SAN FRANCISCO ESTUARY INSTITUTE

REGIONAL MONITORING PROGRAM FOR TRACE SUBSTANCES

Re-design Process of the San Francisco Estuary Regional Monitoring Program for Trace Substances (RMP) Status & Trends Monitoring Component for Water and Sediment

Sarah Lowe, Bruce Thompson, Rainer Hoenicke, Jon Leatherbarrow,
San Francisco Estuary Institute, Oakland, CA

Karen Taberski,
San Francisco Bay Regional Water Quality Control Board, Oakland, CA

Robert Smith, Consultant, Ojai, CA

Don Stevens Jr.,
Oregon State University, Corvallis, OR

CONTRIBUTION
NO. 109
JULY
2005

Table of Contents

Chapter 1 – Introduction

Bruce Thompson, SFEI
Karen Taberski, Regional Water Quality Control Board
Rainer Hoenicke, SFEI
Sarah Lowe, SFEI

The San Francisco Estuary Regional Monitoring Program for Trace Substances (RMP) was recently redesigned to be more responsive to management and regulatory needs. The purpose of this report is to document the process, decisions, and rationale for the RMP redesign efforts between 1997 and 2002.

RMP History

Prior to the late 1980s, there was no formal, sustained monitoring of water quality in the San Francisco Estuary. As a consequence, there was little information about the spatial and temporal distribution of contamination, toxicity, and bioaccumulation of contaminants in the Estuary. As the lead agency for the state for implementing the federal Clean Water Act (CWA) and California's Porter Cologne Water Quality Control Act in the Estuary, the San Francisco Bay Regional Water Quality Control Board (Regional Board) needed to determine whether water quality objectives were being met, and if "beneficial uses" of the Estuary were being protected. Therefore, the Regional Board began conducting pilot studies for a regional water quality monitoring program in 1989, under the State's Bay Protection and Toxic Cleanup Program (BPTCP) and several US EPA grants. The objectives of those studies were to:

1. Determine if water quality objectives were being met and if beneficial uses were being protected
2. Develop a CWA Section 303(d) list (of impaired water bodies) based on sound scientific data
3. Prioritize Regional Board actions to address the most important water quality issues
4. Evaluate management actions by identifying and evaluating trends in contamination
5. Develop a design for a long-term monitoring program

Pilot monitoring stations were subjectively chosen to represent the geographical extent of the Estuary using locations where previous data had already been collected and away from point discharges. The results of these pilot studies were summarized in a series of reports (Edmonds *et al*., 1994; Flegal *et al*., 1994, 1996).

In 1993 the San Francisco Estuary Project (National Estuary Program) completed its Comprehensive Conservation and Management Program (CCMP, 1992) that called for regional monitoring of many features, including contamination in the Estuary. Using that call for monitoring and information from the Pilot Studies, the Regional Board and Estuary dischargers and users agreed to implement the RMP in 1993. The San Francisco

Estuary Institute was given the task of administering and managing the RMP on behalf of the Regional Board and the participants. The RMP's original objectives were to:

1. Obtain high quality baseline data describing concentrations of contaminants in the San Francisco Estuary sediment and water
2. Determine seasonal and annual trends
3. Develop a dataset that could be used to evaluate long-term trends
4. Determine if water and sediment quality are in compliance with regulatory guidelines

The original RMP monitoring design was developed by Regional Board staff and included many of the Pilot monitoring stations. Several stations were moved farther out into the Estuary channel to better capture baseline conditions. Sites were also located at the confluence of the Delta and in the Napa River. In addition to the Status and Trends monitoring component of the RMP, a program of pilot and special studies was implemented. Pilot studies were intended to assess new components or indicators for use in the RMP and special studies were aimed at developing better monitoring methods or improving interpretation of RMP data.

During the first five years, the RMP made several adjustments to the monitoring program in response to the Regional Board's need to obtain a clearer understanding of Estuary processes:

- In 1994 several new sites were added to fill geographic gaps, including two sites in the southern sloughs of the Estuary under NPDES permit conditions of the cities of San Jose and Sunnyvale.
- The aquatic toxicity component of the RMP showed that aquatic toxicity was more prevalent during the wet season and therefore possibly related to storm water run-off events, which led the RMP to implement a pilot program of episodic aquatic toxicity monitoring in several tributaries during seasonal stormwater runoff events in 1996.
- Fish tissue contaminant sampling was added in 1997.

Except for those changes, it was decided to continue using the original Status and Trends monitoring design for five years in order to have sufficient data, from several water-year types, by which the monitoring design could be evaluated.

In 1997, following five years of RMP monitoring, a comprehensive Program review was conducted (Bernstein and O'Connor, 1997). A key recommendation of the review panel was to reconsider program objectives and the justification for the RMP monitoring design. The RMP redesign process was implemented in response to those recommendations.

RMP Redesign Process

The RMP redesign process was guided by the methodology outlined by the National Research Council (1990) and included several steps:

Identify Regulatory Needs. Identification of the Regional Board's information needs was accomplished using *focusing questions*. This was arguably the most important and critical step of the redesign process. The Regional Board staff worked together to provide statements of information needs and to identify a set of questions of interest to the Board. In order to comply with the Clean Water Act, the Regional Board must prepare an impaired water bodies 303(d) list every four years. About this time, the regulatory paradigm was changing to Total Maximum Daily Load (TMDL) based actions, which greatly influenced the Regional Board's focusing questions and the new direction of the RMP.

Revised Objectives. The Regional Board's focusing questions helped identify a revised set of objectives for RMP:

1. Describe patterns and trends in contaminant concentration and distribution;
2. Describe general sources and loadings (inputs) of contamination to the Estuary,
3. Measure the effect of contaminants on selected parts of the Estuary ecosystem,
4. Compare monitoring information to relevant water quality objectives and other guidelines, and
5. Synthesize and distribute information from a range of sources to present a more complete picture of the sources, distribution, fate, and effects of contaminants in the Estuary ecosystem.

Management questions. Each objective has a set of subordinate management questions that linked the objectives to the revised monitoring design. These specific management questions incorporated key agreements among all RMP participants about the new direction of the RMP. The management questions are listed in Appendix 1.

Workgroups. The five-year review recommendations and the management questions indicated that the RMP needed to provide information about contaminant sources, loadings, processes, and effects in order to guide regulation. As a result, several technical workgroups were established to consider the Status and Trends monitoring design options and pilot and special study priorities. These included workgroups that focused on: 1) contaminant sources, pathways and loadings, 2) chlorinated hydrocarbons, 3) pesticides, 4) sediment, and 5) bivalve bioaccumulation studies. The workgroups evaluated data from the RMP and other sources. All workgroups submitted reports that summarized the status of the data and recommended changes that would address the new objectives of the RMP.

Prioritization of redesign activities. The workgroups produced numerous recommendations for redesign of the Status and Trends component of the program and for new pilot and special studies. The Regional Board staff prioritized these recommendations from a management perspective. The Regional Board's focusing questions, RMP objectives, and management questions emphasized the need to protect water quality and beneficial uses consistent with needs for the 303(d) and 305(b) process. The US EPA's guidance for Clean Water Act monitoring recommended a probabilistic approach to those activities (EPA, 2001).

Considering that guidance and recommendations of the workgroups, it was agreed that the original design of the Status and Trends component did not provide a spatially balanced sample of water and sediment quality status in the Estuary. Therefore, the decision was made to investigate the use of a probabilistic monitoring design that included a complete coverage of the Estuary (including both shallow water and deeper channel sites). Such a design would provide the RMP participants with statistically defensible inferences about the status and trends of contaminant concentrations in the various Regions of the Estuary.

The above steps set the stage for the RMP redesign. A Design Integration Work Group (the Re-design Work Group) was established in 1997 following the RMP's five-year review to guide and oversee the actual redesign process. The Re-design Work Group's charge was to evaluate the existing Status and Trends monitoring design, apply what had been learned about contaminant trends in the Estuary to date, and develop a design that would fit the new, revised RMP objectives. Work group participants included Bay Area scientists from research institutions and universities, sub-contracting statisticians, RMP participants (from the municipal, stormwater, industrial, and dredging communities), the Regional Board, and SFEI.

The redesign process took several years because it required considerable discussion, data analysis, and thoughtful decisions by all involved. It could not have happened without the dedication and participation of the Regional Board staff, RMP Participants, SFEI staff, and participants in the Re-design Work Group (Appendix 2). The new spatially randomized sampling of status and trends in the Estuary began in the summer of 2002.

References:
CCMP, 1992. Comprehensive Conservation and Management Plan. http://www.abag.ca.gov/bayarea/sfep/reports/ccmp/

Edmunds, J.L., Cole, B.E., Cloern, J.E., Caffrey, J.M., and A.D. Jassby. 1995. *Studies of the San Francisco Bay, California, estuarine ecosystem. Pilot regional monitoring program results, 1994*. US Geological Survey Open-File Report 95-378

Flegal, A. Russel, R. W. Risebrough, B. Anderson, J. Hunt, S. Anderson, J. Oliver, and M. Stephenson. 1994. San Francisco Estuary pilot regional monitoring program: sediment studies. Final Report to the San Francisco Bay Regional Water Quality Control Board, Oakland CA.

Flegal, A.R., S.A. Sanudo-Wilhelmy, I. Rivera-Duarte, P.R. Ritson, and M.R. Gordon. 1996. Metals contamination in San Francisco Bay waters: historic perturbations, contemporary concentrations, and future considerations. *In* S*an Francisco Bay: The Ecosystem*. J.T. Hollibaugh, ed. American Association for the Advancement of Science, pp. 173-188.

Bernstein, B, and J. O'Connor, 1997. Five-Year Program Review Regional Monitoring Program for Trace Substances in the San Francisco Estuary. SFEI RMP Contribution 28. SFEI, Oakland CA

National Research Council, 1990. Managing Troubled Waters. The role of Marine Environmental Monitoring. National Academy Press, Wash. DC. 125 pp.

EPA, 2001. Robert H. Wayland III (Director). 2002 Integrated Water Quality Monitoring and Assessment Report Guidance. Office of Wetlands, Oceans and Watersheds. November 19, 2001 URL: http://www.epa.gov/owow/tmdl/2002wqma.html

Chapter 2 - Development of the Estuary's Stratification Scheme for the Status and Trends Monitoring Component and Other Re-design Considerations

Sarah Lowe, SFEI
Robert Smith, Ojai, CA.
Jon Leatherbarrow, SFEI
Bruce Thompson, SFEI

Introduction

The original RMP Status and Trends monitoring design emphasized repeated sampling of "background" stations along the spine of the Estuary for both spatial and temporal trend analyses. The sampling locations were chosen for either or both of two reasons: 1) they were spatially consistent with stations monitored by the CA State Mussel Watch Program since 1986 and 2) they were far from potential contaminant sources. Based on nearly ten years of seasonal contaminant monitoring at those stations, the RMP provided information to estimate "background" concentrations of many regulated contaminants in the Estuary, which made it possible for the Regional Board to update the region's 303(d) list and begin to focus on more complex management issues such as implementing the TMDL process for targeted contaminants.

The Re-design Work Group began by identifying specific management questions (see Appendix 1) to be addressed in the new design. Questions such as; "what proportion of the Estuary is above the regulatory water quality guidelines?", "what are the sources, spatial distributions, trends, and fate of targeted contaminants within the sub-regions of the Estuary?", and "are current regulatory actions effective in managing those contaminants?" To answer these types of questions, the Re-design Work Group recognized that a spatially randomized design would meet the RMP objectives better than a historic station design historically employed by the RMP and that the Estuary should be stratified into geographical sub-regions based on distinct hydrodynamic and sediment quality properties observed in different regions of the Estuary. This led to an evaluation of hydrology and water quality, sediment dynamics and sediment quality, and ecological patterns observed in the Estuary. The information was used to create a regional stratification scheme to use in developing the new RMP Status and Trends sampling design.

This chapter describes the approach used to develop the geographic sub-regions of the Estuary used by the new RMP Status and Trends component. It also describes specific decisions made about the new design, including the choice of sampling design, sample allocation within sub-regions, sample timing, geographical extent of the sampling, continuity between the historical and the new design, and identification of confounding factors.

Approach

As an initial step the Re-design Work Group decided to perform several tasks to develop the best estimate of the Estuary's hydrographic sub-regions:

1. Review existing sub-regions defined by various agencies or for other purposes
2. Survey local scientists for their professional recommendations
3. Evaluate empirical data by performing statistical cluster and graphical analyses of water and sediment quality data

Results from these three tasks were mapped to evaluate 1) the overlap or separation of Estuary sub-regions determined by each method, 2) the extent of seasonal variability in water and sediment quality parameters, 3) and spatial differences between sub-regions based on hydrology and sediment characteristics and ecological distributions.

A series of presentations were made to the Re-design Work Group for discussion and guidance. Final decisions on the selection of the new boundaries of the hydrologic regions in the Estuary were made by consensus.

Existing Regional Schemes

The previously existing regional scheme for the Estuary, outlined in the San Francisco Bay Basin Plan (Basin Plan, 1995) and used by the Regional Board for regulatory purposes, may not identify sub-regions of distinct hydrodynamic and sediment properties in the Estuary (Figure 2.1). This scheme was developed using USGS drainage basin maps that border the Estuary and did not take into account hydrodynamic processes within the Estuary (Gunther, 1987). The Basin Plan scheme divides the Estuary into seven sub-regions bounded by the major Bay Area bridges.

A more detailed plan for dividing up sub-regions of the Estuary was proposed by Gunther (1987) for the purpose of investigating ecological distributions and contaminant transport processes and was based on depth, location of major outfalls, geography, and hydrology of the Estuary (Figure 2.2). Gunther's scheme further divided the sub-regions used by the Regional Board, which were largely based on the boundaries formed by Bay Area bridges, into a total of 37 subdivisions: Lower/South Bay (12 divisions), Central Bay (11 divisions), San Pablo Bay (10 divisions), and Suisun Bay (4 divisions). The Work Group decided that this level of division was too much for the amount of sampling resources the program had for both the water and sediment monitoring programs. They also thought that while this level of resolution may be appropriate for intensified sediment studies in the Estuary, water masses in the Estuary are highly dynamic, strongly influenced by tidal action and seasonal freshwater flows, and therefore may not be appropriate for intensified water quality studies. It was noted that the EMAP stratified random sample design can be further stratified within each defined sub-region on the basis of depth or other attributes and that this can occur at the time of data analyses. As a result of these discussions the Work Group decided that the new random sample design for water and sediment quality should have similar regional boundaries.

Other existing regional schemes included a study developed by the Sanitary Research Laboratory of the University of California as part of the Comprehensive Study of San Francisco Bay in 1969 (Pearson, Storrs, and Selleck, 1969). The UC study presented no documentation regarding the rationale behind the developed segmentation scheme and is not currently used by any agency or organization, and therefore was not evaluated further for this review (Gunther, 1987). USGS and the State Water Resources Control Board watershed drainage-basin maps for the U.S. and California were also considered. However, the boundaries in those maps represent large-scale, land-based divisions with little detail of the Estuary and were not useful in the RMP Re-design effort.

Survey for Local Professional Recommendations

Several Bay Area scientists with expertise on the Estuary's hydrology and biological communities, and familiar with the RMP monitoring program, provided their professional recommendations by responding to the following question: "Using your experience and judgment how would you recommend that we subset the Bay into regions for long-term monitoring?" Five scientists responded to our survey, which asked them to identify geographical boundaries on a map (Figure 2.3). Survey responses included two based on hydrologic studies, one on both hydrologic and biological studies, one based on sediment dynamics, and one based on biology. The basis for each scientist's recommendation is summarized in Table 1. The recommended sub-regions were mapped, and included in the stratification evaluation process.

Table 1. Summary of survey responses from Bay Area scientists about how they would stratify the Estuary's sub-regions.

Name	Affiliation	Expertise	Opinion on Geographical Boundaries Were Based On
Dr. Jon Burau and Dr. Gross	USGS and Stanford	Hydrology Models	Studies that showed that density driven thermohaline circulation patterns within the northern Estuary are bounded mainly by bathymetric sills or shoals (Burau *et al.,* 1998), and that the Southern Estuary can be divided into three distinct sub-regions (Gross *et al.*, 1999).
Dr. Dave Schoellhamer	USGS	Suspended Sediment Circulation Studies and Modeling	His view was that the hydrological boundaries in the Estuary are located at the major geographical constrictions, which have a large effect on circulation. (Schoellhammer *et al.*, 1997, Schoellhammer, 2001; Buchanan and Schoellhamer, 1998).
Dr. Alan Jassby	USGS	Hydrology Studies and Long-term Monthly Water Quality Monitoring	Long-term monitoring of monthly water quality attributes using high-resolution sampling transects along the "spine" of the Estuary (extending from Coyote Creek in the Lower South Bay into the Sacramento River in the North Bay-Delta). (Jassby, and Powell, 1994; Jassby, Cole, and Cloern, 1997).
Dr. Wim Kimmerer	San Francisco State University	Hydrodynamics and Biology	Studies of zooplankton distributions, and the Estuary's hydrology based on his participation in development of the X2 salinity standard (Kimmerer and Schubel, 1994; Jassby *et al.,* 1995; Kimmerer, Burau, and Bennett, 1998).
Dr. Bruce Thompson	San Francisco Estuary Institute	Regional Monitoring Program and Benthic Community Studies	Studies of benthic community attributes within various regions of the Estuary. Five distinct benthic assemblages have been identified in the Estuary that shift geographically with seasonal changes (Thompson *et al.*, 2000).

Cluster and Graphic Analyses

Numerical Cluster and simple graphical analyses were performed using water-quality and sediment-quality data from RMP Status and Trends monitoring, Local Effects Monitoring Program (LEMP), Department of Water Resources (DWR), and the Bay Protection and Toxic Cleanup Program (BPTCP) collected between 1989 and 1998. Datasets included: RMP and LEMP water data collected from 1989 – 1998 (576 samples), and sediment data from the RMP, LEMP, DWR, and BPTCP collected from 1991 – 1998 (945 samples).

The Cluster analyses identified spatially coherent and temporally persistent Estuary sub-regions based on similarities in water quality and sediment quality attributes. Contaminant data were not considered in this evaluation because contaminant distributions can be affected by potential sources that could confound identification of regions with relatively homogeneous hydrologic and sediment properties. Six water quality attributes were used in the cluster analyses for water: salinity, dissolved oxygen (DO), dissolved organic carbon (DOC), temperature, total suspended solids (TSS), and pH (Figure 2.4). Two sediment quality attributes were evaluated in the cluster analyses for sediment: percentage of fine sediments (< 63 µm, %Fines), and total organic carbon (%TOC). The cluster analyses were conducted three ways: pooling all available data and separately for the wet and dry seasons.

Graphic analyses were conducted for water and sediment by first plotting salinity vs. temperature (for water) and % TOC vs. % Fines (for sediment), grouping samples with similar characteristics, and then mapping the results in order to visually identify similar sub-regions.

Best Estimate of the Estuary's Sub-regions

Professional Opinions

The professional recommendations among the five responding scientists had good agreement for several sub-regions in the Estuary coinciding largely with 4 large geographical constrictions with some interesting refinements on the hydrologic boundaries by Jon Burau, Alan Jassby, and Bruce Thompson (see Figure 2.3):

1) The Carquinez Strait defined the western reaches of the Suisun Bay region, with some difference of opinion about where the boundary with the Rivers was (probably reflecting the highly dynamic characteristics of that region due to climatology and delta outflow). The Regional Board recommends Chips Island as the eastern boundary of Suisun Bay, which agreed with Dave Shoellhamer's suggestion.
2) Pt. San Pablo and San Pedro defined San Pablo Bay region. However, Jon Burau suggested that the Carquines Strait hydrology extends into San Pablo Bay in the deeper waters towards Pinole Point. Alan Jassby suggested that hydrologic characteristics of the San Pablo Bay region actually reach as far south as Angel Island.
3) The shallows of San Bruno Shoals defined the southern reaches of the Central Bay Region. However, Jon Burau characterized a different Central Bay region that combined most of San Pablo Bay and the Central Bay region to the Bay Bridge. Bruce Thompson suggested that the benthic community characteristics of the Central Bay region actually extend as far south as the San Mateo Bridge.
4) The geographic constriction at the Dumbarton Bridge defined the northern boundary of the Lower South Bay Region.

