc, DDT, DDE

### Newport Bay

Bay is approximately 4 miles long by three to one-half mile wide with one ocean inlet and two main freshwater sources (stormwater channels). Watershed is approximately 150 sq. miles, mostly urbanized/commercial with some agriculture and industry.

1. Lower Newport Bay - Urbanized setting, over 10,000 recreational boat slips, 9 boatyards. Dredged to moderate depth, main channel deeper. Good to moderate tidal mixing.

   Problems: cadmium, copper, lead, zinc, chlordane, PCB, tributyltin, endosulfan

2. Upper Newport Bay - State Ecological Reserve, estuarine wetlands. Main channel of moderate depth with mud flats in end of bay, moderate tidal mixing. Periodically dredged to remove trapped sediment.

   Problems: cadmium, lead, endosulfan, DDT

## Regional Monitoring Goals and Objectives

1. Identify toxic hot spots
2. Determine if level of toxicity impairs beneficial uses of water bodies
3. Identify probable sources of toxic pollutants

## Regional Questions

1. Which data analysis method should be used for the existing data and sites in our region?
2. Would a general weight of evidence approach work better than a strict set of criteria for designating toxic hot spots?
3. How does seasonal sediment deposition and removal affect toxicity results?
4. How should toxicity and chemical data collected over several years be interpreted in conjunction with seasonal sediment depositions?
5. If reference sites are used, should they be located within the general area or from a "clean" site outside the area if the area exhibits elevated levels for many constituents?
6. Are the porewater toxicity tests that have been preformed acceptable if the test organism does not naturally live in sediments?
7. What are ways to differentiate between natural variations in benthic community populations and anthropogenic induced impacts?
8. What conclusions can be made from a site with slightly elevated levels for a few constituents and high mortality on porewater toxicity test results?

## Data Available

### Anaheim Bay/Huntington Harbor Complex

| Source | Media/Tests | Results |
|---|---|---|
| Orange County EMA, 1979-95 | Water column, limited sediments | Background info |
| State Mussel Watch, 1983-94 | Bioaccumulation | Potential THS identified |

| | | |
|---|---|---|
| Toxic Substances Monitoring Program, 1983-94 | Bioaccumulation | Characterization |
| USFWS & USN, 1989 | Bioaccumulation | Unknown |
| Consultants Reports, 1992-93 | Sediments,Chem bioaccumulation | Potential THS identified |
| BPTCP/NOAA, 1992 | Sediments/Toxicity Chem | Screening/Potential THS identified |
| BPTCP Benthic Community Analysis, 1992 | Benthic Community | Characterization |
| BPTCP Screening, 1992 | Sediments/Toxicity Chem | Screening/Potential THS identified |
| BPTCP Screening, 1993 | Sediments/Toxicity Chem | Screening/Potential THS identified |
| BPTCP Confirmation, 1994 | Sediments/Toxicity? Chem | Confirmation of THS? |

## Newport Bay

| | | |
|---|---|---|
| Orange County EMA, 1979-95 | Water column, limited sediments | Background info |
| Seapy, 1981 | Benthic Community | Characterization |
| State Mussel Watch, 1983-94 | Bioaccumulation | Potential THS identified |
| Toxic Substances Monitoring Program, 1983-94 | Bioaccumulation | Characterization |
| Butler, 1988 | Benthic Community | Characterization |
| Rhine Channel Fish Tissue, 1992 | Bioaccumulation | Potential THS identified |
| BPTCP Screening, 1994 | Sediments/Toxicity? Chem | Screening/Potential THS identified |

REGIONAL WATER QUALITY CONTROL BOARD
SAN DIEGO REGION
(REGION 9)

REGIONAL SUMMARY

## San Diego Region Bays and Estuaries

### San Diego Bay

- Approximately 12 nautical miles in length and one-half to two miles wide
- Rainfall along coast about 10 inches per year, November to April
- Ship channel extends well into the southern area
- Population tributary to Bay maybe three-quarters of a million people
- Industrial activity goes back 100 years, with heavy military, aircraft, and shipbuilding activities since about 1940, and 50 million gallons per day of sewage discharges until 1963
- Each of the areas listed below represent approximately one-third of the Bay surface area

North Bay: Good tidal mixing, deeper, sandy bottoms, heavily developed shoreline, and heavy commercial and industrial activity

- Depths 8-41 ft with some deep scour areas, area mostly dredged
- Water temperatures about 16C in winter to 19C in summer
- Shoreline: Maybe 5,000 recreational and smaller commercial vessel slips, Naval Air Station and Submarine Base, and residential and commercial areas
- Runoff: 47 storm drains at least 30 inches in diameter
- Problems: Copper in marinas, PCB spills

Central Bay: Moderate tidal mixing, warmer water, area dredged to moderate depths, with heavy industrial activity

- Depths 5-38 ft with a narrow channel at northern end, mostly dredged
- Water temperatures intermediate between north and south Bay
- Shoreline: Maybe 2,000 recreational vessel slips and about 100 commercial and U.S. Navy ships, Naval Amphibious Base

and Naval Station, with four shipyards and heavy industrial and urban uses

- Runoff: Three creeks, 16 storm drains at least 30 inches in diameter

- Problems: Sediment oil deposits from spills along eastern shore near shipyards, copper from ship antifouling paints, PCB spills

South Bay: Poor tidal mixing, water warmed by power plant, mostly shallow

- Depths 1-18 ft with area mostly undredged

- Water temperatures up to about 21C in summer

- Shoreline: Maybe 1,000 recreational vessel slips, some industrial uses, two rivers tributary, remnant salt marshes, salt ponds

- Runoff: Two controlled rivers with relatively little flow, one creek, and only 3 storm drains at least 30 inches in diameter

- Problems: Copper concentrate, now cleaned up, deposited at marine terminal in National City

**Mission Bay**

- Approximately two nautical miles square

- Bay dredged to 8-12 feet over entire area

- Good tidal mixing in west Bay, poor in east

- Two creeks tributary to east Bay

- Shoreline: Maybe 2,000 recreational and party boat slips, residential and commercial uses, small remnant salt marsh in northeastern portion

- Problems: Copper from antifouling paints

**Dana Point Harbor, Oceanside Harbor, and Del Mar Boat Basin at Camp Pendleton**

- Small harbors dredged to accommodate small vessels

- Shoreline: Marinas and boat repair facilities

- Problems: Copper from antifouling paints and oil

**Coastal Lagoons (17)**

- Mouths intermittently closed with fluctuating salinities in lagoons, except Agua Hedionda (always open) and Buena Vista (converted to freshwater lake)
- Shoreline: Usually undeveloped with agricultural and light residential uses
- Problems: Tijuana receives Mexican sewage; Buena Vista, Batiquitos, San Elijo, San Dieguito, and Los Penasquitos have sewage sludge deposits.

**San Diego Region Monitoring Goals and Objectives**

- Identify known and potential toxic hot spots (but not at certain locations under previous San Diego Regional Board cleanup orders)
    - Identify chemicals causing toxicity and geographic extents and depths of chemicals
    - Identify probable sources of toxic pollutants and estimate probable contributions toward creation of toxic hot spots by each source
    - Estimate effects of causative agents on beneficial uses
- (If feasible:)
  Confirm at certain locations whether toxic hot spots exist after cleanups of toxic wastes (at certain boat yards, off storm drains, and at a copper concentrate transfer area)
- Review data to determine priority rankings of toxic hot spots

**San Diego Region Results**

**Known toxic hot spots:** None

**Potential toxic hot spots:**

24 in San Diego Bay (15 from R. Swartz' amphipod toxicity, 7 from Fish and Game sediment chemistry sampling, 4 from storm drain sediment chemistry sampling)
2 in Dana Point Harbor
2 in Oceanside Harbor

## San Diego Bay Questions

1. Should the graphical method be used for data analysis to designate toxic hot spots?

2. Are the northern, central, and southern parts of San Diego Bay so different that reference sites need to be located within these areas?

3. Are pollutants in urban runoff dispersed so well that the effects cannot be measured at the points of entry?

4. Can recent discharges of PCBs and PAHs in sediments be differentiated from historic discharges?

5. Are PAH deposits under the site of the 10th Avenue Marine Terminal from a turn-of-the-century coal degasification plant entering San Diego Bay at levels which could cause toxicity?

6. Do sediments near boat yards and shipyards exhibit greater toxicity or show other detrimental effects than sediments at marinas and moorings where underwater hull cleaning takes place?

7. Can known toxic hot spots still be designated in areas where high percentages of sediment fines are found and where Rhepoxynius data are therefore excluded?

8. Does San Diego Bay have a characteristic toxicity pattern which sets it off from other bays due to its history of sewage discharges, industrial discharges, oil spills, and urban runoff?

9. Does waste heat from the South Bay Power Plant influence toxicity in the southern part of San Diego Bay?

**TOXICITY AND BIOACCUMULATION STUDIES IN THE SAN DIEGO REGION**

| WATER BODY AND STUDY | GOALS | RESULTS |
|---|---|---|
| DANA POINT AND OCEANSIDE SMALL CRAFT HARBORS: | | |
| State Mussel Watch | Trends | |
| MISSION BAY: | | |
| State Mussel Watch | Characterization and trends | Results also used for landfill surveillance |
| Bay Protection, 1993-94, State Water Board | Characterization | |
| SAN DIEGO BAY: | | |
| State Mussel Watch, 1977-present, State Water Board | Characterization and trends | Tissues show relatively high PCBs and PAHs; results used to locate Bay Protection sampling sites |
| San Diego Bay Cleanup Project, 1987-90, Regional Water Board | Identification of sources of waste | Sediment chemistry results used to locate Bay Protection sampling sites |
| Status and Trends, 1987-present, National Mussel Watch and Benthic Surveillance | Characterization and trends | Tissue results confirmed San Diego Bay had elevated levels of PCBs and PAHs |
| Amphipod toxicity sampling, 1987, USEPA, Rick Swartz | Characterization using Rhepoxynius toxicity | Potential toxic sites identified |
| Bay Protection, screening, 1992-93, State Water Board | Characterization of toxicity near known sources of waste identified by S.D. Bay Cleanup Project, and identification of reference sites | Potential toxic sites identified |

| | |
|---|---|
| Bay Protection toxicity confirmation, 1993-94; State Water Board; NOAA Status and Trends; and USEPA EMAP | Confirmation of toxic hot spots by resampling at 1992-93 hits, and introduction of randomly-placed sites using stratified sampling |

CALIFORNIA BAY PROTECTION AND TOXIC CLEANUP PROGRAM
TECHNICAL QUESTIONS FOR THE
SCIENTIFIC PLANNING AND REVIEW COMMITTEE

## General Questions

Questions on general program topics, such as experimental design, sampling strategies, and reference site selection, are covered in additional documentation. The questions below are specific to individual laboratories conducting analyses for the program.

## Toxicity Testing

### Background

Toxicity testing, using a suite of organisms and protocols, has been used to screen potential hot spots and reference sites. Toxicity tests are also included as part of the "confirmation" phase of the program. If significant toxicity ("associated with toxic pollutants") is observed at least twice in samples from a given site, then that site can be considered a hot spot under the BPTCP hot spot criteria. Toxicity testing methods are described in the BPTCP QAPP.

### Questions

1. What criteria should determine toxicity test selection: comparability with other programs, sensitivity, precision, logistics, cost, matrix (solid vs. pore water), others?

2. How important is it to have complete data sets with all sites tested with all species?

3. What criteria should determine when new techniques should be incorporated: State Board Ocean Plan listing, logistical advantages, increased sensitivity or sub-lethal endpoints (especially in solid-phase tests), increased tolerance to non-anthropogenic factors?

4. Should new species or protocols be avoided for the sake of consistency?

5. Should the use of pore water tests be limited because of concerns over ecological relevance or sample handling artifacts, or does their sensitivity and usefulness in TIEs and sediment quality objectives development outweigh those concerns?

6. How does frozen pore water storage affect test results?

7. What are the effects of storage time for fresh pore water?

8. If delays of two days to two weeks are anticipated between sediment collection and pore water testing, should pore water be extracted immediately and then stored, or should solid sediment be stored and pore water extracted immediately before testing?

9. What negative controls are necessary in pore water (lab seawater vs. Reference site pore water)?

10. Should controls be included for all sample manipulations, such as: travel controls, extraction controls, freezing controls (if samples are frozen), brine controls, dilution water controls and/or sample bottle controls?

11. How would multiple controls be used in the statistical analysis of the data?

12. Could all of the above controls be satisfied with pore water from a good reference site?

13. Would multiple reference sites with various grain size, TOC, etc., be necessary for pore water controls?

14. In solid phase tests, are home sediment controls sufficient for comparisons against test sites to determine significant toxicity in this program?

15. If reference sites are necessary for comparisons against test sites to determine significant toxicity, what constitutes an appropriate reference site, and how many are necessary? (This question belongs with the group of issues that must be addressed in the larger context of additional analyses, such as benthic ecology and sediment chemistry.)

16. Should positive control results (reference toxicant test LC50s) be required to fall within a specified range for test acceptability? If control charts are used, must a test LC50 fall within the bounds of 2 standard deviations for the concurrent test of a sediment sample to be acceptable?

17. Can we and/or should we measure pore water DOC to interpret pore water toxicity in relation to measured chemical concentrations?

18. Ammonia and Hydrogen Sulfide

    Background: Ammonia and sulfide are currently measured at the beginning and end of each toxicity test in both overlying water and in interstitial water centrifuged from test sediment. It appears that measuring

overlying water underestimates levels of the two compounds to which the organisms are exposed, and that measuring interstitial water may overestimate exposure, as animals probably avoid high concentrations by inhabiting shallow oxidized layers in the test containers. We are currently sampling for ammonia and sulfide by taking water from as close as possible to the sediment/water interface (<0.5 cm).

Questions on sulfide and ammonia in toxicity tests:

a. How can we best sample test containers to measure concentrations of sulfide and ammonia to which the organisms are exposed

b. Are there reversals or non-monotonic trends in toxicity of ammonia or sulfide?

c. Are some organisms more sensitive to the ammonium ion than to unionized ammonia, and should both be reported?

## Benthic Community Analyses

### Background

Benthic ecological assessments have been used in the BPTCP "confirmation" phase, after sites have been selected based on past data and toxicity screening. Analyses have focused on indicator species, with diversity, abundance, and biomass also evaluated.

Three to five replicate cores have been collected at each site.

### Questions

1. Is there a single index, or should a single index be developed, to describe the condition of a site in terms of its relative ecological degradation?

2. Should choice of indices be based on correlations with chemistry? What other criteria are appropriate?

3. What is the minimum number of field replicates necessary to adequately characterize the condition of the benthic community structure at a site?

4. If cleanup plans are implemented, can benthic community analysis be used in recolonization studies to monitor site recovery after cleanup? If so, what is the best method, and

what studies should be currently undertaken to assist in the study?

5. What seasonal factors need to be considered in planning benthic studies? How are these addressed in long-term program planning?

## Trace Organics Chemistry

### Background

Trace organic compounds have been measured in bulk sediment (not pore water) at selected sites as part of the "confirmation" phase of the program, after sites have been selected based on past data and toxicity screening. The presence of "toxic pollutants" must be demonstrated in order for a site to be considered a BPTCP hot spot. Compounds on the NOAA analyte list are currently measured. Specific analytical methods are described in the BPTCP QAPP.

### Questions

1. Are the analytical techniques adequate to satisfy program goals?

2. Should organic compounds be measured in pore water?

3. Should the number of compounds analyzed for be increased?

4. Should effort be directed toward identification of unknown peaks?

5. Are detection limits adequate for program goals and to allow meaningful correlations with chemistry?

## Trace Metals Chemistry

### Background

Trace metals have been measured in bulk sediment and occasionally pore water at selected sites as part of the "confirmation" phase of the program, after sites have been selected based on past data and toxicity screening. The association with "toxic pollutants" must be demonstrated in order for a site to be considered a BPTCP hot spot. Aluminum, antimony, arsenic, cadmium, chromium, copper, iron, lead, manganese, mercury, nickel, selenium, silver, tin, and zinc are currently measured. Specific analytical methods are described in the BPTCP QAPP.

Questions

1. Are the analytical techniques adequate to satisfy program goals?

2. Should trace metals be measured in pore water routinely?

3. Should the number of metals analyzed for be increased or decreased?

4. Should the laboratory be analyzing AVS routinely?

5. Are detection limits adequate for program goals and to allow meaningful correlations with chemistry?

## Biomarkers

Background

A number of biomarkers have been analyzed in special BPTCP studies. Biomarkers may be used to demonstrate environmental "impairment", which, in association with elevated contaminant concentrations in tissues, can lead to hot spot designation. To date, the program has supported work on the following biomarkers:

1. Heat stress proteins in mussels - Brenda Sanders

2. Reporter gene system (luciferase) - UC San Diego and Jack Anderson

3. Cytogenetics - mitotic aberrations in sea urchin embryos - MPSL/UC Santa Cruz

4. Histopathology - gonadal/somatic indices - EROD in fish tissues - UC Davis and Bob Spies

Questions

1. Which biomarkers are most appropriate for the goals of the program?

2. How are candidate biomarkers affected by sampling techniques and other artifacts?

3. What level of within-site precision is necessary for defensible results?

4. How do the effects of temperature, salinity, food availability and seasonal physiological cycles affect the validity of biomarker results?

5. What additional QAQC is necessary for biomarker studies?

6. Are biomarkers necessary? Given the high number of toxicity hits found so far, is the increased sensitivity of biomarkers necessary if their interpretation is difficult to support in a regulatory context?

## Natural Toxins and Unknowns

### Background

Many sites investigated so far have relatively low concentrations of measured contaminants, yet demonstrate toxicity to test organisms. Ammonia, sulfide, and grain size are measured routinely, but often do not account for toxicity. Natural toxins or unmeasured contaminants may be responsible, and their analysis may facilitate interpretation of the relationships between chemistry and toxicity. Chromatographs often show large peaks for unknown chemicals. There has been no effort to date to evaluate natural toxins or unknown chemicals as part of the BPTCP.

### Questions

1. How much effort should the BPTCP invest in natural toxins and unknown chemicals?

2. What is the best and/or most cost effective approach to investigate natural toxins?

3. What are the best analytical techniques?

4. How can we distinguish between natural and anthropogenic biological effects?

## Statistics

### Background

For the purposes of hot spot designation, the BPTCP must demonstrate significant biological impacts. It may not be sufficient from a regulatory perspective to show simply that a "sample" is significantly toxic relative to a control. More likely it will be necessary to demonstrate that a "site" is significantly more toxic than unimpaired sites or reference conditions. Further, while toxicity data can be analyzed from a single station using laboratory replicates, benthic community data may need to be analyzed on the basis of multiple stations (field replicates) to characterize a site. The BPTCP needs to

establish precise statistical definitions for what is toxic and/or impaired.

Questions

1. How can toxicity and benthic community data be integrated to demonstrate significant differences among sites (with or without field replication)?

2. Is comparison against laboratory negative controls (such as home sediment or laboratory seawater) sufficient to indicate significant toxicity of test sites?

3. Is comparison against a single reference site sufficient to indicate significant toxicity of test sites?

4. What are the best methods for incorporating natural variability among sites (in the absence of pollution) into the determination of significant toxicity?

5. What other sources of variability must be incorporated into statistical methods (field replicate variability, temporal variability, between site variability)?

6. Does a "reference envelope" approach account for all applicable variation, and is such an approach appropriate for this program?

7. Is the Hampel Outlier Identifier method a preferable approach for discriminating between reference sites and toxic sites?

8. What is a reference site? (Note: This issue is and has been taken up in a broader context elsewhere.)

9. We are now conducting site-by-site separate variance t-tests instead of Analysis of Variance and Dunnett's test. Is this the best method? What other possibilities exist (especially when trying to incorporate natural between-site variance)?

10. Along with simple test-specific significance, should we include protocol "detectable difference" criteria based on the cumulative 90% MSD(as suggested by Thursby)?

11. When conducting a large number of individual correlations, must the significance level (alpha) be adjusted to account for the possibility of attaining significant correlation values by random chance? If so, how?

12. Are multivariate techniques (such as principle components analysis) appropriate for the BPTCP efforts to associate chemistry and biological effects?

13. Given program objectives, available funding, and potential for legal scrutiny, what is the most appropriate sampling design for identifying toxic hot spots? For developing sediment quality objectives? For monitoring cleanup efforts? For monitoring bay and estuarine areas that are clean or potentially subject to future pollution?