Water Analyses

Cluster analysis results for water are presented in Figure 2.4. Seasonal clustering patterns show a large Central Bay sub-region (predominantly marine) is more extensive than defined by the Regional Board. During the dry season it extends northward to the naturally constricting peninsulas north of the Richmond Bridge: Point San Pablo and Point San Pedro. During the wet season, fresh water flow moves this boundary southward towards Angel Island. The southern boundary of the Central Bay marine region extends to San Bruno Shoals during the dry season, and shifts further south during the wet season to the natural constrictions at Dumbarton Bridge.

Dry weather graphical analysis results were identified by graphing salinity vs. temperature results for available the dry season water data, and delineating regions with similar salinity and temperatures (Figure 2.5). Results were similar to the Basin Plan sub-regions, except that there was no boundary at the San Mateo Bridge.

Sediment Analyses

Dry weather cluster analysis results for sediment quality attributes showed more distinct regions than when the cluster analyses were run using wet weather data. Results indicated that the majority of the Bay consisted of greater than 60% Fine sediments and less than 4% TOC (Figure 2.6). Bay margins had highly variable sediment grain size, and generally had higher TOC than the rest of the Estuary.

Dry weather graphical analysis results for sediment were grouped based on similar regimes for % Fines and % TOC (Figure 2.7). These results showed partitioning between fine-grained sediments (> 75% Fines) and slightly sandier sediments (between 50 and 75% Fines) in San Pablo Bay, the lower region of the Central Bay, and in the extreme South Bay. Table 1 shows the sediment characteristics of the corresponding partitions in Figure 2.7.

Table 1. Graphical Analysis Results of sediment characteristics.

Partition Color	% Fines	%TOC
Black	75–100	<4
Blue	50–75	<2
Green	0–50	<2
Yellow	70–100	4–7
Pink	50–70	2–4

Figure 2.8 summarizes the recommendations of the five Estuary scientists, the dry-season water cluster analysis, and the graphical analyses for both water and sediment.

Stratifying the Estuary into Sub-regions and Defining the Sample Frame

Using a "weight-of-evidence" approach, the regions recommended by the Estuary scientists, and the results of the cluster and graphical analyses were summarized. The results are presented in Table 2, which counts the number of times a regional boundary was suggested for each analysis method. There were five expert opinions by Estuary scientists so there were five times (number of possible hits) any regional boundary could have been identified for that analysis method; the water cluster analysis was performed three ways (using all the data, wet-season only and dry-season only data) for a total of three possible hits; there was only one possible hit each for both the water and sediment graphical analyses. The total number of times each regional boundary was suggested was tallied in the right-hand column. The Re-design Work Group reviewed and discussed the results and proposed a new set of regional boundaries for use by the RMP Status and Trends program for water and sediment monitoring (denoted by ✕).

Table 2. Summary of the number of times each regional boundary was suggested for the various evaluation methods. ✕ represents proposed boundary by the Re-design Work Group for the new sub-regions. * represents an existing boundary defined in Basin Plan, 1995.

Regional Boundary	Expert Opinion	Water Cluster	Water Graphical (dry season)	Sediment Graphical (dry season)	Total	
Number of possible hits	**5**	**3**	**1**	**1**	**10**	
Chipps Island *	1	0	1	0	**2**	
Benicia Bridge*	5	2	1	0	**8**	✕
Carquinez Bridge*	0	2	0	0	**2**	
Carquinez Straight (west end)	3	1	1	1	**6**	✕
Pt. Pinole	1	0	0	1	**2**	
Pt. San Pablo	2	2	1	0	**5**	✕
Richmond Bridge*	0	0	0	0	**0**	
Angel Island	1	1	0	0	**2**	
Bay Bridge*	1	0	1	1	**3**	
San Bruno Shoal	4	2	0	1	**7**	✕
San Mateo Bridge*	1	0	0	0	**1**	
Dumbarton Bridge	4	0	0	0	**4**	✕
Sloughs	0	3	1	1	**5**	✕

This exercise identified new regional boundaries for the Estuary in the Central Bay and South Bay sub-regions (Figure 2.9). The Central Bay region was originally bounded at the Richmond Bridge and the Bay Bridge. The new northern boundary was expanded northward to the natural constriction between Point San Pedro and Point San Pablo, and southward to the natural shallows formed by San Bruno Shoals. In the South Bay, another natural constriction at the location of the Dumbarton Bridge formed the new boundary between the South Bay and the Lower South Bay regions.

Results from both cluster and graphical analyses for both water and sediment data showed that the RMP sampling stations located near the mouths of major tributaries and along the margins of the Estuary were distinct from the surrounding Estuary. These included the following sampling stations: Petaluma, Napa, San Joaquin, and Sacramento Rivers, and the two Southern Slough stations (San Jose and Sunnyvale). The Re-design Work Group decided that these locations were influenced by localized conditions and did not represent the general status of the sub-regions. While the mouths of tributaries are important transitions zones that warrant further study, the Work Group agreed to exclude targeted sampling at the mouths of the Petaluma and Napa Rivers in the new Status and Trends sampling design since the focus of the new design was to evaluate long-term spatial and temporal trends of water and sediment contamination in the major sub-regions of the Estuary.

The Re-design Work Group recommended that RMP studies and/or monitoring of contaminants in the margins and tributaries of the Estuary be further developed through the workgroup process (i.e., the Sources, Pathways, and Loadings Work Group, toxicity Work Group) and implemented through RMP pilot and special studies.

The Re-design Work Group felt it was important to maintain some continuity between the two sampling designs, and some site-specific information on long-term trends at a subset of historical RMP sites for both water and sediment (described in more detail below).

Although random samples were allocated into the Carquinez Strait region, it was not included in the new design because the region is a narrow corridor that is very difficult to sample (due to the swift currents in that region), and the Work Group felt that monitoring in both Suisun Bay and San Pablo Bay, would provide enough information to infer if the Carquinez Strait was a region of concern.

Final Decisions and Further Design Considerations

In consideration of the new RMP objectives, the Re-design Work Group chose a **spatially balanced probabilistic sampling design** developed for the federal Environmental Monitoring and Assessment Program (EMAP) (Stevens, 1997; Stevens and Olsen, 1999; Stevens and Olsen, 2000). The design provides good spatial coverage in each sub-region and can be used to make inferences about contaminant conditions in the Estuary at different spatial and/or temporal scales. The design allows for several years of sampling to be combined to get regional or Estuary-wide averages. Samples are proportionally allocated to all areas, so shallow areas and deep-channels within each region could be "back-stratified" for analyses at a later date. See Chapter 4 for a detailed discussion of the probabilistic sampling design.

The Re-design Work Group further defined the **sample frame** by evaluating both logistical and environmental factors. Considering that high turbidity results in the shallow areas due to localized processes, such as tidal action and boat draft, water samples would be randomly allocated into each sub-region up to a minimum depth of 3 feet mean lower low water (MLLW). Sediment samples would be randomly allocated into each sub-region up to a minimum depth of 1 foot MLLW. For the purposes of "back-stratifying" (during data analyses) the channels would be defined to be the portions of the Estuary that are deeper than 6 feet at MLLW.

The Re-design Work Group chose to **continue to monitor several historical-stations** from the RMP sampling design (since 1993) to provide continuity and transition between sampling designs, and to continue to monitor potential contaminant contributions from the Delta. Three historical RMP water sites, the Golden Gate Bridge "background" station (BC20) and two tributary stations near the confluence of the Sacramento and San Joaquin Rivers (BG20 and BG30) were retained in the new design because they are located at the 'boundary' of the Estuary (between the Estuary and the ocean and the

Estuary and the Central Valley Delta[1]). Seven sediment sites were continued in the new design, one from each of the five targeted monitoring regions and the two historical Rivers sites as Estuary/Delta boundary sentinels. The sediment sites within each targeted region were chosen because they were sites with long-term synoptic chemistry and toxicity measures.

The Re-design Work Group **focused the design on characterizing both the spatial extent of contamination in the Estuary and long-term trends**. Long-term contaminant trends are not clearly discernable, even with a decade of RMP monitoring data, due to high variability observed in contaminant concentrations throughout the Estuary. In order to maximize spatial coverage in each sub-region and in an effort not to increase the cost of the Status and Trends monitoring program, the Re-design Work Group decided to reduce sampling to annual dry-season sampling while maintaining the total number of samples collected annually. While some contaminants have higher ambient concentrations during the winter with increased runoff, the Work Group felt that, although this was an important consideration for environmental managers, the need for increased spatial sampling of the Estuary was more important for the Status and Trends program. The Work Group recommended that seasonal contaminant issues continue to be addressed, through additional RMP pilot and special studies (a process whereby the Technical Review and the Steering Committees prioritize proposed Pilot and Special study topics for funding on an annual basis). Furthermore, the Re-design Work Group recommended that the RMP's Sources, Pathways and Loadings Work Group pursue seasonal contaminant issues.

Characterizing long-term temporal trends in sediments is addressed in two ways: 1) by maintaining historical sites as mentioned above, and 2) repeat sampling at a subset of randomly allocated stations in each sub-region. Because water characteristics are in constant movement at any specific location in any given sub-region between years, no repeat water sampling was warranted. Temporal trends in water will be determined by using the set of random samples collected over time. See Chapter 4 for further explanation.

A final consideration in water quality monitoring was that tidal action stirs up sediments increasing total suspended solids (TSS) concentrations in the water column. TSS influences total water column contaminant concentrations, adding to the variability in long-term trend analyses. The Re-design Work Group decided that, when planning the water-sampling cruise, the RMP should attempt to collect water samples from shallow stations at slack tide to minimize the variability associated with TSS. They acknowledged, however, the added costs of prioritizing this request (sampling of shallows could only take place two times each day), and therefore only suggested it as a

[1] In 2003, the RMP Technical Review Committee and Steering Committee approved adding two additional historic RMP water sampling stations (Yerba Buena Island (BC10) and Dumbarton Bridge (BA30)) to the water sampling design, because those stations (along with Sacramento (BG20)) are currently used by the Regional Board in calculations related to the National Pollutant Discharge Elimination System (NPDES) Permitting Program. The random sample size for the South Bay and Lower South Bay were reduced by one sample each (to 9 and 5 samples respectively) so that this change made no extra cost to the Status and Trends program.

goal. The Re-design Work Group also addressed this issue by limiting the water sampling frame to the 3-foot MLLW depth contour of the Estuary.

Summary

The RMP Re-design Work Group developed a new sampling frame for water and sediments for the Status and Trends component that included identifying distinct sub-regions within the Estuary and outlining the spatial and temporal objectives of the new design. The Work Group also provided technical rationale throughout the decision making process. The work group process was based on consensus building and it helped to clarify the roles of the three main RMP components: the Status and Trends component, Pilot Studies, and Special Studies. The Regional Board accepted the newly defined sub-regions for use in the RMP. However, these regional boundaries are not being used for regulatory purposes at this time.

The next step in developing the probabilistic sampling design for the RMP, and the topic of the next chapter, was to determine how many samples to allocate into each sub-region based on fiscal considerations and the amount of statistical power that might be achieved per various sample sizes. A rigorous power analyses was performed that compared existing water and sediment contaminant information to various guidelines for key contaminants of concern in the Estuary.

References

Basin Plan, 1995. Water Quality Control Plan (Basin Plan) for the San Francisco Bay Basin (1995). URL: http://www.swrcb.ca.gov/rwqcb2/basinplan.htm

Buchanan, P.A., and Schoellhamer, D.H., 1998, Summary of suspended-solids concentration data in San Francisco Bay, California, water year 1996: U.S. Geological Survey Open File Report 98-175, 59 p.

Burau, Jon R., J.W. Gartner, M. Stacy. 1998. Results from the Hydrodynamic Element of the 1994 Entrapment Zone Study in Suisun Bay By Jon R. Burau, Jeffrey W. Gartner, and Mark Stacey US. Geological Survey, Sacramento, CA.

Gross, E.S., Koseff, J.R. Koseff, and S.G. Monismith, 1999. Three-dimensional salinity simulations in South San Francisco Bay. J. Hyd. Div. ASCE , 125(11), pp. 1199-1209.

Gunther, A.J., 1987. Segmentation of the San Francisco Bay-Delta. Aquatic Habitat Institute. Richmond, CA. 22pp.

Jassby, A. D., and T. M. Powell. 1994. Hydrodynamic influences on interannual chlorophyll variability in an estuary: Upper San Francisco Bay-Delta (California, USA). Estuarine Coastal and Shelf Science 39:595-618.

Jassby, A. D., W. J. Kimmerer, S. G. Monismith, C. Armor, J. E. Cloern, T. M. Powell, J. R. Schubel, and T. J. Vendlinski. 1995. Isohaline position as a habitat indicator for estuarine populations. Ecological Applications 5:272-289.

Jassby, A. D., B. E. Cole, and J. E. Cloern. 1997. The design of sampling transects for characterizing water quality in estuaries. Estuarine Coastal and Shelf Science 45:285-302.

Kimmerer, W.J. and J.R. Schubel. 1994. Managing freshwater flows into San Francisco Bay using a salinity standard: results of a workshop. Pp. 411-416 In K.R. Dyer and R.J. Orth (eds.), Changes in fluxes in estuaries. Olsen and Olsen, Fredensborg, Denmark.

Kimmerer, W.J, J. R. Burau, and W.A. Bennett. 1998. Tidally-oriented vertical migration and position maintenance of zooplankton in a temperate estuary. Limnology and Oceanography 43:1697-1709.

Pearson, E.A., Storrs, P.N., and Selleck, R.E., 1969, A comprehensive study of San Francisco Bay. Final Report, volume III. Waste discharges and Loadings. SERL Rept. No. 67-3, Sanitary Engineering Research Laboratory, University of California, Berkeley, CA. 98pp

Schoellhamer, D.H., Cacchione, D.A., Rubin, D.M., Cheng, R.T., Dingler, J.R., Jaffe, B.E., and Stumpf, R.P., 1997, Sediment-Transport Research in San Francisco Bay: Proceedings of USGS Sediment Workshop, Harpers Ferry, West Virginia, February 4-7, 1997.

Schoellhamer, D.H., 2001, Influence of salinity, bottom topography, and tides on locations of estuarine turbidity maxima in northern San Francisco Bay, in McAnally, W.H. and Mehta, A.J., ed., Coastal and Estuarine Fine Sediment Transport Processes: Elsevier Science B.V., p. 343-357.

Stevens, Jr., Don.L. 1997. Variable density grid-based sampling designs for continuous spatial populations. Environmetrics 8:167-195.

Stevens, Jr., Don.L. and A.R. Olsen. 1999. Spatially restricted surveys over time for aquatic resources. Journal of Agricultural, Biological, and Environmental Statistics 4:415-428.

Stevens, Jr., Don.L. and A.R. Olsen. 2000. Spatially-restricted random sampling designs for design-based and model-based estimation. In Accuracy 2000: Proceedings of the 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences. Delft University Press, The Netherlands, pp. 609-616.

Thompson B., S. Lowe, and M. Kellogg. 2000. Results of the Benthic Pilot Study, 1994-1997. San Francisco Estuary Institute, Oakland. CA. URL: http://www.sfei.org/sfeireports.htm

Figure 2.1 – San Francisco Estuary hydrographic regions from the 1995 Basin Plan for regulatory purposes.

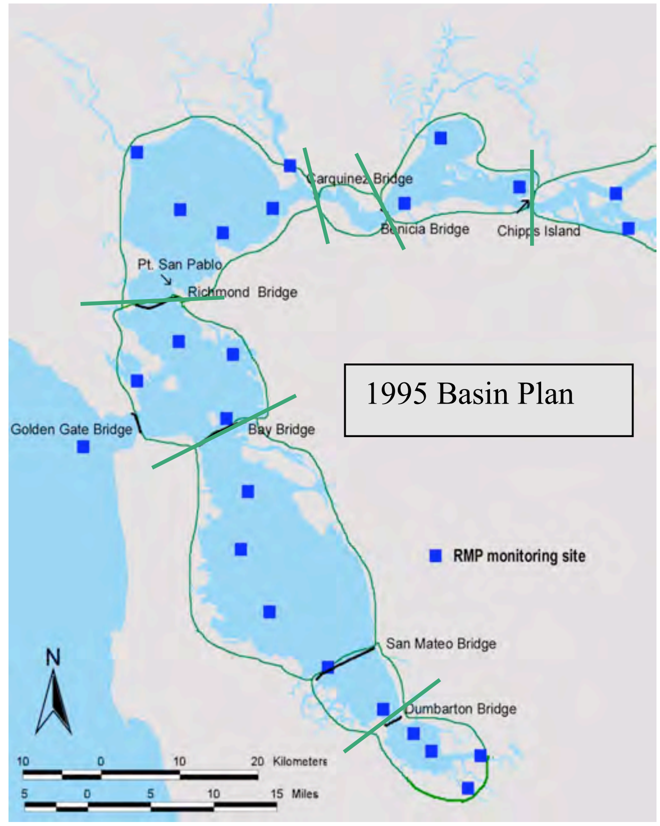


Figure 2.2 – Further divisions of the San Francisco Estuary sub-regions based on depth, location of major tributaries, geography and hydrology.

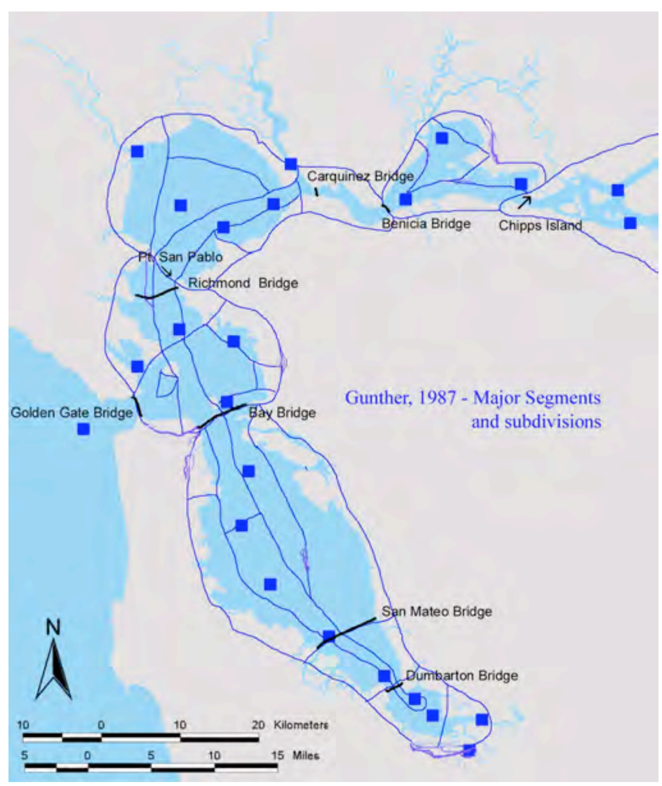


Figure 2.3 – San Francisco Estuary sub-regions based on hydrology and biological communities.

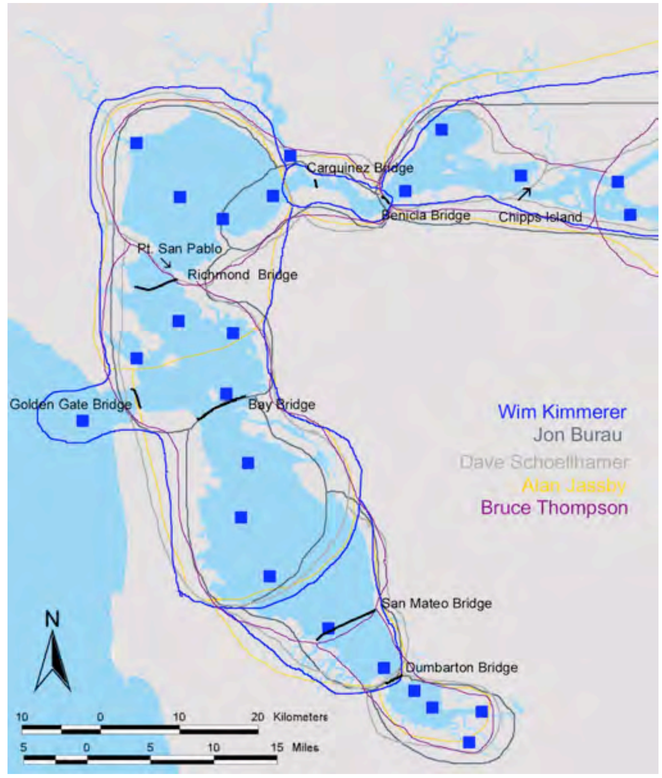


Figure 2.4 – San Francisco Estuary hydrographic regions based on cluster analyses using six water quality attributes: temperature, DO, DOC, total suspended solids, pH and salinity. Analyses were performed using only wet or dry season data and for all data combined.

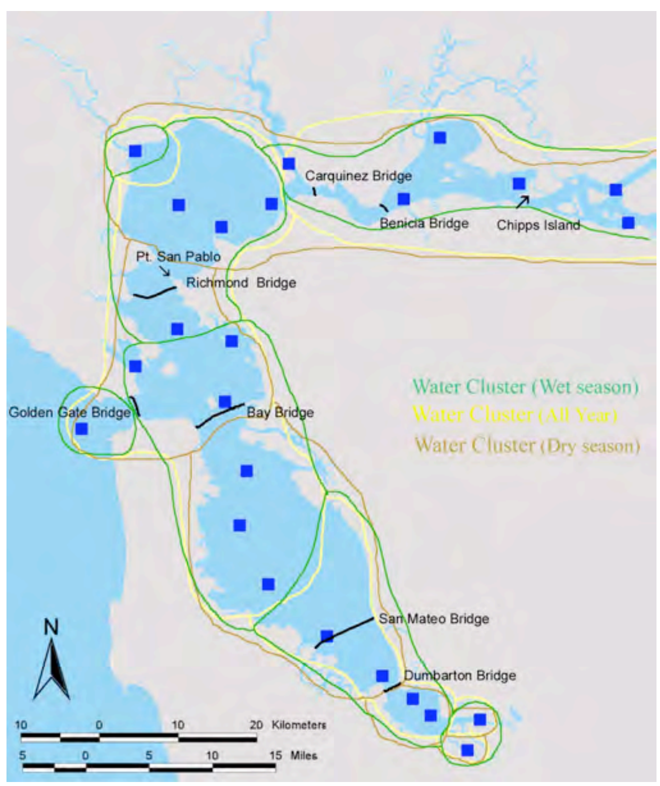


Figure 2.5 – San Francisco Estuary hydrographic regions based on graphical analysis using salinity and temperature to group sites. Grapical analyses were performed on dry season data only.

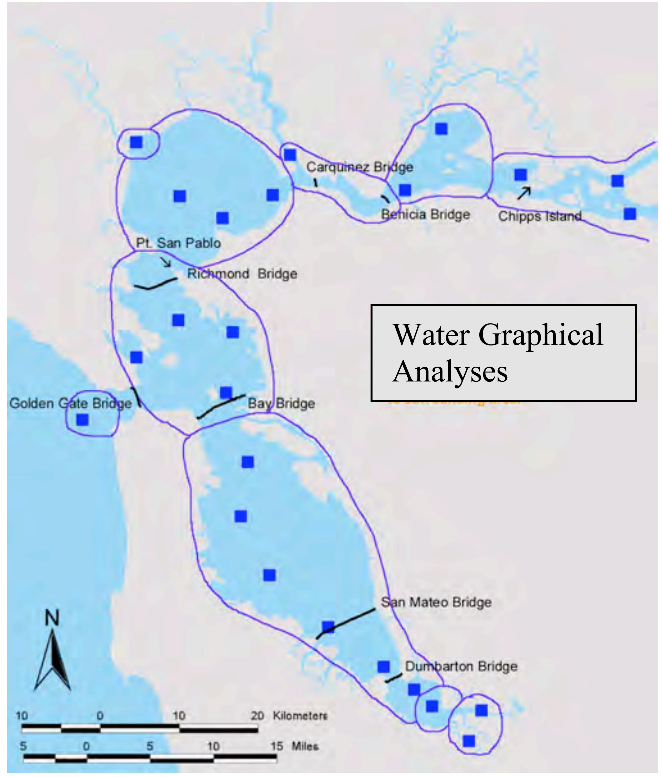


Figure 2.6 – San Francisco Estuary sub-regions based on sediment cluster analysis using two sediment quality attributes, % Fines and TOC, in the dry season.

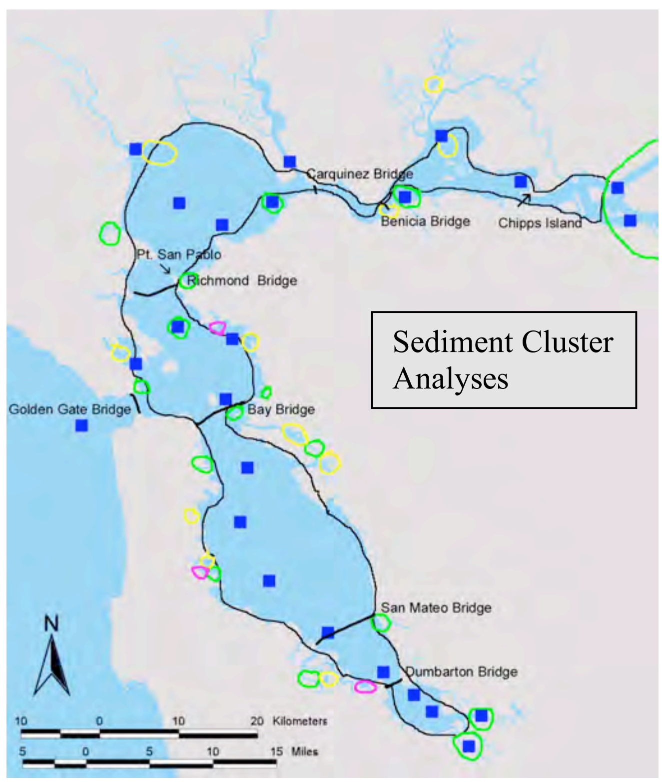
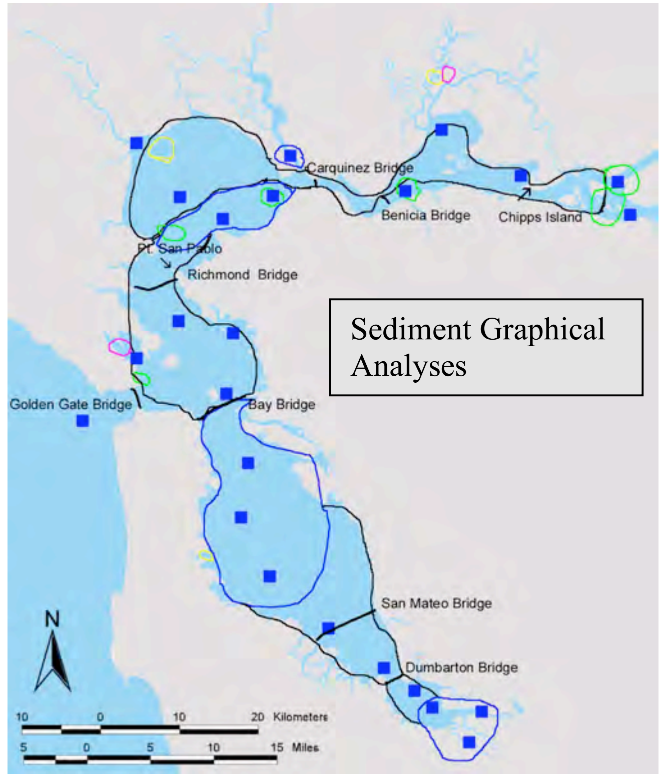


Figure 2.7 – Dry season graphical analyses results for the San Francisco Estuary using two sediment quality attributes: % Fines and TOC.



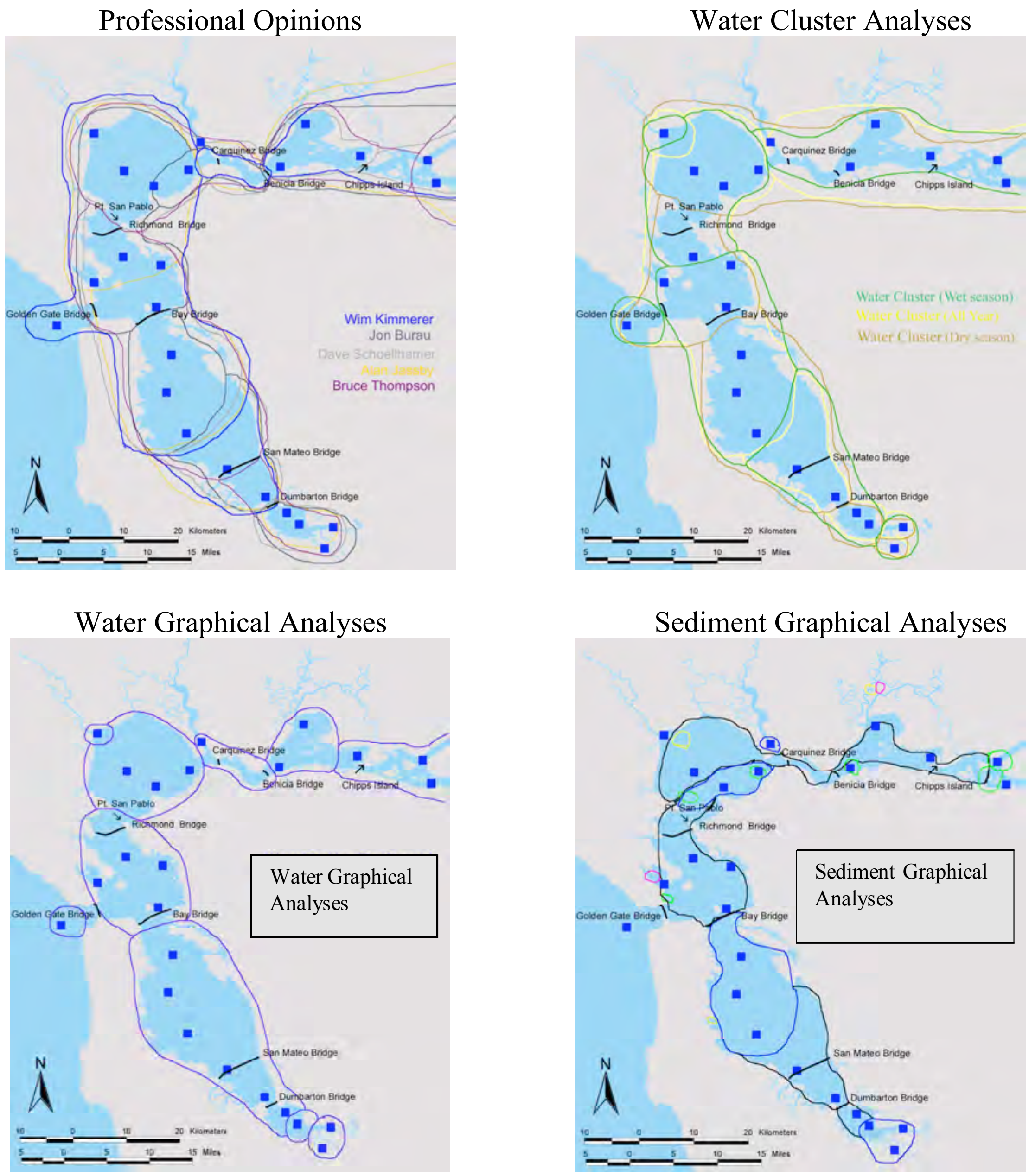


Figure 2.8 – Summary of the sub-regions from the recommendations of five Estuary scientists, the dry-season water cluster analysis, and the graphical analyses for both water and sediment.

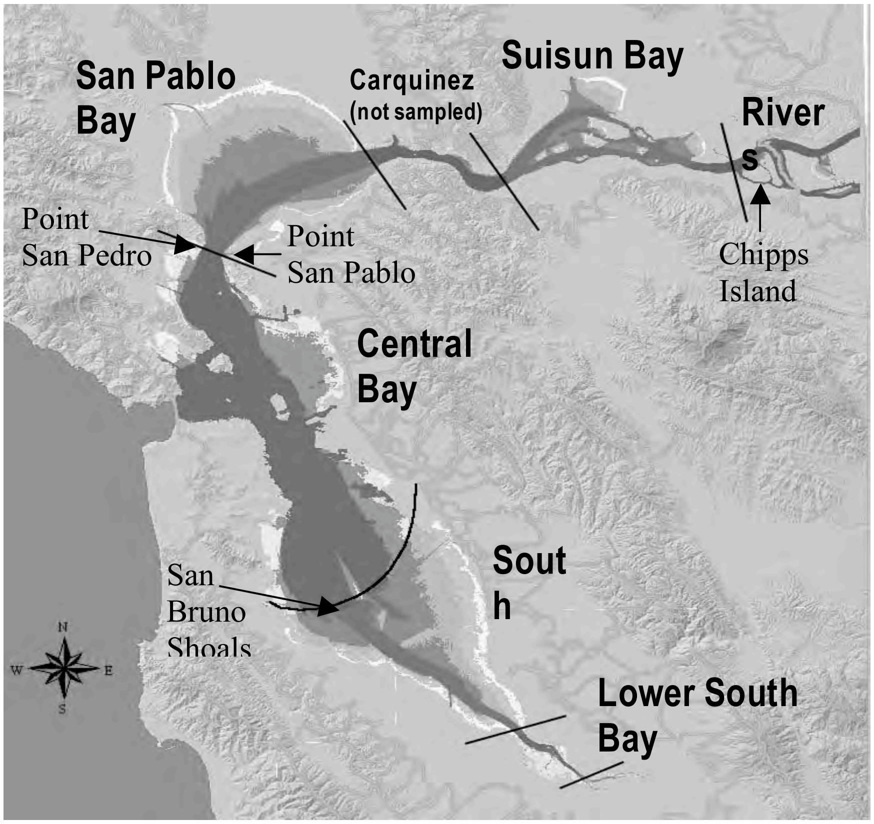


Figure 2.9 – Final sub-regions recommended by the Re-design work group for the new RMP Status & Trends sampling of Estuary water and sediment.

Sub-Region	Boundary location
Rivers	*Upstream from Chipps Island*
Suisun Bay	*Chipps Island to Benicia Bridge*
Carquinez	*Benicia Bridge to Carquinez Strait* (region not to be sampled)
San Pablo Bay	*Carquinez Strait to Point San Pablo/Pedro*
Central Bay	*Point San Pablo/Pedro to San Bruno Shoal*
South Bay	*San Bruno Shoal to Dumbarton Bridge*
Lower South Bay	*South of Dumbarton Bridge*

Chapter 3 - Application of Statistical Power Analysis for Determining Sample Sizes for the Random Sampling Design

Robert Smith, Ojai, CA
Sarah Lowe, SFEI

Introduction

Chapter 2 describes a new sample frame for the new RMP stratified random sampling design. Within this design, it is necessary to choose sample sizes sufficiently large for meeting the RMP objectives. This chapter uses statistical power analyses with available RMP data to determine the sample size for each Estuary sub-region based on comparing specific contaminants of concern to water quality and sediment quality guidelines. This exercise will examine the relationship between sample sizes and the power of the test and will be used to guide decisions in choosing adequate sample sizes for each sub-region in the Estuary for the purposes of the monitoring program.

One of the RMP objectives is to determine the status of chemical contaminants in the Bay. An important management question here is whether chemicals in the water column and sediments are found in sufficiently high concentrations to cause significant environmental harm. Ideally, the sample sizes chosen for the sampling design should be able to address this question.

The concentrations of chemical contaminants vary in time and space in the Estuary. To provide focus on the sub-regions of the Estuary and specific periods of time, our main interest here is whether *the **mean** chemical concentrations for a **sub-region*** (defined in Chapter 2) *exceed water quality (or sediment quality) **guideline** concentrations **during a survey period***. The guidelines used in these analyses were chosen by the Re-design Work Group based on consensus of their relevance to the current conditions in the Estuary. These guidelines are commonly-used standards that have been determined to be potentially harmful to living organisms when exceeded.

The assumed underlying mean contaminant concentration for each sub-region, computed from the monitoring data, will only be an estimate of the true regional mean since measurements from only a relatively small number of locations within each sub-region will be used to generalize to the status of the entire sub-region. A statistical test will be applied to determine whether an assumed underlying regional mean for a chemical constituent is above or below guideline values indicating potential environmental harm. Given the uncertainty in our regional mean estimates, a statistical test is used to control the rates of making wrong conclusions when comparing the estimates with the guidelines. The uncertainty associated with the assumed underlying regional means will decrease as the sample size in the sub-region increases. The power analyses will estimate the effects of different sample sizes on the sensitivity of the statistical test to distinguish between real differences (from guideline values) and differences due to random background noise. Such information is critical in designing a monitoring program, since it can indicate the level of effort that may be required to produce meaningful information.

Methods

Guidelines

The guidelines used in these analyses were the U.S. Environmental Protection Agency's proposed California Toxics Rule (CTR, U.S. EPA, 1997), the Effects Range Low and Effects Range Medium (ERL and ERM; Long et al. 1995, Long et al. 1998), and the mercury TMDL (Total Maximum Daily Load) sediment target developed by the Regional Board (Johnson & Looker, 2004). For a chemical in the water column, a Water Quality Criterion (WQC) indicates the concentration where a chemical may begin to cause harm to the exposed organisms. In sediment, the chemical concentration corresponding to the ERL has been associated with sediment toxicity (in bioassay tests) about 10% of the time, and the sediment chemical concentration corresponding to the ERM has been associated with toxicity about 50% of the time. The Regional Board proposed a new sediment target for total mercury in the Estuary as part of the TMDL process (Johnson & Looker, 2004). Sediment concentrations were also compared to this mercury TMDL target.

Data

For each sub-region and chemical constituent of interest, the power analyses require input of an assumed underlying mean and variance of the chemical measurements. We have no way of knowing the actual assumed underlying means and variances, therefore historical estimates computed from data collected for the RMP Pilot and Status and Trends program between 1989 and 1998 were used. It should be noted that a random sampling design is required to assure that the estimates of the means and variances are unbiased, but the RMP sampling design was not based on random sampling. Therefore, the estimates used may be biased and lead to inaccurate power estimates. However, using these potentially biased estimates is better than proceeding without any information on the potential sensitivity of our sampling design. More appropriate power tests can be applied in the future when data are available from the new random monitoring design described in Chapter 4.

Initially the power analyses were run for both water and sediment contaminants using dissolved and total concentrations for all ten RMP metals (Ag, As, Cd, Cu, Fe, Hg, Mg, Pb, Se, and Zn) and total (dissolved + particulate) concentrations for the major organic contaminants (total Chlordanes, DDTs, PCBs, and PAHs) measured by the RMP. Results from those analyses are reported in separate tables (by contaminant for both water and sediment) in Appendix 3A. Results from the initial analyses indicated that, for many contaminants, relatively few water samples per sub-region were needed to achieve 80% or greater power. That meant that the mean contaminant concentrations in many sub-regions were clearly above or below the guidelines. In an effort to simplify the power evaluation process (where dozens of parameters were initially evaluated) the Re-design Work Group chose to focus only on those contaminants currently of most concern in the Estuary that demonstrated some change in power with a relatively small change in sample size. Through presentations of the initial results and discussions with the Re-design Work Group, four water and sediment contaminants were chosen by the Re-design

Work Group for reasons outlined below for use in the final analyses: dissolved copper (for water); copper, total mercury, and total PAHs (for sediment).

Dissolved copper was an ideal surrogate contaminant for the water column, because it was generally found at ambient levels near or above the WQC guideline and therefore most sensitive in the power analyses. Additionally it was an important contaminant of concern for management reasons. For sediments, mercury, copper, and total PAH were selected. Mercury was selected because TMDL development for mercury is more advanced compared to other 303(d) listed contaminants. Copper was selected for sediment power analyses because the South Bay Copper Action Plan highlighted the need to understand sediments as a potential source of dissolved copper to overlying waters (Copper Action Plan, 2001). For this, we need to be able to detect changes in the concentration of copper in sediments. Total PAH's were selected as organic constituents that are close to guideline levels, and therefore potentially sensitive to the sampling design. PAHs are also on the TMDL "watch-list" because of their high concentrations near guideline levels. Impairment by bioaccumulative substances, such as PCBs, is assessed using bivalve and fish tissue concentrations, so PCBs were not used in the power analysis for sediments.

The statistical test

The tests addressed whether, within a survey period, the mean concentration of a chemical constituent in a sub-region was above the guideline for causing potential environmental harm. The null hypothesis of the statistical test addressing this question is that the actual mean concentration (μ) is *above* the guideline value. The test procedure involves estimating the regional mean ($\overline{x}$) and its one-tailed upper 95% confidence interval. If the upper confidence interval of the mean crosses the guideline value, then the null hypothesis is accepted (Figure 3.1).