## CONTRIBUTORS TO THE BRIEFING DOCUMENT

| | |
|---|---|
| Agenda | Gita Kapahi[1] |
| Program Summary | Craig J. Wilson[1], Gita Kapahi |
| Monitoring Program Design | Craig J. Wilson, Max Puckett[2], Rusty Fairey[3], John Hunt[4], Brian Anderson[4], Mark Stephenson[2] |
| Issue Papers | |
| Toxicity | John Hunt, Brian Anderson |
| Chemical Association | Craig J. Wilson, Rusty Fairey, Max Puckett, Mark Stephenson |
| Benthic Impact | Carrie Bretz[5] |
| Sample Collection | Craig J. Wilson, Rusty Fairey |
| Screening and Confirmation | Craig J. Wilson, Rusty Fairey, Kim Paulson[2] |
| Biological Methods | John Hunt, Brian Anderson |
| Chemical Methods | Mark Stephenson, Jon Goetzl[2], John Newman[4], Kim Paulson |
| Regional Summaries | |
| Region 1 | Bruce Gwynne[6] |
| Region 2 | Karen Taberski[7], Peter Otis[7] |
| Region 4 | Shirley Birosik[8] |
| Region 5 | Bill Croyle[9], Chris Foe[9] |
| Region 8 | Scott Dawson[10] |
| Region 9 | Pete Michael[11] |

[1]State Water Resources Control Board

[2]Department of Fish and Game

[3]San Jose State University

[4]University of California, Santa Cruz

[5]Moss Landing Marine Laboratory

[6]Regional Water Quality Control Board, North Coast Region

[7]Regional Water Quality Control Board, San Francisco Bay Region

[8]Regional Water Quality Control Board, Los Angeles Region

[9]Regional Water Quality Control Board, Central Valley Region

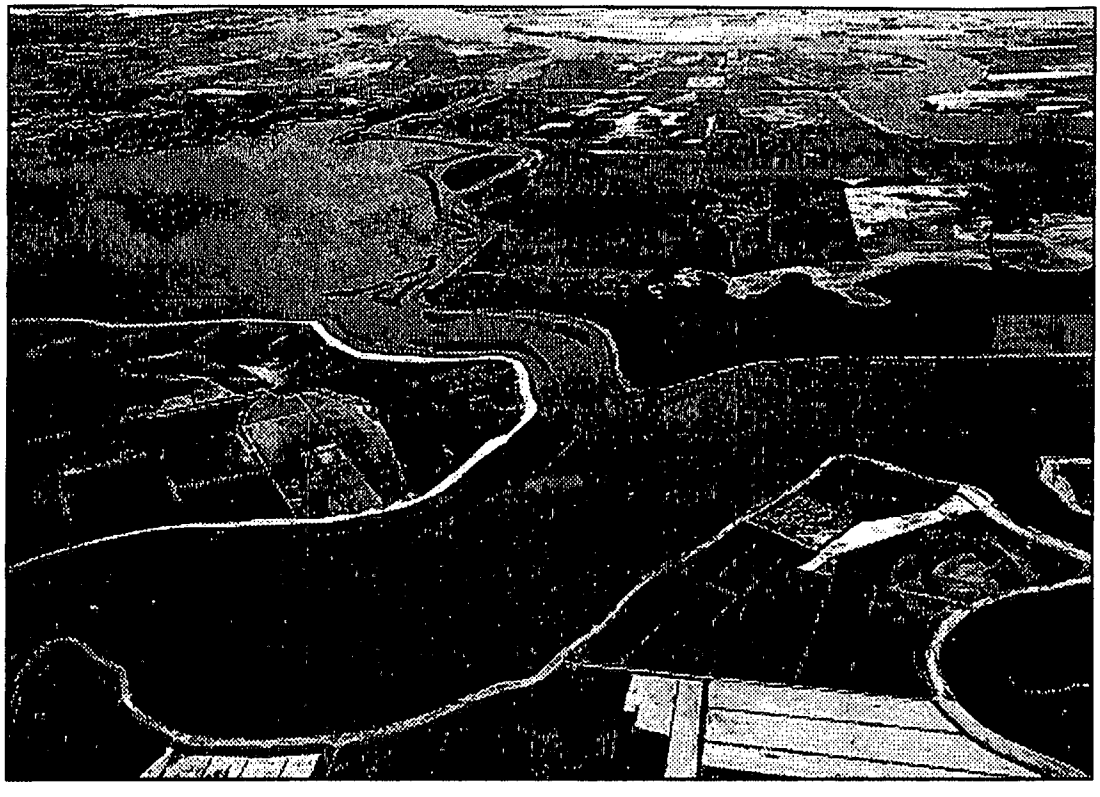
[10]Regional Water Quality Control Board, Santa Ana Region

[11]Regional Water Quality Control Board, San Diego Region

# A P P E N D I X  B

Scientific Planning and Review Committee
Briefing Document for Recommendations
on the Bay Protection and Toxic Cleanup Program
Monitoring Activities

May 1996



# Scientific Planning and Review Committee Briefing Document for Recommendations on the Bay Protection and Toxic Cleanup Program Monitoring Activities

**May 1996**

Department of Fish and Game
Regional Water Quality Control Boards
State Water Resources Control Board

STATE OF CALIFORNIA
STATE WATER RESOURCES CONTROL BOARD
REGIONAL WATER QUALITY CONTROL BOARDS
DEPARTMENT OF FISH AND GAME

# SCIENTIFIC PLANNING AND REVIEW COMMITTEE:

# BRIEFING DOCUMENT FOR RECOMMENDATIONS ON THE BAY PROTECTION AND TOXIC CLEANUP PROGRAM MONITORING ACTIVITIES

MAY 1996

# PREFACE

This briefing document was developed to assist the Scientific Planning and Review Committee (SPARC) in preparing for a technical workshop to review the State Water Resources Control Board's Bay Protection and Toxic Cleanup Program (BPTCP) monitoring activities. It contains a summary of the SPARC recommendations on questions posed at the SPARC meeting held on April 12 and 13, 1995 as well as descriptions of the specific issues SPARC will consider and comment on at the Workshop scheduled for May 15-17, 1996.

MEMBERS OF THE SCIENTIFIC PLANNING AND REVIEW COMMITTEE

Dr. Rick Swartz
Environmental Protection Agency
2111 S.E. Marine Science Drive
Newport, OR 97365-5260

Ms. Rachel Friedman-Thomas
Washington State Department of Ecology
P.O. Box 47600
Olympia, WA 98504-7600

Dr. Bruce Thompson
San Francisco Estuarine Institute
1301 South 46th Street
180 Richmond Field Station
Richmond, CA 94804

Dr. Mel Suffet
Environmental Science and
Engineering Program
University of California, Los Angeles
10833 Le Conte Avenue
Los Angeles, CA 90024-1772

Dr. John Knezovich
Health & Ecological Assessment
Division L-453
Lawrence Livermore Laboratory
Livermore, CA 94550

Dr. Don Stevens
Environmental Statistics
and Aquatic Monitoring
ManTech Environmental
Research Services, Inc.
200 S.W. 35th Street
Corvallis, OR 97333

Dr. Ed Casillas
Environmental Conservation Division
National Marine Fisheries Service
2725 Montlake Blvd. East
Seattle, Wash 98112

TABLE OF CONTENTS


PREFACE . . . . . i

MEMBERS OF THE SCIENTIFIC PLANNING AND REVIEW COMMITTEE . . . ii

PURPOSE OF THE TECHNICAL WORKSHOP . . . . . 1
Focus of the Workshop . . . . . 1
Contents of the Briefing Document . . . . . 1

AGENDA . . . . . 3

MONITORING ACTIVITIES COMPLETED BY THE
BAY PROTECTION AND TOXIC CLEANUP PROGRAM
FY 1995-1996 . . . . . 5
San Diego Bay Report . . . . . 5
Small Bays and Estuaries Pilot Study . . . . . 5
San Francisco Bay Fish Contaminant Study . . . . . 6
San Francisco Bay Reference Site Study . . . . . 7
Stockton Urban Stormwater Runoff (Region 5) . . . . . 8
Cache Creek mercury mass loading study (Region 5) . . . 9
References . . . . . 10

SCIENTIFIC PLANNING AND REVIEW COMMITTEE
RECOMMENDATIONS . . . . . 12
Issue 1. Toxicity . . . . . 12
Selection of Reference Sites Within Each Region . . 13
Issue 2. Association of Chemistry and Biological
Effects . . . . . 14
Issue 3. Benthic Impacts . . . . . 14
Replication of Benthic Ecological Analysis . . . . 15
Issue 4. What is the most appropriate sampling design . 15
Field Replication . . . . . 15
Issue 5. Toxic Hot spot designation (Screening and
Confirmation approach) . . . . . 15
Issue 6. Appropriate Biological Methods . . . . . 16
Biomarkers . . . . . 17
Issue 7. Appropriate Chemical Methods . . . . . 17
Metals . . . . . 17
Organics . . . . . 18
Overall summary of SPARC recommendations . . . . . 18
Region-specific SPARC Recommendations . . . . . 20
Region 1 . . . . . 20
Region 2 . . . . . 20
Region 5 . . . . . 20

REVISED STATEWIDE MONITORING DESIGN . . . . . 21
Legislative Mandate . . . . . 21
Specific Definition of a Toxic Hot Spot . . . . . 21
Proposed Specific Definition . . . . . 21
Potential Toxic Hot Spot . . . . . 22

Candidate Toxic Hot Spot . . . . . 22
Known Toxic Hot Spot . . . . . 25
Monitoring Program Objectives . . . . . 26
Review of Preliminary Studies and Research . . . . . 26
Biological Monitoring Methods . . . . . 27
Chemical Methods . . . . . 27
Sampling Strategy . . . . . 27
Screening Sites and Confirming Toxic Hot Spots . . 27
A Battery of Screening Tests . . . . . 30
Site Selection . . . . . 33
Reference sites . . . . . 34
Toxicity Screening . . . . . 36
Confirmation (i.e., Qualification as Candidate Toxic Hot Spots) . . . . . 37
Quality Assurance . . . . . 38
References . . . . . 38

TOXICITY ISSUE PAPER . . . . . 41

Statistical Methods for Distinguishing Sites of Concern Using the Reference Envelope Approach:
Issues to be Resolved . . . . . 41
Rationale for the Reference Envelope Approach . . . . . 41
Reference Envelope Statistical Method . . . . . 41
Sampling Design . . . . . 42
Statistical Test . . . . . 42
Issues Regarding Use of the Reference Envelope Approach . . . . . 43
Reference Site Selection Criteria . . . . . 43
Chemistry Criteria . . . . . 43
Benthic Ecology Criteria . . . . . 45
Questions Regarding Reference Site Selection Criteria . 45
How Many Reference Sites are Necessary? . . . . . 45
How Often Must Reference Sites Be Sampled? . . . . 46
How Many Reference Populations Are Necessary for Analysis of California Sites? . . . . . 46
Interactions Between Policy and Science . . . . . 47
What Should Be Done When the Analysis Doesn't Work? . . . . . 47
References . . . . . 47

CHEMICAL ASSOCIATION ISSUE PAPER . . . . . 49

Use of Sediment Quality Guidelines . . . . . 49
ERM and PEL Quotients . . . . . 49
Chemical Association Issues . . . . . 50

FRESHWATER STUDIES ISSUE PAPER (REGION 5) . . . . . 51

Water Column Toxicity Issues . . . . . 51
Delta Toxicity and Metals Monitoring . . . . . 51
Issues and Questions . . . . . 52

BIOACCUMULATION ISSUE PAPER . . . . . 54

The Use of Bioaccumulation Monitoring in the Bay Protection and Toxic Cleanup Program . . . . . 54
Background . . . . . 54
Existing California Bioaccumulation Data . . . . . 54
Bay Protection Toxic Cleanup Program Data . . . . . 55
Sample Design Questions . . . . . 56
Analysis Questions . . . . . 58
Relationship between Bioaccumulation/Human Health & Aquatic Toxicity . . . . . 59
Interpretation/designation of Hot Spot Questions . . . . 59

BENTHIC COMMUNITY ISSUE PAPER . . . . . 61

Characterization of Benthic Community Degradation for San Diego Report . . . . . 61
Data Analyses and Interpretation . . . . . 61

Characterization of Benthic Community Degradation for Southern California Coastal Lagoons Project Report . . . 65
Benthic Index . . . . . 65
Community Data . . . . . 65
Indicator species . . . . . 66
Calculation of the Benthic Index . . . . . 66
Other Benthic Community Issues . . . . . 66
References . . . . . 67

WEIGHT-OF-EVIDENCE ISSUE PAPER . . . . . 69

Application of Weight of Evidence Approach for San Diego Bay . . . . . 69

CONTRIBUTORS TO THE BRIEFING DOCUMENT . . . . . 75

BAY PROTECTION AND TOXIC CLEANUP PROGRAM
SCIENTIFIC PLANNING AND REVIEW COMMITTEE

# PURPOSE OF THE TECHNICAL WORKSHOP

The Bay Protection and Toxic Cleanup Program (BPTCP) is a Statewide Program legislatively mandated to identify toxic hot spots in the enclosed bays and estuaries of each of the seven coastal regions of the State. Once toxic hot spots are identified, each coastal Regional Water Quality Control Board is legislatively mandated to develop Toxic Hot Spot Cleanup Plans specifying where and how each identified toxic hot spot will be remediated. The major focus of the Program to date has been monitoring to identify polluted sites.

The BPTCP is sponsoring this workshop to provide a forum for the review of studies performed by the BPTCP. The studies will be reviewed by experts in the fields of toxicology, benthic ecology, organic and inorganic chemistry, program implementation and direction, experimental design, statistics, and bioaccumulation.

The purposes of this workshop are to (1) add and modify, as needed, the SPARC recommendations received at the workshop held on April 12 and 13, 1995 and (2) review the reports developed by the BPTCP, and (3) receive specific advice on appropriate methods for evaluating the monitoring data collected.

## Focus of the Workshop

1. Review and incorporation of the SPARC recommendations into the Statewide monitoring approach.

2. Interpretation of toxicity data collected.

3. Interpretation of the benthic community data collected.

4. Setting priorities using a weight-of-evidence approach.

5. Review of the studies of water column toxicity and chemistry in the Central Valley Region.

6. Completion of the discussion on organic chemistry methods.

7. The use of bioaccumulation monitoring techniques.

## Contents of the Briefing Document

For each of these topics, a brief issue paper outlining the approaches the BPTCP has taken is presented. In addition to the

issue papers the recommendations from the April 1995 SPARC meeting are listed and the revised monitoring approach is presented.

Each of the topics presented in this document could take several days of discussion to fully evaluate and assess. It is the intent of this workshop that SPARC hear the approaches being pursued by the Program and comment on their appropriateness and usefulness. The SPARC is charged with determining if the approaches the Program is taking are scientifically appropriate and, if not, what approaches the Program should use.

BAY PROTECTION AND TOXIC CLEANUP PROGRAM
SCIENTIFIC PLANNING AND REVIEW COMMITTEE
TECHNICAL WORKSHOP

MAY 15-17, 1996

MOSS LANDING MARINE LABORATORIES SHORE STATION
AND MOSS LANDING CHAMBER OF COMMERCE BUILDING

MOSS LANDING, CALIFORNIA

AGENDA

***WEDNESDAY, MAY 15, 1996: Moss Landing Laboratories Shore Station--North***

| | |
|---|---|
| 1:00 p.m. to 2:00 p.m. | Register |
| 2:00 p.m. to 5:00 p.m. | Review of SPARC 1995 recommendations and overview of BPTCP progress to date |

***THURSDAY, MAY 16, 1996: Moss Landing Chamber of Commerce Building***

| | |
|---|---|
| 8:00 a.m. to 8:15 a.m. | Welcome |
| 8:15 a.m. to 8:30 a.m. | Introductions |
| 8:30 a.m. to 9:00 a.m. | Overview, previous SPARC recommendations and reports completed |
| 9:00 a.m. to 10:30 a.m. | Interpretation of toxicity data<br>Reference envelope<br>80% of Controls<br>Others |
| 10:30 a.m. to 10:45 a.m. | Break |
| 10:45 a.m. to 12:00 noon | Interpretation of toxicity data (continued) |
| 12:00 noon to 1:00 p.m. | Lunch |
| 1:00 p.m. to 2:30 p.m. | Interpretation of chemistry data<br>ERM, ERL<br>PEL, TEL<br>Quotients<br>AETs |

| | |
|---|---|
| 2:30 p.m. to 2:45 p.m. | Break |
| 2:45 p.m. to 5:30 p.m. | Water column toxicity,<br>Bioaccumulation of pollutants,<br>Organic chemistry methods |

***FRIDAY, MAY 17, 1996: Moss Landing Chamber of Commerce Building***

| | |
|---|---|
| 8:00 a.m. to 8:30 a.m. | Welcome |
| 8:30 a.m. to 10:00 a.m. | Interpretation of benthic community data<br>Benthic index development<br>Assessment of degraded conditions |
| 10:00 a.m. to 10:15 a.m. | Break |
| 10:15 a.m. to 12:00 noon | Weight of Evidence approach<br>Comprehensive interpretation of data<br>Setting priorities for sites |
| 12:00 noon to 1:00 p.m. | Lunch |
| 1:00 p.m. to 3:00 p.m. | Weight-of-evidence approach (continued) |
| 3:00 p.m. to 3:15 p.m. | Break |
| 3:15 p.m. 4:30 p.m. | Wrap-up: SPARC recommendations |

MONITORING ACTIVITIES COMPLETED BY THE
BAY PROTECTION AND TOXIC CLEANUP PROGRAM
FY 1995-1996

As part of the legislative mandates of the Program, the BPTCP has implemented regional monitoring programs to identify toxic hot spots (this work is described in SWRCB et al., 1995). Regional monitoring efforts are being implemented in all seven coastal Regions (SWRCB, 1993; SWRCB et al., 1995). Several reports have been completed in the last year.

Each of the reports completed have been submitted to the SPARC for review. A brief description of each of the reports is presented below.

**San Diego Bay Report**

Three-hundred and fifty stations have been sampled and data analyzed. The first draft of the report was completed by DFG in February, 1996 (Fairey et al., in review).

In this study, San Diego Bay, Mission Bay and the Tijuana River Estuary were sampled. Two sampling designs were used: directed point sampling and stratified random sampling. Measurements of sediment toxicity, benthic community structure and chemicals present in the sediments were made. Three stations were found to satisfy the conditions listed in the definition of a toxic hot spot (DWQ/SWRCB, 1995). Eighty-four other stations were identified to be of moderate and low concern.

**Small Bays and Estuaries Pilot Study**

The NOAA/EMAP/SWRCB Small Bays and Estuaries pilot study was initiated in March 1995 (SWRCB et al., 1994; SWRCB and NOAA, 1993). This study is a cooperative effort between the SWRCB, NOAA and the EPA Environmental Monitoring and Assessment Program. The draft report on this study is undergoing internal review (Anderson et al., in review).

The pilot study has seven objectives:

1. Estimate with known confidence the percent of degraded fine-grained sediment area in Southern California small bays and estuaries using several critical threshold values of toxicity, benthic community analysis, and chemistry.

2. Produce a map of the data collected for sediment toxicity, benthic community analysis and chemistry.

3. Identify a set of sites that should be revisited for confirmation as either toxic hot spots or reference sites.

4. Assess the effectiveness of locating toxic hot spots and reference sites (for which prior knowledge of likely impacts exists) or random sampling throughout the set of water bodies.

5. Assess the concordance of two solid phase sediment toxicity tests over a range of substrate, salinity, and toxicant concentration conditions.

6. Develop a benthic index for interpretation of benthic data.

7. Identify which of the measured toxicants are most associated with toxic response.

**San Francisco Bay Fish Contaminant Study**

The draft of this report was released for public review in December 1994. The final report was released at the end of June 1995 (RWQCB et al., 1995). The comprehensive human health risk analysis to be conducted by OEHHA using the study results is currently in-progress, and is expected to require several months. As a result of the data, OEHHA issued an interim health advisory for fish consumption in San Francisco Bay in December 1994.

This study (RWQCB et al., 1995) was conducted to measure contaminant levels in fish caught and consumed by anglers in San Francisco Bay. The main objectives of the study were to identify, to the maximum extent possible, the chemicals, species and geographical areas of concern in San Francisco Bay. This study was designed in a coordinated effort between OEHHA, DFG, the Department of Health Services, environmental groups and anglers. Thirteen fishing piers were sampled for fish with a small habitat range. Other regions of the Bay were sampled for fish that had a larger habitat range. The species of fish that were collected were white croaker (which was the highest priority fish based on its feeding behavior and lipid content), shiner surfperch, walleye surfperch, leopard sharks, brown smoothhound sharks, striped bass, sturgeon and halibut. Pilot Study Screening Values based on the consumption rate of 30 grams per day were used to screen the data for potential chemicals of concern. Results showed that:

1. The EPA guidance document, Guidance For Assessing Chemical Contaminant Data For Use In Fish Advisories- Volume 1- Fish Sampling And Analysis (EPA 823-R-93-002, 1993), was an effective tool for designing the pilot study and analyzing data collected from the San Francisco Bay study.

2. Based on EPA screening values six chemicals or chemical groups were identified as potential chemicals of concern in San Francisco Bay. They were PCBs, mercury, dieldrin, total DDT, total chlordane and the dioxin/furans.

3. High levels of the pesticides dieldrin, total DDT and total chlordane were most often found in fish from the North Bay.

4. Levels of PCBs, mercury and the dioxin/furans were found at concentrations exceeding EPA screening values throughout the Bay.

5. Fish with high lipid content (croaker and shiner surfperch) in their muscle tissue generally exhibited higher organic contaminant levels. Fish with low lipid levels (halibut and shark) generally exhibited lower organic contaminant levels.

6. Of the Bay fish collected, white croaker consistently exhibited the highest tissue lipid concentrations. Lipophilic PCBs and pesticides concentrated to the highest levels in the muscle tissue of these fish.

7. Mercury levels were found to be the highest in the two shark species collected; the leopard shark and the brown smoothhound shark. Both the sharks and white croaker exhibit increasing mercury concentration with increasing fish size indicating bioaccumulation of this metal in Bay area fish.

8. Vallejo-Mare Island was the sampling location from which fish most often exhibited high levels of chemical contaminants. Oakland Inner Harbor also exhibited a high incidence of tissue contamination.

**San Francisco Bay Reference Site Study**

The main purposes of this study (Hunt et al., in review) are to: (1) identify sediment reference sites in San Francisco Bay to use in toxicity tests, (2) recommend sediment toxicity test protocols to use in monitoring sediment toxicity in San Francisco Bay, (3) develop Sediment Toxicity Identification (TIE) protocols that can be used in San Francisco Bay and (4) identify the cause of toxicity at previously identified reference sites. For this study five potential sediment reference sites were chosen. Two sites were in San Pablo Bay, one site was in the Central Bay and two sites were in the South Bay. Chemical analysis has been conducted at all sites that do not show toxicity. Sediment samples from Tomales Bay and several contaminated sites were also collected. All potential reference sites had three field replicates. In addition, all potential reference sites, except those in the South Bay, were sampled three times during the year during different hydrographic conditions. Since the most likely locations to find reference sites were in San Pablo and the Central Bay, those sites were chosen first. Since these sites seemed to be good reference sites based on results from two sampling events, additional sites were chosen in the South Bay. Between seven to nine toxicity tests were performed on each

sample. These tests were: (1) the 10 day solid phase amphipod test using Eohaustorius, (2) the 10 day solid phase amphipod test using Ampelisca, (3) the 10 day amphipod test using Eohaustorius in undisturbed cores, (4) the 10 day amphipod test using Eohaustorius in pore water, (5) the bivalve larvae development test in pore water, (6) the urchin larvae development test in pore water, (7) the urchin larvae development test using a sediment/water interface exposure, (8) the Neanthes growth and survival test and (9) a 10 day solid phase test using Nubelia. Toxicity tests were dropped out of the study based on the level of control survival, performance at reference sites and sensitivity to contaminated sites.

The first step in this project was to develop Sediment TIE protocols for the 10 day amphipod test, the bivalve larvae development test and the urchin larvae development test. When all laboratory tests were completed including pore water extraction experiments, testing the sensitivity of the various organisms to TIE manipulations and spiking experiments, the field portion of the study began. Samples were collected at the reference sites with enough field replication to try to determine field variability and during different hydrologic conditions to try to determine seasonal variability. By collecting the samples in this way we hoped to identify reference sites, determine the variability at those sites for statistical purposes, and identify sediment toxicity tests that perform well at reference sites but are sensitive to contaminated sites. Once reference sites are identified, testing of these sites will continue and data will be added to develop a "reference envelope" for these sites. In addition, we performed the amphipod test with undisturbed cores and the urchin test using a sediment/water interface to evaluate the environmental relevance of the standard amphipod and urchin tests. These tests could possibly be used in confirming toxic hot spots.

When samples were found to be toxic, a TIE was performed using the pore water test that showed the toxicity. The first two field TIEs were performed on sediment from Islais Creek, where the City of San Francisco has had their main outfall for decades, and on Tomales Bay sediment. After removing ammonia and hydrogen sulfide from the Islais Creek sample, toxicity remained. After running TIEs on both samples results seemed to indicate that in both samples toxicity was being caused by a polar organic degradation product. Additional work has been performed to try to extract and identify the cause of this toxicity. A draft report on this study is currently available.