Figure 3.1. Illustration of the statistical test comparing the mean chemical concentration of a sub-region with a guideline value. The $\overline{x}$ represents the estimated value of a regional mean, and the

error bar above $\overline{x}$ is based on the one-tailed upper 95% confidence limit of the mean. The Null Hypothesis (H_0) is that the underlying regional mean (μ) is above the guideline value.

This null hypothesis is conservative in the direction of environmental protection in that the regional mean is not considered below the guideline until the upper confidence bound of the mean is below the guideline value. Due to sampling error, we don't know the value of the actual regional mean, but we expect that the upper 95% confidence bound of the mean will exceed the actual mean (μ) about 95% of the time. This means that we do not accept that the mean concentration is below the guideline until we are about 95% certain that we are correct.

A Statistical Test for Temporal Trend

The statistical test comparing the regional means with the guidelines are useful for determining which constituents are sufficiently below the guideline to not be of concern. When constituent means are above or very near the guideline, measures for lowering the level of the constituent in the environment may be implemented. In this case, it will be of interest to observe whether the level of the constituent is decreasing over time. Here a statistical test for temporal trend would be appropriate. Fryer and Nicholson (1993) describe a statistical model detecting linear temporal trends and provide a method for computing the associated power. In this report, we do not compute the power for detecting trends, but will wait until data from a few future surveys are available, and we will be able to compute better estimates of the required within- and between-year variance components.

Power analysis

The power of a statistical test is the relative frequency that the test will reject the null hypothesis when the null hypothesis is actually false. For example, if the power of our statistical test (Figure 3.1) is computed to be 80%, then we expect that we would reject the null hypothesis (i.e., the actual chemical mean concentration for a sub-region is above the guideline) 80% of the time when the assumed underlying mean concentration for a region is actually below the guideline.

The specific power calculations are described in more detail in Appendix 3B. Here we discuss the more important inputs to the power calculations. The power of a particular test can depend on several factors, including the details of the sampling design and statistical model, the nominal type-1 error rate, the sample size, the magnitude of the differences to be detected by the test, and the underlying magnitude of the pertinent random error in the data. The sampling design and the statistical model are already established, and the underlying error variability in the data is not under our control. The remaining factors affecting the power are discussed separately below.

Nominal type-1 error rate

The nominal type-1 error is the maximum acceptable rate of rejecting the null hypothesis when it is actually true. For our test, a type-1 error would involve assuming that the actual chemical mean (μ) is below the guideline when it is actually above the guideline. The type-1 error rate has been set by the choice of the 95% confidence intervals in our statistical test. Potential type-1 errors are associated with the 5% of the time that the upper confidence interval bound does not exceed the actual underlying sub-region mean (μ). Thus, our maximum type-1 error rate is 5%.

We have chosen to fix the type-1 error rate at the relatively low value of 5% to minimize possible environmental harm caused by assuming mean chemical concentrations are below the guideline when in fact they are not.

Increasing the nominal type-1 error rate will increase the power of the test, since the null hypothesis will be rejected more often with a higher type-1 error rate.

The sample size

Larger sample sizes will lead to narrower confidence intervals of the mean in our statistical test. Shorter confidence intervals will allow for more frequent rejection of the null hypothesis and in turn will provide higher power for the test. The main purpose of this exercise is to study the relationship between sample sizes and the power of the test. This information will be useful in choosing sample sizes that produce sufficient power for the purposes of the monitoring program.

The magnitude of the differences to be detected by the test

The farther the actual mean chemical concentration (μ) for each sub-region is below the guideline concentration, the greater will be the power of the test. When computing the power of a test, we need to provide a value for the assumed underlying sub-regional mean (μ in Appendix 3B). We do not know the value of the underlying mean, but for the purpose of the power tests, we obtained a ball-park value from the available historical data. We used the following rules for choosing the underlying sub-region mean. 1) If the historical regional mean was above the guideline or less than 10% below the guideline value, then we assumed that the underlying regional mean was 10% below the guideline value. 2) If the historical regional mean was more than 10% below the guideline value, then the computed historical mean was used.

Underlying mean values (μ) less than the 10% below the guideline were not used in these analyses because unreasonably large sample sizes are required to obtain sufficient power as μ and the guideline value converge. The 10% value was (subjectively) chosen as the smallest reasonable amount of change that could usually be detected without unreasonable sample sizes.

Given the null hypothesis that we chose, we accepted the fact that the regional mean contaminant values need to get somewhat below the guideline values before we will consistently reject the null hypothesis and consider the chemicals to be below their guideline levels. This is the price we pay for our more environmentally protective null hypothesis.

Results

The estimated power associated with different sample sizes are shown in Tables 3.1 to 3.3. For estimating the power associated with sample size in water samples from the various sub-regions dissolved copper was the target contaminant. Dissolved copper concentrations in the Estuary are generally found at ambient levels near or above the WQC guideline making copper a good surrogate for other contaminants in this exercise.

Table 3.1 shows the percent power achieved with between two and ten water samples for dissolved copper compared to the WQC in the major sub-regions of the San Francisco Estuary during the wet season (February), receding flow period (April), and the dry season (August). Samples from the Rivers region were compared to the freshwater WQC while samples from the other regions were compared to the saltwater WQC. In the northern and central Estuary greater than 80% power was achieved with relatively few samples in the dry season (2-4 samples per region). This is because the historical means were well below the WQC. In the South Bay region of the Estuary, where the historical mean approached the WQC, even with ten samples we can not achieve 80% power. In the Lower South Bay region 6 samples would achieve 80% power when sampled during the dry season.

Table 3.1. Estimated percent power with water sample sizes of 2 to 10 for ***dissolved copper*** and the ***WQC guideline*** in the major sub-regions of the San Francisco Estuary. Both the historical mean (Historical Mean) and the assumed underlying mean (μ) are shown. When the historical mean is < 10% below the guideline the historical mean is used for μ . Otherwise μ is assumed to be 10% below the guideline. Power computations were based on lognormal (L) and normal (N) distributions, respectively. Seasons: wet-(Nov-March); rec-(receding flow in April); dry-(May-Oct). Concentrations are in µg/L.

Dissolved Copper						**% Power at Sample Size of**								
Sub-Region	Season	Distribution	Historical Mean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	dry	L	1.8	1.8	9	100								
Rivers	rec	L	1.7	1.7	9	100								
Rivers	wet	L	2.0	2.0	9	96	100							
Suisun Bay	dry	L	2.0	2.0	3.1	57	98	100						
Suisun Bay	rec	L	1.9	1.9	3.1	92	100							
Suisun Bay	wet	L	2.0	2.0	3.1	63	99	100						
San Pablo Bay	dry	L	1.9	1.9	3.1	29	60	82	92	97	99	100		
San Pablo Bay	rec	L	1.8	1.8	3.1	28	57	78	90	95	98	99	100	
San Pablo Bay	wet	L	1.9	1.9	3.1	27	56	77	89	95	98	99	100	
Central Bay	dry	L	1.4	1.4	3.1	43	86	98	100					
Central Bay	rec	L	1.4	1.4	3.1	51	94	100						
Central Bay	wet	L	1.5	1.5	3.1	40	83	97	100					
Golden Gate	dry	N	0.6	0.6	3.1	89	100							
South Bay	dry	N	2.9	2.8	3.1	13	20	28	35	41	48	53	58	63
South Bay	rec	N	2.7	2.7	3.1	15	24	34	43	51	58	64	70	74
South Bay	wet	N	2.1	2.1	3.1	67	100							
Lower South Bay	dry	L	3.6	2.8	3.1	22	42	60	73	82	89	93	96	97
Lower South Bay	rec	L	3.6	2.8	3.1	12	19	26	32	39	44	49	54	59
Lower South Bay	wet	L	2.3	2.3	3.1	67	100							

Table 3.2 shows the percent power achieved for three targeted contaminants with between two and ten sediment samples in the major regions of the San Francisco Estuary. Copper (Cu), total mercury (Hg), and total PAH (TPAH) concentrations were compared to the sediment ERL guideline in this exercise. Even with ten samples per region for the two metals evaluated, generally less than 50% power was achieved. This is because the mean regional contaminant concentrations in the Estuary were close to the ERL guidelines making it difficult to determine that the mean concentrations were different from the guideline values given the observed regional variations. For TPAH three or fewer samples achieved at least 80% power in most regions except San Pablo Bay and Central Bay. The reason those regions had low power for TPAH was because of the extremely high observed variances in those regions.

Table 3.2. Estimated percent power with sediment sample sizes of 2 to 10 for ***copper (Cu), total mercury (Hg), and total PAHs (TPAH)*** and the ***ERL guideline*** in the major sub-regions of the San Francisco Estuary. Both the historical mean (Historical Mean) and the assumed underlying mean (μ) are shown. When the historical mean is < 10% below the guideline the historical mean is used for μ . Otherwise μ is assumed to be 10% below the guideline. Power computations were based on lognormal (L) and normal (N) distributions, respectively. All estimates are for the dry season only. Cu and Hg are measured in mg/Kg dry weight, and TPAH is measured in µg/Kg dry weight.

						% Power at Sample Size of								
Sub-Region	Parameter	Distri-bution	Historical Mean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers	Cu	N	29	29	34	11	16	21	26	30	35	39	43	46
	Hg	L	0.11	0.11	0.15	9	11	13	16	18	20	22	24	26
	TPAH	L	72	72	4022	56	97	100						
Suisun Bay	Cu	L	31	31	34.0	7	7	8	9	10	10	11	11	12
	Hg	N	0.2	0.1	0.15	6	7	7	7	8	8	9	9	9
	TPAH	N	423	423	4022	99	100							
San Pablo Bay	Cu	L	39	31	34	7	8	9	10	10	11	12	13	13
	Hg	L	0.22	0.14	0.15	6	7	8	9	9	10	11	11	12
	TPAH	L	890	890	4022	18	32	46	58	67	75	81	86	90
Central Bay	Cu	N	39	31	34.0	7	7	8	9	10	10	11	12	12
	Hg	L	0.21	0.14	0.15	7	8	9	10	10	11	12	13	13
	TPAH	N	2920	2920	4022	10	13	17	20	23	26	29	32	35
South Bay	Cu	N	38	31	34.0	16	28	39	50	59	66	73	78	82
	Hg	L	0.29	0.14	0.15	11	15	20	24	29	33	36	40	44
	TPAH	L	2084	2084	4022	31	65	86	95	98	99	100		
Lower South Bay	Cu	N	40	31	34	12	18	23	29	34	39	44	48	53
	Hg	L	0.33	0.14	0.15	12	19	25	31	37	43	48	53	57
	TPAH	L	1881	1881	4022	45	89	99	100					

Copper, total mercury, and total PAH concentrations in sediment were also compared to the sediment ERM guidelines in this exercise. However, the results were not considered by the Re-design Work Group when formulating the final sample size recommendations because the historical mean concentrations were well below guideline levels and greater than 80% power was achieved with only 2 samples per region.

Table 3.3 shows the percent power achieved with between two and ten sediment samples for total mercury (Hg), normalized by percent fines, and the TMDL guideline for the major regions of the San Francisco Estuary. Results show that even with ten samples per sub-region very little power is achieved with the exception of Suisun Bay.

Table 3.3. Estimated percent power with sediment sample sizes of 2 to 10 for ***total mercury (Hg)*** normalized by percent fines and the ***TMDL target*** in the major regions of the San Francisco Estuary. Both the historical mean (Historical Mean) and the assumed underlying mean (μ) are shown. When the historical mean is < 10% below the guideline the historical mean is used for μ . Otherwise μ is assumed to be 10% below the guideline. All estimates are for the dry season only. Units are in mg/Kg dry weight.

				% Power at Sample Size of								
Sub-Region	Historical Mean	μ	TMDL	2	3	4	5	6	7	8	9	10
Rivers	0.35	0.35	0.4	7	9	10	11	12	13	14	15	16
Suisun Bay	0.29	0.29	0.4	17	30	43	54	64	71	78	83	87
San Pablo Bay	0.40	0.36	0.4	6	7	7	8	8	9	9	10	10
Central Bay	0.36	0.36	0.4	7	8	10	11	12	13	13	14	15
South Bay	0.37	0.36	0.4	11	16	22	27	31	36	40	44	48
Lower South Bay	0.44	0.36	0.4	8	10	12	13	15	16	18	19	21

Recommendations

Based on general discussions and the power analysis results, the Re-design Work Group made decisions about the sample size allocations for each sub-region in the new sampling design. The recommendations were based on the following considerations: 1) areas of greater regulatory concern (e.g. South Bay, Lower South Bay), 2) statistically adequate power for key contaminants (Tables 3.1-3.3), and 3) funding constraints.

The Work Group decided on 28 randomly allocated water samples and 40 randomly allocated sediment samples to be sampled annually during the dry season with additional sampling at a subset of historical RMP sites (as mentioned in the previous chapter). The sediment sample size per region was limited to a maximum of 8 by funding considerations and not as a result of achieving adequate power.

Table 3.4 summarizes the recommended sample sizes by region, and Table 3.5 gives the rationale for the sample size allocations. In interpreting Tables 3.1-3.3 for making sample-size recommendations, we considered power around 80% (or higher) to be an acceptable amount of certainty.

Table 3.4. Recommended water and sediment sample sizes for the major sub-regions of the San Francisco Estuary.

Sub-Region	**Water Column**	**Sediment**
Napa River	0	0
Rivers (historic stations)	2	2
Suisun Bay	4	8
Carquinez Strait	0	0
San Pablo Bay	4	8
Central Bay	4	8
South Bay	10	8
Lower South Bay	6	8
Additional historic RMP stations	1	5
Total	**31**	**47**

Table 3.5. Recommended sample size for each targeted sub-region for water and sediment monitoring in the San Francisco Estuary and workgroup rationale. It was decided that a subset of historic RMP sites would continue to be monitored to provide program continuity in the long-term trends dataset started in 1993.

Sub-Region	Sample Type	Sample Size	Rationale
Rivers	Water	0	Only two sampling units needed to achieve > 80% power for dissolved copper. The workgroup decided to only sample the two historical Rivers sites in this region for fiscal reasons. It was also acknowledged that this region is a major pathway for contaminant loading to the Estuary from the Central Valley and that this region was largely monitored by the Central Valley Regional Board. Additional pilot and/or special studies could augment sampling in this region if warranted.
	Sediment	0	For copper and mercury, an inordinate number of sampling units were required to obtain adequate power for the ERL and TMDL guidelines.
	Historical	2	Keep the two historic water and sediment sites. Adjustments should be made to ensure that each sample is representative of each river system.
Suisun Bay	Water	4	For WQC guideline, three sampling units were sufficient to achieve > 98% power for dissolved copper. However, workgroup members were more comfortable with a minimum of four samples per region.
	Sediment	8	With the mercury TMDL guideline, 8 or 9 sampling units would achieve substantial power. However, with the ERL guidelines for copper and mercury, an inordinate number of sediment sampling units are needed to achieve substantial power.
	Historical	1	Keep the Grizzly Bay (BF20) station. This site has had on-going water, sediment, water & sediment toxicity, and bioaccumulation studies (test organism = clams). It has also been used in toxicity TIE studies in recent years.
Carquinez Straight	Water and Sediment	0	Workgroup members decided not to monitor in this region as it is a logistically difficult region to sample and both the inputs (Suisun Bay) and the outputs (San Pablo Bay) would be monitored.
San Pablo Bay	Water	4	For the WQC guideline, five sampling units are sufficient to achieve > 80% power for dissolved copper. A study is currently underway to characterize the status of dissolved copper in San Pablo Bay. Information from this study would adequately compliment RMP monitoring efforts.
	Sediment	8	For all chemical constituents and guidelines, an inordinate number of sediment samples are needed to achieve high power.
	Historical	1	Keep the Pinole Point station (BD30/BD31). This site has had on-going water, sediment, water toxicity, and bioaccumulation studies (test organism = mussels). Mussels have been the bioaccumulative test organisms at this site and could be the species used for site comparison purposes to other monitoring sites (except possibly Grizzly Bay where mussels may not survive even in the dry season).

Table 3.5 (cont.)

Sub-Region	Sample Type	Sample Size	Rationale
Central Bay	Water	4	For the WQC guideline, three sampling units are sufficient to achieve > 95% power for dissolved copper. However, workgroup members were more comfortable with a minimum of four samples per region.
	Sediment	8	For the ERL and TMDL guidelines, an inordinate number of sediment sampling units are needed to achieve high power.
	Historical	1	Keep Yerba Buena Island (BC10/BC11). This site has had on-going water, sediment, sediment toxicity, and bioaccumulation studies (test organism = mussels). Additionally, this site has been the focus of recent PCB studies by Jay Davis of SFEI.
South Bay	Water	10	For the WQC guideline with dissolved copper, reasonable power obtained with ten sampling units.
	Sediment	8	For the ERL guidelines for copper and TPAH there was reasonable power with eight sampling units. For mercury and the TMDL guideline, moderate power with eight sampling units.
	Historical	1	Keep Redwood Creek (BA40/BA41). This site has had on-going water and sediment toxicity studies, along with bioaccumulation studies (test organism = mussels).
Lower South Bay	Water	6	For the WQC guideline, high power in wet and dry seasons with six sampling units.
	Sediment	8	For mercury with the ERL and TMDL guidelines, an inordinate number of sampling units needed for high power.
	Historical	1	Keep Coyote Creek (BA10/BA10). This site has had on-going water, and sediment studies, along with bioaccumulation studies (test organism = clams).

References

Copper Action Plan, 2001. Environmental Services Department. City of San Jose San Jose/Santa Clara Water Pollution Control Plant. San Jose, CA
URL: http://www.ci.san-jose.ca.us/esd/PDFs/Copper%20Action%20Plan%20Report%20-%20March%202001.pdf

Fryer R. J. and M. D. Nicholson. 1993. The power of a contaminant monitoring programme to detect linear trends and incidents. International Council for the Exploration of the Sea (ICES) Journal of Marine Science, 50:161-168.

Long, E. R., L. J. Field, and D. D. MacDonald. 1998. Predicting toxicity in marine sediments with numerical sediment quality guidelines. *Environmental Toxicology and Chemistry* 17:714-727.

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.

Johnson B., and R. Looker, 2004. Mercury in San Francisco Bay Total Maximum Daily Load (TMDL) Proposed Basin Plan Amendment and Staff Report. California Regional Water Quality Control Board San Francisco Bay Region. Oakland, CA

U.S. EPA. 1997. The California Toxics Rule (CTR): Establishment of Numeric Criteria for Priority Toxic Pollutants for the State of California. Federal Register: Vol. 62, No. 150. August 1997. U.S. Environmental Protection Agency.

Chapter 4 – SAMPLING DESIGN AND SPATIAL ALLOCATION

Don Stevens, Oregon State University

Abstract

The objective of this work is to design an environmental sampling program for the San Francisco Estuary Regional Monitoring Program (RMP). The approach includes using a probability based sampling method to optimize the identification of patterns and trends in contaminant concentration and distribution in water and sediment of the San Francisco Estuary. The RMP sampling design is based on the Generalized Random Tessellation Stratified design (GRTS) used by EPA's Environmental Monitoring and Assessment Program (EMAP). This type of probability sampling structure is appropriate for the RMP's goal of determining both status and trends in the Estuary by removing the bias of historic sampling sites and providing a spatially balanced random sample of the Estuary, with increased resolution over time.

A properly designed probability sample should be guided by prior knowledge about the population in question and should address the goals of the sampling. The identification of the geographic sub-regions in the Estuary, described in chapter 2, incorporates prior information about the Estuaries hydrology into the sample design. The outer boundary of the Estuary was defined at the 3-foot and 1-foot depth contours for water and sediment sample populations respectively (based on mean lower low tide bathymetry information). Defining the target population (water and fine-grained sediments), and distributing the sampling sites within each sub-region with spatially balanced regularity creates a sampling structure that addresses both the study objectives (status & trends) and prior information. The number of sampling sites selected within a sub-region, for a discrete sampling period, varies based on results from the power analyses, described in chapter 3, and decisions by the Re-Design work group.

The GRTS design is such that sampling sites within a sub-region and across sub-regions are not independent of one another. Inferences can be made about the condition of each sub-region or about the whole Estuary. Additionally, the sampling sequence is designed to provide a spatially balanced random sample of each sub-region (or the Estuary as a whole) at any given time in the monitoring effort.

Trends monitoring is usually accomplished by taking repeated measures of the target population at the same sites over time. This is especially true if sites within the target population retains much of their identity over time (e.g., sediment) but not so if the population exhibits no temporal identity (e.g., water). Because of this, the water sampling design add new sites every year (up to 100 total randomly allocated sites per sub-region) with no repeat sampling of the randomly allocated sites. The sediment sampling design also adds new sites every year (up to 72 total randomly allocated sites per sub-region), but in addition has sites that are revisited. The revisit sites panels of two randomly choses sites each that are visited annually, once every 5 years, once every 10 years, or once every 20 years. The visit interval was chosen based on the assumption that trends in contaminant concentrations in the Estuary are relatively small compared to spatial and inter-annual variation. As a result it will take at least 20 years of sampling to realize the utility for trend detection of the revisit patternin the design . All areas of fine-grained sediments (for the target sediment population) within the Estuary are not known prior to sampling. Therefore random sites selected for sampling may need to be discarded due to unsuitable substrate. Over-sample (replacement) sites have been provided to replace the discarded sites.

A major advantage of using a probability based sampling method is that a variety of statistical tools may be used for data analysis and population inference. These statistical tools include pure design-based survey methodology, model-assisted design-based methods, classical model-based methods, and Bayesian techniques. All tools can be used for both the sediment and water populations with the caveat that sediment analysis has additional techniques due to the rotating/repeating panel structure.

Targeted versus Probability Samples for Monitoring

The RMP monitors contaminant concentrations in water, sediment, and fish and shellfish tissue in San Francisco Estuary. One of the objectives of the RMP is to describe patterns and trends in contaminant concentration and distribution in the Estuary. Because the contaminant concentration over the entire Estuary cannot be directly observed, we must infer its characteristics based on observation of some subset of the Estuary, that is, a sample. The only two scientifically defensible methods for extrapolating from a subset to an entire population are (1) base the inference on explicit specification of the relationship between the subset and the entirety; and (2) select the subset via a probability sample and use design-based inference methods.

Alternative (1) is usually referred to as "model-based inference", where "model" is used as a generic term for a description of the relationship between the subset and the population. The model may be something as simple as an assumption that the observations are "representative" of the population to the extent that the mean of the sample should be close to the mean of the population. In one approach, the sites are selected based on their anticipated ability to reflect regional characteristics. The site features used in site selection may be physical characteristics, spatial pattern, expected sensitivity to stress, anticipated exposure level, or any other aspect that might influence the response of the site to known or suspected environmental stresses. The National Oceanic and Atmospheric Administration's (NOAA) National Status and Trends Program (NS&T) is an example of such an approach to site selection. The sites for the NS&T were explicitly selected `...to be "representative" of large coastal areas and to avoid small-scale patches of contamination' (O'Connor, 1990). The model may also be a complex set of equations that use sophisticated statistical techniques (spatial statistics, analysis of variance, regression, principal components, ordination) to infer population characteristics from the sample.

Model-based inference is very common, and it plays a critical role in advancing the state of scientific knowledge. In a very real sense, the model is often the embodiment of our current state of theoretical understanding. The model describes the relationships between the response and those attributes that we can measure and that we believe influence the response.

The advantage of model-based inference is that it enables very general and precise inference from limited data. For example, if we assume that the population response follows a normal distribution, we need only estimate the mean and standard deviation to infer the shape of the entire population distribution. Using only the mean and the standard deviation, we can tell how much of the population meets some regulatory criterion. In a sense, the inference "borrows strength" from the model: the model structure provides the framework for the inference, and the precision of the inference is judged relative to the model.

The difficulty with model-based inference stems from the same basis as its advantage: the structure of the model. If the model is not a faithful description of reality, the population inference based on the model may have little resemblance to the true

population value. Moreover, because precision is judged relative to the model, there may be no indication that the inference is substantially in error. For example, the model may fit observed data very well, resulting in apparently high precision for population estimates. However, if the observed data are from a select subset of the population that conforms to the model while the rest of the population does not, the extrapolation and its apparent precision may be substantially in error. In any case, inference to an associated population rests on the assumption that the behavior of the selected sites reflects and is typical of the behavior of the population. In a long-term monitoring program aimed at status and non-specific change in a spatially distributed population subjected to non-uniform stresses, that assumption does not seem tenable under a judgmental sample selection protocol. The National Research Council (NRC), in commenting on NOAA's NS&T program, stated:
"There is some question about how representative any isolated site can be of wider regional conditions when it is located in an area where there may be a range of pollution concentration or considerable local variability in the processes controlling the transport and distribution of contaminants. These limitations lead to the conclusion that the NS&T Program may be more useful in measuring temporal trends at individual stations than in assessing national status of the marine environment or in comparing the extent and severity of pollution among regions in any precise way. This limitation, in turn, can lead to misinterpretation of the significance of the program's findings." (NRC, 1990)

The remaining alternative for expanding sample data to a population is to pick the sample using a probability-based selection method, and then apply inference techniques consistent with the selection method. This is usually referred to as “design-based” inference because the resulting inference draws its generality and validity from the design, and not from any presumption of the correctness of a model or representiveness of the sample sites. Provided the design is properly applied, the resulting inference is unassailable and irrefutable. (See Hansen et al. (1983), Särndal (1978), or Smith (1976) for discussions of the issues involved in design-based versus model-based inference. These issues were also discussed in a spatial context by de Gruijter and Ter Braak (1990) and Brus and de Gruijter (1993).). A probability sample allows the use of both design-and model-based analyses. Moreover, even if the model-based analysis does not make explicit use of the probability structure of the sample, the model-based inference is strengthened by the characteristics of a probability sample.

Some states have also applied probability sampling in parallel to purposive sampling, in some cases with eye-opening results. For example, Jacobs and Cooney compare the results of probability and purposive sampling of Oregon Coastal streams to estimate the Coho salmon escapement. The purposive sampling based estimates drastically over estimated the size of the Coho salmon population. (Jacobs and Cooney, 1995) Peterson, Urquhart, and Welch (1999) contrasted the distribution of turbidity inferred from a convenience sample of lakes in the eastern US with the corresponding inference from a probability sample, and found substantial discrepancies.

Hansen et al., (1983) makes several relevant points in their discussion of design-versus model-based inference. One that is particularly relevant for an environmental monitoring program is that a probability sample permits inferences that are free of even the appearance of subjectivity. A probability sample from an explicitly defined resource population is a means of certifying that the data collected are free from any selection bias, conscious or not. Furthermore, analysis methods that are as free as possible from the appearance of subjectivity are also available under a design-based protocol

Design Considerations

A key property of a probability sample is that every element in the population has some chance of being included in the sample. If this were not the case, then some parts of the population might as well not exist, since no matter what, their condition could have no influence on estimates of population characteristics. This property has a side benefit, in that it forces an explicit and complete definition of the population being described. The definition of the target population should incorporate both the physical attributes of the population and the operational aspects of the sampling method. The target population must conform to our scientific conceptualization and include only those elements that can actually be sampled. The definition is used to create a construct, called a *population frame*, which contains all of the population of interest and makes it available for sampling.

A common misconception is that a probability sample is the same as a "simple random sample", where all prior information and knowledge about the population are ignored, and sample points are selected independently of one another. On the contrary, a properly designed probability sample can and should incorporate prior knowledge and insight into the population. Moreover, the design should be explicitly directed to the objective of the study.

One method of focusing a sample on project goals and incorporating prior information is stratification, which partitions the sample among disjoint subpopulations. Stratified sampling is frequently used to ensure that subpopulations of particular interest are allocated sufficient samples. A sample can also be structured to take advantage of other population features, e.g., spatial pattern, and to provide higher statistical power to meet project objectives, e.g., use of repeated measures to detect trend or change. Relevant population features, project objectives and the statistical techniques used to focus the sample are discussed in the following.

Spatial extent of target population and sample frame

The objective of the RMP is to assess status and trends for both water and sediment in the San Francisco Estuary. The outer boundary of the Estuary was defined by mean low low tide bathymetry. Different depth contours were used for the water column and sediment samples. The water column population was defined by the 3-foot depth contour, and the sediment population by the 1-foot depth contour.

Furthermore, the Estuary was partitioned into several sub-regions (Central Bay, Lower South Bay, San Pablo Bay, South Bay, Southern Slough, and Suisun) for which individual assessments were needed. (The process used to define these sub-regions is described in Chapter 2 of this report). The definition of the sub-regions was incorporated into a Geographic Information System (GIS) coverage. Two additional sub-regions in the partition (Carquinez Straits, and Rivers) were defined but no random sampling is planned in those regions at this time. Because each sub-region is a separate reporting unit, the design was constrained to ensure a specified number of sample sites within each region (as described in Chapter 3). Conceptually, the sub-regions are similar to strata.

Spatial regularity of sample points

Most environmental populations have appreciable spatial structure. Nearby elements can interact with one another, and tend to be influenced by the same set of natural and anthropogenic factors. We assume that both the water and sediment populations will exhibit substantial spatial pattern. Whether the pattern is an irregular mosaic or a smooth gradient, a sample point pattern that is more or less spatially regular will tend to be more efficient (lower variance for the same number of samples) than a completely random sample. (See,for example, Munholland and Borkowski (1996), Breidt (1995), Iachan (1985), Olea (1984), Bellhouse (1977), Dalenius *et al.* (1961), Matérn (1960), Das (1950), Quenouille (1949), Cochran (1946)).

The extreme case of a completely regular sample, e.g., at the nodes of a regular, systematic grid, works very well for populations with smooth gradients, but does not work well for periodic patterns or irregular mosaics. The US Environmental Protection Agency's Environmental Monitoring and Assessment Program (EMAP) (Overton, White, and Stevens, 1991; Larsen, *et al.*, 1991; Larsen, *et al.*, 1994; Stevens, 1994; Herlihy, *et al.*, 2000;) uses a sample design called a Generalized Random Tessellation Stratified design (GRTS) (Stevens, 1997; Stevens and Olsen, 1999; Stevens and Olsen, 2000, Stevens and Olsen, 2002, in review) to achieve a spatially-balanced point distribution that is nonetheless random. The GRTS design captures much of the potential efficiency of a completely regular design for any spatially patterned response.

Briefly, the GRTS design achieves a random, nearly-regular sample point pattern via a random function that maps 2-dimensional space, e.g., a square, onto a 1-dimensional line. The function is defined recursively in a manner that preserves some 2-dimensional proximity relationships; in particular, the images of two points that are close together in 2-space will tend to be close together in the 1-dimensional (linear) space. A systematic sample is selected in the linear space, and the sample points are mapped back into 2-space via the inverse of the random function. The resulting sample will be nearly regular in 2-space because of the proximity-preserving property of the random function. Sampling of an irregular 2-dimensional object, such as San Francisco Estuary, is accomplished by enclosing the object in a square, and then discarding points outside the object. Details of the construction of the random function and sample selection are provided in Stevens and Olsen (2004).

The Estuary sub-regions are conceptually similar to strata in that they form separate reporting units, and have specified sample sizes. However, in the GRTS design, the requirement for specified sample sizes within segments is satisfied by using variable probability instead of formal stratification. This is accomplished by differentially weighting intervals on the linear image of the Estuary that correspond to the various segments. One of the major differences between this method and stratification is that the location of sample points is not independent across sub-region boundaries. The spatial regularity requirement extends across regional boundaries, so that the location of sample points in one sub-region influences the location of sample points in adjacent sub-regions. If the regional boundaries were physical barriers that precluded any interaction of water or sediment on opposite sides of the boundary, then strict stratification would be appropriate. However, this is not the case, so extending the regularity requirement across the boundary will result in a more precise inference to the entire estuary.

Sampling fine-grained sediments

The sediment sample is restricted to fine-grained sediments, and the extent of fine-grained sediment in the Estuary is not precisely known. A common sampling technique in a case where the target population cannot be precisely delimited is to select a sample from a frame that is known to include all of the target population, plus possibly some non-target elements; and then to discard non-target points as they are identified. In the present case, in order to ensure that all fine-grained sediments were subject to being sampled, potential sample points were selected using a GIS coverage of the entire Estuary, including both fine- and coarse-grained sediments. All potential sample points will be visited, and those that do not have fine-grained sediment will be discarded. The final sample will then consist of the potential sample points that fall on fine-grained sediments.

The GRTS design has a feature called sequential spatial balance that makes it particularly suitable for the RMP sediment samples. The recursive property of the random map from 2-space to the line can be used to order the selected sample points in such a way that the collection of the first m points of an n-point sample ($m \le n$) is in itself a nearly regular but random sample. Furthermore, the first m points in fine-grained sediments will be a nearly regular sample from the population of fine-grained sediments. (Stevens and Olsen,2004). Thus, even though the extent of the target population is not known prior to sample selection, the resulting sample will be evenly distributed over the actual extent.

Provision for Trend Detection

In many studies designed to assess trend, there is some value in repeated measurements of the same population element over time. This advantage of repeated measurements accrues because of the potential for eliminating a component of within-population variation. It follows that repeated measurements are useful if a population element retains much of its identity through time. However, a study designed to assess status through time will generally be more powerful if new sites are used at every visit cycle.

A class of sampling designs referred to as “sampling with partial replacement designs” or “rotating panel designs” has evolved to meet these conflicting attributes. Binder and Hidiroglou (1988) review some of these alternative design and analysis approaches for sampling a population through time. Duncan and Kalton (1987) describe the characteristics of sampling designs through time, especially as they apply to human populations. Skalski (1990) recommends sampling with partial replacement designs for long-term environmental monitoring. The essential feature of rotating panel or partial replacement designs is some organized schedule of revisiting some units or sites from the sample, dropping others, and adding new ones as time passes.

For both water and sediment samples, spatial location will identify and label population elements, so a rotating panel design is useful if the population element at location "*x*" in one year is strongly related to the population element at "*x*" in another year. The sampling design is based on the assumption that this is true for the sediment samples, but not for the water samples.

The sediment sampling design is a rotating panel design with a base period of five years. A power analysis, described in Chapter 3, indicated that eight sites per sub-region per year would provide sufficient sensitivity. The eight sites are arranged in a rotating panel design, with four panels of two sites each being visited each year. The design has one panel that is visited every year, five panels that are visited on a 5-year rotation, ten panels that are visited on a 10-year rotation, and twenty panels that are visited on a 20-year rotation. One panel of each rotation period is visited each year.

The selection of the 5-year base rotation period was based on the assumption that any trend in the sediment indicators is small relative to spatial and inter-annual components of variation. Thus, the design provides only minimal revisit information in the first five years of the sampling, and complete revisit pattern will require at least 20 years to realize. This may seem like an unreasonably long time, but the unfortunate truth is that to detect changes that are small relative to the total variation in the population, a large number of years are required (Urquhart and Kincaid, 1999). If for some reason the sampling program is not continued for that length of time, trend estimates will still be possible, but the full power of the designs to detect trend will not be available.

Assignment of Sediment Sample Points to Panels

Thirty-six panels are required for the study. As was noted above, points from a GRTS sample can be ordered in such a way that any sequence of m consecutive points of an n-point sample ($m \leq n$) is in itself a nearly regular but random sample. The GRTS-ordered sample points were assigned to the required 36 panels sequentially, that is, the first 2 points to panel 1, the next two to panel 2, and so on. In order to preserve spatial balance through time, the panels were assigned a rotation period in sequence as panels were visited. For example, panels 1 through 4 are to be visited in year 1 of sampling (eight sites total per sub-region). In the second year of sampling, panel 1 is re-visited, and panels 5, 6, and 7 are added. The added panels are designated as a 5-year, 10-year, and 20-year panel, respectively. A similar procedure is repeated for years 3 through 5. In year 6, the annual panel and the 5-year panel from year 1 are revisited, and new 10-year and 20-year panels are added. In year 11, the annual panel, the 5-year panel, and the 10-year panel from year 1 are revisited, and a new 20-year panel is added. In year 21, the entire sequence of visits starts over with the panels visited in year 1.

Following this assignment scheme and revisit schedule ensures that each annual sample is reasonably well-distributed over the target extent, i.e., over the area of the segment. Table 1 gives the assignment of rotation period to panel and Table 2 gives the visit schedule by panel number.

The basic rotating panel structure requires 72 points in each segment to complete 20 years of sampling. However, we anticipate that some of the first 72 points will be unsampleable, most likely because of falling on coarse-grained sediment. In order to allow "replacement" of unsuitable points, approximately 140 samples were selected in each segment. The first 72 sediment points were assigned to the panel structure specifying a visit schedule, given in Tables 1 and 2. The remaining 68 points are reserve points to be used in order in the event some of the first 72 points are not sampleable, e.g., because the sediment is coarse-grained. These 68 points should be used as needed in the order they were provided.

Table 1: Panels in each rotation period

Rotation Period	Panel Number
Annual	1
5-Year	2,6,8,12,14
10-Year	4,5,10,11,16,17,20,21,24,25
20-Year	3,7,9,13,15,18,19,22,23,26,27,28,29,30,31,32,33,34,35,36

Table 2: Visit schedule by panel number

Panel Number					Panel Number				
Year 1	1	2	3	4	Year 11	1	2	4	27
Year 2	1	5	6	7	Year 12	1	5	6	28
Year 3	1	8	9	10	Year 13	1	8	10	29
Year 4	1	11	12	13	Year 14	1	11	12	30
Year 5	1	14	15	16	Year 15	1	14	16	31
Year 6	1	2	17	18	Year 16	1	2	17	32
Year 7	1	6	19	20	Year 17	1	6	20	33
Year 8	1	8	21	22	Year 18	1	8	21	34
Year 9	1	12	23	24	Year 19	1	12	24	35
Year 10	1	14	25	26	Year 20	1	14	25	36

Water Column Sample

The water column sample was selected using the 3-foot depth contour as the definition of the sample frame. We believe there will be substantial spatial pattern in the responses measured on the water column, but that a particular point will not retain its identity from year to year because of mixing. Thus, there is no benefit from revisiting the same site, but there is a potential benefit from a spatially regular sample. Again, the GRTS design was used to select water column samples.

The power analysis in Chapter 3 indicated that not all sub-regions required the same number of samples. We can accommodate the varying nominal sample sizes and at the same time allow for changing sample sizes by drawing more than the required number and utilizing the sequential spatial balance property of the GRTS design. Approximately 100 samples were selected in each region. Within each region, the sample sites were arranged in an ordered list that gives sequential spatial balance, that is, at any point in the list, the sites up to that point constitute a spatially-balanced sample of the target domain. The points should be used in order as required, e.g., if 10 samples are required in the first year, the first 10 points in the ordered list should be used. The sample in year 2 would then use samples beginning with point number 11. The samples selected in year 2 would still be a spatially balanced sample even if a different number of points were used, e.g., 15 sites instead of the 10 sampled in year 1.

Data analysis and population inference

An advantage of a probability sample is that the full range of statistical tools are available for analysis, including pure design-based survey methodology, model-assisted design-based methods, classical model-based methods, and Bayesian techniques. In this section, we will sketch out some possible analyses that might be performed with the data resulting from the redesigned RMP. We will begin with the design-based analyses inasmuch as analysis of a complex survey may be unfamiliar to most environmental scientists.

The population status description will be accomplished using the usual EMAP descriptive analyses (Diaz-Ramos, Stevens, and Olsen, 1996). The general approach is sketched briefly here, with details in a separate technical report *Estimation of Means, Totals, and Distribution Functions from Probability Survey Data* by Don Stevens available on the web at: http://www.sfei.org/sfeireports.htm.

Design-Based Population Status Estimation

The annual samples from the overall GRTS design are very similar to the samples that might result from a spatially–stratified design. Within each sub-region, the sample is equiprobable, so that the naïve estimators of sub-region characteristics are "correct" estimators. Thus, for example, the simple mean of the annual values within a sub-region is an unbiased estimator of the sub-region mean value. The design needs to be taken into account when (1) data is combined across sub-regions, say to estimate an Estuary-wide mean value, or (2) when the variance (standard error, confidence interval) of an estimate is being computed. The difference comes about because of the spatially-balanced feature of the GRTS design, and because of the disparity in area between the sub-regions.

Heuristically, we can think of each sample point as representing a certain amount of area within a sub-region. For the annual sediment samples, for example, each of the 8 points represents 1/8 of the area of the sub-region. In survey sampling terminology, the area that each point represents is its *weight*. If we let y_{ij}, i = 1, …, 8, be the response at the 8 sediment sample sites in sub-region j, the general design-based estimator of the mean value is $\overline{Y}_j = \frac{\sum w_{ij} y_{ij}}{\sum w_{ij}}$, where w_{ij} is weight for point i in sub-region j, in this case the sub-region area divided by 8. If we let A_j be the sub-region area, so that $w_{ij} = A_j / 8$, then the estimator reduces to the usual estimator of the mean: $\bar{Y}_j = \frac{\sum y_{ij}}{8}$.