**Stockton Urban Stormwater Runoff (Region 5)**

The primary objective of the work is to identify pollutants present in Stockton wet weather urban runoff which cause toxicity in water samples collected from waterways located in the Southern

Delta. Limited testing occurred last year at Stockton which confirmed that runoff from the City was also toxic. Little work has been done on urban runoff linking the responsible pollutant(s) and the observed toxicity. The number of pollutants typically present in urban runoff is extensive and it is not possible to adequately assess toxicity with standard, concurrent chemical analyses. Bioassays and toxicity identification evaluations (TIEs) must be conducted to determine the responsible chemicals. In addition, the toxicity monitoring program at Stockton last year noted suppressed dissolved oxygen levels in water samples collected from Smith Canal, the Calaveras River and Five Mile Slough after the first rainfall event of the year. Board staff and local residents reported observing dead catfish, bass and carp in these waterways. Fish mortality from low oxygen levels would also have occurred in the bioassays had they not been continuously aerated. Continuous aeration is not a normal procedure in these tests. Apparently the dissolved oxygen problem occurs almost annually at Stockton and has repeatedly been reported to the Department of Fish and Game. It is not known whether the oxygen suppression results from biological or chemical oxygen demand nor how extensive (temporally and spatially) the problem is.

This study has two objectives: to identify the specific pollutants present in Stockton urban runoff causing toxicity in bioassays and to identify both spatially and temporally the extent of the oxygen sag. A secondary objective will be to identify whether the oxygen suppression is the result of elevated biological or chemical oxygen demand.

<u>Cache Creek mercury mass loading study (Region 5)</u>

The Central Valley trace metal monitoring program element has three objectives: to define the extent of metal criteria exceedances throughout the Delta, to determine the extent of metal associated toxicity throughout the Delta; and to determine the metal (mostly mercury) loading patterns to the Delta. The latter emphasizes the importance of storm events. Two patterns have emerged after more than two years of study. First, no incidents of toxicity have been linked to metal exceedances. Some exceedances of criteria have occurred but generally appear to be limited to storm events. Second, large amounts of mercury (greater than 95 percent of the annual load) is transported into the Estuary during winter high flow periods. At this time the concentration of mercury exceeds the EPA recommended freshwater criteria of 12 ng/l. Normal dry weather mercury concentrations in the Sacramento River and Delta are between 2 and 4 ng/l. During wet weather water from the Sacramento Valley enters the Delta through both the Sacramento River and the Yolo Bypass (Prospect Slough). Wet weather high flow mercury levels in the Sacramento River ranged between 15 and 40 ng/l and in Prospect Slough between 30 and 600 ng/l. Concentrations as far downstream

as the City of Martinez have been measured at 16 ng/l. The Prospect Slough data suggest a potentially significant source in the Bypass. Follow-up studies of the major inputs to the Bypass found that the Cache Creek watershed was the probable source. Mercury concentrations in the Creek ranged between 600 and 2200 ng/l. High mercury levels were also detected in some other Coast Range creeks discharging to the Sacramento River upstream of the Feather River. All these sources are outside the Delta but are probably responsible for the mercury human health advisory for consumption of fish caught in the Sacramento-San Joaquin Delta Estuary. Follow-up work proposed this coming winter to confirm the mercury sources detected in winter 1995 and to begin evaluating the feasibility of mercury abatement projects. We propose concentrating on Cache Creek for an evaluation of how to proceed with mercury abatement work. If successful, we will use the information gained on Cache Creek to evaluate abatement work on other coastal creeks which contribute elevated mercury loads to the Estuary.

**References**

Anderson et al. In Review. Measures of bioeffects associated with toxicants in small bays and estuaries of southern California (Pilot Study).

DWQ/SWRCB. 1995. Draft Functional Equivalent Document for Implementation of the Bay Protection and Toxic Cleanup Program. Sacramento, CA. 358 pp.

Fairey et al. In Review. Chemistry, toxicity and benthic community conditions in sediments of the San Diego Bay region. 162 pp. + 5 Appendices.

Hunt et al. In Review. Evaluation of sediment reference sites and toxicity tests in San Francisco Bay.

Regional Water Quality Control Board, San Francisco Bay Region; State Water Resources Control Board, California Department of Fish and Game. 1995. Contaminant levels in fish tissue from San Francisco Bay. Oakland, CA. 150 pp.

SWRCB. 1993. Staff Report: Status of the Bay Protection and Toxic Cleanup Program. Sacramento, CA. 230 pp. + 6 appendices.

SWRCB and NOAA. 1993. Measures of Bioeffects Associated with Toxicants in Southern California: Year Three Proposal to Continue a Cooperative Agreement. State Water Resources Control Board and National Oceanic and Atmospheric Administration. State

Water Resources Control Board. Division of Water Quality. Sacramento, CA.

SWRCB, RWQCBs and DFG. 1995. Scientific Planning and Review Committee Briefing Document for the Bay Protection and Toxic Cleanup Program. Sacramento, CA. 89 pp.

SWRCB and the U.S. Environmental Protection Agency Environmental Monitoring and Assessment Program. 1994. Measures of Bioeffects Associated with Toxicants in Small Bays and Estuaries of Southern California (Pilot Study). Proposal for a Cooperative Agreement. 27 pp. + 3 Appendices.

## SCIENTIFIC PLANNING AND REVIEW COMMITTEE RECOMMENDATIONS

An overview of the BPTCP along with its goals and activities was presented at the April 12 and 13, 1995 meeting. The workshop focused on discussion of the following questions identified by the State and Regional Boards and the Department of Fish and Game:

1. What is toxic?

2. How should we show association between toxicity, benthic community, etc. and chemical concentrations?

3. What is a benthic impact?

4. Should we use a probability-based sampling design (random sampling) or directed point sampling approach (i.e. based on best professional judgment)?

5. Should we use a screening and confirmation approach?

6. What biological methods should we use?

7. What chemical methods should we use?

The SPARC provided recommendations to improve the BPTCP monitoring program and specifically addressed the seven questions that needed to be resolved. Further comments and suggestions will be considered and incorporated as they are provided by SPARC.

The SPARC recommendations from the April 1995 meeting follow:

**Issue 1. Toxicity**

a. The selection of toxic and reference sites will ultimately be a policy decision based on best available scientific approaches for determining biological response.

b. The reference envelope approach is preferred over simple comparison to laboratory controls, and there is agreement that this is the statistical approach to pursue for determining the level of toxicity suitable for meeting toxic hot spot toxicity criterion.

c. All toxicity data should be normalized to laboratory controls to account for any variation in laboratory factors or test organism condition.

d. Compare test site response to large reference envelope population from a comprehensive data base of reference site results for the protocol used.

e. Compare test site response to reference envelope population from samples collected concurrently with test samples.

f. A site is toxic if it falls below the reference envelope lower bounds for both the reference site data base and concurrent samples.

g. If a site is toxic relative to the large reference envelope population from the comprehensive database, but concurrent reference site results are also low, the site should be revisited.

Selection of Reference Sites Within Each Region

Some level of pollution will always be unavoidable. However, reference sites should be selected through the following process:

a. Reference sites should not include those sites where toxicity is observed in association with pollution. Common sense and knowledge of local conditions should be used in order to avoid areas known to be disturbed or polluted.

b. Randomly sample the rest of the water body, conducting analyses of chemistry, benthic community structure, and toxicity.

c. Allow trained benthic ecologists to select the sites that have moderate to high species richness, abundant presence of amphipods or other indicator species, and any other indicator of ecological health that can be argued convincingly.

d. Evaluate the chemistry data and narrow the sites to those that do not exceed more than one upper value (such as PEL or ERM) for existing chemistry guidelines.

e. Evaluate the toxicity data and eliminate only those sites that have extremely high toxicity, as determined by a qualified toxicologist, not by a priori criteria.

f. Once reference sites are chosen they are sampled along with test sites. Include the new reference site toxicity results

in the reference envelope regardless of the magnitude of the toxicity response. The reference envelope toxicity result will fall where it may.

g. Compile a data base of toxicity responses from appropriately selected reference sites, and include past and current reference site data in the reference envelope. Allow the number of data points in the reference envelope to grow as more studies are completed in the area.

**Issue 2. Association of Chemistry and Biological Effects**

a. Causal relationships are more powerful than correlations in providing evidence of links between pollutant concentrations and biological effects.

b. Development of spiked bioassay data is recommended to allow unit approach to identifying chemicals responsible for observed effects.

c. Simultaneous Extracted Metals and Acid Volatile Sulphides (SEM/AVS) data is essential for understanding metal effects.

d. Measurement of Dissolved Organic Carbon (DOC) is recommended to help understand organic and metal bioavailability.

e. The effect of oxidation state of chemical compounds should be investigated.

f. Pore water toxicity and chemistry are valuable in determining causal relationships.

g. It is recognized that sorbed pollutants may become bioavailable after ingestion and metabolism.

h. Professional judgement and knowledge of local conditions should be used to decide how best to allocate resources to determine causal relationships.

i. The Program should use all available criteria and biological measurements in assessing the relationships between chemistry and biological effects (i.e., use weight of evidence approach).

**Issue 3. Benthic Impacts**

No single index is defensible in a regulatory setting. A site should be characterized as "healthy", "intermediate", or "degraded" based on the best professional judgement of a

qualified ecologist, using whatever methods are most appropriate to the site.

Replication of Benthic Ecological Analysis

An analysis of existing data should be conducted to determine benthic replication, keeping in mind the types of analyses that can be done with benthic data, the cost of the analysis and benefits derived. Do not replicate unless there is a clear reason to do so.

**Issue 4. What is the most appropriate sampling design**

a. During the screening phase, sampling should incorporate a stratified random design in order to provide an opportunity to find unknown toxic hot spots.

b. Confirmation phase sampling should be based on grids covering the site of concern, with random placements of stations within grid blocks.

c. Grids should be configured to match site characteristics.

d. Temporal variations should be accounted for with repeated sampling at locations at least one meter apart.

e. Spatial and temporal scales should be based on knowledge of the site.

Field Replication

a. Random sampling over suitable sized grids may be preferable to replication. There is no need to replicate unless there is a clear and defensible reason why.

b. It would be best to conduct statistical analysis of past data to determine replication needs for future work.

**Issue 5. Toxic Hot spot designation (Screening and Confirmation approach)**

a. A three tiered data analysis approach should be used. This would include chemical, toxicity, and benthic community analyses. Having hits in all three components of a triad analysis, would classify a site as a worst case toxic hot spot. Hits on fewer than all three would result in classification as a site to be concerned about. All sites could be ranked in this way.

b. Under the BPTCP, the screening phase would consist of using either toxicity or benthic community analysis or chemistry or bioaccumulation data or some combination of all of these. Screening should be flexible, designed to fit the Regional Board's needs. Analysis in this phase should be done only when needed to provide sufficient information to convince the Regional Boards to list or consider the site as a priority site of concern for further action. A hit in either of these analyses would elicit concern, trigger confirmation phase monitoring under the BPTCP and/or perhaps prompt a specific Regional Board to pursue some other type of regulatory review action. It would be very important to involve potential responsible parties as early in the process as possible and coordinate studies and funding.

c. The confirmation phase should consist of toxicity and chemistry and benthic community analyses on a previously visited site of concern or wherever previous evidence indicates a site may be impacted. A confirmatory hit in toxicity, benthic community structure; or all three analyses performed during this phase would classify a site as a worst case toxic hot spot, assuming that there was a hit registered during screening. This phase could also include intensive investigations to identify causal relationships, and intensive grid sampling necessary to show gradients and spatial extent.

d. Allow for a mechanism for de-listing sites if intensive studies prove preliminary designation was in error.

e. It is important to focus on the most impacted sites for successful toxic hot spot designation and application of regulatory actions.

**Issue 6. Appropriate Biological Methods**

a. Use the amphipod 10 day solid phase test and the sea urchin 96 hour larval development test in pore water for screening sites.

b. Use the amphipod solid phase test, the sea urchin larval development test in pore water, and the sea urchin larval development test at the sediment water interface (SWI) for confirmation. (A sensitive chronic test, such as the 28 day protocol for Leptocheirus, or tests using resident species may also be useful for confirmation).

c. Centrifuge pore water for bioassay testing. Frozen storage is probably acceptable if necessary.

d. Pore water dilutions are not necessary for screening, but do provide additional information for confirmation.

e. Pore water toxicity coupled with chemical analyses may be useful for establishing relationships between chemistry and biological effects.

f. Use of the Neanthes test should be discontinued because it provides no additional information beyond that provided by the amphipod and sea urchin protocol.

g. Studies should be conducted to investigate whether inhibition of embryo/larval development in pore water or solid phase (SWI) exposures can be correlated, or,is associated with ecological perturbation, such as impacts on benthic community structure.

Biomarkers

a. Biomarker analyses are currently difficult to interpret in terms of ecological effects. These types of analyses should not be used for toxic hot spot designation at present.

b. Biomarker analyses may be useful in monitoring cleanup activities to determine if there is continued exposure to pollutants.

Bioaccumulation

Recruit the services of a bioaccumulation expert into SPARC and examine how bioaccumulation can be used in the BPTCP.

**Issue 7. Appropriate Chemical Methods**

Metals

a. Perform SEM/AVS.

b. Use performance-based approach rather than rigid USEPA protocols.

c. Do bulk-phase metals in screening.

d. Do pore water metals to help determine causality for confirmation and cleanup planning

e. Preserve original samples for pore water chemistry.

f. Sediment samples can be frozen for a year for chemical analysis.

Organics

The April 1995 meeting ended before the organic chemical methods could be fully discussed. Nevertheless, similar recommendations to metal chemical methods were made. Further examination of this topic is scheduled for the next SPARC meeting.

a. The analyte list should be expanded to include Diazinon and other organophosphate pesticides

b. Use performance-based approach rather than rigid USEPA protocols.

c. Do bulk-phase organics in screening.

d. Do pore water organics to help determine causality for confirmation and cleanup planning

e. Preserve original samples for pore water chemistry.

f. Sediment samples can be frozen for a year for chemical analysis.

**Overall summary of SPARC recommendations**

a. Base program decisions on defensible science to provide common ground for all participants and interested parties.

b. Prepare workplans in advance to allow adequate scientific review, efficient allocation of funds, and timely reporting.

c. Use a carefully considered weight-of-evidence approach to accomplish program goals.

d. Include a bioaccumulation expert on the SPARC panel and examine how bioaccumulation can be used in the BPTCP. Thought should be given to reconciling the two different aspects of toxic hot spot designation: human health risk vs. observed ecological effects.

e. Food web models are not sophisticated enough to allow development of sediment quality criteria based on fish tissue concentrations. The mobility of most fish species limits utility for designation of toxic hot spots on a reasonable scale.

f. Site specific investigations are necessary for toxic hot spot designations. Focus immediately on sites most likely to be successfully designated as a toxic hot spot, and demonstrate program capacity for restoring environmental value to polluted sites.

g. Regional Boards must have more authority and take more responsibility for the planning of work in their respective regions. Local knowledge should be used to focus on the most relevant sites and analyses.

h. In designating toxic hot spots, follow a three-tiered approach: (1) carry out a flexible screening phase using any analysis of the triad or bioaccumulation technique (or); (2) a confirmation phase using all triad analyses (and); (3)intensive site specific studies demonstrating spatial extent, and causal relationships between pollutants and observed biological effects. It is very important to bring the potential responsible parties into the process as early as possible.

i. Confirmation and intensive cleanup studies should use a stratified random sampling design, with grids of suitable size to cover the area of concern. Field replication of all measures (toxicity, chemistry, benthic community structure, and bioaccumulation) should only be used when there is a clear and valid reason.

j. Statistical significance of toxicity should be determined based on a comparison to a reference envelope.

k. Benthic community degradation should not be based on a single index. A single community index is too easily discredited. Benthic community degradation should be based on convincing evidence determined on a site specific basis by a qualified ecologist.

l. Performance-based chemistry should be used.

m. Pore water toxicity, concurrent chemistry and spiked assays may be useful to determine associations between pollutants and biological effects. Correlations are not nearly as convincing in demonstrating associations. A TIE approach would also provide evidence of cause-effects relationships but should be used judiciously because of cost.

n. SEM/AVS are recommended for all samples.

o. Statewide and site-specific chemical objectives should be pursued.

p. Bioavailability concerns complicate interpretation of solid-phase sediment toxicity testing in evaluating the relationships between pollutant and biological effects.

q. Solid-phase sediment toxicity testing is useful for sediment quality assessment and toxic hot spot designation.

**Region-specific SPARC Recommendations**

Region 1

If local problems can be identified without toxicity screening then proceed to use the available resources as effectively as possible.

Bioaccumulation data may be appropriate to identify problem chemicals, biological exposure and potential sources of pollution in Region 1.

Biological effects measurements (toxicity screening or benthic community analysis) should be considered in cases where unknown toxic hot spots are present.

Region 2

Sampling should be done in a way to avoid mixing oxic and anoxic sediment regardless of sampling depth. Do the experiment necessary to show the effects of changes in oxidation state on toxicity and toxicity/chemistry relationships.

Use appropriate amphipod species based on knowledge of species tolerance limits to ammonia, salinity, and grain size.

Determine how to include bioaccumulation data into toxic hot spot screening.

Region 5

Pursue monitoring of pesticide degradation products.

Request that the SWRCB, Regional Boards, and Federal agency executive management agree to coordinate monitoring programs and share information from studies in the Bay-Delta. Also that the two Regional Boards pursuing BPTCP work in the Bay-Delta coordinate in the planning and monitoring work.

## REVISED STATEWIDE MONITORING DESIGN

This section comprises the revised Statewide monitoring approach for the BPTCP. The section was taken from the SPARC briefing document (SWRCB, et al. 1995) and revised to incorporate the SPARC recommendations. Revisions are included in the text in ~~strikeout~~ (deletions) and ***bold italics*** (additions).

### Legislative Mandate

Section 13391.5 of the Water Code defines toxic hot spots as "...locations in enclosed bays, estuaries, or adjacent waters in the 'contiguous zone' or the 'ocean' as defined in Section 502 of the Clean Water Act (33. U.S.C. Section 1362), the pollution or contamination of which affects the interests of the State, and where hazardous substances have accumulated in the water or sediment to levels which (1) may pose a substantial present or potential hazard to aquatic life, wildlife, fisheries, or human health, or (2) may adversely affect the beneficial uses of the bay, estuary, or ocean waters as defined in the water quality control plans, or (3) exceeds adopted water quality or sediment quality objectives."

### Specific Definition of a Toxic Hot Spot

One of the most critical steps in the development of toxic hot spot cleanup plans is the identification of hot spots. Once they are identified the parties responsible for the sites could be liable for the cleanup of the site or further prevention of the discharges or activities that caused the hot spot. Because the cost of cleanup or added prevention could be very high, the SWRCB is considering categorizing toxic hot spots to distinguish between sites with little information (potential toxic hot spots) and areas with significantly more information (candidate toxic hot spots).

### Proposed Specific Definition

Although the Water Code provides some direction in defining a toxic hot spot, the definition presented in Section 13391.5 is broad and somewhat ambiguous regarding the specific attributes of a toxic hot spot. The following specific definition provides the RWQCBs with a specific working definition and a mechanism for identifying and distinguishing between "potential," "candidate" and "known" toxic hot spots. A Candidate Toxic Hot Spot is considered to have enough information to designate a site as a Known Toxic Hot Spot except that the candidate hot spot has not been approved by the appropriate Regional Water Quality Control Board. Once a candidate toxic hot spot has been adopted into a

toxic hot spot cleanup plan then the site shall be considered a known toxic hot spot and all the requirements of the Water Code shall apply to that site.

a. Potential Toxic Hot Spot

The Water Code requires the identification of suspected or "potential" toxic hot spots (Water Code Section 13392.5). Sites with existing information indicating possible impairment, but without sufficient information to be classified further as a "candidate" or "known" toxic hot spot are classified as "potential" toxic hot spots. Four conditions sufficient to identify a "potential" toxic hot spot are defined below. If any one of the following conditions is satisfied, a site can be designated a "potential" toxic hot spot:

1. Concentrations of toxic pollutants are elevated above background levels, but insufficient data are available on the impacts associated with such pollutant levels to determine the existence of a known toxic hot spot;

2. Water or sediments which exhibit toxicity in screening tests or tests other than those specified by the State or Regional Boards;

3. Toxic pollutant levels in the tissue of resident or test species are elevated, but do not meet criteria for determination of the site as a known toxic hot spot, tissue toxic pollutant levels exceed maximum tissue residue levels (MTRLs) derived from water quality objectives contained in appropriate water quality control plans, or a health advisory for migratory fish that applies to the whole water body has been issued for the site by OEHHA, DHS, or a local public health agency, the waterbody will be considered a potential toxic hot spot. Further monitoring is warranted to determine if health warnings are necessary at specific locations in the waterbody.

4. The level of pollutant at a site exceeds Clean Water Act Section 304(a) criterion, or sediment quality guidelines or EPA sediment toxicity criteria for toxic pollutants.

b. Candidate Toxic Hot Spot:

A site meeting any one or more of the following conditions is considered to be a "candidate" toxic hot spot.

1. The site exceeds water or sediment quality objectives for toxic pollutants that are contained in appropriate water quality control plans or exceeds water quality criteria promulgated by the U.S. Environmental Protection Agency.

   This finding requires chemical measurement of water or sediment, or measurement of toxicity using tests and objectives stipulated in water quality control plans. Determination of a toxic hot spot using this finding should rely on recurrent measures over time (at least two separate sampling dates). Suitable time intervals between measurements must be determined.

2. The water or sediment exhibits toxicity associated with toxic pollutants ***that is significantly different from the toxicity observed at reference sites (i.e., when compared to the lower confidence interval of the reference envelope),*** based on toxicity tests acceptable to the State Water Resources Control Board or the Regional Water Quality Control Boards.

   To determine whether toxicity exists, recurrent measurements (at least two separate sampling dates) should demonstrate an effect. Appropriate reference and control measures must be included in the toxicity testing. The methods acceptable to and used by the BPTCP may include some toxicity test protocols not referenced in water quality control plans (e.g., the Bay Protection and Toxic Cleanup Program Quality Assurance Project Plan). Toxic pollutants should be present in the media at concentrations sufficient to cause or contribute to toxic responses in order to satisfy this condition.

3. The tissue toxic pollutant levels of organisms collected from the site exceed levels established by the United States Food and Drug Administration (FDA) for the protection of human health, or the National Academy of Sciences (NAS) for the protection of human health or wildlife. When a health advisory against the consumption of edible resident non-migratory organisms has been issued by OEHHA or DHS, on a site or waterbody, the site or waterbody is automatically classified a "candidate" toxic hot spot if the chemical contaminant is associated with sediment or water at the site or water body.

Acceptable tissue concentrations are measured either as muscle tissue (preferred) or whole body residues. Residues in liver tissue alone are not considered a suitable measure for known toxic hot spot designation. Animals can either be deployed (if a resident species) or collected from resident populations. Recurrent measurements in tissue are required. Residue levels established for one species for the protection of human health can be applied to any other consumable species.

Shellfish: Except for existing information, each sampling episode should include a minimum of three replicates. The value of interest is the average value of the three replicates. Each replicate should be comprised of at least 15 individuals. For existing State Mussel Watch information related to organic pollutants, a single composite sample (20-100 individuals), may be used instead of the replicate measures. When recurrent measurements exceed one of the levels referred to above, the site is considered a known toxic hot spot.

Fin-fish: A minimum of three replicates is necessary. The number of individuals needed will depend on the size and availability of the animals collected; although a minimum of five animals per replicate is recommended. The value of interest is the average of the three replicates. Animals of similar age and reproductive stage should be used.