If we want to estimate the mean for a region that crosses several sub-regions, then the samples that come from different sub-regions must be appropriately weighted and the general formula must be used. For example, suppose we wanted to estimate the mean value for the combined area of two sub-regions, which for convenience we label sub-regions 1 and 2. The appropriately weighted estimator is $\bar{Y}_{12} = \frac{\sum_{j=1}^{2} \sum_{i=1}^{8} w_{ij} y_{ij}}{\sum_{j=1}^{2} \sum_{i=1}^{8} w_{ij}}$.

We noted above that design-based variance estimators should be used in place of the usual estimators found in most statistical software packages. While this statement is true, the common estimator can be used, but will give a conservative result, that is, it will over-estimate the variance. This happens because the spatially-constrained sample will have lower variance than a completely random sample. In order to capture the reduced variance, the design-based variance estimator must be used (Stevens and Olsen, 2003). Formulas and computational details are given in a separate technical report (Stevens, 2002).

Both the water column and sediment samples can be analyzed using the same general techniques, but specific techniques are available to use the rotating panel structure of the sediment sample. The discussion here will concentrate on the sediment sample analyses, since they will be more complex than for the water column samples.

The rotating panel can be viewed as a "multi-phase" design, consisting of a number of design phases nested within one another. From this perspective, the entire collection of 72 points is a single sample. Each point in the collection (the first phase) is visited at least once in the 20-year duration of the design. A subset (the second phase sample) of the first phase is visited at least twice over 20 years. The second phase consists of those sites in the 10-year, the 5-year, and the annual panels. The third phase of the design is those sites visited at least 4 times in 20 years, and consists of the 5-year and annual panels. Finally, the fourth phase consists of the annual panels, i.e., those that are visited 20 times during the 20 years. Thus, each successive phase is a subset of the preceding phase

The advantage of this viewpoint is that a variety of analysis techniques is available for multi-phase designs (see, for example, Sarndal, Swensson, and Wretman, 1992). Within the multi-phase paradigm, for example, we have (1) composite estimators of status, which use prior years data to improve current years estimators, (2) multi-phase regression, which uses ancillary information that need not be available on all phases. This last technique could be a very powerful way of describing regional trend.

Composite estimators use the revisit pattern to incorporate data from previous years in annual status estimates. The basic concept is that we use a model to predict the response for location *s* at time $t+1$ based on the response at time *t*. The model estimates from all sites visited at time *t* are used to predict the mean at time $t+1$. Residuals (the difference between the observed and predicted values) at the re-visit sites are used to estimate the difference between the model-based mean and the true mean, and the estimated difference is used to correct the model-based mean. This correction term makes the adjusted model-based mean design-unbiased. We get a second design-unbiased estimate from the sites visited only at time $t+1$. Any convex combination of the two estimates of the mean is also design-unbiased. If the model prediction is reasonably accurate, then the composite estimator should have lower variance than the estimator based solely on the sites observed at time $t+1$.

The usual application relies on the correlation between revisited sites, but we could extend that to draw on the space-time covariance model, or more generally, any model that predicts a response at time *t* and location *s* based on data from previously observed times and locations. This can be done so that the resulting estimates remain design-based/model-assisted, so that even if we don't get the model exactly right, our estimates are still design-unbiased. The utility of these estimators is that they "borrow strength" from data that are nearby in space or time. The result is that we get more precise, that is, lower variance, estimators.

Another approach to annual status estimation is via small-area estimation. The underlying concept here draws on spatial correlation to "borrow strength" from adjacent regions. Thus, we could increase the effective sample size in a sub-region, for instance, by using some nearby observations in adjacent sub-regions, again with an appropriate weight determined by a correlation model.

Estimation of a proportion, e.g., the proportion of area that meets some criterion, is essentially the same as estimation of mean value. We estimate the proportion by replacing the observed response with an indicator function that is equal to 1 if the criterion is met, and is equal to 0 if not. The mean value of the indicator function is then the proportion that we want. With only 8 values per sub-region, of course, we can't get much resolution. In the first several years, then, reasonable estimates of proportions will only be available for the entire Estuary or a combination of several sub-regions. Once more data is available, various methods for borrowing strength from previous years or adjacent areas become available.

Design-based Trend description

The rotating panel of the sediment sample provides more options for trend detection than does the water column sample. Even so, only minimal trend information will be available for the first five years of the design. Site-specific trend can be estimated at the annual sites; however, the utility of such estimates will be limited by the small number of annual sites. Once the design has been in place long enough to get some re-visit information from the 5-year sites, more possibilities for estimating trend are available.

A design-based approach to describing regional trend is to characterize the population of site-specific trends. Suppose we observe a site once a year for 20 years. We could describe the trend at that site by a statistic such as the slope coefficient from a linear regression of the response on year number. We can regard the slope itself as a response variable, and think about having that available at every point in the region, say at every point in the Central Bay sub-region. The population of slope coefficients characterizes trend for the region, and we can summarize that population in a variety of ways, just as we would summarize any other population response. For example, we can calculate statistics such as the mean and variance; the population distribution function; or the percentage of the population that has a positive trend.

Besides the slope statistic based on all 20 years of data, we could also calculate a "5-year" slope based on observations 5 years apart, e.g., on observations from years 1, 6, 11, and 16, or 2, 7, 12, and 17. We would not expect to get the same number as for the 1-year interval data (the 1-year slope), but over the entire population, we would expect the 5-year slope to be strongly correlated with the 1-year slope. We can calculate the 1-year slope only for the sites that are observed every year, but we can calculate the 5-year slope for both annual and 5-year sites. The idea behind multi-phase regression is that we use the 5-year slope to predict the 1-year slope for the 5-year sites. The predictions are then used together with the observed 1-year slopes to get more precise estimates for the population of 1-year slopes. We can also carry the analysis a step further, by estimating 10-year slopes, and predicting 1-year slopes from the observed 10-year slopes. Eventually, if we go back to the 20-year sites beginning in year 21, we could also incorporate 20-year slopes.

Multi-phase regression works with any estimator of slope, that is, we are not limited to using linear least-squares regression. For example, there are several non-parametric alternatives (the Sen slope (Sen, 1968), or various robust/resistant estimators, such as lowess (Cleveland, 1979)). One could even use a model-based estimator of slope. The key to applying multi-phase regression to describe trend is that we regard the trend estimator based on the complete record of annual observations as the site response. Observations at sites not observed every year are used as ancillary variables to predict the response at those sites. Multi-phase regression provides the analytic framework in which to develop the population-level estimates of trend.

Model-based Analysis

The data set will support a rich variety of model-based methods. Certainly classical spatial statistics models will be a topic to investigate. Initially, only information on spatial variance will be available. The design points will provide an ideal data set for estimating field properties, such as the semi-variogram. Over time, the repeat visits of the annual sites and the 5- and 10-year sites will permit estimation of the properties of a space-time random field.

A common approach to fitting such a space-time model is to use a technique such as regression to incorporate co-measured explanatory variables (for example, sediment characteristics, water depth, or loading estimates) and to model space-time dependency through an appropriate covariance structure (the "random field" approach that is central to spatial statistics and geostatistics). More generally, we split the overall response into a mean structure component and a covariance structure component. The mean structure models the influence of the co-measured explanatory variables, and the covariance structure models the lack of independence in the residuals due to influence of latent variables. Of these two steps, getting the covariance right is arguably the most critical, and the hardest to do. Simplifying assumptions that are often adopted are stationarity, isotropy, and separability, that is, that the covariance between the residuals at two space-time points depends on the spatial distance and the time difference between the points, but not on the actual locations, the direction from one point to the other, or the particular time. See, for example, the papers by Cressie and Majure (1997) or Carroll (1998). One practical result of applying these assumptions is that the amount of data required to estimate the covariance is greatly reduced.

The covariance is typically incorporated into the model via the semi-variogram (Cressie, 1993), which describes the variance of points separated in time and space. Each pair of observations in a data set contributes a spatial separation and a time separation; for example, the same point in space observed at two different times will have zero spatial separation, but non-zero time separation. From the sediment data set that we will (eventually!) have for each sub-region (8 observations per year for 20 years) spread over time and space), we will end up with 160(159)/2 distinct space by time separation pairs. The empirical semi-variogram is usually obtained by binning the separation pairs, and calculating the variance of the response difference within each bin. For example, we might bin the data into a two-way table, with each cell in the table defined by a time and distance separation. If we used a 1-year time interval and a 1 km distance interval, then each bin would contain point pairs that were the same number of years apart, and within 1 km of each other. For instance, one bin would have all points between 1 and 2 km apart, and observed 7 years apart. A good data set to estimate the semi-variogram has many pairs of points in each cell of that two-way table, and preferably, approximately the same number in each cell. A conventional rule of thumb is that each cell should have 30 or more pairs of points with approximately the same spatial and temporal separation, for example, at least 30 pairs of points that are between 1 and 2 km apart and observed 7 years apart. One advantage of the rotating panel design is that it guarantees a reasonably even distribution of points in the cells of the two-way table.

In the last ten years, there has been a huge increase in the use of hierarchical models to analyze dependencies among variables. The basic idea of hierarchical models is that the observed data are modeled as a function depending on unknown parameters and unobserved random errors. The random errors in turn are modeled as depending on unknown parameters. In the Bayesian version, the unknown parameters are given a prior distribution, and the observed data are used to calculate the posterior distribution. The techniques have been used to develop statistical models that accommodate missing data, temporal and spatial dependencies, latent variables, multiple response variables, and nonlinear functional forms. Several years ago, such models were not feasible, because the parameter estimation was too computation intensive. Several recent theoretical developments (most notably, the so-called Gibbs sampler and Markov Chain Monte Carlo, Gelfand, 2000) have dramatically reduced the computational burden. In the last five years, hierarchical models have been widely applied to environmental data, especially data with space/time components.

Spatial Displays

Several techniques for spatial display of the survey data are possible. For example, location and response magnitude can be shown by plotting site location using symbols whose size is proportional to response. The spatial balance property makes the data set well-suited to surface estimation, so that a response can be represented as contour lines or as a 3-dimensinal perspective plot. The surface can be estimated using a variety of interpolation or smoothing techniques, for example, polynomial interpolation or spatial splines. One could also use a spatial statistics approach, e.g., kriging, to predict the surface. Displays that are more dramatic can be obtained using color. One possibility is to color the Voronoi polygons around each point using a gradation of shade to correspond to response magnitude. (The Voronoi polygons for a set of points $\{ s_1, s_2, \ldots, s_k \}$ are a partition of the domain spanned by the points such that the i^{th} polygon is the collection of points in the domain that are closer to s_i than to any other s_j in the set.)

References

Bellhouse, D.R. 1977. Some optimal designs for sampling in two dimensions. *Biometrika* **64**: 605–611.

Binder, D.A., and M.A. Hidiroglou. 1988. Sampling in Time. In: Handbook of Statistics Vol. 6, Sampling. P.R. Krishnaiah and C.R. Rao, eds., pp. 187-211. Elsevier Science Publishers B.V. Amsterdam, The Netherlands.

Breidt, F.J. 1995. Markov chain designs for one-per-stratum sampling. *Survey Methodology* **21**:1; 63-70.

Brus, D.J. and J.J. de Gruitjer. 1993. Design-based versus model-based estimates of spatial means: theory and application in environmental soil science. *Environmetrics* **4**: 123–152.

Carroll, S.S. 1998. Modelling abiotic indicators when obtaining spatial predictions of species richness. *Environmental and Ecological Statistics* **5**: 257-276.

Cassel, C.M., C.E. Särndal and J.H. Wretman. 1977. *Foundations of Inference in Survey Sampling*. New York: John Wiley & Sons.

Cleveland, W.S. 1979. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* **74**: 829-836.

Cochran, W.G. 1977. *Sampling Techniques*, Third Edition. John Wily & Sons: New York, NY.

Cotter, J., and J. Nealon. 1987. Area Frame Design for Agricultural Surveys. Area Frame Section, Research and Applications Division, National Agricultural Statistics Service, U.S. Department of Agriculture.

Cressie, N.A.C. 1993. *Statistics for Spatial Data*. John Wiley & Sons, New York. 885p.

Cressie, N. and J.J. Majure. 1997. Spatial-temporal statistical modeling of livestock waste in streams. *Journal of Agricultural, Biological, and Environmental Statistics*. **2**: 24-47.

de Gruijter, J.J. and C.J.F. Ter Braak. 1990. Model free estimation from survey samples: a reappraisal of classical sampling theory. *Mathematical Geology* **22**: 407–415.

Dalenius, T., J. Hájek, and S. Zubrzycki. 1961. On plane sampling and related geometrical problems. *Proceedings of the 4th Berkeley Symposium on Probability and Mathematical Statistics* **1**: 125–150.

Das, A.C. 1950. Two-dimensional systematic sampling and the associated stratified and random sampling. *Sankhya* **10**: 95–108.

Diaz-Ramos, S., D.L. Stevens, Jr, and A.R. Olsen. 1996. *EMAP Statistical Methods Manual.* Environmental Monitoring and Assessment Program, National Health and Environmental Effects Research Laboratory, Office of Research and Development, Corvallis, OR.

Duncan, G.J., and G. Kalton. 1987. Issues of design and analysis of surveys across time. *International Statistical Review* **55**: 97-117.

Gelfand, A.E. 2000. Gibbs Sampling. *Journal of the American Statistical Association* **95**: 1300-1304.

Hansen, M.H., W.G. Meadow, and B.J. Tepping. 1983. An evaluation of model-dependent and probability sampling inferences in sample surveys. *Journal of the American Statistical Association*, **83**: 776-793.

Herlihy, A.T., D.P. Larsen, S.G. Paulsen, N.S. Urquhart, and B.J. Rosenbaum. 2000. Designing a spatially balanced, randomized site selection process for regional stream surveys: the EMAP mid-Atlantic pilot study. *Environmental Monitoring and Assess*ment **63**: 95-113.

Iachan, R. 1985. Plane Sampling. *Statistics and Probability Letters* **50**: 151–159.

Jacobs, S.E. and C.X. Cooney. 1995. *Improvement of methods used to estimate the spawning escapement of Oregon coastal natural coho salmon*. Oregon Department of Fish and Wildlife.

Larsen, D.P., K.W. Thornton, N.S. Urquhart, and S.G. Paulsen. 1994. The role of sample surveys for monitoring the condition of the nation's lakes. *Environmental Monitoring and Assessment* **32**: 101-134.

Larson, D.P., D.L. Stevens, A.R. Selle, and S.G. Paulsen. 1991. Environmental Monitoring and Assessment Program - Surface Waters: A Northeast Lakes Pilot. *Lake & Reserv. Management* **7**(1):1-11. ERL-COR-1300J.

Munholland, P.L., and J.J. Borkowski. 1996. Simple Latin square sampling +1: A spatial design using quadrats. *Biometrics* **52**: 125-136.

Matérn, B. 1960. *Spatial Variation*. Stockholm, Sweden: Meddelanden från Statens Skogsforskningsinstitut.

N.R.C. 1990. *Managing Troubled Waters: The Role of Marine Environmental Monitoring*. Washington, D.C: National Academy Press.

O'Connor, T.P. 1990. *Coastal Environmental Quality in the United States, 1990. Chemical Contamination in Sediment and Tissues. A Special NOAA 20th Anniversary Report*. Rockville, Maryland: National Oceanic and Atmospheric Administration, Ocean Assessments Division.

Olea, R.A. 1984. Sampling design optimization for spatial functions. *Mathematical Geology* **16**: 369–392.

Overton, W.S. and S.V. Stehman. 1993. Properties of designs for sampling continuous spatial resources from a triangular grid. *Communications in Statistics Part A: Theory and Methods*, **22**: 2641–2660.

Overton, W.S., D. White, and D.L. Stevens, Jr. 1991. *Design Report for EMAP: Environmental Monitoring and Assessment Program*. U.S. EPA/600/3-91/053 ERL-COR-696.

Paulsen, S.G., R.M. Hughes, and D.P. Larsen. 1998. Critical elements in describing and understanding our nation's aquatic resources. *Journal of the American Water Resources Association* **34**: 995-1005.

Peterson, S.A., N.S. Urquhart, and E.B. Welch. 1999. Sample Representativeness: A Must for Reliable Regional Lake Condition Estimates. *Environmental Science & Technology* **33**: 1559-1565.

Quenouille, M.H. 1949. Problems in plane sampling. *Annals of Mathematical Statistics* **20**: 335–375.

Särndal, C. 1978. Design-based and model-based inference for survey sampling. *Scandinavian Journal of Statistics* **5**: 27-52.

Sen, P.K. 1968. Estimates of regression coefficient based on Kendall's tau. *Journal of the American Statistical Association* **63**: 1379-1389.

Skalski, J.R. 1990. A design for long-term status and trends monitoring. *Journal of Environmental Management* **30**: 139-144.

Smith, T.H. 1976. The foundations of survey sampling: a review. *Journal of the Royal Statistics Society A* **139:** 183-204.

Stevens, Jr.; D.L. 1997, Variable density grid-based sampling designs for continuous spatial populations. *Environmetrics*, **8**: 167-195.

Stevens, D.L., Jr., and T.M. Kincaid. 1997. Variance estimation for subpopulation parameters from samples of spatial environmental populations. *Proceedings of the American Statistical Association Section on Statistics and the Environment*, American Statistical Association, Alexandria, VA.

Stevens, Jr., D.L. and A.R. Olsen. 1999. Spatially restricted surveys over time for aquatic resources. *Journal of Agricultural, Biological, and Environmental Statistics* **4**: 415-428.

Stevens, Jr., D.L., and A.R. Olsen. 2000. Spatially-restricted random sampling designs for design-based and model-based estimation. In *Accuracy 2000: Proceedings of the 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*. Delft University Press, The Netherlands. pp. 609-616.

Stevens, Jr., D. L. and A. R. Olsen. 2003. Variance estimation for Spatially Balanced Natural Resource Samples. *Environmetrics* **14**:593-610.

Stevens, Jr., D.L., and A.R. Olsen. 2004. Spatially-Balanced Sampling of Natural Resources. *Journal of the American Statistical Association* **99**:262-277

Stevens, Jr., D.L., 2002. Estimation of Means, Totals, and Distribution Functions from Probability Survey Data, SFEI Techical Report, submitted to: San Francisco Estuary Regional Monitoring Program for Trace Substances, June 20, 2002. San Francisco Estuary Institute, Oakland, CA.

Urquhart, N.S., and T.M. Kincaid. 1999. Trend detection in repeated surveys of ecological responses. *Journal of Agricultural, Biological, and Environmental Statistics* **4**: 404-414.

Chapter 5 - Design Integration into the RMP

Bruce Thompson, SFEI
Sarah Lowe, SFEI

The RMP provides scientifically sound data and interpretation on the status and trends of contaminants of concern to the Regional Board for environmental management purposes (SFEI, 1998). The revised objectives, described in Chapter 1 (page 4), provide the framework for all RMP effort, which are managed by several science-based work-groups, and two oversight committees (the Technical Review and the Steering Committees (TRC and SC respectively). The workgroups meet on a periodic basis to help develop the program, and the two committees provide oversight on a quarterly basis.

There are three major components of the RMP:

1) Status and Trends program includes
    a. the new spatially balanced, random sampling design (and historic sampling) of Estuary water and sediment for targeted contaminants, toxicity, and ancillary parameters (sampled annually)
    b. the Fish Contamination Study targeting important sport fisheries in the Estuary for Hg and organic contaminants (sampled triennially)
    c. the bivalve bioaccumulation monitoring program targeting organic contaminants that bioaccumulate (sampled annually)
    d. the Episodic Toxicity Monitoring Program that samples Estuary tributaries during storm events to investigate the potential for aquatic and sediment toxicity (sampled annually during the wet season)
2) Pilot Studies that are designed to evaluate new monitoring “tools” that, if effective, may be incorporated into the Status and Trends component; and
3) Special Studies that are developed to address specific management questions or provide additional scientific information used to help interpret contaminant data collected by the Status and Trends component.

The new Status and Trends sampling design meets the RMP’s revised objectives 1 and 4 (see Chapter 1, page 4). It provides the scientific data to estimate the spatial and temporal distribution of water and sediment contamination within each sub-region, and in the Estuary as a whole. The random sampling will allow us to determine if the mean contaminant concentration within a sub-region is above a regulatory guideline or other effects thresholds. Data collected can be divided according to the shallow and deep channels (delineated by the six-foot contour at mean lower low water (MLLW)) to determine what proportion of the shallow versus deep channels is impaired. These types of inferences were not possible with the original, historical RMP sampling design.