4. Impairment measured in the environment is associated with toxic pollutants found in resident individuals.

Impairment means reduction in growth, reduction in reproductive capacity, abnormal development, histopathological abnormalities.~~, or identification of adverse effects using biomarkers.~~ Each of these measures must be made in comparison to a reference condition where the endpoint is measured in the same species and tissue is collected from an unpolluted reference site. Each of the tests shall be acceptable to the SWRCB or the RWQCBs.

Growth Measures: Reductions in growth can be addressed using suitable bioassay acceptable to the State or Regional Boards or through measurements of field populations.

Reproductive Measures: Reproductive measures must clearly indicate reductions in viability of eggs or offspring, or reductions in fecundity. Suitable measures include: pollutant concentrations in tissue, sediment, or water which have been demonstrated in laboratory tests to cause reproductive impairment, or significant differences in viability or development of eggs between reference and test sites.

Abnormal Development: Abnormal development can be determined using measures of physical or behavioral disorders or aberrations. Evidence that the disorder can be caused by toxic pollutants, in whole or in part, must be available.

Histopathology: Abnormalities representing distinct adverse effects, such as carcinomas or tissue necrosis, must be evident. Evidence that toxic pollutants are capable of causing or contributing to the disease condition must also be available.

~~Biomarkers: Direct measures of physiological disruption or biochemical measures representing adverse effects, such as significant DNA strand breakage or perturbation of hormonal balance, must be evident. Biochemical measures of exposure to pollutants, such as induction of stress enzymes, are not by themselves suitable for determination of "candidate" toxic hot spots. Evidence that a toxic pollutant causes or contributes to the adverse effect are needed.~~

5. Significant degradation in biological populations and/or communities associated with the presence of elevated levels of toxic pollutants.

   This condition requires that the diminished numbers of species of individuals of a single species (when compared to a reference site) are associated with concentrations of toxic pollutants. The analysis should rely on measurements from multiple stations. Care should be taken to ensure that at least one site is not degraded so that a suitable comparison can be made.

   In summary, sites are designated as "candidate" hot spots after generating information which satisfies any one of the five conditions constituting the definition.

c. Known Toxic Hot Spot:

A site meeting any one or more of the conditions necessary for the designation of a "candidate" toxic hot spot and has gone through a full State or Regional board hearing process, is considered to be a "known" toxic hot spot. A site will be considered a "candidate" toxic hot spot until approved as a known toxic hot spot in a Regional Toxic Hot Spot Cleanup Plan by the Regional Water Quality Control Board and approved by the State Water Resources Control Board.

## Monitoring Program Objectives

The four objectives of BPTCP regional monitoring are:

1. Identify locations in enclosed bays, estuaries, or the ocean that are potential or candidate toxic hot spots. ***Potential toxic hot spots are defined as suspect sites with existing information indicating possible impairment (criteria above) but without sufficient information to be classified further as a candidate toxic hot spot.***

2. Determine the extent of biological impacts in portions of enclosed bays and estuaries not previously sampled (areas of unknown condition);

3. Confirm the extent of biological impacts in enclosed bays and estuaries that have been previously sampled; and

4. Assess the relationship between toxic pollutants and biological effects.

## Review of Preliminary Studies and Research

Each of the seven RWQCBs participating in the program has assembled information that was used to develop a preliminary list of potential and candidate toxic hot spots (SWRCB, 1993). ***Further monitoring will be initiated by each RWQCBs participating in the program by preparing a toxic hot spot monitoring identification workplan identifying sites suspected to be impaired by pollutants or sites already identified as areas of concern. The workplan will specify if the sampling is for screening or confirmation and should include a list of types of analyses to be performed at each site. The workplan information will be assembled with input from Department of Fish and Game, SWRCB, and OEHHA staff based on the knowledge of local conditions and best professional judgement plus any pertinent scientific information obtained through either previous BPTCP screening or confirmation results or through information provided by other monitoring programs.***

## Biological Monitoring Methods

The tests listed in Table 1 are acceptable to measure water and sediment toxicity. Other tests may be added to the list as deemed appropriate by the State or Regional Water Boards provided the tests have a detailed written description of the test method; inter-laboratory comparisons of the method; adequate testing with water, wastewater, or sediments; and measurement of an effect that is clearly adverse and interpretable in terms of beneficial use impact.

## Chemical Methods

The BPTCP measures a variety of organic and inorganic pollutants in estuarine sediments (Stephenson et al. 1994). The BPTCP requires its laboratories to demonstrate comparability continuously through strict adherence to common Quality Assurance/Quality Control (QAQC) procedures, routine analysis of certified reference materials, and regular participation in an on-going series of inter-laboratory comparison exercises (round-robins). This is a "performance-based" approach of quality assurance.

The method used by the BPTCP are those used in the NOAA National Status and Trends Program (Lauenstein et al. 1993) and the methods documented in the DFG QAQC Manual (DFG, 1992). Under the BPTCP performance-based chemistry QA program, laboratories are not required to use a single, standard analytical method for each type of analysis, but rather are free to choose the best or most feasible method within the constraints of cost and equipment.

## Sampling Strategy

### Screening Sites and Confirming Toxic Hot Spots

~~In order to identify known toxic hot spots a two tier process was used. The first tier was a screening step where at least two toxicity tests were used at a site (Tables 2 and 3).~~ ***In order to identify toxic hot spots a two step process is used. Both steps are designed around a three tiered analysis approach (Triad analysis) plus an optional bioaccumulation information component. The Triad analysis consists of toxicity tests as listed in Table 2 (results from tests in Table 1 are also acceptable), benthic community analysis as characterized by the best professional judgement of the scientists performing the analysis, and performance-based chemical analysis for metals and organic chemicals. Screening and confirmatory phase toxicity tests specifically used by the BPTCP are listed in Table 2. Data***

*collected in the screening and confirmation phases are listed on Table 3.*

***The first step is a screening phase that consists of measurements using toxicity tests <u>or</u> benthic community analysis <u>or</u> chemical tests <u>or</u> bioaccumulation data to provide sufficient information to list a site as a potential toxic hot spot or a site of concern.*** Sediment grain size, total organic carbon (TOC) and $H_2S$ concentration are measured to differentiate pollutant effects found in screening tests from natural factors. ~~Chemical analyses (metals and organics) were performed on a subset of the screening samples.~~

***A positive result or an effect in any of the triad tests would trigger the confirmation step (depending on available funding).*** ~~If effects were found at sites by these screening steps, some sites were retested (depending on available funding) to confirm the effects.~~ ***The confirmation phase consists of performing all components of the triad analysis: toxicity, benthic community analysis, and chemical analysis, on the previously sampled site of concern or wherever previous evidence indicates a site may be impacted. A candidate THS is a station that has significant effect measured in the toxicity tests or benthic community analysis coupled with chemistry information that shows that pollutants could contribute to the observed effects. A hit in toxicity and chemistry, benthic community analysis and chemistry or all three components of the triad analysis would classify a site as a candidate toxic hot spot (as described in the candidate THS criteria listed above).*** ~~In the confirmation step measurements were replicated and compared to reference sites or conditions. Chemical measurements (metals, organics, TOC, $H_2S$) and other factors (e.g., sediment grain size) were measured. Measurements of benthic community structure and, if needed, bioaccumulation were also made.~~

Table 1
Water and Sediment Toxicity Tests That Meet
the Criteria For Acceptability

| Type of Toxicity Test | Organism Used | | Reference |
|---|---|---|---|
| | Common Name | Scientific Name | |
| Solid Phase | Amphipod | Rhepoxynius | ASTM, 1993 |
| Sediment | Amphipod | Eohaustorius | ASTM, 1993 |
| | Amphipod | Ampelisca | ASTM, 1993 |
| | Amphipod | Hyalella | ASTM, 1993 |
| | ***Sea Urchin*** | ***Strongylocentrotus*** | ***Anderson et al., 1995*** |
| | ~~Polychaete~~ | ~~Neanthes~~ | ~~Johns et. al., 1990~~ |
| Sediment Pore Water~~*~~ | Bivalve larvae | Crassostrea | ASTM, 1993 |
| | | Mytilus | ASTM, 1993 |
| | Abalone larvae | Haliotis | Anderson et al., 1990 |
| | Echinoderm fertilization | Strongy-locentrotus | Dinnel et al., 1987; with modification by EPA, 1992 |
| | Giant kelp | Macrocystis | Anderson et al., 1990 |
| | Red alga | Champia | Weber et al., 1988 |
| | Fish embryos | Atherinops | Anderson et al., 1990 |
| | | Menidia | Middaugh et al., 1988 |
| | | Pimephales | Spehar et al., 1982 |
| | Cladoceran | Daphnia | Nebecker et al., 1984 |
| | | Cereodaphnia | Horning and Weber, 1985 |
| Ambient Water | Bivalve larvae | Crassostrea | ASTM, 1993 |
| | | Mytilus | ASTM, 1993 |
| | Abalone larvae | Haliotis | Anderson et al., 1990 |
| | Echinoderm fertilization | Strongylocen-trotus | Dinnel et al., 1987; with modifications by EPA, 1992 |
| | Giant kelp | Macrocystis | Anderson et al., 1991 |
| | Red alga | Champia | Weber et al., 1988 |
| | Mysid | Holmesimysis | Hunt et al., 1992 |
| | Fish embryos | Atherinops | Anderson et al., 1990 |
| | | Menidia | Middaugh et al., 1988 |
| | | Pimephales | Spehar et al., 1982 |
| | Fish larvae | Atherinops | Anderson et al., 1990 |
| | | Menidia | Peltier and Weber, 1985<br>Weber et al., 1988 |
| | | Pimephales | Peltier and Weber, 1985<br>Weber et al., 1988 |
| | Cladocerans | Daphnia | Nebecker et al., 1984 |
| | | Cereodaphnia | Horning and Weber, 1985 |

~~*Pore water tests (other than amphipods) alone can not be used to designate a candidate toxic hot spot.~~

Table 2

Screening and Confirmation Tests for
Toxic Hot Spot Identification

| Test Organism | Type | End Point |
|---|---|---|
| Rhepoxynius, Eohaustorius (Amphipod) | Solid Phase | Survival *(10 day)* |
| ~~Haliotus, Mytilus, Crassostrea~~ | ~~Overlying water~~ | ~~Shell development~~ |
| Strongylocentrotus (Sea urchin) | Sediment pore water<br>Sediment/water Interface<br>(Confirmation only) | *72-96 hour* ~~Fertilization~~ development~~, and/or anaphase aberration~~ |
| ~~Neanthes (Polychaete worm)~~ | ~~Bedded sediment~~ | ~~Survival and growth~~ |

**A Battery of Screening Tests**

Selecting a battery of toxicity screening tests (Table 2) can improve cost-effectiveness by expanding the range of potential impacts to be evaluated. Although recurrent toxicity must be demonstrated to qualify a site as a "candidate" toxic hot spot, the degree of certainty for each of the measurements does not necessarily have to be equivalent. The cost of confirming toxicity at a site can be prohibitively high, especially if it includes a large number of field replicates and extensive reference site testing. The screening tests should allow for a relatively rapid lower cost assessment of the site. ***Toxicity screening test should include an amphipod solid phase test and a sea urchin larval development test in pore water. Confirmation toxicity test should include an amphipod solid phase test, a sea urchin larval development test using pore water, and a sea urchin larval development at the sediment/water interface (Tables 2 and 3).***

~~Even though the list of acceptable tests is long (see Table 1), the State and Regional Water Boards have used between two and four tests to screen sites (Table 2). For all screening, at least one amphipod test was performed. Other tests were performed as needed depending on funding availability, the needs of collaborators (such as the National Oceanic and Atmospheric Administration or the EPA Environmental Monitoring and Assessment Program), test organisms sensitivity to the~~

Table 3

Types of Data *To Be* Collected in Regional Monitoring Programs for the Identification of Toxic Hot Spots

| Type of Data | Screening | Confirmation |
|---|---|---|
| Toxicity testing | Suite of ~~4~~ ***2*** tests (see Table 5) | ~~Repeat of positive results~~ ***Suite of 3 tests*** |
| Field replicates | None | if needed |
| Lab replicates | Five | Five |
| Reference sites | Reference Envelope | Reference Envelope |
| Physical analysis | Grain size | Grain size |
| Chemical analyses | Ammonia, hydrogen sulfide, TOC, pesticides, PCB, PAH, TBT, metals, ***AVS/SEM*** | Ammonia, hydrogen sulfide, TOC, pesticides, PCB, PAH, TBT, metals, ***AVS/SEM*** |
| Benthic community analysis | Optional | ~~Required~~ ***Optional*** |
| Bioaccumulation | Occasionally | Occasionally (sites with no pre-existing bioaccumulation data) |

Table 4

Sequence of Tasks for Designating Toxic Hot Spots

| | |
|---|---|
| 1. | Select toxicity screening sites. |
| 2. | Sample screening sites. |
| 3. | Conduct battery of two toxicity screening tests; or Benthic community analysis; or Chemical analysis; or bioaccumulation. ~~analyze~~ ***measure*** for hydrogen sulfide, ammonia, TOC, and grain size. |
| 4. | Determine whether quality assurance requirements have been met. |
| 5. | Report on Items 3 and 4. |
| 6. | Select ~~and match~~ hits and ~~potential~~ reference ***envelope*** sites. ~~for ammonia, hydrogen sulfide, and grain size.~~ |
| ~~7~~. | ~~Conduct metals and organic chemical analysis on subset of screening sites from Item 6.~~ |
| ***7***~~8~~. | Determine whether quality assurance requirements have been met. |
| ***8***~~9~~. | Report on Item~~s~~ 7 ~~and 8~~. |
| ***9***~~10~~. | Select sites ~~and toxicity tests~~ for confirmation and reference ***envelope*** sites. |
| ***10***~~11~~. | Sample confirmation and reference ***envelope*** sites. |
| ***11***~~12~~. | Conduct ~~subset of the~~ battery of ~~toxicity~~ tests which were screening hits; analyze for hydrogen sulfide, TOC, ***DOC,*** and conduct benthic community analysis. |
| ***12***~~13~~. | Conduct ***bulk phase, pore water or both,*** metals and organic chemical analyses, ***plus SEM/AVS***. |
| ***13***~~14~~. | Determine whether quality assurance requirements have been met. |
| ***14***~~15~~. | Report on Items ~~12~~ ***11*** through ~~15~~ ***14***. |
| ***15***~~16~~. | Conduct statistical and other analyses to determine whether sites qualify as candidate toxic hot spots. |

~~pollutants expected to be present, and the media (bedded sediment or pore water) thought to be contaminated.~~

## Site Selection

Two somewhat different approaches ~~were~~ ***are*** used in BPTCP monitoring. Six of the coastal RWQCBs have used a design that combines toxicity testing, chemical analysis, and benthic community analysis in a two-phased screening-confirmation framework (Tables 3 and 4).

The Central Valley RWQCB, with jurisdiction over the Sacramento-San Joaquin Delta, has designed its program to respond to Delta conditions and to the water quality problems characteristic of that area. Fresh water toxicity testing combined with water chemistry (metals and pesticides) constitutes the main program components. Sediment toxicity testing could be added to the monitoring design at a later stage.

Four different categories of sites have been identified for sampling in the BPTCP monitoring program: (1) potential toxic hot spots base on existing information, (2) high risk sites *of* ***concern*** based on existing information ***and local knowledge of the area,*** (3) stratified random sites, and (4) reference sites ***to be included in the reference envelope.***

Potential toxic hot spots are the highest priority sites because some ~~indications~~ ***information*** already exists that these sites have a pollution-related problem. The~~se~~ data ***associated with these*** sites indicate ~~are typically sites with information available on~~ chemical contamination of mussel tissue, data documenting water and sediment toxicity, measurements of metals or organic chemicals in sediments, ~~and~~ or ~~occasionally~~, biological impairment. These sampling efforts are typically point estimates.

There are many other sites that are considered "high risk" ***sites of concern*** even though we have no monitoring information to support this contention. High risk sites are locations where a nearby activity (such as marinas, storm drains, and industrial facilities) are thought to be associated with a certain risk of toxicity. The measurements at high risk sites ***of concern*** are either point estimates or selected probabilistically ***or suspected problem sites on the basis of local knowledge.***

When little is known about the quality of a waterbody segment, the monitoring efforts should use a stratified, random sampling approach. ***This would be used during the screening phase in order***

*to provide the opportunity of finding new toxic hot spots and* ~~*as well as These random sites are useful*~~ ***help*** in determining the quality of larger areas in the State's enclosed bays and estuaries. This probabilistic approach will allow for the State and Regional Water Boards to make better estimates of area (percentage) of water bodies that is impacted. The State and Regional Water Boards have used the techniques used by the U.S. Environmental Protection Agency's Environmental Monitoring and Assessment Program (SWRCB et al. 1994).

Reference sites

~~Locating reference sites requires identification and testing of a variety of potential reference sites encompassing the expected range of grain size, TOC, and other characteristics. Existing data sets that describe chemical contamination, grain size, and TOC at marine and estuarine sites are reviewed. Since these sources yield an insufficient number of sites, fine grained areas presumed to be relatively free of contamination are also examined. These sites may likewise prove to be rare, so sites with chemicals present, but experiencing low energy tidal flushing, will also be sampled. Sites with previous indication of no pollution, and those lacking sediment toxicity measurements will also be sampled. Finally, random selection of sites (as described above) may prove useful in locating reference sites.~~

Locating reference sites requires identification and testing of a variety of potential sites encompassing the expected range of grain size, TOC, and other characteristics. ***In selecting a reference site common sense and knowledge of local conditions should always be used to avoid areas known to be disturbed or polluted. Some criteria to consider in defining a reference condition are as follow:***

1. ***High amphipod abundance***
2. ***High species richness***
3. ***Sediment tolerance for non-treatment effects ($NH_3$, $H_2S$, grain size temperature, salinity, etc., (Table 5) above or below which biological effects could be attributed relatively to pollutant toxicity.***
4. ***Sites with low Chemistry (below median values ERM, PEL, etc.)***

***After excluding known unacceptable areas the remaining water bodies are randomly sampled (screening phase tests or existing information can be used). The samples are analyzed for chemistry, toxicity, and benthic ecology. The chemistry data is evaluated in order to select the sites that do not exceed more than one upper value for existing chemistry criteria. The***

*Table 5. Non-Treatment Limits for 10-d sediment toxicity tests with Ampelisca abdita, Eohaustorius estuarius, Leptocheirus plumulosus, or Rhepoxynius abronius (U.S. EPA. 1994).*

| Parameter | *Ampelisca abdita* | *Eohaustorius estuarius* | *Leptocheirus plumulosus* | *Rhepoxynius abronius* |
|---|---|---|---|---|
| Temperature(°C) | 20 | 15 | 25 | 15 |
| Overlying Salinity (%) | >10 | 0-34 | 1.5-32 | <25 |
| Grain Size (% silt/clay) | >10 | full range | full range | <90 |
| Ammonia (total mg/l,pH 7.7 | <30 | <60 | <60 | <30 |
| Ammonia (UI[1] mg/L,pH 7.7 | <0.4 | <0.8 | <0.8 | <0.4 |
| Sulfides | NA | NA | NA | NA |

[1] UI = unionized ammonia

*toxicity data is evaluated to eliminate those sites that have extremely high toxicity. Finally, the reference envelope sites are chosen on the basis of moderate to high species richness, abundance of amphipods or other indicator species, and any other indicator of ecological health that can be argued convincingly.*

Once reference sites are chosen for a particular area they are re-sampled along with the test sites during the confirmation phase.

*Determination of toxic hot spots will be achieved by comparing the test site toxicity response against a sufficiently large reference envelope of a population of reference site responses. The reference envelope will include results from all reference sites in a particular area, past and present. The reference envelope approach, currently under development, will be used to determine whether the level of toxicity exceeds the lower confidence interval of the reference envelope. As more reference site toxicity results become available more will be known on the range of organism responses found within a reference site condition. This will provide a better tool for determining differences between the toxicity response at reference sites relative to the level of toxicity responses at impacted sites.*

Toxicity Screening

All tests include controls which ~~were~~ *are* conducted in media known to exert minimal stress on test organisms. Both positive (toxicant present) and/or negative (toxicant absent) controls ~~were~~ *are* used to ensure that test organisms are responding within expected limits (Table 3).

The screening step begins with the collection of a single field sample from each site (Table 4, Steps 1 and 2). Five laboratory replicates are required to accommodate statistical comparison with the control. Ammonia and hydrogen sulfide analyses are performed on the media of all tests (Table 4, Step 3) to determine their relative contribution to any observed toxic effects. Grain size and TOC values are determined on all sediment samples to evaluate the response of the organisms to these factors. ~~Although the lack of field replicates restricts statistical comparisons with other sites, this approach allows the BPTCP to test more locations for toxicity within the allocated funding.~~ *Screening can include the use of chemistry, toxicity tests, benthic community structure analysis, or bioaccumulation monitoring. The analysis is designed to be flexible, and to fit the Regional Board's needs to provide sufficient information to warrant listing a site as a potential toxic hot spot or pursue some other type of regulatory action.*

All these data, along with an assessment of quality assurance performance, are reviewed. Toxicity hits and potential reference ***envelope*** sites are selected and matched for ammonia, hydrogen sulfide, grain size, and TOC. Sites with hits in either one of the tests performed are candidates for re-sampling during the confirmation phase.

**Confirmation (i.e., Qualification as Candidate Toxic Hot Spots)**

Some of the screening sites (Table 4, Steps *9* ~~10~~ and *10* ~~11~~) with at least one positive test result will be revisited to evaluate the recurrent nature of the toxicity, impacts on the benthic community or high concentrations of specific pollutants. This requires repeat testing of potential toxic hot spots by ***performing the three components of the triad analysis: toxicity, benthic community analysis, and chemistry. This phase could also include intensive investigations to identify causal relationships and grid sampling to show gradients and spatial extent.*** ~~to ensure that toxicity was present or absent. Confirmation testing was more intensive because of (1) addition of field replicates (three to a site), (2) comparison to reference sites (unless water toxicity is the focus), and (3) benthic community analysis (Table 3).~~

~~For each positive toxicity test at a screening site, confirmation was performed for the same test.~~ ~~Generally,~~ Benthic analysis ~~was also performed and~~ ***will*** *be* added to an ever-enlarging nearshore benthic community database which will be periodically evaluated to determine whether impacted and non-impacted sites can be distinguished (Table 4, Step ***11*** ~~12~~). ~~When either recurrent toxicity was is demonstrated with a positive confirmation test or benthic impacts are were suspected, chemical analysis were also performed (Table 4, Step 13).~~ Careful review of all quality assurance procedures ~~was~~ is conducted and, upon approval, will be followed by statistical analysis of the data. ~~Compared to screening, this analysis will be is more comprehensive. and will include measures of field variability in toxicity, benthic data, and reference site conditions.~~

Once both toxicity and benthic impacts have been confirmed through comparison with an appropriate reference ~~site~~ ***envelope*** ~~and appear to be due to human causes~~ the site will be declared a candidate toxic hot spot. When toxicity is present but benthic impacts are lacking, careful analysis will be performed to determine whether the two results are in conflict. Similarly, when toxicity is not demonstrated but benthic impacts are observed, careful review will be conducted to determine whether the same explanation prevails or whether some factor other than

toxicants may be responsible. In either case, decisions about a particular site will be based upon best scientific judgement after careful consideration of the evidence gathered. Further characterization of the site (such as areal extent, range of effects, and source determination) will be described in the cleanup plan ~~and is not intended (unless samples are collected using a random or stratified random design) under this phase of the program.~~

## Quality Assurance

The BPTCP Quality Assurance Project Plan (Stephenson et al. 1994) presents a systematic approach that has been implemented within each major data acquisition and data management component of the program. Basic requirements specified in the QAPP are designed to: (1) ensure that collection and measurement procedures are standardized among all participants; (2) monitor the performance of the various measurement systems being used in the program to maintain statistical control and to provide rapid feedback so that corrective measures can be taken before data quality is compromised; (3) assess the performance of these measurement systems and their components periodically; and, (4) verify that reported data are sufficiently complete, comparable, representative, unbiased, and precise.

## References


Anderson, B.S., J.W. Hunt, S.L. Turpen, A.R. Coulon, M. Martin, D.L. McKeown and F.H. Palmer. 1990. Procedures manual for conducting toxicity tests developed by the marine Bioassay Project. California State Water Resources Control Board 90-10WQ. Sacramento, California. 113 p.

***Anderson, S.L., J.P. Knezovich, J. Jelinski, D.J. Steichen. 1995. The utility of pore water toxicity testing for development of site-specific marine sediment quality objectives for metals, Report LBL-37615 UC-000, Lawrence Berkeley National Laboratory, University of California, Berkeley, CA.***

ASTM. 1987. Standard practice for conducting static acute tests with larvae of four species of bivalve molluscs. Procedure E724-80. Annual book of ASTM standards; water and environmental technology. Vol 11.4: 382-388. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1993. Designation E 1367: Standard guide for conducting 10-day static sediment toxicity tests with marine and estuarine amphipods. Volume 11.04: Pesticides; resource recovery; hazardous substances and oil spill responses; waste disposal; biological

effects. Annual book of standards; water and environmental technology. American Society for Testing and Materials, Philadelphia, PA.

Chapman, P.M. and J.D. Morgan. 1983. Sediment bioassay with oyster larvae. Bull. Environs. Centime. Toxicol. 31: 438-444.

Department of Fish and Game. 1992. Laboratory Quality Assurance Program Plan. Environmental Services Division. Sacramento, CA.

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

Dinnell, P.J., J. Link, and Q. Stober. 1987. Improved methodology for sea urchin sperm cell bioassay for marine waters. Arch. Envir. Cont. and Toxicol. 16: 23-32.

Di Toro, D.M., J.D. Mahony, D.J. Hansen, K.J. Scott, M.B. Hicks, S.M. Mayr, and M.S. Redmond. 1990. Toxicity of Cadmium in Sediments: The Role of Acid Volatile Sulfide. Environmental Toxicology and Chemistry, Vol. 9, pp. 1487-1502. 1990.

Environmental Protection Agency (EPA). 1992. Sea urchin (Strongylocentrotus purpuratus) fertilization test method. Final Draft. Gary A. Chapman, U.S. Environmental Protection Agency. ERL - Pacific Ecosystems Branch, Newport, Oregon.

Horning, W.B. II and C.I. Weber (eds.). 1985. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms. Environmental Monitoring and Research Laboratory - Cincinnati Office of Research and Development. U.S. Environmental Protection Agency, Cincinnati, Ohio. EPA/600/4-85/014.

Hunt, J.W., B.S. Anderson, S.L. Turpen, H.R. Barber, D.L. Johns, D.M., T.C. Ginn and D.J. Reish. 1990. The juvenile neanthes sediment bioassay. Puget Sound Notes, No. 24, U.S. EPA, Seattle, WA.

Lauenstein, G.L., A.Y. Cantillo, and S. Dolvin (eds.) 1993. A compendium of methods used in the NOAA National Status and Trends Program. National Ocean Service, Office of Ocean Resources Conservation and Assessment, Silver Springs, MD.

Long, E.R. and L. Morgan 1990. The potential for Biological Effects of Sediment-sorbed Contaminants Tested in the National Status and Trends Program. NOAA Technical Memorandum NOS OMA 52.

McKeown D., F.H. Palmer and M. Martin. 1991. Marine Bioassay Project Sixth Report: Interlaboratory comparisons and protocol development with four marine species. Report #91-21-WQ. State Water Resources Control Board, California.

Middaugh, D.P., M.J. Hemmer, and E.M. Lores. 1988. Teratological effects of 2,3-dinitrophenol, produced water, and naphthalene on embryos of the inland silverside Menidia beryllina. Dis. Aquat. Org., 4,53 - 65.

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

Peltier, W.H., and C.I. Weber. 1985. Methods for measuring the acute toxicity of effluents to freshwater and marine organisms. Environmental Monitoring and Support Laboratory - Cincinnati Office of Research and Development. U.S. Environmental Protection agency, Cincinnati, Ohio. EPA/600/4-85/013.

Spehar, R.L., D.K. Tanner and J.H. Gibson. 1982. Effects of kelthane and pydrin on early life stages of fathead minnows (Pimephales promelas) and amphipods (Hyalella azteca). In J.G. Peatson, R.B. Foster and W.E. Bishop, eds., Aquatic Toxicity and Hazard Assessment: Fifth Conference. STP 766. American Society for Testing and Materials, Philadelphia, PA. pp. 234-244.

State Water Resources Control Board. 1993. Staff Report: The Status of the Bay Protection and Toxic Cleanup Program. Sacramento, CA. 231 pp + 5 Appendices.

State Water Resources Control Board, National Oceanic and Atmospheric Administration, and U.S. Environmental Protection Agency Environmental Monitoring and Assessment Program. 1994. Technical proposal for a cooperative agreement between the State Water Resources Control Board, and the U.S. Environmental Protection Agency Environmental Monitoring and Assessment Program. Sacramento, CA. 26 pp + 3 appendices.

State Water Resources Control Board, Regional Water Quality Control Boards, Department of Fish and Game. 1995. Scientific Planning and Review Committee Briefing Document for the Bay Protection and Toxic Cleanup Program. Sacramento, CA. 89 pp.

Stephenson, M., M. Puckett, N. Morgan, and M. Reid. 1994. Bay Protection and Toxic Cleanup Program Quality Assurance Project Plan. 12 Sections and 1 appendix.

***United States Environmental Protection Agency. 1994. Methods for Assessing the Toxicity of Sediment Associated Contaminants with Estuarine and Marine Amphipods. U.S. EPA Office of Research and Development. Washington, D.C. 20460. EPA/600/R-94/025.***

TOXICITY ISSUE PAPER

**Statistical Methods for Distinguishing Sites of Concern Using the Reference Envelope Approach: Issues to be Resolved**

At the April 12-13, 1995 meeting of the BPTCP SPARC, the committee expressed general support for the Reference Envelope statistical approach presented by Bob Smith. This approach is based on the use of toxicity test data from reference sites to describe a population of values indicative of general ambient conditions in a given water body or group of water bodies. From this population of values, a tolerance limit can be calculated that serves as a cutoff for determining significant toxicity relative to the general condition of the water body. This approach is described below, followed by a list of issues that need to be resolved in order to apply this approach in study plans under consideration by the program.

## Rationale for the Reference Envelope Approach

For the purposes of identifying sites of concern in the BPTC Program, it is necessary to distinguish between sites where toxicity is clearly indicative of localized pollution and sites where test results are more characteristic of background response. Since samples from a group of study sites would be expected to produce some level of variation in toxicity test response even in the absence of pollution, a method is required to determine what level of test response is significantly greater than expected of samples representing general water body conditions. In many heavily urbanized estuaries, it is probable that all sites have some degree of contamination and some resulting potential for causing adverse biological effects. However, logistical constraints require that efforts be focused on sites where it can be convincingly demonstrated that observed toxicity is due to localized pollution rather than to background variability. In this context, the terms "background", "ambient" or "reference" are defined as representative of general water body conditions, rather than conditions thought to exist prior to anthropogenic influence.

## Reference Envelope Statistical Method

The concept of the reference envelope is described here, as taken from Bob Smith's notes from the previous SPARC meeting. A manuscript containing more details is available from Bob Smith upon request.

**Sampling Design**

An effectively random sample of a population of locations (stations) representative of the "natural background" of indicator values for the area of interest is required. This "natural background" may contain some toxicity or contamination, e.g., Tomales Bay or San Pablo Bay. The chosen hot spots should be "hotter" than the background condition, since it is not practical to remediate very large areas, nor is it legally defensible to penalize someone for local toxicity no worse than that found in the larger area in general.

The random sample of stations will be used to characterize what will be called a reference population. In a statistical test, potential hot spots will be compared to this reference population.

**Statistical Test**

We would like to see if potential hot spots are unusual (in the direction of toxicity or "badness") compared to the reference population. We can use a statistical test to estimate if a potential hot spot is outside a chosen percentile of the reference population distribution (in the direction of toxicity). The percentile chosen for the test would reflect how "unusual" relative to the reference population a station must be in order to be declared a hot spot. For example, if considering % survival for a bioassay test, one might pick the 1st percentile. This would mean that a station would have to be associated with % survival lower than 99% of the reference population in order to be called a hot spot.

The statistical test is used to identify an indicator value (e.g., a % survival value) that can be used as a cutoff or threshold to distinguish between the reference population and a hot spot (as far as the indicator is concerned). A one-tailed tolerance interval bound will accomplish this. The tolerance interval is based on the variance of the random sample of reference stations, and will therefore incorporate the important sources of natural variation among station locations. The tolerance interval also accommodates the uncertainty involved in estimating the mean and variance of the reference population and the test stations.

The computed tolerance interval bound is equivalent to the edge of a "reference envelope", thus this is called the reference envelope approach. This implies that the reference population is largely contained within a figurative reference envelope, and outliers (potential hot spots) are found outside the envelope.

We can compute the toxicity level that will cover the $p^{th}$ percentile 1 minus alpha proportion of the time as the lower bound (L) of a tolerance interval (Vardeman 1992) as follows:

$$L = X_r - [\, g_{a,p,n} \cdot S_r \,]$$

where $X_r$ is the mean of the sample of reference stations, $S_r$ is the standard deviation of the toxicity results among the reference stations, and n is the number of reference stations. The g values can be obtained from tables in Hahn and Meeker (1991) or Gilbert (1987). S contains the within- and between-location variability expected among reference locations. If the reference stations are sampled at different times, then S will also incorporate between-time variability. We call L the "edge of the reference envelope" because it represents a cutoff toxicity level we will use to distinguish toxic from non-toxic sediments. The value used for p will depend on the level of certainty needed for a particular regulatory situation.

The population of reference values and estimates of the $p^{th}$ percentile of the reference distribution are shown in Figure 1.

## Issues Regarding Use of the Reference Envelope Approach

### Reference Site Selection Criteria

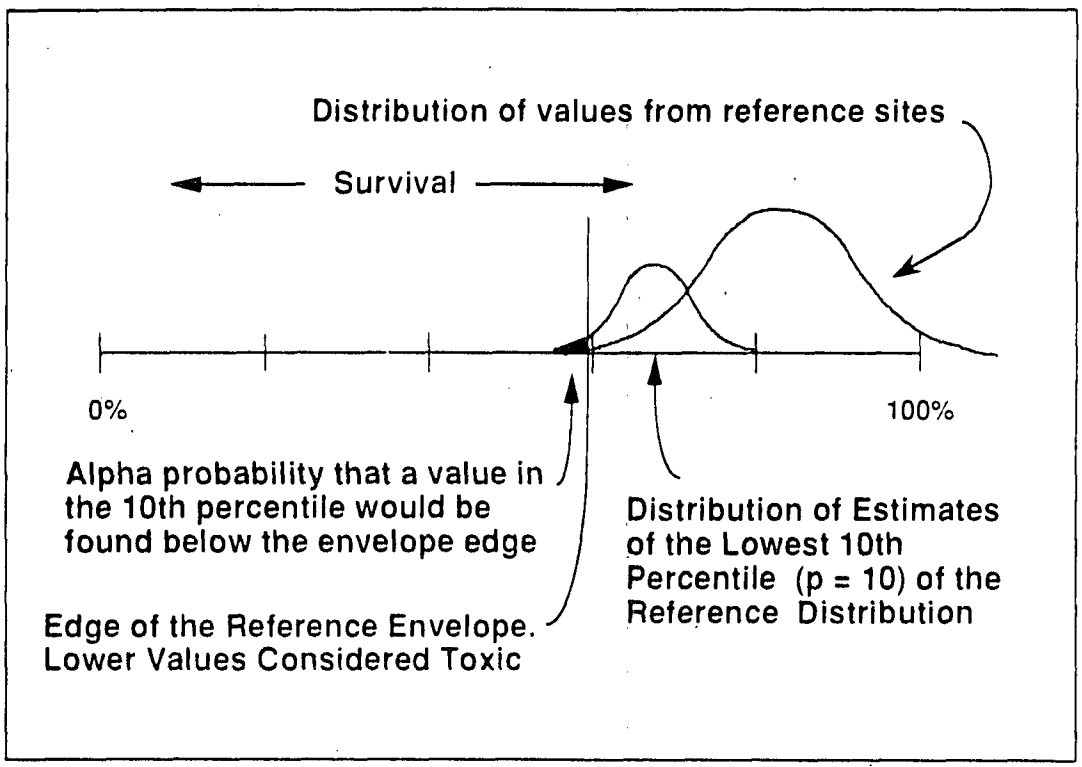
At the April 12-13, 1995 meeting of the BPTCP SPARC, the committee decided that reference sites should be chosen based on data from chemistry and benthic community analyses that indicate low levels of pollution and lack of impacts to benthic communities. It is assumed that both reference and test sites must have physical parameters of grain size and salinity within the tolerances of the test organisms.

### Chemistry Criteria

In the BPTCP San Diego Bay study and Southern California Coastal Lagoons study, sites were eliminated from consideration as reference sites if any chemicals for which ERM and/or PEL values have been derived exceeded either of those values. In the Southern California Coastal Lagoons Study, DDT and DDT metabolite concentrations above the ERM were allowed if they were below the SEC concentration derived by MacDonald (1994).

In the San Francisco Bay Reference Site Study, all reference sites exceeded the PEL value for chromium and the ERM and PEL values for nickel. Nickel is ubiquitous in San Francisco Bay, and has been shown to be toxic only at pore water concentrations much higher than corresponding ERM values (Anderson et al.,

Figure 1. Schematic illustration of the method for determining the lower tolerance interval bound (edge of the reference envelope) to determine sample toxicity relative to a percentile of the reference site distribution.



1995). AVS/SEM data were not available to evaluate potential impacts of these metals. Total DDT was found to be above ERM and PEL values at one field replicate of one candidate reference site. The concentration of dibenz(ah)anthracene barely exceeded the PEL in one sample at another candidate reference site but this concentration was half of the ERM.

**Benthic Ecology Criteria**

In studies of San Diego Bay and the Southern California Coastal Lagoons, sites were classified as "degraded", "undegraded", or "transitional" based on the total number of species per station, the total number of individuals per station, the number of crustacean species per station, and the presence of indicator species (either positive or negative). Sites classified as "undegraded" were eligible for use as reference sites.

Benthic ecology data were not used in the selection of reference sites in San Francisco Bay because of the magnitude of seasonal fluctuations in species composition and the impact of invading exotic species.

**Questions Regarding Reference Site Selection Criteria**

1. Are the chemistry selection criteria appropriate?

2. How should elevated concentrations of DDT and nickel be evaluated?

3. Should AVS/SEM ratios be used in place of ERM or PEL values in determining metals concentrations allowable at reference sites?

4. How should test organisms tolerances to ammonia and hydrogen sulfide be factored into reference site selection?

5. Are the benthic selection criteria appropriate?

6. Should site toxicity data be considered at all? If so, how?

7. How should site location with respect to pollution sources and other "common sense" considerations be factored into the selection process?

**How Many Reference Sites are Necessary?**

The reference envelope approach provides a tolerance limit that serves as the threshold for toxicity test results. Percent survival below the tolerance limit indicates significant sample toxicity. The calculation of this tolerance limit is influenced by the reference population mean and variance, and by the number of reference sites. The more reference sites available, the

tighter the distribution, and the higher the tolerance limit (assuming high survival in reference site samples).

The effect of reference sample size on calculation of the tolerance limit is indicated by the table of g values. As presented above, the tolerance limit is calculated by subtracting the product of the reference population variance and the appropriate g statistic from the reference population mean. Therefore, the lower the g value, the closer the tolerance limit will be to the population mean. For an alpha value of 0.05 and a p value of 10% (lowest 10th percentile), g varies with n as follows:

| n: | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 15 | 20 | 30 | 50 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| g: | 20.6 | 8.2 | 4.2 | 3.4 | 3.0 | 2.8 | 2.6 | 2.5 | 2.4 | 2.1 | 1.9 | 1.8 | 1.6 | 1.5 |

***Question:*** How many reference sites are necessary to adequately compute the tolerance limit?

**How Often Must Reference Sites Be Sampled?**

The main question here is, if we were to decide that 10 reference sites are necessary, must all ten reference sites be sampled and tested every time a test sample is analyzed, or can historic data be included in the population of reference values used to make the comparison?

If previous data can be used, how should this be done?

1. Create a data set of reference site toxicity values appropriate for each type of test condition (salinity, grain size, physical features). This would generate a single tolerance limit that could always be used as a toxicity threshold.

2. Create a reference data set as above, but add new concurrent reference site data each sampling run.

3. Compare test site data against both concurrent and historic reference site data. Would a site be toxic if it were outside <u>either</u> the concurrent or historical set of reference values, or would it need to be outside <u>both</u>?

4. What details need to be worked out from a statistical perspective to allow comparisons against historical reference data?

**How Many Reference Populations Are Necessary for Analysis of California Sites?**

How closely do reference sites need to match test sites in terms of:

1. Grain size

2. TOC

3. Salinity

4. Physical Environment (e.g., coastal lagoon, human-made harbor, estuary, open bay, etc.)

5. Human Environment (e.g., dredging history, history of pollutant inputs).

Are multiple reference populations necessary within a single water body, such as San Francisco or San Diego Bay?

**Interactions Between Policy and Science**

How should policy and scientific perspectives be reconciled in the following areas:

1. Selection of reference sites?

2. Choice of p values in calculating reference envelope tolerance limits?

3. Identification of toxic sites?

**What Should Be Done When the Analysis Doesn't Work?**

In some cases, variability among reference site responses to some protocols can lead to very low tolerance limits. In the San Diego study, tolerance limits for some pore water tests were below zero. In such cases, significant toxicity would be impossible to detect, regardless of test sample response. Can test data be used in hot spot designation under such circumstances?

**References**

Anderson, S.L., J.P. Knezovich, J. Jelinski, D.J. Steichen. 1995. The utility of pore water toxicity testing for development of site-specific marine sediment quality objectives for metals, Report LBL-37615 UC-000, Lawrence Berkeley National Laboratory, University of California, Berkeley, CA.

Hahn, G.J. and W.Q. Meeker. 1991. Statistical intervals. A guide for practitioners. A Wiley-Interscience Publication. John Wiley & Sons, Inc. New York. 392 pp.

Smith, R.W. 1995. Southern California Regional Monitoring Project The Reference Envelope Approach to Impact Monitoring.

Report to USEPA, Region IX. EcoAnalysis: Ojai, California 67 pp.

Smith, R.W. 1995. The reference envelope approach to impact monitoring. EcoAnalysis, Ojai, California.

Vardeman, S.B. 1992. What about the other intervals? The American Statistician 46(3): 193-197.

CHEMICAL ASSOCIATION ISSUE PAPER

## Use of Sediment Quality Guidelines

### ERM and PEL Quotients

In the San Diego and S.Cal. EMAP reports, comparisons of the data to effects-based numerical guidelines (TELs & PELs; ERLs & ERMs) were made to relate sediment pollution to a national scale. Additionally, these guidelines were used to identify individual chemicals of concern for sediment quality management within both regions. Also, a new technique was used for rankings and comparisons using ERM-quotients (ERMQ) and PEL-quotients (PELQ). These were summations of chemical concentrations, divided by their respective ERM or PEL value, for the 30 chemicals for which guidelines have been developed. Cases where levels of measured chemicals were below the analytical method detection limit (MDL), a value of one-half the MDL was used for summations. This is a simple approach for addressing overall chemical pollution where there are multiple pollutants at a station, and is in addition to the standard chemical by chemical approach. Synergistic effects are possible, but not implied by the quotient summations, therefore, this method should be recognized only as a ranking scheme meant to better focus management efforts on interpretation of ambient sediment chemistry data.

Interpretations using ERM and PEL quotients were limited to statistical analysis within this dataset because the approach has not been formally presented in other report. A number of San Diego Bay Region stations were characterized by low levels of pollution so the data set was not normally distributed. A root x+0.5 transformation was applied to achieve a normal distribution. Using the CHI-square test with 90% confidence interval for the 229 stations on which chemical analysis was performed, stations with an ERMQ > 14.6 or a PELQ > 16.3 were found to fall above this confidence interval. Points falling above the 90% confidence interval have a very low probability of being from the same theoretical random distribution as those falling within the interval. Although these values of 14.5 and 16.3 cannot be considered threshold levels with proven ecological significance, they are useful for regional comparative purposes. In the San Diego data set, forty-one stations exhibited ERM or PEL summary quotient levels exceeding the confidence interval cutoffs. Of these forty-one stations, twelve received benthic community analysis, all which were found to have degraded benthic communities. All forty-one stations were tested for *Rhepoxynius* toxicity, of which 29% demonstrated significant toxicity, at the 48% limit established by the reference envelope method. This difference in biological response to pollutants, between benthic community structure and bioassays, is a topic SPARC may wish to

discuss. These differences may be explained by long term exposure to pollutants in the benthic community relative to short term (10 day) pollutant exposure in bioassay tests. Use of the ERM and PEL quotients appear to give a worthwhile representation of overall chemical pollution and were used in both reports for station rankings and characterizations.

**Chemical Association Issues**

1. Is the use of summary quotients an acceptable data analysis technique?

2. Is the 90% confidence interval an appropriate cutoff value?

3. What variables make summary quotients better predict benthic community degradation than amphipod toxicity.

FRESHWATER STUDIES ISSUE PAPER (REGION 5)

## Water Column Toxicity Issues

As part of the Bay Protection Toxic Cleanup Program, Regional Board staff have established a monitoring program in the Delta to determine if Delta waters exceed either the narrative toxicity objective or water quality criteria for metals The focus has been on water column testing instead of sediment testing because previous work has demonstrated that acute toxicity is common in surface water samples collected from the Central Valley Region. Before sediment testing can be addressed we need a solid understanding of water column toxicity issues. In addition to emphasizing water column over sediment issues, we have focused on linking toxicity detections with Toxicity Identification Evaluations (TIEs). Once the chemical responsible for toxicity has been identified we then focus on the chemical, its source, its impact on the Delta and a cleanup strategy.