A subset of the fixed historical RMP sampling sites from each region are maintained within the Status and Trends component to serve as a link between the old and new sampling designs, and to maintain the long-term temporal trends for contaminants compiled at those sites. Bivalve bioaccumulation monitoring will continue at the original historical sites under the Status and Trends component during the dry season only. The program has been investigating the possibility of deploying only one species throughout the Estuary in an effort to make the results from the various regions of the Estuary more comparable. Oysters (*Crasostrea gigas*) and mussels (*Mytilus califorianus*) have been deployed in at several sites in side-by-side studies of survival and growth. Mussels have a wide tolerance for salinity (15 – 35 psu) so they are a good species for deploying throughout the Estuary. Since sampling will occur only during the summer, the mussel will be deployed at all sites except the two River sites where resident clams (*Corbicula fluminia*) will be collected.

Given budgetary constraints, increased spatial coverage of the Estuary was considered higher priority than maintaining seasonal sampling, even though the historic RMP data showed that sediment and water are sometimes more contaminant impacted during the wet season. It was decided that annual dry season sampling (when conditions in the Estuary are most "stable") would provide a good framework for collecting both spatial and long-term trends data for regulatory management purposes. The RMP work-groups, and oversight committees would address the seasonal aspect of contaminant loadings and potential biological effects within Pilot and Special studies.

Sources, pathways, and loadings of contaminants to the Estuary (objective 2 of the revised RMP objectives described in Chapter 1, page 4) are addressed through the RMP's Sources, Pathways and Loadings Work Group where local scientists and RMP participants meet to review current study results, evaluate information gaps, and help design new pilot or special studies submitted to the TRC for review, prioritization and budgeting. For example, the RMP is currently studying particle bound contaminant loads to the Estuary during seasonal, high-flow events. Another study, identified by the work-group but not yet funded, includes a study to estimate the active sediment layer in the Estuary. Pilot and special studies provide empirical data used by the RMP's contaminant modeling group (a component of the RMP's data integration effort) to refine their contaminant fate models. These models provide regulatory managers with a science-based "tool" to evaluate the past and potential efficacy of contaminant regulation in and around the Estuary.

Contaminant effects on selected parts of the Estuary ecosystem (objective 3; Chapter 1, page 4) are being addressed by several RMP components. The Exposure and Effects Pilot Study, another Work Group driven study, is a five-year pilot study that began in 2001, is underway to identify appropriate indicators of contaminant effects on the local estuarine ecology. The study includes:

- comparing the relative sensitivity of resident and laboratory invertebrates to key toxic contaminants,
- benthic community assessments,
- reproductive effects, bioaccumulation of contaminants, and general health evaluations in fish, birds, and seals.

The goal of the study is to develop a "tool-box" of locally tried and tested indicators of ecosystem health that can be used to monitor the Estuary at different ecological scales (metabolically, individually, population-wise, and community-wise).

The RMP completed a special study in 2002 to identify new and previously unmonitored organic contaminants in water, sediment, and bivalve tissue samples (Oros, 2002). Results of this screening effort identified several classes of organic contaminants that were added to the RMP Status and Trends analyte list in 2002 to further investigate the presence of these contaminants in the Estuary and determine if any are at levels that cause concern.

As a result of a recent regulatory mandate, the RMP conducted a two-year screening study (began in 2002) to determine if any contaminants currently listed in the California Toxics Rule (CTR), but not routinely measured by the RMP, are found at levels in the Estuary that cause concern.

The Status and Trends component of the RMP samples sediments and conducts laboratory toxicity testing on amphipods and bivalves to measure potential biological effects in the Estuary. The new random sampling design will allow for estimating what proportion of the estuarine sediments is toxic to laboratory test organisms. Historic aquatic toxicity testing in the RMP has shown that Estuary waters, collected during the dry season, were generally not toxic to laboratory invertebrates. To evaluate if this absence of toxicity was true for the shallows as well (since the old monitoring design sampled only in the deeper channels of the Estuary) the RMP sampled the shallow regions of the Estuary in 2002. The results were negative leading to the decision to monitor aquatic toxicity in the Status and Trends program on a five-year monitoring cycle. However, potential sources of wet season aquatic toxicity continue to be studied through the Episodic Toxicity Monitoring Program, which samples several tributaries during major flow events. This program began as a pilot study in 1996 and is now incorporated into the Status and Trends component of the RMP. The RMP's Toxicity Work Group, a subgroup of the Exposure and Effects Pilot Study, provides guidance for this program and the RMP's toxicity monitoring effort as a whole.

Another long-term Status and Trends monitoring component of the RMP is the Fish Contamination Study, a triennial bioaccumulation monitoring effort of key sport fish in the Estuary. Fish are sampled at several locations within the Estuary and are analyzed for a suite of bioaccumulative contaminants of concern (mercury, PCBs, DDTs, and other organic contaminants). Results from this study are evaluated against fish consumption guidelines and are used by the regulatory agencies in considering public fish consumption advisories.

Status and Trends monitoring results are compared to relevant water quality objectives and other guidelines annually (revised objective 4, Chapter 1) in several RMP products: the Annual Monitoring Results (includes data summaries), the Pulse of the Estuary (a public outreach summary on contaminants in the Estuary), and Technical Reports (detailed scientific documents of various RMP efforts). All these documents are available on the web at http://www.sfei.org.

RMP results are synthesized and distributed (revised objective 5, Chapter 1) in a variety of formats (some mentioned above) including in publications, formal and informal presentations (e.g. the RMP Annual Meetings, local, national and international conferences, and by private request), and on the web at http://www.sfei.org. A range of information sources and collaborations are actively pursued in an effort to present a more complete picture of the sources, distribution, fate, and effects of contaminants in the Estuary ecosystem. Information and collaboration efforts are often made with other Bay Area scientists at universities (e.g., University of California at Berkeley, University of California at Davis, University of California at Santa Cruz, and San Francisco State University), government agencies (e.g., USGS, Regional Board, and US EPA), and regional programs (e.g., CEP, NOAA, and EMAP). Recent emphasis has been made to present RMP results in formats that appeal to a larger audience ranging from short newsletter articles to detailed technical documents.

The impetus for redesigning the RMP's Status and Trends component was the five-year program review that prompted a careful reconsideration of the program's Objectives and Questions. It was concluded that the original design could not answer the key management and regulatory questions of the Regional Board and the RMP participants. The redesign of the RMP was a unique process dictated by the information needs of the Regional Board, the lead regulatory agency responsible for protecting San Francisco Estuary water quality. The direct involvement by the Regional Board staff in the redesign process was critical. The RMP could not have been successfully redesigned without their thoughtful and considered statements of information needs and focusing questions.

It is important to acknowledge the role of subjective and professional judgment in monitoring design. The use of data analysis and evaluations were shown above to inform those decisions, but almost all final decisions about monitoring design required integrating several different lines of information and making informed decisions. To that extent, the use of workgroups composed of experts and experienced participants assured that issues were well discussed and that decisions made were supported and justifiable with strong science.

All RMP Participants have agreed that the new RMP design will lead to an increased understanding of the sources, fate, and effects of contamination in the Estuary, which will result in more informed and effective regulatory decisions.

Importance of Pilot and Special Studies

The redesigned Status and Trends component will not sample important seasonal or episodic events (e.g., high freshwater inflows, spills, etc.) that may affect Estuary water quality. The RMP Pilot and Special studies are an important component of the RMP that are intended to fill many of the gaps of understanding the processes and mechanisms that influence water quality, and to develop improved indicators for the RMP. Understanding the sources, transport, fate, and effects of contamination requires a different approach to sampling and analysis than Status and Trends monitoring. Depending on the information needed, either sampling studies or manipulative experiments designed to answer specific questions will be required. Information obtained from these studies will help understand contamination in the Estuary and how best to intercede to protect beneficial uses and meet water quality objectives. They will also help interpret the RMP Status and Trends monitoring data by placing them in the context of information gained from the special studies.

The RMP work groups have been charged with identifying and scoping the most important studies needed to address the Objectives and Questions for the RMP. The Sources Pathways and Loadings Work Group is designing a new RMP component focused on understanding space and time scales of contaminant loading through gradient studies at the confluence of major or representative tributaries. The Contamination Fate Work Group is considering studies to understand transformations and transport of contamination. The Exposure and Effects Work Group is evaluating and testing new indicators of contamination effects that will eventually provide a weight-of-evidence for, or against contaminant effects at several trophic levels and ecological compartments. Another important part of the RMP is the collaboration and interaction with other agencies, such as United States Geological Survey (USGS), and Interagency Ecological Program (IEP), and local universities working on similar water quality questions. The RMP will support, or defer to such programs as needed to obtain information needed to answer the Management Questions.

Since the completion of the RMP redesign, another component to water quality management has been implemented. The Clean Estuary Partnership (CEP) is a cooperative partnership among the Regional Board and several Estuary permittees. This partnership is focusing on policy decisions needed to conduct and implement TMDLs in the Estuary. Policy decisions are not part of the RMP, but the RMP's data collection and interpretation are necessary for regulatory purposes. Thus, the RMP and CEP are working closely together to provide necessary scientific information to address the Clean Water Act and TMDL regulation.

References

SFEI. 1998. Management Questions Guiding the Regional Monitoring Program for Trace Substances--First Edition, 1998. San Francisco Estuary Institute. http://www.sfei.org/rmp/documentation/management_q.html

Oros, Daniel R. 2002. Identification and Evaluation of Unidentified Organic Contaminants in the San Francisco Estuary. RMP Technical Report: SFEI Contribution 45. San Francisco Estuary Institute, Oakland, CA,

APPENDIX 1
Addressing Management Questions

The RMP redesign process began with articulation of Objectives and Management Questions, thus, it is appropriate to recount how the new RMP design will address each Management Questions. Answering the Management Questions will require both objective and subjective evaluations.

1. Compare Monitoring Data.

1a. Which contaminants should be monitored?
RMP data will provide information about which have highest concentrations. Combined with literature information about concentrations that may pose the highest probability for adverse effects, the Regional Board can prioritize the contaminants for which information is most urgently needed.

1b. How do RMP data compare with relevant water, sediment, and tissue quality guidelines?
The RMP has done this annually since its inception. Objective comparisons of monitoring measurements and water quality objectives, or sediment quality guidelines, and other appropriate screening values will be reported in the RMP annual data reports.

1c. How do the various Estuary reaches compare to each other, in time and space, relative to water, sediment and tissue guidelines?
The new geographical sub-regions of the Estuary and the probabilistic monitoring design will allow for making objective comparisons (mean and variance) of individual contaminants between sub-regions using simple statistical tests.

Describe patterns and trends

2a. How do contaminants levels change over the long term?
Plots of contaminant trends over time (using historic and new monitoring data) in each sub-region can be evaluated using linear regression analysis to test for significant trends. Confounding factors (e.g. salinity, TOC) can be statistically factored out of the regressions.

2b. Can those changes be linked to changes in inputs to the Estuary?
This evaluation will be more subjective. The RMP is not monitoring effluent or storm water runoff, but data are available. Appropriately matching data types will be imprecise. However, long-term trend analyses should reflect major changes in management of water quality in the Estuary. Specific studies will be needed to evaluate effectiveness of individual waste water improvement projects.

2c. What is the relationship between pollutant trends and patterns seen in the “spine” of the Estuary and those in the shallow margins?

The new probabilistic monitoring design will permit such statistical comparisons using post-stratification of samples.

2d. How are spatial and temporal patterns and trends in contaminants affected by estuarine processes?
This is a special study question that may be addressed by analyzing existing data, future RMP special studies, or through collaboration with others.

Describe general sources and loadings

3a. What proportion of the contaminants in each Estuary reach are contributed by point source outfalls, storm drains, large and small tributaries, harbor activities including dredging, atmospheric deposition, and historic deposits?
Loadings estimates were made for the State Board in 1999 (Davis *et al*., 2000). Current RMP mass balance studies and CEP sponsored loading studies are addressing this question. The Sources, Pathways, and Loadings Workgroup is developing monitoring and study plans for future RMP implementation.

3b. How do contaminants move and transform after they enter the Estuary?
The Contaminant Fate Workgroup will determine information needs and prioritize studies that address this question over the next several years.

3c. At what spatial and temporal resolution should loadings to the Estuary and changes in upstream contaminant inputs due to pollution prevention efforts be monitored?
This question is being considered by SPL Workgroup.

3d. What are the background concentrations of contaminants in the Estuary from natural sources?
RMP data was used to determine Ambient (background) sediment contaminant concentrations in 19xx (Smith and Riege, 19xx). The USGS's sediment coring studies (ref) estimated pre-industrial concentrations in a few locations (ref).

Measure contaminant effects

4a. Which contaminants bioaccumulate in estuarine organisms to levels of concern?
The RMP monitors fish and bivalve tissues. Special studies are also underway examining Hg and PCBs in Estuarine food chains. The Exposure and Effects Pilot Study (EEPS) Workgroup is also considering additional bioaccumulation monitoring using fish and birds in the future.

4b. What is the spatial and temporal extent of toxicity in the Estuary?
The RMP has conducted aquatic and sediment toxicity testing since 1993. The new probabilistic monitoring design will provide unbiased spatial extent estimates. Episodic aquatic toxicity testing will occur following storm events, and sediment toxicity testing

will be conducted each summer, but this monitoring may not provide adequate temporal extent assessments. Additional special studies may be needed to relate toxicity to loadings events or to sources of contamination.

4c. Which contaminants cause effects in the Estuary?
Demonstrating cause due to contamination is very difficult and usually requires special studies. Special studies conducted on causes of sediment toxicity have helped narrow the potential causative agents down in some areas Thompson *et al.* 1999; Phillips *et al.* ms). The EEPS Workgroup has began pilot studies to identify the best indicators of contaminant effects for future RMP Status and Trends monitoring and additional special studies to determine cause.

Synthesize information

5a. Provide periodic interpretation and synthesis on selected contaminant-related topics. The RMP has done this with annual and technical reports, newsletters, and peer reviewed publications since its inception. The nine-year RMP synthesis slated as a special study in 2003 will assess selected RMP data.

5b. Describe and distribute key RMP findings to a variety of audiences.
RMP annual and technical reports, newsletters, presentations at conferences and symposia, and peer reviewed publications reach a variety of audiences. RMP contractors and collaborators also produce reports and publications using RMP data.

5c. Assess the use of RMP data and information in decision making.
This is an ongoing activity. Some examples include the use of RMP data in TMDLs, San Francisco Airport Runway EIR/EIS, dredging assessments, and development of sediment quality guidelines.

The above listings show that nearly all RMP Management Questions have been addressed, or are in the process of being addressed, in some manner. It is anticipated that the RMP will continue to remain adaptive in considering necessary changes in the future. As new information from the RMP workgroups, pilot and special studies, and CEP mandated studies emerge, the RMP will respond by adopting new components and indicators that provide sound data. Most importantly, through the active involvement of the Regional Board staff, the RMP participants, and SFEI and its contractors, the appropriate checks and balances, ranges of opinion, and expertise will provide defensible data for regulation and TMDL implementation.

APPENDIX 2

List of participants in the Design Integration Workgroup

Name	Company
Andy Gunther	AMS
Bob Smith	Consultant
Anke Mueller-Solger	DWR
Don Stevens	OSU
Andy Jahn	Port of Oakland
Jim McGrath	Port of Oakland
Karen Taberski	SFBRWQCB
Dyan Whyte	SFBRWQCB
Rainer Hoenicke	SFEI
Sarah Lowe	SFEI
Bruce Thompson	SFEI
Jay Davis	SFEI
Trish Mulvey	SFEI Board
Genine Scelfo	UCSC-ETOX
Jim Delorey	USACE
Tara Schraga	USGS

APPENDIX 3A
Power Analyses Results Tables

Water data were summarized for three seasons: winter (wet (Jan/Feb)), spring (when wet season flows are receding (rec (April)), and fall (during the dry season (July/Aug)). Sediment data from both the wet and dry seasons were combined providing an annual estimated mean. Both the historical mean (HisMean) and the assumed underlying mean (μ) are shown. When the historical mean is < 10% below the guideline the historical mean is used for μ . Otherwise μ is assumed to be 10% below the guideline. Either the water quality criterion (WQC) or the Effects Range Low (ERL)) was used in the power calculations. Data distributions were evaluated and log transformed (if warranted) prior to power calculations (N = normal distribution, L= log transformed). See Chapter 3 for further explanations.

Appendix Table 3A.1. Power results for dissolved metals. Empirical data used in these calculations were RMP data measured between 1993 and 1998. Units are in ug/L.

Dissolved Silver						**% Power at Sample Size of**								
Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	N	0.003	0.003	3.4	100								
Rivers	rec	N	0.002	0.002	3.4	100								
Rivers	dry	N	0.001	0.001	3.4	100								
Suisun Bay	wet	N	0.002	0.002	1.9	100								
Suisun Bay	rec	N	0.001	0.001	1.9	100								
Suisun Bay	dry	N	0.001	0.001	1.9	100								
San Pablo Bay	wet	L	0.002	0.002	1.9	100								
San Pablo Bay	rec	L	0.001	0.001	1.9	100								
San Pablo Bay	dry	L	0.004	0.004	1.9	100								
Central Bay	wet	L	0.002	0.002	1.9	100								
Central Bay	rec	L	0.002	0.002	1.9	100								
Central Bay	dry	L	0.005	0.005	1.9	97	100							
Golden Gate	dry	N	0.002	0.002	1.9	100								
South Bay	wet	L	0.003	0.003	1.9	100								
South Bay	rec	L	0.002	0.002	1.9	100								
South Bay	dry	L	0.010	0.010	1.9	100								
Lower South Bay	wet	L	0.003	0.003	1.9	100								
Lower South Bay	rec	L	0.003	0.003	1.9	100								
Lower South Bay	dry	L	0.007	0.007	1.9	100								

Dissolved Arsenic						% Power at Sample Size of								
Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	N	1.4	1.4	150	100								
Rivers	rec	N	1.4	1.4	150	100								
Rivers	dry	N	1.9	1.9	150	100								
Suisun Bay	wet	N	1.4	1.4	36	100								
Suisun Bay	rec	N	1.4	1.4	36	100								
Suisun Bay	dry	N	2.1	2.1	36	100								
San Pablo Bay	wet	L	1.5	1.5	36	100								
San Pablo Bay	rec	L	1.6	1.6	36	100								
San Pablo Bay	dry	L	2.5	2.5	36	100								
Central Bay	wet	L	1.6	1.6	36	100								
Central Bay	rec	L	1.6	1.6	36	100								
Central Bay	dry	L	2.2	2.2	36	100								
Golden Gate	dry	N	1.6	1.6	36	100								
South Bay	wet	L	2.1	2.1	36	100								
South Bay	rec	L	2.2	2.2	36	100								
South Bay	dry	L	3.3	3.3	36	100								
Lower South Bay	wet	L	2.1	2.1	36	100								
Lower South Bay	rec	L	2.6	2.6	36	100								
Lower South Bay	dry	L	4.2	4.2	36	100								

Dissolved Cadmium						% Power at Sample Size of								
Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	L	0.01	0.01	2.2	100								
Rivers	rec	L	0.01	0.01	2.2	100								
Rivers	dry	L	0.02	0.02	2.2	100								
Suisun Bay	wet	L	0.01	0.01	9.3	100								
Suisun Bay	rec	L	0.01	0.01	9.3	100								
Suisun Bay	dry	L	0.03	0.03	9.3	100								
San Pablo Bay	wet	N	0.04	0.04	9.3	100								
San Pablo Bay	rec	N	0.05	0.05	9.3	100								
San Pablo Bay	dry	N	0.10	0.10	9.3	100								
Central Bay	wet	L	0.04	0.04	9.3	100								
Central Bay	rec	L	0.05	0.05	9.3	100								
Central Bay	dry	L	0.10	0.10	9.3	100								
Golden Gate	dry	N	0.08	0.08	9.3	100								
South Bay	wet	L	0.07	0.07	9.3	100								
South Bay	rec	L	0.06	0.06	9.3	100								
South Bay	dry	L	0.12	0.12	9.3	100								
Lower South Bay	wet	L	0.07	0.07	9.3	100								
Lower South Bay	rec	L	0.07	0.07	9.3	100								
Lower South Bay	dry	L	0.13	0.13	9.3	100								