### Delta Toxicity and Metals Monitoring

Bioassay monitoring has been conducted using the EPA freshwater three species protocols. In the first 18 months of the program (1993-94), the bioassay program was composed of two parts: a multi seasonal, fixed station monitoring effort, and a series of special studies designed to follow-up on high priority incidences of toxicity. During this first year, 24 sampling sites were located in the Delta representing all major riverine inputs (the Sacramento, San Joaquin, and Mokelumne Rivers), sites along the major channels carrying this water across the Delta to the pumps or to San Francisco Bay, a number of back slough sites draining urban and agricultural areas along the Delta's periphery and a number of agricultural drains from representative Delta Islands. In addition to toxicity testing, seven of these sites were monitored for metals. Finally, when toxicity was detected, archived samples were submitted for pesticide analyses (because in previous testing toxicity had always been linked to pesticides). On samples which exhibited acute toxicity to *Ceriodaphnia*, TIEs were performed to identify the specific pesticide responsible for the mortality.

The testing has identified bioassay mortality in the Sacramento River (fathead minnow), in the south Delta sloughs surrounding the City of Stockton during rainfall events (*Ceriodaphnia*), in the Sacramento and San Joaquin Rivers after application of orchard dormant spray insecticides (January and February; *Ceriodaphnia*) and in Delta back sloughs during the early irrigation season (spring and early summer; *Ceriodaphnia*) . In addition, algal toxicity has been detected in the south Delta during the period when fish barriers restrict flow in this

region, in the sloughs surrounding Stockton following rainfall events and in Delta back sloughs. However, in all cases additional information is needed either to better define the hot spot incidents, aid in development of cleanup plans or help prioritize future work. Not all incidents of toxicity have been linked to a chemical, but in the instances where TIE has been conducted, only pesticides have been identified as toxicants. No metal toxicity has been detected. Staff is currently attempting to link specific pesticides to specific agricultural processes.

Metals monitoring at the fixed sites suggests that exceedances of EPA metals criteria are uncommon. Peak metal values occur during storm events when increased flows cause an increase in total suspended sediment. The most significant finding is the high load of mercury entering the Delta from both the Sacramento River and the Yolo Bypass. Detailed studies in 1994-95 and 1995-96 have determined that Cache Creek is a major source of mercury to the Bypass. Preliminary work in the watershed suggests that the loads originate after heavy rains in an unaccessible 20 mile reach of Cache Creek canyon downstream of Clearlake and Indian Valley Reservoir and upstream of Bear Creek. Follow-up studies next rain season should identify the source. In addition Region 5 has funded UC Davis to collect aquatic organisms from the Cache Creek watershed with an emphasis on the area that appears to export large amounts of mercury to ascertain local aquatic bioavailability.

**Issues and Questions**

1. At the upcoming Scientific Planning and Review Committee Meeting, Central Valley Regional Board staff will present an over view of our three year program outlining the approach taken to identify water column "hot spots". Does this approach make sense?

2. Should water column "hot spots" be defined or treated differently than sediment "hot spots"?

3. We have used the toxicity testing approach as a screen for surface water problems. Several pesticides have been identified. These pesticides are additive in their toxicity (all are acetylcholinesterase inhibitors) and are frequently found concurrently or sequentially. Cumulatively, are they impacting the health of the Delta? What studies could be done to answer this question?

4. Are there areas where the SPARC Committee thinks we need more information? That is, should our remaining resources be focused on more toxicity testing, chemistry work or defining the duration, magnitude and frequency of these pesticide pulses?

5. Long lived fish in the Estuary have elevated mercury body burdens. This has resulted in a fish mercury health advisory. It is clear that large amounts of mercury are still entering the Estuary and it seems possible with more work to identify sources and develop detailed mercury load estimates both in Cache Creek and elsewhere upstream on the Sacramento River. Local bioavailability can also be assessed. However it is unclear how to ascertain the bioavailability of the various sources of mercury once in the Estuary. This is important as the State has limited funds and mercury abatement work should focus on those sites which result in the greatest amount of both local and estuarine bioavailable mercury. Does the SPARC have suggestions on how to proceed with ascertaining the degree to which the various mercury sources are bioavailable once in the Estuary?

BIOACCUMULATION ISSUE PAPER

## The Use of Bioaccumulation Monitoring in the Bay Protection and Toxic Cleanup Program

### Background

The Bay Protection Toxic Cleanup Program (BPTCP) legislation mandates in part that the State Water Board develop and adopt Sediment Quality Objectives (SQO) and base these objectives "on human health risk assessment if there is a potential for exposure of humans to pollutants through the food chain to edible fish, shellfish, or wildlife". Human exposures do occur for those chemicals that bioaccumulate and SQO were to be developed for highly bioaccumulative chemicals. Although the development of numeric SQO is on hold, public health is still being protected in the BPTCP by occasional monitoring of sport fish for bioaccumulative chemical residues, by narrative SQO and by identifying and designating toxic hot spots based on their established potential for human health risk. This potential is recognized based on the existence of a fish consumption advisory on a waterbody. Such advisories are based on the analysis of metals and organic compounds in muscle tissue of sport fish from the waterbody. Thus, fish bioaccumulation data are currently used in the BPTCP to support narrative SQO and the identification of hot spots, and could be used for future development of numeric SQO.

### Existing California Bioaccumulation Data

Some potentially relevant bioaccumulation data exists from the California State Mussel Watch (MWP) and Toxic Substances Monitoring Programs (TSMP). Department of Fish and Game (DFG) carries out the sampling and analysis of both programs and reports the result to the State and Regional Water Board(s). Both programs are focused on monitoring known or suspected water impacts not on the overall assessment of statewide water quality. In addition, Region 2 (San Francisco Bay) has developed a regional monitoring program to identify long term trends in water quality in their region. This program also gathers bioaccumulation data from transplanted mussels.

Mussel Watch has been in existence since 1977 and uses transplants of marine mussel species (Mytilus sp.) and also freshwater clams (Corbicula flumiea)to monitor trace elements and organic compounds in state water bodies. Bags of mussels are hung in the water column for 2 to 6 month exposures. Soft body parts are collected and frozen for analysis without depuration. Composites, not individuals, are analyzed. Soft body parts excluding gonads are used for trace metal analysis. And soft body parts including gonads are used for analysis of organic

compounds. Much of the monitoring has occurred in bays and estuaries. Sampling sites are usually determined by Regional Water Board staff with knowledge of local water bodies. Sites are not necessarily sampled on a seasonal, yearly, or repeat schedule in most regions. The regional monitoring program in Region 2 does include sampling twice a year at established stations.

TSMP was initiated in 1976 to monitor trace elements and organic compounds (e.g., pesticides and PCBs) in endemic fish and other aquatic life (e.g., crayfish) in fresh, estuarine and marine waters in California. A variety of fish species have been sampled during the history of the program. Fish of various sizes may be sampled and effort is made to collect the same species at multiple stations. Samples may be composites of whole fish, fish livers, or fillets. Much of the monitoring has occurred in inland lakes, rivers and estuaries. Again Regional Water Board staff are instrumental in determining monitoring sites, but no consistent schedule of repeat sampling has been used throughout the state.

Three regional fish sampling efforts separate from MWP and TSMP have been undertaken in Santa Monica Bay, San Diego Bay and Monterey Bay for the purpose of evaluating the human health risks of eating fish from these areas. Sampling was done by DFG and fillet samples of sport fish species of legal size were analyzed for metals and organics in these studies. As a result of these studies a number of fish consumption advisories have been issued and many are still in force. Occasional studies of sport fish have also been done in freshwater lakes in the state, and consumption advisories are still in force for several sites.

There is no state-wide program to regularly monitor sport fish for chemical residues in tissue. When monitoring data are available the Office of Environmental Health Hazard Assessment (OEHHA) evaluates the health risk of eating fish from the sampled location and issues consumption advisories if the potential risk is excessive. Fish tissue data may be from studies commissioned by industry, city, county, state or federal agencies and programs.

**Bay Protection Toxic Cleanup Program Data**

A pilot study of sport fish contamination in San Francisco Bay was undertaken by the San Francisco Regional Water Quality Control Board in 1994. This study sampled a number of representative sport fish species caught and consumed in the San Francisco Bay area. This study analyzed metals and organic compounds in composite fillet samples taken primarily from legal sized fish sampled near fishing piers or other locations where people fish. As a result of a preliminary analysis of this data an interim fish consumption advisory was issued for the whole of

San Francisco Bay by OEHHA. A comprehensive risk assessment of this data is being conducted by OEHHA.

While analysis of the data show that several chemicals have accumulated to levels of potential health concern in some bay caught fish it has been difficult to determine if certain sites are potentially more contaminated based on the tissue residue data alone. Limited sample size and the effects of confounding biological factors (e.g., lipid level and fish size) have complicated this effort.

**Sample Design Questions**

1. Should probability based sampling be used for fish tissue sampling? What about for transplants of mussels, etc.?

   Site-directed sampling has typically been used for both. Fish are sampled at fishing sites and mussels tend to be placed near suspected pollutant sources.

2. Are separate screening and confirmation sampling recommended for bioaccumulation?

   For human health concerns once you get multiple composites from a single waterbody with elevated tissue concentrations of pollutant an advisory may be issued. This would often occur before or without 'confirmation' sampling.

   'Confirmation' (second samples at a later date) fish samples are often of species not sampled and analyzed in the first sampling event. Different species sampled in the second event are targeted to see if they are as contaminated as those species in the first sampling.

3. Should temporally separated samples be used? required? When they exist how should they be interpreted? Are they meaningful for identifying hot spots over a 1 year time-frame, a 5 year time-frame, 10 years?

   Mussel Watch sample data may be available from 5 or more years ago. This may be useful to choose target locations for sampling but should these data be used to designate hot spots? How should they be interpreted?

   Fish advisories may still be in effect based on sample data from 5 or more years ago. There is no regular repeat sampling of water bodies for which an advisory has been issued, although this type of program is currently being planned for San Francisco Bay. The advisories are in force until data is gathered that shows tissue levels have been

reduced to a level that is no longer a potential health risk.

Fish samples from one sampling event may actually reflect exposure averaged over multiple years.

4. Are physically separated samples recommended? When they exist how should they be interpreted, i.e., at what distance do they represent a single sediment exposure source or multiple exposure sources?

   Transplanted mussels reflect a more discrete exposure.

   Resident fish may reflect exposures from more than one discrete location and prior years.

5. Should whole fish samples be taken in addition to fillet samples?

   Muscle tissue samples are necessary for human health risk assessment interpretations. Skin and fat may be removed. These samples are not very useful for wildlife risk assessment.

6. Are 'replicate' samples of a single species necessary?

   Sampling protocols to collect data for health advisories typically try for multiple composites (2 minimum) of the most abundant fish species or one that is frequently consumed.

   Different sized composites are used for different fish species, e.g., for San Francisco Bay study:

| | | |
|---|---|---|
| shiner surf perch: | | 20 |
| croaker: | | 5 |
| large surf perch (e.g. white) | | 5 |
| shark | 3 | |
| halibut | 3 | |
| sturgeon | | 3 |
| striped bass: | | 3 |

   Always target legal sized fish caught and eaten by sport fishers. Seldom sample fish caught exclusively by wildlife.

7. Should transplanted mussels be depurated before the chemical concentration is determined? Some of chemical level determined prior to depuration may be in gut, not absorbed and accumulated in tissue. A special BPTCP study in San Francisco Bay to evaluate the MWP protocol showed no significant difference between non-depurated and depurated mussels.

Are water filtering mussels (depurated or non-depurated) a useful measure of bioavailability and/or indicative of a potential or candidate hot spot in sediment? Would mollusks that filter sediment be better?

8. Should resident species other than fish (e.g., resident crabs, clams, mussels, etc.) be used for monitoring bioaccumulation? Some of these species are collected recreationally. Are these potentially more useful than transplanted mussels for identifying hot spots? Region 1 (North Coast) has used resident species for monitoring in Humboldt Bay.

9. How should bioaccumulation sampling be incorporated in the overall screening sampling design? In most cases at present sediment samples are collected, toxicity tests are run, and sediment is archived for later chemistry. Bioaccumulation samples (especially fish), like other chemistry analyses, are expensive to run and have additional costs and difficulties associated with sampling. However, there is concern that we may be missing locations where significant bioaccumulation is occurring without toxicity.

10. Are fish studies, which by there nature reflect a larger area, more useful to identify hot spots than mussel studies, which reflect a smaller area?

    Would it be useful to attempt a general screen of fish contamination for major bays and estuaries? This might be used to establish 'background levels' and sampling could be independent of toxicity sampling. Or should the focus be on linking fish sample locations to toxicity sample sites?

11. Can bioaccumulation measures stand alone to determine hot spots or do they need to be linked to other biological indicators (e.g., toxicity, benthic community analysis, or biomarkers)?

**Analysis Questions**

1. Is the reference site concept applicable to fish?

   It is not used in human health risk assessment, although some consideration may be included of 'background' tissue concentrations. Should we be trying to sample background chemical levels for sport fish? Could this information be used to determine hot spots (e.g., defined as locations above background) or clean spots that are of no further concern?

2. Can Mussel Watch data be used as substitute or adjunct to AVS and SEM to show evidence that metals are bioavailable?

**Relationship between Bioaccumulation/Human Health & Aquatic Toxicity**

1. Do the hot spots based on aquatic life (primarily determined by toxicity) and human health criteria (primarily determined by exposure potential) lead to different sorts of hot spots?

   Those defined by aquatic life have the potential to be in a smaller discrete area and short lived.

   Those defined by human health are potentially hard to focus on a discrete area and may reflect longer lived conditions and possibly deeper sediment.

2. Do we need to reconcile these differences? Are the differences useful?

3. Can we associate the two results? How? Via sediment chemistry, biomarkers, stomach content chemistry, chemistry on younger (non-legal size) fish, or etc.? Is this association necessary or merely satisfying?

**Interpretation/designation of Hot Spot Questions**

1. Should mussel transplants in the water column be used for designating sediment hot spots? Can they be incorporated as part of the weight-of-evidence for designating a hot spot?

   Existing Mussel Watch data are often used by Regional Boards to help pick sample sites.

2. Should resident mussels or other invertebrate species be given equal or greater weight compared to transplanted mussels for designating sediment hot spots? Should sediment filtering/ingesting/dwelling species be given greater weight?

3. Should 'migratory' fish be used to designate potential or candidate hot spots?

   Presently migratory fish are used to designate potential hot spots and non-migratory fish to designate candidate hot spots. At present we treat anadromous fish (e.g., striped bass and salmon) as migratory and de-emphasize their use for designating hot spots. But striped bass spend a lot of time in San Francisco Bay and do show the same chemical contaminants as other species in San Francisco Bay.

Fish behavior is hard to characterize. Species may spend different portions of their life history in different habitats and at different trophic levels. The size of the area over which they forage may vary. Still fish in different areas (especially in southern California) have shown different chemical levels.

So what is a workable definition of migratory vs. a non-migratory fish species? How often does a fish have to leave a waterbody to be considered migratory? For how long? At what life stage? Should we make these distinctions and use them for sampling design and program goals?

4. Is it necessary to have sediment chemistry in addition to tissue bioaccumulation data to designate a human health toxic hot spot? Should this be part of the weight-of-evidence?

5. What tissue chemistry values can be used to interpret bioaccumulation data for hot spot designation? FDA action levels? National Academy of Sciences (1973, guidelines for fish-eating birds and mammals)? SWRCB Maximum Tissue Residue Levels from the Inland Surface Waters and Bays and Estuaries Plans? US EPA 'screening values' from Guidance for Assessing Chemical Contaminant Data for use in Fish Advisories?

BENTHIC COMMUNITY ISSUE PAPER

**Characterization of Benthic Community Degradation for San Diego Report (excerpted from the San Diego Report text)**

## Data Analyses and Interpretation

The identification of benthic degraded and undegraded habitat (as determined by macrobenthic community structure) was conducted using a cumulative, weight-of-evidence approach. Tests were employed without prior knowledge or integration of results from laboratory exposures or chemical analyses. Analyses were performed to identify relationships between community structure within and between each station or site. This included diversity/evenness indices, analyses of habitat and species composition, construction of dissimilarity matrices for pattern testing, assessment of indicator species and development of a benthic index, cluster and ordination (multidimensional scaling) analyses. Initially, a triangular correlation matrix was produced from species density data from each site using the Systat® statistical program. From this matrix several tests for association of variables were performed. The tests employed are common in marine and estuarine benthic community analyses and are well-documented in the literature (Field *et al.*, 1982; Pearson *et al* 1983; Swartz *et al.*, 1985; Gray, 1989; Clark and Ainsworth, 1993). Classification analysis was employed to demonstrate site-related community patterns such as species dominance. Cluster analysis is a multivariate procedure for detecting natural groupings in data, and, for our purposes, data were grouped by average similarities in total composition and species abundance (Krebs, 1989). The average-linkage method calculates similarity between a pair of cluster groups as the average similarity among entities in the two groups. Species information is used to compute similarity index values. Grouped stations were clustered at a conservative distance limit of 50-60% similarity, however, this level is purely arbitrary. Because classification analyses have the tendency to force data into artificially distinct groups, another method (*e.g.*, multi-dimensional scaling) was used to confirm the validity of group clusters and site similarity. Ordination analyses are useful because it enables one to see multidimensional gradients in data rather than just groupings (Smith, personal communication).

Multi-dimensional scaling (MDS) is used extensively in the analyses of benthic communities, particularly in estuarine and marine pollution studies. MDS is a procedure for fitting a set of points in space such that the distance between points correspond to a given set of dissimilarities. This technique is more flexible than principal co-ordinate analyses when handling the

large number of zero counts generally characteristic of species-samples matrices. Nonmetric MDS analyses were performed using Systat®. For a detailed account of MDS statistical procedures, see Clarke and Ainsworth (1993) and Warwick and Clarke (1993). Inferences from the resultant ordination are also presented. It is important to note that, as with cluster analyses, MDS results are not definitive and must be used in conjunction with additional ecological information. MDS results are based on total species number and numbers of individuals. Inferences from the resultant ordination are also presented.

After classification and ordination patterns were determined, the raw data were reevaluated to assess which species may have influenced the observed patterns. Indicator species were then selected on the basis of a literature review (*i.e.*, distribution, life history strategies and habitat preference), by recommendations from other experienced benthic taxonomists, and review of the raw data. Initially, community analyses were conducted as a per "site" comparison. Later, it was decided analyses also be expanded to a per "station" comparison to produce a more definitive data set for the reference pool. The extended analysis of station variability was performed using the benthic index.

Benthic assemblages have many attributes which make them reliable and sensitive indicators of the ecological condition in estuarine environments. The following procedure summarizes the construction and application of the benthic index used to reliably discriminate between degraded and undegraded conditions at sites in the San Diego Bay Region. Although there are problems with trying to simplify complex biological communities, we attempted to develop a quantitative method which creates a partition between degraded and undegraded areas. Polluted sites can not be conclusively identified using results from benthic community analyses alone, but these analyses impartially describe "environmentally stressed" areas. This benthic index is based on species (indicators), and group (general taxa) information. The index also evaluates community parameters such as species richness, abundance and presence of pollution indicators, which identify the extremes of the community characteristics. Sites are ranked according to these extremes and are represented by a single value. In general, decreasing numbers of species, increasing numbers of individuals, and decreasing diversity values are common responses observed near polluted areas. These trends are incorporated into the index. One of the important restrictions with the existing method is it evaluates this limited San Diego Bay benthic data set when dividing groups for categorization. Construction and subsequent validation of this simplified benthic index are loosely based on criteria developed by several agencies, including USEPA-EMAP and SCCWRP. However, the benthic index developed by USEPA-EMAP (Weisberg *et al.*, 1993) included several environmental variables in its construction

(*e.g.* dissolved $O_2$), while the index for San Diego Bay data used only biological parameters. Briefly, the following major steps are followed in constructing and validating this benthic index:

1. Degraded and undegraded (*i.e.*, reference condition) stations are identified on the basis of measured environmental and biological variables.

2. A list of "candidate" parameters is developed using species abundance data. The list included metrics having ecological relevance (*e.g.*, species diversity indices, etc.) which potentially may be used to discriminate between degraded and reference areas.

3. A value for each candidate parameter (*i.e.*, diversity, abundance, taxonomic composition) is calculated for each station (*e.g.*, total species per station, total individuals per station, total crustaceans species per station, total number of polychaete individuals, total amphipods per station, etc.).

4. Range of values per metric is determined (lowest to highest value).

5. Quartiles from that range are determined.

6. Ranking within quartiles are assigned: upper range quartile=2, lower range quartile=0, middle quartile=1. Apply these calculations on the metrics from step 3.

7. The index is defined by values of 0, 1, or 2. A value of 0 defines the degraded (detectable stress) stations(s), and 2 identifies environmentally undegraded stations(s). Stations with an index value of 1 are considered transitional communities, which are neither degraded nor reference stations. Transitional stations have species or other parameters which indicate both degraded and undegraded habitats. These stations are investigated further to determine the cause of ambiguity of the transitional status.

8. Relative abundance of indicator species (both degraded and undegraded habitat indicators) per station is assessed.

A primary concern regarding the benthic index is how well it fulfills the objective of discriminating among degraded and

undegraded estuarine conditions. This simplified version forms the basis for ongoing iterative procedures involved in construction of an index. This index will include a variety of indicator values (Bascom *et al.*, 1978; Kerans *et al.*, 1994; EcoAnalysis *et al.*, 1995) for future applications of the assessment of benthic community structure. The following sections report results of benthic community analyses based solely on composition and abundance of macrobenthic species from sediment cores throughout San Diego Bay and its vicinity. Environmental parameters (*e.g.*, total organic carbon levels and sediment grain size range) and other factors capable of influencing benthic composition were examined, but not evaluated in conjunction with the data presented here. Those data are examined later in sections which address correlative analyses.

In this study, bioeffects are required to be demonstrated in relation to properly selected reference sites and to occur in association with significant pollutant levels. The following evidence for undegraded (possible reference) and degraded (possible contaminated) sites was based on benthic community "quality" at each site and station. Benthic community structure was evaluated as an indicator of environmentally degraded or undegraded areas and not as a pollution or contamination indicator. Benthic reference sites were determined predominantly by analyses of specific indicator species and groups (*e.g.*, amphipods). These species are generally not found in polluted or disturbed areas.

It is our intention in this section to clearly describe the condition of macrobenthic communities from sampling areas. Definitions of degraded, transitional, and undegraded used in this section are adopted from several papers (Bascom *et al.*, 1978; Pearson and Rosenberg, 1978; Schindler, 1987; Swartz *et al.*, 1985; Underwood and Peterson, 1988). Although the boundaries set in Bascom *et al.* (1978) were based on food supply and not on toxicants, the same general principles apply to this study. In benthic analyses, the term "degraded" does not refer to a community response to significant levels of toxic chemicals. Degraded areas are those which contain significant numbers of opportunistic species, in the absence of non-opportunistic species, and have relatively low species diversity. Correlations are later used to determine if community profiles are influenced by chemistry or by natural environmental disturbances. Sites and stations which are categorized as "undegraded" have high species diversity, high proportional abundance of amphipods and other crustaceans, while noting there are a few exceptions to this rule (*e.g.*, *Grandidierella japonica*, etc.). Undegraded areas generally contain species which are known to be sensitive to pollutants. Transitional sites and stations are those which are not confidently partitioned into the other two categories. These areas may solicit further study. Overall, an integration of data from laboratory exposures, chemical analyses, and benthic

community assessments provide strong complementary evidence of the degree of pollution-induced degradation in aquatic communities. The following data analyses were conducted on a per site basis using sample replicates (n=5) at each sampling location. (Table 6). An analysis also was performed using per station data (n=1) and is presented later in this section. Tests included classification and ordination analyses, diversity measurements, construction of a benthic index, and assessment of indicator species.

## Characterization of Benthic Community Degradation for Southern California Coastal Lagoons Project Report (Summarized from internal draft report text)

The methodology described in the previous section (above) which was employed for the San Diego Report for classifying benthic community degradation was refined for use in classifying benthic communities at sites in the Southern California Coastal Lagoons Project. A brief summary of this revised methodology for the Southern California Coastal Lagoons Project is described below.

### Benthic Index

The benthic index used in this study is a refined version of the index used in the San Diego BPTCP Report. It combines the use of benthic community data with the presence of positive or negative indicator species to give a measure of the relative degree of degradation of the benthic fauna. It does not require the presence of uncontaminated reference stations, and does not refer to data beyond that collected in this study. Other benthic indices often rely on apriori assumptions, particularly the presence of uncontaminated reference sites, which can lead to false results if the assumptions are not met.

### Community Data

Two aspects of the community data were used in the benthic index: the total number of species, and the number of crustacean species. An increase in species richness is one of the most long-standing indicators of healthy environments. While a variety of indices have been developed to quantify species richness in absolute terms, for a study limited in spatial scale, as was this one, total number of species is as valid as any.

Crustaceans are generally more sensitive to environmental contaminants than most other components of the infauna, particularly polychaetes and bivalves. Species and numerically abundant crustacean faunas on the Pacific coast of the U.S. are generally only found in uncontaminated environments, making the number of crustaceans species an important indicator of overall environmental health.

## Indicator species

Eleven of the 168 total species were chosen as indicators species. The bioindicators were chosen based on a review of pertinent literature, known habitat preferences and life history, their abundance over all of the stations, and on discussions with experienced ecologists. The 3 negative indicator species are highly opportunistic annelids which thrive in disturbed, polluted, or marginal environments, and are generally not found in mature, undisturbed communities. The 8 positive indicator species consist of 1 polychaete, 2 bivalves, and 5 crustaceans, and are generally not found in polluted habitats.

## Calculation of the Benthic Index

Based on the previous work, it was determined that three levels of index classification would give sufficient resolution to detect possible impacted areas, while being robust enough to reduce false positives. Accordingly, for Total Fauna and Number of Crustacean Species, the total range for the 43 stations were determined. After outliers were removed, the ranges of each were divided into thirds. Those with the lower third were ranked as "1", in the middle third as "2", and in the upper third as "3". For example, the range of crustacean species was 0-15. Station 95004 had 6 crustacean species, so was given a crustacean index of "2". The Total Fauna and Crustacean values were calculated for each station. These two numbers represent two-thirds of the Benthic Index for each station.

The Indicator indices were based mostly on presence or absence, with abundance of negative species given additional weight. Stations were given a negative Indicator Index of "1" if they contained at least two of the 3 negative species, and had at least one species in the middle third of the range. Stations were given a Positive Indicator Index of "3" if at least 3 of the 8 positive species were present. Stations not ranked either "1" or "3" were ranked "2". There were no stations with an overlap of the positive and negative indicators indices.

To determine the overall benthic index, the Total Fauna, Crustacean Species, and Indicator Species indices were averaged. This resulted in a range of 1 (most impacted) to 3 (cleanest) with 5 gradations between.

## Other Benthic Community Issues

- Benthic community composition/summary parameters at a location can be well characterized with many fewer replicates than it takes to level out a species-area curve. What about use of optimization and power analyses (based on variance components estimated from the available data)? Also, what about utilization of number

of species per grab in order to reduce the number of replicates necessary for statistical tests? Bob Smith comments from SPARC 1995.

**References**


Bascom, W., A.J. Mearns and J.Q. Word. 1978. Establishing boundaries between normal, changed, and degraded areas. Southern California Coastal Water Research Project, annual report, p. 81-94.

Clarke, K.R. and M. Ainsworth. 1993. A method of linking multivariate community structure to environmental variables. Mar. Ecol. Prog. Ser. 92: 205-219.

Ecoanalysis Inc. and Southern California Coastal Water Research Project. 1995. Southern California Regional Monitoring Project: Numerical tools for assessing benthic monitoring data. U.S.EPA Region IX, NPDES Program- Regional Marine Monitoring Project, draft report.

Field, J.G., K.R. Clarke and R.M. Warwick. 1982. A practical strategy for analyzing multispecies distribution patterns. Mar. Ecol. Prog. Ser. 8: 37-52.

Gray, J.S. 1989. Effects of environmental stress on species rich assemblages. Biol. J. Linn. Soc. 37: 19-32.

Kerans, B.L. and J.R. Karr. 1994. A benthic index of biotic integrity (B-IBID) for rivers of the Tennessee Valley. Ecological Applications 4(4): 768-785.

Krebs, C.J. 1989. Ecological Methodology. Harper & Row, publ., NY.

Pearson, T.H., J.S. Gray and P.J. Johannessen. 1983. Objective selection of sensitive species indicative of pollution-induced change in benthic communities. 2. Data analyses. Mar. Ecol. Prog. Ser. 12: 237-255.

Pearson, T.H. and R. Rosenberg. 1978. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. Oceanogr. Mar. Biol. Annu. Rev. 16: 229-311.

Schindler, D.W. 1987. Detecting ecosystem responses to anthropogenic stress. Can. J. Fish. Aqua. Sci. 44: 6-25.

Swartz, R.C., D.W. Schultz, G.R. Ditsworth, W.A. DeBen and F.A. Cole. 1985. Sediment toxicity, contamination, and macrobenthic communities near a large sewage outfall. ASTM STP 865. T.P. Boyle, ed., American Society for Testing and Materials, Phil. pp. 152-175.

Underwood, A.J. and C.H. Peterson. 1988. Towards an ecological framework for investigating pollution. Mar. Ecol. Prog. Ser. 46: 227-234.

Warwick, R.M. and K.R. Clarke. 1993. Comparing the severity of disturbance: a meta-analysis of marine macrobenthic community data. Mar. Ecol. Prog. Ser. 92: 221-231.

Weisberg, S.B., J.B. Frithsen, A.F. Holland, J.F. Scott, J.K.Summers, H.T. Wilson, R.M. Valente, D.G. Heimbuch, J. Gerritsen, S.C. Schimmel, and R.W. Latimer. 1993. EMAP-Estuaries Virginian Province 1990 Demonstration Project Report. EPA 600/R-92/100. U.S. Environmental Protection Agency, Environmental Research Laboratory, Narragansett, RI.

WEIGHT-OF-EVIDENCE ISSUE PAPER

## Application of Weight of Evidence Approach for San Diego Bay

One of the primary goals of the BPTCP is to establish state guidelines under which contaminated or toxic stations can be designated "toxic hot spots". These guidelines are currently being developed based on data collected throughout the State. Although final guidelines are contingent upon further data analysis, the "toxic hot spot" definition currently utilized by the BPTCP, requires that one or more of the following criteria must be met:

1. The water or sediment exhibits toxicity associated with toxic pollutants, based on toxicity tests acceptable to the SWRCB or the RWQCB. To determine whether toxicity exists, recurrent measurements (at least two separate sampling dates) should demonstrate an effect.

2. Significant degradation in biological populations and/or benthic communities associated with the presence of elevated levels of toxic pollutants.

3. The site exceeds water or sediment quality objectives for toxic pollutants which are contained in appropriate water quality control plans, or exceeds water quality criteria promulgated by the U.S. Environmental Protection Agency.

4. The tissue toxic pollutant levels of organisms collected from the site exceed levels established by the United States Food and Drug Administration (FDA) for protection of human health, or the National Academy of Sciences (NAS) for the protection of human health or wildlife.

Because tissue residues were not analyzed in San Diego Bay (and most BPTCP data sets), criteria are generally limited to the first three. Satisfying any one of these criteria can designate a site a "toxic hot spot". Satisfying more than one criterion and the severity demonstrated within each criterion determines the weighting for which qualitative rankings can be made. In the San Diego report, stations were not be designated as "toxic hot spots", because this designation is still under evaluation and development by the BPTCP. Instead, stations were be prioritized for further evaluation for hot spot status. This priority was be classified as high, moderate, low, or no action and is to be used by State and Regional Water Board staff to direct further investigations at these stations. Each station receiving a high to low priority ranking meets one or more of the first three criteria established above. Those meeting all three criteria were designated as the highest priority for further action.

San Diego stations were evaluated for repeat toxicity (criterion 1) using the reference envelope method, the most conservative measure developed. Only those stations which demonstrated amphipod survival less than 48% in repeated tests, without confounding ammonia, hydrogen sulfide or grain size effects, were considered to exhibit repeat toxicity hits. Because only one critical value could be determined for any of the dilutions of the pore water bioassays, pore water toxicity results were not evaluated for repeat toxicity when prioritizing stations.

Stations with repeat toxicity and elevated chemistry and/or degraded benthic communities, were assigned moderate or high priority. Stations with repeat toxicity, but lacking elevated chemistry or degraded benthic communities, were assigned low priority (Table 6 - Repeat Toxicity Hits).

Stations with only a single toxicity hit were also considered a moderate or high priority, when associated with elevated chemistry and/or degraded benthic communities. Stations with a single toxicity hit, but lacking elevated chemistry or degraded benthic communities, were assigned a low priority (Table 6 - Single Toxicity Hits).

Twenty-two stations demonstrated repeat or single toxicity hits but were given a "no action" recommendation at this time (Table 6). These stations had measured grain size, hydrogen sulfide or ammonia concentrations which confound interpretation of the bioassay test results. Chemistry levels were low, or not analyzed, and the benthic community was undegraded or transitional, where sampled. These results provided little or no evidence that these stations should be prioritized for hot spot status.

Stations were evaluated for benthic community condition using the benthic index. Stations determined to be degraded, with elevated chemistry and/or toxicity, were assigned a moderate or high priority. Stations determined to be degraded, but which did not demonstrate elevated chemistry or toxicity, were assigned a low priority. Transitional and undegraded stations were not considered a priority unless chemical or toxicity results initially prioritized the stations. (Table 7 - Degraded Benthics)

Stations were evaluated for elevated chemistry (criterion 3) using an ERM quotient >14.6 or a PEL quotient >16.3. It was determined these values are statistically above the 90% confidence interval of summary quotients from all San Diego stations analyzed. These quotients were used to identify stations where multiple pollutants were near or above established ERM and PEL guidelines (Table 7-Chemistry-Summary Quotients). 100% of the stations analyzed for benthics were found to be degraded when chemical analysis demonstrated a summary ERM quotient above 14.6. Although the 21 stations in Table 7

# TABLE 6
## FUTURE INVESTIGATION PRIORITY LIST FOR SAN DIEGO BAY REGION

| Station # | Station | IDORG | Leg | Fines | Ammoni | Rhepox. Survival | >4x ERM or >5.9x PEL | ERMQ | PELQ | Benthics | Comments | PRIORITY |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **REPEAT TOXICITY HITS** | | | | | | | | | | | |
| 90009.0 | 28 SWARTZ | 158 | 7 | 64.00 | 0.002 | 0.00 | Chlordane, DDT | 26.77 | 29.37 | not analyzed | ELEVATED CHEM | HIGH |
| 90009.0 | 28 SWARTZ (7TH ST CHANNEL Q1) | 893 | 23 | 23.84 | 0.016 | 5.00 | Chlordane | 12.56 | 15.64 | DEGRADED | TRIAD HIT | HIGH |
| 93228.0 | SEVENTH ST CHANNEL Q1 (x6) | 895 | 23 | 60.67 | 0.010 | 2.00 | Chlordane | 40.15 | 48.55 | DEGRADED | TRIAD HIT | HIGH |
| 93179.0 | NAVAL SHIPYARDS O3 (x1) | 797 | 19 | 79.01 | 0.539 | 20.00 | | 17.49 | 23.47 | not analyzed | ELEVATED CHEM | HIGH |
| 93179.0 | NAVAL SHIPYARDS O3 (x1)-REP 1 | 1122 | 27 | 79.81 | 0.059 | 44.00 | | 18.59 | 23.45 | not analyzed | ELEVATED CHEM | HIGH |
| 90043.0 | CORONADO WHARF | 192 | 12 | 36.00 | 0.684 | 29.00 | | 2.94 | 3.72 | not analyzed | NH3 > .6 | NO ACTION |
| 90043.0 | CORONADO WHARF-REP 1 | 1156 | 28 | 20.77 | 0.423 | 33.00 | | 2.02 | 2.82 | not analyzed | | LOW |
| 90043.0 | CORONADO WHARF-REP 2 | 1157 | 28 | 77.38 | 0.224 | 43.00 | | 11.95 | 15.56 | not analyzed | MODERATE CHEM | LOW |
| 90030.0 | BF SCHROEDER SITE F | 179 | 12 | 94.00 | 0.066 | 47.00 | PAHs | 17.61 | 26.93 | not analyzed | FINES > 90%, ELEVATED CHEM | MODERATE |
| 90030.0 | BF SCHROEDER SITE F | 749 | 16 | 92.00 | 0.204 | 43.00 | | not analyzed | not analyzed | not analyzed | FINES > 90% | NO ACTION |
| 93122.0 | SOUTH SHORE-CORONADO DD3 (x1) | 725 | 16 | 89.00 | 0.740 | 23.00 | | 7.22 | 9.69 | not analyzed | NH3 > 0.6, MODERATE CHEM | LOW |
| 93122.0 | S.S.- CORONADO DD3 (x1) REP 1 | 1013 | 24 | 65.52 | 0.463 | 33.00 | | 5.33 | 7.34 | not analyzed | | LOW |
| 90036.0 | STORM DRAIN- ROHR CHANNEL | 185 | 5 | 64.00 | 0.894 | 27.00 | | 2.99 | 4.02 | not analyzed | NH3 > 0.6 | NO ACTION |
| 90036.0 | STORMDRAIN EA (ROHR CH.) REP 3 | 1024 | 24 | 25.65 | 0.119 | 1.00 | | 1.74 | 2.27 | not analyzed | | LOW |
| 90036.0 | STORMDRAIN EA (ROHR CH.) REP 1 | 1022 | 24 | 24.00 | 0.136 | 0.00 | | 2.46 | 2.80 | not analyzed | | LOW |
| 93125.0 | SILVER STRAND FF4 (x4) REP 1 | 1016 | 24 | 22.68 | 0.514 | 38.00 | | 2.21 | 3.05 | not analyzed | | LOW |
| 93125.0 | SILVER STRAND FF4 (x4) REP 2 | 1017 | 24 | 15.44 | 0.720 | 22.00 | | 2.00 | 2.76 | not analyzed | NH3 > 0.6 | NO ACTION |
| 93125.0 | SILVER STRAND FF4 (x4) REP 3 | 1018 | 24 | 19.05 | 0.484 | 22.00 | | 2.25 | 3.22 | not analyzed | H2S HIGH | NO ACTION |
| 93158.0 | SOUTH BAY GG1 (x1) REP 1 | 1035 | 24 | 53.67 | 0.043 | 33.00 | | 2.98 | 4.18 | not analyzed | | LOW |
| 93158.0 | SOUTH BAY GG1 (x1) REP 2 | 1036 | 24 | 62.76 | 0.108 | 39.00 | | 3.18 | 4.36 | not analyzed | | LOW |
| 93158.0 | SOUTH BAY GG1 (x1) REP 3 | 1037 | 24 | 51.74 | 0.072 | 46.00 | | 2.61 | 3.66 | not analyzed | | LOW |
| 90024.0 | SDNI-N1 | 173 | 7 | 69.00 | 0.684 | 40.00 | | 5.60 | 7.86 | not analyzed | NH3 > 0.6 | NO ACTION |
| 90025.0 | SDNI-N5 | 174 | 7 | 73.00 | 0.925 | 7.00 | | 5.15 | 7.60 | not analyzed | NH3 > 0.6 | NO ACTION |
| 93188.0 | CARRIER BASE V1 (x2) | 806 | 19 | 40.85 | 2.593 | 37.00 | | 3.87 | 5.42 | not analyzed | NH3 > 0.6 | NO ACTION |
| 90025.0 | SDNI-N5 (CARRIER BASE V2) | 899 | 23 | 75.96 | 0.643 | 37.00 | | 5.23 | 7.20 | UNDEGRADED | NH3 > 0.6 | NO ACTION |
| 93232.0 | CARRIER BASE V2 (x7) | 1001 | 23 | 63.79 | 0.773 | 35.00 | | 5.22 | 7.46 | UNDEGRADED | NH3 > 0.6 | NO ACTION |
| 90057.0 | 5 SDG&E | 206 | 12 | 98.00 | 0.011 | 25.00 | | 2.72 | 3.98 | not analyzed | FINES > 90% | NO ACTION |
| 90057.0 | 5 SDG&E REP 1 | 1019 | 24 | 98.45 | 0.046 | 41.00 | | 3.71 | 5.18 | not analyzed | FINES > 90% | NO ACTION |
| 90057.0 | 5 SDG&E REP 2 | 1020 | 24 | 97.80 | 0.011 | 39.00 | | 3.37 | 4.74 | not analyzed | FINES > 90% | NO ACTION |
| 90057.0 | 5 SDG&E REP 3 | 1021 | 24 | 97.22 | 0.032 | 31.00 | | 3.31 | 4.68 | not analyzed | FINES > 90% | NO ACTION |
| | | | | | | | | | | | | |
| | **SINGLE TOXICITY HITS** | | | | | | | | | | | |
| 90002.0 | 12 SWARTZ(DOWNTOWN ANCH)-REP 1 | 878 | 22 | 48.25 | 1.836 | 15.00 | Chlordane | 30.74 | 38.46 | DEGRADED | TRIAD HIT | HIGH |
| 90007.0 | 25 SWARTZ | 156 | 7 | 67.00 | 0.004 | 37.00 | Mercury | 13.57 | 16.62 | not analyzed | ELEVATED CHEM, SITE DEGRADED IN LEG 23 | MODERATE |
| 90008.0 | 27 SWARTZ | 157 | 7 | 66.00 | 0.010 | 29.00 | | 5.73 | 8.55 | not analyzed | SITE DEGRADED IN LEG 23 | MODERATE |
| 90022.0 | P SWARTZ | 171 | 7 | 87.00 | 0.008 | 38.00 | | 12.88 | 18.52 | not analyzed | ELEVATED CHEM, SITE DEGRADED IN LEG 22 | MODERATE |
| 93181.0 | NAVAL SHIPYARDS O6 (x1) | 799 | 19 | 89.12 | 0.042 | 45.00 | | 15.51 | 21.44 | not analyzed | ELEVATED CHEM | MODERATE |
| 93210.0 | NAVAL BASE/SHIPYARDS O4 (x1) | 863 | 22 | 48.75 | 0.775 | 37.00 | | 16.15 | 17.76 | DEGRADED | NH3 > 0.6, ELEVATED CHEM | MODERATE |
| 90010.0 | 31 SWARTZ | 159 | 6 | 85.00 | 1.291 | 39.00 | | not analyzed | not analyzed | not analyzed | NH3 > 0.6, SITE DEGRADED IN LEG 23 | LOW |
| 90039.0 | CL | 188 | 12 | 24.00 | 0.090 | 38.00 | Chlordane, DDT | 13.86 | 17.71 | not analyzed | ELEVATED CHEM | MODERATE |
| 93178.0 | NAVAL SHIPYARDS O2 (x1) | 796 | 19 | 55.32 | 0.350 | 20.00 | | 10.85 | 14.40 | not analyzed | | LOW |
| 93166.0 | NAVY ESTUARY G2 (x1) | 779 | 18 | 23.62 | 1.129 | 20.00 | | 5.81 | 7.79 | not analyzed | NH3 > 0.6 | NO ACTION |
| 93118.0 | TIJUANA R. ESTUARY HH1 (x2) | 713 | 15 | 60.00 | 0.187 | 30.00 | DDE | 5.80 | 6.66 | not analyzed | ELEVATED CHEM | MODERATE |
| 90018.0 | D DE LAPPE | 748 | 16 | 42.00 | 0.039 | 19.00 | | not analyzed | not analyzed | not analyzed | | LOW |
| 90023.0 | NM SANDBAG | 172 | 7 | 27.00 | 0.378 | 32.00 | | 3.12 | 4.55 | not analyzed | | LOW |
| 90050.0 | 10 SWARTZ | 199 | 7 | 81.00 | 0.004 | 47.00 | | 4.19 | 6.41 | not analyzed | | LOW |
| 90051.0 | 16 SWARTZ (INTERCONT. MARINA) | 818 | 20 | 41.33 | 3.340 | 1.00 | | 3.49 | 4.80 | TRANSITIONAL | NH3 > 0.6, SITE TRANSITIONAL IN LEG 20 | LOW |
| 90055.0 | 43 SWARTZ | 204 | 7 | 64.00 | 0.075 | 37.00 | | 4.12 | 5.71 | not analyzed | | LOW |
| 90102.0 | HARBOR BRIDGE 71A | 256 | 7 | 75.00 | 0.113 | 14.00 | | 2.69 | 3.77 | not analyzed | | LOW |
| 90104.0 | WEST BASIN ENTRANCE (71C) REF | 275 | 12 | 74.00 | 1.046 | 13.00 | | 3.43 | 4.88 | not analyzed | NH3 > 0.6 | NO ACTION |
| 93106.0 | MISSION BAY A2 (x1)-REP 2 | 1102 | 27 | 94.52 | 0.106 | 25.00 | | 3.49 | 4.43 | not analyzed | FINES > 90% | NO ACTION |
| 93117.0 | SAN DIEGO RIVER B2 (x2) | 1029 | 24 | 92.05 | 0.110 | 0.00 | | not analyzed | not analyzed | not analyzed | FINES > 90% | NO ACTION |
| 93119.0 | TIJUANA R. ESTUARY HH1 (x1) | 714 | 15 | 84.00 | 0.224 | 22.00 | DDE, DDT | not analyzed | not analyzed | not analyzed | ELEVATED CHEM | MODERATE |
| 93127.0 | SOUTH BAY GG2 (x1) | 1028 | 24 | 43.93 | 0.096 | 47.00 | | not analyzed | not analyzed | not analyzed | | LOW |
| 93128.0 | SOUTH BAY GG5 (x1) | 1033 | 24 | 96.80 | 0.031 | 27.00 | | not analyzed | not analyzed | not analyzed | FINES > 90% | NO ACTION |
| 93132.0 | CORONADO CAYS T3 (x1) | 1025 | 24 | 90.97 | 0.004 | 47.00 | | not analyzed | not analyzed | not analyzed | FINES > 90% | NO ACTION |
| 93138.0 | SHELTER ISLAND E3 (x2) | 741 | 16 | 60.00 | 0.020 | 29.00 | | 3.29 | 4.38 | not analyzed | | LOW |
| 93148.0 | CHANNEL-CORONADO Y1 (x2) | 751 | 16 | 23.00 | 0.525 | 47.00 | | 2.71 | 3.47 | not analyzed | | LOW |
| 93154.0 | NORTH SHORE-MOUTH CC4 (x1) | 763 | 17 | 32.94 | 0.836 | 31.00 | | not analyzed | not analyzed | not analyzed | NH3 > 0.6 | NO ACTION |
| 93159.0 | SOUTH BAY GG3 (x1) | 768 | 17 | 58.87 | 0.675 | 21.00 | | not analyzed | not analyzed | not analyzed | NH3 > 0.6 | NO ACTION |
| 93174.0 | TIJUANA R. ESTUARY HH3 (x2) | 787 | 18 | 70.63 | 0.282 | 6.00 | | 4.60 | 5.58 | not analyzed | | LOW |
| 93175.0 | TIJUANA R. ESTUARY HH3 (x3) | 788 | 18 | 92.67 | 0.141 | 10.00 | DDE, DDT | not analyzed | not analyzed | not analyzed | FINES > 90%, ELEVATED CHEM | MODERATE |
| 93219.0 | SWEETWATER CH. JJ1 (x1)-REP 2 | 876 | 22 | 60.74 | 0.319 | 31.00 | | 2.25 | 3.06 | TRANSITIONAL | | LOW |

# TABLE 7
## FUTURE INVESTIGATION PRIORITY LIST FOR SAN DIEGO BAY REGION

| Station # | Station | IDORG | Leg | Fines | Ammoni | Rhepox. Survival | >4x ERM or >5.9x PEL | ERMQ | PELQ | Benthics | Comments | PRIORITY |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **DEGRADED BENTHICS** | | | | | | | | | | | |
| 90007.0 | 25 SWARTZ (NAVAL BASE/SY O10) | 887 | 23 | 81.62 | 0.014 | 86.00 | | 11.74 | 15.74 | DEGRADED | | LOW |
| 93223.0 | NAVAL BASE/SHIPYARD O10 (x2) | 888 | 23 | 85.99 | 0.016 | 79.00 | | 14.27 | 20.25 | DEGRADED | ELEVATED CHEM | MODERATE |
| 93224.0 | NAVAL BASE/SHIPYARD O10(x6) | 889 | 23 | 48.07 | 0.010 | 90.00 | Zinc | 10.56 | 15.36 | DEGRADED | ELEVATED CHEM | MODERATE |
| 93211.0 | NAVAL BASE/SHIPYARDS O4 (x2) | 864 | 22 | 70.59 | 0.158 | 86.00 | ntimony, Copper, PC | 24.89 | 29.83 | DEGRADED | ELEVATED CHEM | MODERATE |
| 90021.0 | K SWARTZ (NAVAL BASE O4) | 862 | 22 | 69.03 | 0.060 | 93.00 | | 10.55 | 14.81 | DEGRADED | | LOW |
| 90006.0 | 23 SWARTZ (NAVAL BASE O7) | 865 | 22 | 63.34 | 0.054 | 92.00 | Chlordane | 18.15 | 23.62 | DEGRADED | ELEVATED CHEM | MODERATE |
| 93212.0 | NAVAL BASE/SHIPYARDS O7 (x1) | 866 | 22 | 32.88 | 0.026 | 91.00 | Chlordane | 10.29 | 13.52 | DEGRADED | ELEVATED CHEM | MODERATE |
| 93213.0 | NAVAL BASE/SHIPYARDS O7 (x4) | 867 | 22 | 69.06 | 0.010 | 94.00 | Chlordane | 21.00 | 27.21 | DEGRADED | ELEVATED CHEM | MODERATE |
| 93227.0 | SEVENTH ST CHANNEL O1 (x5) | 894 | 23 | 53.40 | 0.076 | 79.00 | Chlordane | 14.49 | 18.73 | DEGRADED | ELEVATED CHEM | MODERATE |
| 93206.0 | DOWNTOWN PIERS K1 (x11) | 848 | 21 | 56.03 | 0.048 | 95.00 | PAHs | 17.08 | 29.59 | DEGRADED | ELEVATED CHEM | MODERATE |
| 90004.0 | 15 SWARTZ (G ST. PIER MARINA) | 849 | 21 | 67.23 | 0.220 | 77.00 | | 8.37 | 11.32 | DEGRADED | | LOW |
| 93207.0 | G ST. PIER MARINA L1 (x4) | 850 | 21 | 79.29 | 0.173 | 89.00 | | 7.91 | 10.65 | DEGRADED | | LOW |
| 90022.0 | P SWARTZ (NAVAL BASE O12) | 868 | 22 | 88.09 | 0.061 | 91.00 | PAHs | 16.64 | 23.33 | DEGRADED | ELEVATED CHEM | MODERATE |
| 93214.0 | NAVAL BASE/SHIPYARDS O12 (x3) | 869 | 22 | 56.64 | 0.031 | 93.00 | | 7.88 | 10.92 | DEGRADED | | LOW |
| 93215.0 | NAVAL BASE/SHIPYARDS O12 (x4) | 870 | 22 | 64.17 | 0.017 | 88.00 | | 6.20 | 8.92 | DEGRADED | | LOW |
| 90008.0 | 27 SWARTZ (NAVAL BASE/SH O13) | 890 | 23 | 59.15 | 0.008 | 92.00 | | 7.23 | 10.36 | DEGRADED | | LOW |
| 93225.0 | NAVAL BASE/SHIPYARD O13 (x1) | 891 | 23 | 74.94 | 0.013 | 81.00 | | 12.03 | 17.33 | DEGRADED | ELEVATED CHEM | MODERATE |
| 93226.0 | NAVAL BASE/SHIPYARD O13 (x3) | 892 | 23 | 79.38 | 0.019 | 91.00 | | 10.82 | 15.91 | DEGRADED | | LOW |
| 90010.0 | 31 SWARTZ (MARINE TERMINAL R3) | 896 | 23 | 38.75 | 0.077 | 86.00 | | 2.78 | 4.11 | DEGRADED | | LOW |
| 93229.0 | MARINE TERMINAL R3 (x1) | 897 | 23 | 69.13 | 0.109 | 70.00 | PAHs | 14.55 | 22.94 | DEGRADED | ELEVATED CHEM | MODERATE |
| 93230.0 | MARINE TERMINAL R3 (x3) | 898 | 23 | 76.64 | 0.056 | 63.00 | | 7.77 | 11.51 | DEGRADED | | LOW |
| 93116.0 | SAN DIEGO RIVER B1 (x4)-REP 1 | 881 | 22 | 44.01 | 0.216 | 92.00 | | 4.87 | 5.90 | DEGRADED | | LOW |
| 93116.0 | SAN DIEGO RIVER B1 (x4)-REP 2 | 882 | 22 | 92.30 | 0.098 | 92.00 | Chlordane | 9.12 | 11.87 | DEGRADED | ELEVATED CHEM | MODERATE |
| 93116.0 | SAN DIEGO RIVER B1 (x4)-REP 3 | 883 | 22 | 92.25 | 0.162 | 78.00 | Chlordane | 12.29 | 15.92 | DEGRADED | ELEVATED CHEM | MODERATE |
| 90028.0 | NSB-M1 (SUB BASE C2) | 871 | 22 | 79.41 | 0.078 | 84.00 | PAHs | 9.71 | 15.88 | DEGRADED | ELEVATED CHEM | MODERATE |
| 93216.0 | SUB BASE C2 (x1) | 872 | 22 | 36.48 | 0.079 | 93.00 | | 3.53 | 5.39 | DEGRADED | | LOW |
| 93217.0 | SUB BASE C2 (x3) | 873 | 22 | 72.12 | 0.074 | 81.00 | | 8.03 | 12.59 | DEGRADED | | LOW |
| 90012.0 | 34 SWARTZ (C.V. YACHT BASIN) | 824 | 20 | 80.17 | 0.334 | 57.00 | | 2.61 | 3.94 | DEGRADED | | LOW |
| 93196.0 | CHULA V. YACHT BASIN S1 (x1) | 825 | 20 | 96.81 | 0.260 | 76.00 | | 4.36 | 6.84 | DEGRADED | | LOW |
| 93197.0 | CHULA V. YACHT BASIN S1 (x3) | 826 | 20 | 94.23 | 0.165 | 79.00 | | 3.37 | 5.00 | DEGRADED | | LOW |
| 90003.0 | 14 SWARTZ (DOWNTOWN PIERS) | 846 | 21 | 59.57 | 0.084 | 70.00 | | 5.46 | 7.51 | DEGRADED | | LOW |
| 93205.0 | DOWNTOWN PIERS K1 (x9) | 847 | 21 | 48.18 | 0.167 | 84.00 | PAH | 5.64 | 8.49 | DEGRADED | ELEVATED CHEM | MODERATE |
| 93107.0 | MISSION BAY A3 (x1)-REP 1 | 853 | 21 | 93.03 | 0.075 | 57.00 | | 5.51 | 6.83 | DEGRADED | | LOW |
| 93107.0 | MISSION BAY A3 (x1)-REP 2 | 854 | 21 | 92.25 | 0.046 | 77.00 | | 6.42 | 7.73 | DEGRADED | | LOW |
| 93204.0 | CORONADO CAYS T2 (x2) | 845 | 21 | 59.85 | 0.062 | 82.00 | | 2.63 | 3.73 | DEGRADED | | LOW |
| 93220.0 | SWEETWATER CH. JJ1 (x8)-REP 3 | 877 | 22 | 36.99 | 0.129 | 81.00 | | 1.78 | 2.45 | DEGRADED | | LOW |
| 93208.0 | G ST. PIER MARINA L1 (x5) | 851 | 21 | 85.24 | 0.064 | 83.00 | | 12.18 | 16.11 | DEGRADED | | LOW |
| | | | | | | | | | | | | |
| | **CHEMISTRY- Summary Quotients** | | | | | | | | | | | |
| 90020.0 | G DE LAPPE | 169 | 12 | 82.00 | 0.020 | 49.00 | | 16.13 | 19.41 | not analyzed | ELEVATED CHEM | MODERATE |
| 90020.0 | G DE LAPPE-REP 1 | 1104 | 27 | 82.53 | 0.086 | 65.00 | | 17.45 | 21.68 | not analyzed | ELEVATED CHEM | MODERATE |
| 90020.0 | G DE LAPPE-REP 2 | 1105 | 27 | 84.43 | 0.087 | 59.00 | | 17.33 | 21.53 | not analyzed | ELEVATED CHEM | MODERATE |
| 90020.0 | G DE LAPPE-REP 3 | 1106 | 27 | 82.37 | 0.049 | 57.00 | | 15.72 | 19.84 | not analyzed | ELEVATED CHEM | MODERATE |
| 90030.0 | BF SCHROEDER SITE F-REP 1 | 1144 | 28 | 93.76 | 0.192 | 70.00 | | 15.76 | 21.77 | not analyzed | ELEVATED CHEM | MODERATE |
| 90030.0 | BF SCHROEDER SITE F-REP 2 | 1145 | 28 | 96.04 | 0.616 | 76.00 | PAHs | 16.58 | 23.52 | not analyzed | ELEVATED CHEM | MODERATE |
| 90030.0 | BF SCHROEDER SITE F-REP 3 | 1146 | 28 | 91.74 | 0.017 | 68.00 | | 17.00 | 22.41 | not analyzed | ELEVATED CHEM | MODERATE |
| 93178.0 | NAVAL SHIPYARDS O2 (x1)-REP 1 | 1119 | 27 | 51.95 | 0.185 | 61.00 | | 15.47 | 19.80 | not analyzed | ELEVATED CHEM | MODERATE |
| 93178.0 | NAVAL SHIPYARDS O2 (x1)-REP 2 | 1120 | 27 | 61.76 | 0.145 | 66.00 | PCBs | 19.38 | 24.82 | not analyzed | ELEVATED CHEM | MODERATE |
| 93178.0 | NAVAL SHIPYARDS O2 (x1)-REP 3 | 1121 | 27 | 46.68 | 0.168 | 67.00 | PCBs | 20.77 | 25.07 | not analyzed | ELEVATED CHEM | MODERATE |
| 93181.0 | NAVAL SHIPYARDS O6 (x1)-REP 1 | 1110 | 27 | 93.71 | 0.071 | 53.00 | | 11.98 | 16.72 | not analyzed | ELEVATED CHEM | MODERATE |
| 93181.0 | NAVAL SHIPYARDS O6 (x1)-REP 2 | 1111 | 27 | 92.52 | 0.021 | 48.00 | | 13.73 | 18.61 | not analyzed | ELEVATED CHEM | MODERATE |
| 93181.0 | NAVAL SHIPYARDS O6 (x1)-REP 3 | 1112 | 27 | 94.34 | 0.037 | 65.00 | | 15.14 | 21.01 | not analyzed | ELEVATED CHEM | MODERATE |
| 90022.0 | P SWARTZ-REP 1 | 1107 | 27 | 84.62 | 0.061 | 58.00 | PAHs | 17.30 | 23.75 | not analyzed | ELEVATED CHEM | MODERATE |
| 90022.0 | P SWARTZ-REP 2 | 1108 | 27 | 80.73 | 0.073 | 61.00 | PAHs | 18.35 | 27.02 | not analyzed | ELEVATED CHEM | MODERATE |
| 90022.0 | P SWARTZ-REP 3 | 1109 | 27 | 87.48 | 0.038 | 54.00 | PAHs | 18.40 | 26.44 | not analyzed | ELEVATED CHEM | MODERATE |
| 93179.0 | NAVAL SHIPYARDS O3 (x1)-REP 2 | 1123 | 27 | 88.89 | 0.049 | 51.00 | | 17.82 | 22.50 | not analyzed | ELEVATED CHEM | MODERATE |
| 93179.0 | NAVAL SHIPYARDS O3 (x1)-REP 3 | 1124 | 27 | 88.24 | 0.115 | 78.00 | Antimony | 22.02 | 25.51 | not analyzed | ELEVATED CHEM | MODERATE |
| 93184.0 | NAVAL SHIPYARDS O11 (x1) | 802 | 19 | 81.41 | 0.070 | 53.00 | DDT | 20.34 | 27.23 | not analyzed | ELEVATED CHEM | MODERATE |
| 93177.0 | NAVAL SHIPYARDS O1 (x1) | 795 | 19 | 28.88 | 0.023 | 50.00 | PAHs | 11.44 | 18.21 | not analyzed | ELEVATED CHEM | MODERATE |
| 90017.0 | C DELAPPE | 166 | 6 | 71.00 | 0.840 | 64.00 | PAHs | 19.60 | 29.72 | not analyzed | ELEVATED CHEM | MODERATE |
| | | | | | | | | | | | | |
| | **CHEMISTRY- Individual Quotients** | | | | | | | | | | | |
| 93162.0 | SUB BASE C3 (x1) | 775 | 18 | 83.09 | 0.585 | 53.00 | PAHs | 6.10 | 9.35 | not analyzed | ELEVATED CHEM | LOW |
| 93107.0 | MISSION BAY A3 (x1)-REP 3 | 855 | 21 | 94.34 | 0.145 | 73.00 | Chlordane | 9.25 | 11.46 | TRANSITIONAL | ELEVATED CHEM, SITE TRANSITIONAL IN LEG 21 | MODERATE |
| 93221.0 | DOWNTOWN ANCH. J1 (x1)-REP 2 | 879 | 22 | 83.50 | 0.143 | 83.00 | Chlordane | 10.03 | 13.04 | UNDEGRADED | ELEVATED CHEM | LOW |
| 90037.0 | STORMDRAIN EM(GRAPE ST.)-REP 3 | 1161 | 29 | 64.02 | 0.290 | 85.00 | Chlordane | 11.46 | 14.94 | not analyzed | ELEVATED CHEM | LOW |
| 93141.0 | COMMERCIAL BASIN F3 (x1)-REP 3 | 1170 | 29 | 70.09 | 0.057 | 70.00 | Mercury | 10.77 | 13.79 | not analyzed | ELEVATED CHEM | LOW |
| 93116.0 | SAN DIEGO RIVER B1 (x4) | 711 | 15 | 77.00 | 0.137 | 88.00 | Chlordane | not analyzed | not analyzed | not analyzed | ELEVATED CHEM, SITE DEGRADED IN LEG 22 | MODERATE |
| 93120.0 | TIJUANA R. ESTUARY HH2 (x3) | 715 | 15 | 55.00 | 0.087 | 85.00 | DDE | not analyzed | not analyzed | not analyzed | ELEVATED CHEM | LOW |
| 93121.0 | TIJUANA R. ESTUARY HH2 (x5) | 716 | 15 | 59.00 | 0.010 | 85.00 | DDE | not analyzed | not analyzed | not analyzed | ELEVATED CHEM | LOW |
| 93174.0 | TIJUANA R. EST. HH3 (x2)-REP 3 | 1152 | 28 | 91.38 | 0.084 | 80.00 | DDE | 5.75 | 6.34 | not analyzed | ELEVATED CHEM | LOW |

(CHEMISTRY-Summary Quotients) did not have benthic community analysis performed, it is likely that these stations will demonstrate degraded benthic communities, when analyzed. In consideration of this concern, all stations with elevated chemistry, based on summary quotients, were assigned a moderate priority ranking.

In situations where high summary quotient values were not found, but where any single chemical concentration exceeded four times (4x) its associated ERM or 5.9 times (5.9x) its associated PEL, the station was also considered to exhibit elevated chemistry. The 4x and 5.9x cutoffs were not statistically determined using the 90% confidence interval as they were with the summary quotients. Values for individual chemical quotients were not normally distributed and transformations did not improve distributions, so statistical determination of confidence limits was not appropriate. Instead, a qualitative examination of the data set indicated that only in the top 10th percentile of chemical measurements do values exceed four times their respective ERM or 5.9 times their respective PEL (Table 7 - Chemistry-Individual Chemical Quotients). These cutoffs were used to help identify stations where any single chemical was extremely elevated. Stations with elevated individual chemical quotients and evidence of benthic community degradation were assigned a moderate ranking. Stations which exhibited elevated chemistry, but showed no biological effects, were assigned a low priority.

Stations which satisfied all three of the criteria were considered a triad hit and are given the highest priority ranking. These stations demonstrated toxicity in the bioassay tests, benthic community degradation and elevated chemistry. Three stations (representing two sites) fell in this category. Three stations were given a high priority ranking although not all conditions of the triad were met. These stations demonstrated repeated toxicity and elevated chemistry but no benthic analyses were performed. However, benthic data for stations analyzed in the same proximity, or later sampling of the station, led to the concern that these sites would have been found degraded, if analyzed. In addition, chemical summary quotients at these three stations were at levels which suggest probable benthic community degradation, as discussed earlier. These concerns warranted upgrading these three stations from a moderate priority to a high priority. Forty-eight stations were given moderate priorities and fifty-two were given low priorities, based on the methods of prioritization previously discussed.

Stations were prioritized to assist SWRCB and RWQCB staff in meeting sediment quality management objectives for San Diego Bay. These recommendations were based on scientific evaluation of data collected between 1992 and 1994. They are intended to focus future efforts toward scientifically and economically responsible

characterization of locations which have a high probability of causing adverse effects to aquatic life. This report should be evaluated in conjunction with all available information and additional research when management and policy decisions are made by SWRCB and RWQCB staff.

## CONTRIBUTORS TO THE BRIEFING DOCUMENT

| | |
|---|---|
| Agenda | Gita Kapahi[1], Max Puckett[1], Craig J. Wilson[2] |
| SPARC Recommendations | John Hunt[3], Fred LaCaro[2] |
| Revised Monitoring Approach | Fred LaCaro, Craig J. Wilson, Max Puckett, Rusty Fairey[4], John Hunt, Brian Anderson[3] |
| Monitoring Activities | Karen Taberski[5], Bill Croyle[6], Craig J. Wilson |

Issue Papers

| | |
|---|---|
| Toxicity | John Hunt, Brian Anderson |
| Chemical Association | Rusty Fairey |
| Bioaccumulation | Robert Brodberg[7], Bruce Gwynne[8] |
| Freshwater Studies | Val Connor[6], Chris Foe[6], Bill Croyle |
| Benthic Community | Carrie Bretz[5], Jim Oakden[5], Max Puckett |
| Weight-of-Evidence | Rusty Fairey |

[1]State Water Resources Control Board

[2]Department of Fish and Game

[3]University of California, Santa Cruz

[4]San Jose State University

[5]San Francisco Bay Regional Water Quality Control Board

[6]Central Valley Regional Water Quality Control Board

[7]Office of Environmental Health Hazard Assessment

[8]North Coast Regional Water Quality Control Board

[9]Moss Landing Marine Laboratory

# STATE WATER RESOURCES CONTROL BOARD

P.O. BOX 100, Sacramento, CA 95812-0100

Legislative and Public Affairs: (916) 657-1247
Water Quality Information: (916) 657-0687

Clean Water Programs Information: (916) 227-4400
Water Rights Information: (916) 657-2170

## CALIFORNIA REGIONAL WATER QUALITY CONTROL BOARDS

**NORTH COAST REGION (1)**
5550 Skylane Blvd., Ste. A
Santa Rosa, CA 95403
(707) 576-2220

**SAN FRANCISCO BAY REGION (2)**
2101Webster Street, Ste. 500
Oakland, CA 94612
(510) 286-1255

**CENTRAL COAST REGION (3)**
81 Higuera Street, Ste. 200
San Luis Obispo, CA 93401-5427
(805) 549-3147

**LOS ANGELES REGION (4)**
101 Centre Plaza Drive
Monterey Park, CA 91754-2156
(213) 266-7500

**CENTRAL VALLEY REGION (5)**
3443 Routier Road, Suite A
Sacramento, CA 95827-3098
(916) 255-3000

**FRESNO BRANCH OFFICE**
3614 East Ashlan Avenue
Fresno, CA 93726
(209) 445-5116

**REDDING BRANCH OFFICE**
415 Knollcrest Drive
Redding, CA 96002
(916) 224-4845

**LAHONTAN REGION (6)**
2092 South Lake Tahoe Blvd.
South Lake Tahoe, CA 96150
(916) 542-5400

**VICTORVILLE BRANCH OFFICE**
15428 Civic Drive, Ste. 100
Victorville, CA 92392-2383
(619) 241-6583

**COLORADO RIVER BASIN REGION (7)**
73-720 Fred Waring Dr., Ste. 100
Palm Desert, CA 92260
(619) 346-7491

**SANTA ANA REGION (8)**
California Tower
3737 Main Street, Ste. 500
Riverside, CA 92501-3339
(909) 782-4130

**SAN DIEGO REGION (9)**
9771 Clairemont Mesa Blvd., Ste. B
San Diego, CA 92124
(619) 467-2952


**STATE OF CALIFORNIA**
Pete Wilson, Governor

**CALIFORNIA ENVIRONMENTAL PROTECTION AGENCY**
James M. Strock, Secretary

**STATE WATER RESOURCES CONTROL BOARD**
John P. Caffrey, Chair

1/96