Dissolved Chromium

Region	Season	Distribution	HisMean	μ	WQC	% Power at Sample Size of 2	3	4	5	6	7	8	9	10
Rivers	wet	L	0.66	0.66	11	63	99	100						
Rivers	rec	L	0.36	0.36	11	71	100							
Rivers	dry	L	0.33	0.33	11	95	100							
Suisun Bay	wet	L	0.69	0.69	50	99	100							
Suisun Bay	rec	L	0.44	0.44	50	84	100							
Suisun Bay	dry	L	0.23	0.23	50	99	100							
San Pablo Bay	wet	L	0.49	0.49	50	90	100							
San Pablo Bay	rec	L	0.17	0.17	50	100								
San Pablo Bay	dry	L	0.16	0.16	50	100								
Central Bay	wet	L	0.15	0.15	50	100								
Central Bay	rec	L	0.14	0.14	50	100								
Central Bay	dry	L	0.12	0.12	50	100								
Golden Gate	dry	N	0.13	0.13	50	100								
South Bay	wet	L	0.16	0.16	50	100								
South Bay	rec	L	0.14	0.14	50	100								
South Bay	dry	L	0.16	0.16	50	100								
Lower South Bay	wet	L	0.23	0.23	50	99	100							
Lower South Bay	rec	L	0.20	0.20	50	100								
Lower South Bay	dry	L	0.16	0.16	50	100								

Dissolved Copper

Region	Season	Distribution	HisMean	μ	WQC	% Power at Sample Size of 2	3	4	5	6	7	8	9	10
Rivers	wet	L	2.0	2.0	9	96	100							
Rivers	rec	L	1.7	1.7	9	100								
Rivers	dry	L	1.8	1.8	9	100								
Suisun Bay	wet	L	2.0	2.0	3.1	63	99	100						
Suisun Bay	rec	L	1.9	1.9	3.1	92	100							
Suisun Bay	dry	L	2.0	2.0	3.1	57	98	100						
San Pablo Bay	wet	L	1.9	1.9	3.1	27	56	77	89	95	98	99	100	
San Pablo Bay	rec	L	1.8	1.8	3.1	28	57	78	90	95	98	99	100	
San Pablo Bay	dry	L	1.9	1.9	3.1	29	60	82	92	97	99	100		
Central Bay	wet	L	1.5	1.5	3.1	40	83	97	100					
Central Bay	rec	L	1.4	1.4	3.1	51	94	100						
Central Bay	dry	L	1.4	1.4	3.1	43	86	98	100					
Golden Gate	dry	N	0.6	0.6	3.1	89	100							
South Bay	wet	N	2.1	2.1	3.1	67	100							
South Bay	rec	N	2.7	2.7	3.1	15	24	34	43	51	58	64	70	74
South Bay	dry	N	2.9	2.8	3.1	13	20	28	35	41	48	53	58	63
Lower South Bay	wet	L	2.3	2.3	3.1	67	100							
Lower South Bay	rec	L	3.6	2.8	3.1	12	19	26	32	39	44	49	54	59
Lower South Bay	dry	L	3.6	2.8	3.1	22	42	60	73	82	89	93	96	97

Dissolved Nickel

Region	Season	Distribution	HisMean	μ	WQC	% Power at Sample Size of 2	3	4	5	6	7	8	9	10
Rivers	wet	L	1.8	1.8	52	100								
Rivers	rec	L	1.1	1.1	52	100								
Rivers	dry	L	1.0	1.0	52	100								
Suisun Bay	wet	L	1.9	1.9	8.2	90	100							
Suisun Bay	rec	L	1.4	1.4	8.2	99	100							
Suisun Bay	dry	L	1.3	1.3	8.2	99	100							
San Pablo Bay	wet	L	2.7	2.7	8.2	25	50	70	83	91	95	98	99	99
San Pablo Bay	rec	L	1.8	1.8	8.2	62	99	100						
San Pablo Bay	dry	L	1.8	1.8	8.2	80	100							
Central Bay	wet	L	1.7	1.7	8.2	81	100							
Central Bay	rec	L	1.3	1.3	8.2	95	100							
Central Bay	dry	L	1.4	1.4	8.2	86	100							
Golden Gate	dry	N	0.7	0.7	8.2	100								
South Bay	wet	N	2.6	2.6	8.2	100								
South Bay	rec	N	2.4	2.4	8.2	100								
South Bay	dry	N	2.6	2.6	8.2	99	100							
Lower South Bay	wet	L	3.0	3.0	8.2	77	100							
Lower South Bay	rec	L	3.1	3.1	8.2	75	100							
Lower South Bay	dry	L	3.6	3.6	8.2	54	96	100						

Dissolved Lead

Region	Season	Distribution	HisMean	μ	WQC	% Power at Sample Size of 2	3	4	5	6	7	8	9	10
Rivers	wet	N	0.16	0.16	2.5	100								
Rivers	rec	N	0.05	0.05	2.5	100								
Rivers	dry	N	0.05	0.05	2.5	100								
Suisun Bay	wet	L	0.07	0.07	8.1	71	100							
Suisun Bay	rec	L	0.03	0.03	8.1	90	100							
Suisun Bay	dry	L	0.01	0.01	8.1	97	100							
San Pablo Bay	wet	L	0.04	0.04	8.1	86	100							
San Pablo Bay	rec	L	0.01	0.01	8.1	100								
San Pablo Bay	dry	L	0.01	0.01	8.1	100								
Central Bay	wet	N	0.02	0.02	8.1	100								
Central Bay	rec	N	0.01	0.01	8.1	100								
Central Bay	dry	N	0.01	0.01	8.1	100								
Golden Gate	dry	N	0.01	0.01	8.1	100								
South Bay	wet	N	0.04	0.04	8.1	100								
South Bay	rec	N	0.04	0.04	8.1	100								
South Bay	dry	N	0.04	0.04	8.1	100								
Lower South Bay	wet	L	0.06	0.06	8.1	96	100							
Lower South Bay	rec	L	0.07	0.07	8.1	90	100							
Lower South Bay	dry	L	0.06	0.06	8.1	98	100							

Dissolved Zinc						% Power at Sample Size of								
Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	L	1.2	1.2	117	100								
Rivers	rec	L	0.5	0.5	117	100								
Rivers	dry	L	0.5	0.5	117	100								
Suisun Bay	wet	L	1.0	1.0	81	85	100							
Suisun Bay	rec	L	0.5	0.5	81	100								
Suisun Bay	dry	L	0.5	0.5	81	100								
San Pablo Bay	wet	L	1.1	1.1	81	88	100							
San Pablo Bay	rec	L	0.4	0.4	81	100								
San Pablo Bay	dry	L	0.5	0.5	81	100								
Central Bay	wet	L	1.0	1.0	81	100								
Central Bay	rec	L	0.4	0.4	81	100								
Central Bay	dry	L	0.6	0.6	81	100								
Golden Gate	dry	N	0.3	0.3	81	100								
South Bay	wet	L	1.7	1.7	81	99	100							
South Bay	rec	L	0.7	0.7	81	100								
South Bay	dry	L	0.7	0.7	81	100								
Lower South Bay	wet	L	2.7	2.7	81	85	100							
Lower South Bay	rec	L	1.6	1.6	81	86	100							
Lower South Bay	dry	L	1.2	1.2	81	87	100							

Appendix Table 3A.2. Power results for total water column metals and organic contaminants. Empirical data used in these calculations were RMP data measured between 1993 and 1998. Units are in ug/L for metals and pg/L for organics.

Total Silver — **% Power at Sample Size of**

Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	L	0.008	0.008	4.1	100								
Rivers	rec	L	0.012	0.012	4.1	53	96	100						
Rivers	dry	L	0.006	0.006	4.1	100								
Suisun Bay	wet	L	0.010	0.010	2.2	100								
Suisun Bay	rec	L	0.015	0.015	2.2	58	98	100						
Suisun Bay	dry	L	0.013	0.013	2.2	97	100							
San Pablo Bay	wet	L	0.011	0.011	2.2	95	100							
San Pablo Bay	rec	L	0.018	0.018	2.2	79	100							
San Pablo Bay	dry	L	0.016	0.016	2.2	96	100							
Central Bay	wet	L	0.008	0.008	2.2	99	100							
Central Bay	rec	L	0.008	0.008	2.2	69	100							
Central Bay	dry	L	0.014	0.014	2.2	95	100							
Golden Gate	dry	N	0.007	0.007	2.2	100								
South Bay	wet	L	0.014	0.014	2.2	100								
South Bay	rec	L	0.012	0.012	2.2	98	100							
South Bay	dry	L	0.027	0.027	2.2	99	100							
Lower South Bay	wet	L	0.020	0.020	2.2	97	100							
Lower South Bay	rec	L	0.029	0.029	2.2	96	100							
Lower South Bay	dry	L	0.016	0.016	2.2	89	100							

Total Arsenic — **% Power at Sample Size of**

Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	N	2.1	2.1	150	100								
Rivers	rec	N	1.5	1.5	150	100								
Rivers	dry	N	2.1	2.1	150	100								
Suisun Bay	wet	L	2.0	2.0	36	100								
Suisun Bay	rec	L	1.9	1.9	36	100								
Suisun Bay	dry	L	2.9	2.9	36	100								
San Pablo Bay	wet	L	2.2	2.2	36	99	100							
San Pablo Bay	rec	L	2.4	2.4	36	90	100							
San Pablo Bay	dry	L	3.1	3.1	36	94	100							
Central Bay	wet	L	1.8	1.8	36	100								
Central Bay	rec	L	1.6	1.6	36	100								
Central Bay	dry	L	2.3	2.3	36	100								
Golden Gate	dry	N	1.6	1.6	36	100								
South Bay	wet	L	2.4	2.4	36	100								
South Bay	rec	L	2.1	2.1	36	100								
South Bay	dry	L	3.7	3.7	36	100								
Lower South Bay	wet	L	2.5	2.5	36	100								
Lower South Bay	rec	L	2.9	2.9	36	100								
Lower South Bay	dry	L	4.5	4.5	36	100								

Total Cadmium						**% Power at Sample Size of**								
Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	N	0.03	0.03	2.5	100								
Rivers	rec	N	0.02	0.02	2.5	100								
Rivers	dry	N	0.03	0.03	2.5	100								
Suisun Bay	wet	L	0.03	0.03	9.4	100								
Suisun Bay	rec	L	0.03	0.03	9.4	100								
Suisun Bay	dry	L	0.06	0.06	9.4	100								
San Pablo Bay	wet	L	0.05	0.05	9.4	99	100							
San Pablo Bay	rec	L	0.06	0.06	9.4	100								
San Pablo Bay	dry	L	0.10	0.10	9.4	100								
Central Bay	wet	L	0.04	0.04	9.4	100								
Central Bay	rec	L	0.05	0.05	9.4	100								
Central Bay	dry	L	0.10	0.10	9.4	100								
Golden Gate	dry	N	0.08	0.08	9.4	100								
South Bay	wet	N	0.08	0.08	9.4	100								
South Bay	rec	N	0.07	0.07	9.4	100								
South Bay	dry	N	0.13	0.13	9.4	100								
Lower South Bay	wet	L	0.08	0.08	9.4	100								
Lower South Bay	rec	L	0.09	0.09	9.4	100								
Lower South Bay	dry	L	0.13	0.13	9.4	100								

Total Chromium						**% Power at Sample Size of**								
Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	L	6.8	6.8	11	20	37	53	65	75	83	88	92	94
Rivers	rec	L	3.7	3.7	11	41	83	97	100					
Rivers	dry	L	3.9	3.9	11	88	100							
Suisun Bay	wet	L	9.9	9.9	50	79	100							
Suisun Bay	rec	L	8.9	8.9	50	71	100							
Suisun Bay	dry	L	9.3	9.3	50	51	95	100						
San Pablo Bay	wet	L	7.4	7.4	50	41	83	97	100					
San Pablo Bay	rec	L	13.0	13.0	50	29	60	81	92	97	99	100		
San Pablo Bay	dry	L	7.2	7.2	50	36	76	94	99	100				
Central Bay	wet	L	1.8	1.8	50	86	100							
Central Bay	rec	L	1.0	1.0	50	70	100							
Central Bay	dry	L	1.3	1.3	50	90	100							
Golden Gate	dry	N	0.3	0.3	50	100								
South Bay	wet	N	3.9	3.9	50	100								
South Bay	rec	N	4.4	4.4	50	99	100							
South Bay	dry	N	4.5	4.5	50	100								
Lower South Bay	wet	L	4.6	4.6	50	65	99	100						
Lower South Bay	rec	L	9.0	9.0	50	47	91	99	100					
Lower South Bay	dry	L	5.5	5.5	50	66	100							

Total Copper — % Power at Sample Size of

Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	L	4.5	4.5	9.3	46	89	99	100					
Rivers	rec	L	3.7	3.7	9.3	77	100							
Rivers	dry	L	3.3	3.3	9.3	97	100							
Suisun Bay	wet	L	4.9	3.3	3.7	13	20	27	33	39	45	50	55	60
Suisun Bay	rec	L	5.3	3.3	3.7	10	14	18	21	25	28	31	35	38
Suisun Bay	dry	L	5.0	3.3	3.7	8	10	12	14	16	18	19	21	23
San Pablo Bay	wet	L	4.5	3.3	3.7	8	10	12	13	15	17	18	20	21
San Pablo Bay	rec	L	4.8	3.3	3.7	7	8	9	10	11	12	13	14	14
San Pablo Bay	dry	L	4.3	3.3	3.7	7	9	10	11	12	14	15	16	17
Central Bay	wet	L	2.4	2.4	3.7	31	65	86	95	98	99	100		
Central Bay	rec	L	1.8	1.8	3.7	43	86	98	100					
Central Bay	dry	L	2.1	2.1	3.7	30	61	83	93	97	99	100		
South Bay	wet	N	3.5	3.3	3.7	10	15	19	23	27	31	34	38	41
South Bay	rec	N	3.8	3.3	3.7	8	10	11	13	14	16	17	19	20
South Bay	dry	N	4.3	3.3	3.7	9	12	15	18	21	23	26	28	31
Lower South Bay	wet	L	3.3	3.3	3.7	10	14	18	22	26	30	33	36	39
Lower South Bay	rec	L	5.9	3.3	3.7	9	11	14	16	19	21	23	25	27
Lower South Bay	dry	L	6.4	3.3	3.7	13	21	30	37	44	51	56	62	66

Total Mercury — % Power at Sample Size of

Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	L	0.009	0.009	0.012	17	30	43	54	64	71	78	83	87
Rivers	rec	L	0.006	0.006	0.012	23	44	63	77	86	91	95	97	98
Rivers	dry	L	0.005	0.005	0.012	47	91	99	100					
Suisun Bay	wet	L	0.012	0.012	0.025	45	89	99	100					
Suisun Bay	rec	L	0.013	0.013	0.025	27	56	77	89	95	98	99	100	
Suisun Bay	dry	L	0.011	0.011	0.025	27	54	75	87	94	97	99	99	100
San Pablo Bay	wet	L	0.014	0.014	0.025	14	23	31	40	47	54	60	65	70
San Pablo Bay	rec	L	0.015	0.015	0.025	12	17	23	29	34	39	43	48	52
San Pablo Bay	dry	L	0.010	0.010	0.025	25	49	70	83	91	95	97	99	99
Central Bay	wet	L	0.005	0.005	0.025	42	85	98	100					
Central Bay	rec	L	0.003	0.003	0.025	54	96	100						
Central Bay	dry	L	0.004	0.004	0.025	55	97	100						
Golden Gate	dry	N	0.009	0.009	0.025	100								
South Bay	wet	L	0.009	0.009	0.025	30	61	83	93	97	99	100		
South Bay	rec	L	0.008	0.008	0.025	16	27	38	47	56	64	70	76	80
South Bay	dry	L	0.008	0.008	0.025	32	66	87	96	99	100			
Lower South Bay	wet	L	0.016	0.016	0.025	17	29	41	51	61	68	75	80	84
Lower South Bay	rec	L	0.019	0.019	0.025	8	11	13	15	17	19	21	23	25
Lower South Bay	dry	L	0.013	0.013	0.025	24	48	68	81	90	94	97	98	99

Total Nickel						**% Power at Sample Size of**								
Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	L	5.6	5.6	52	69	100							
Rivers	rec	L	3.7	3.7	52	98	100							
Rivers	dry	L	3.3	3.3	52	100								
Suisun Bay	wet	L	6.8	6.8	8.3	15	24	34	43	51	58	64	70	75
Suisun Bay	rec	L	6.6	6.6	8.3	14	22	30	37	44	51	57	62	67
Suisun Bay	dry	L	6.1	6.1	8.3	16	26	37	47	56	63	70	75	80
San Pablo Bay	wet	L	7.9	7.5	8.3	7	8	8	9	10	10	11	12	12
San Pablo Bay	rec	L	7.4	7.4	8.3	7	8	9	10	11	12	13	13	14
San Pablo Bay	dry	L	5.9	5.9	8.3	13	21	29	36	42	49	54	59	64
Central Bay	wet	L	3.0	3.0	8.3	53	96	100						
Central Bay	rec	L	2.1	2.1	8.3	64	99	100						
Central Bay	dry	L	2.7	2.7	8.3	48	92	99	100					
Golden Gate	dry	N	1.1	1.1	8.3	100								
South Bay	wet	N	5.6	5.6	8.3	20	37	53	66	76	83	89	92	95
South Bay	rec	N	4.8	4.8	8.3	24	49	69	82	90	95	97	99	99
South Bay	dry	N	5.9	5.9	8.3	24	47	67	80	89	94	96	98	99
Lower South Bay	wet	L	6.8	6.8	8.3	12	19	25	32	38	43	48	53	57
Lower South Bay	rec	L	8.6	7.5	8.3	7	9	11	12	13	14	16	17	18
Lower South Bay	dry	L	9.0	7.5	8.3	9	11	14	16	18	20	22	24	26

Total Lead						**% Power at Sample Size of**								
Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	L	0.84	0.84	3.2	69	100							
Rivers	rec	L	0.72	0.72	3.2	91	100							
Rivers	dry	L	0.78	0.78	3.2	85	100							
Suisun Bay	wet	L	1.05	1.05	8.5	87	100							
Suisun Bay	rec	L	1.29	1.29	8.5	72	100							
Suisun Bay	dry	L	1.39	1.39	8.5	56	97	100						
San Pablo Bay	wet	L	1.07	1.07	8.5	53	96	100						
San Pablo Bay	rec	L	1.51	1.51	8.5	35	74	93	98	100				
San Pablo Bay	dry	L	1.15	1.15	8.5	46	89	99	100					
Central Bay	wet	L	0.39	0.39	8.5	83	100							
Central Bay	rec	L	0.22	0.22	8.5	79	100							
Central Bay	dry	L	0.42	0.42	8.5	75	100							
Golden Gate	dry	N	0.23	0.23	8.5	100								
South Bay	wet	L	0.73	0.73	8.5	65	99	100						
South Bay	rec	L	0.75	0.75	8.5	65	99	100						
South Bay	dry	L	1.03	1.03	8.5	70	100							
Lower South Bay	wet	L	1.16	1.16	8.5	58	98	100						
Lower South Bay	rec	L	1.92	1.92	8.5	42	85	98	100					
Lower South Bay	dry	L	1.74	1.74	8.5	60	99	100						

Total Selenium						% Power at Sample Size of								
Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	N	0.18	0.18	5	100								
Rivers	rec	N	0.20	0.20	5	100								
Rivers	dry	N	0.39	0.39	5	100								
Suisun Bay	wet	L	0.17	0.17	5	100								
Suisun Bay	rec	L	0.19	0.19	5	100								
Suisun Bay	dry	L	0.25	0.25	5	100								
San Pablo Bay	wet	L	0.18	0.18	5	100								
San Pablo Bay	rec	L	0.24	0.24	5	100								
San Pablo Bay	dry	L	0.27	0.27	5	98	100							
Central Bay	wet	L	0.16	0.16	5	99	100							
Central Bay	rec	L	0.17	0.17	5	100								
Central Bay	dry	L	0.19	0.19	5	94	100							
Golden Gate	dry	N	0.35	0.35	5	100								
South Bay	wet	L	0.25	0.25	5	79	100							
South Bay	rec	L	0.30	0.30	5	94	100							
South Bay	dry	L	0.31	0.31	5	99	100							
Lower South Bay	wet	L	0.38	0.38	5	99	100							
Lower South Bay	rec	L	0.48	0.48	5	92	100							
Lower South Bay	dry	L	0.48	0.48	5	81	100							

Total Zinc						% Power at Sample Size of								
Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	L	6.9	6.9	119	93	100							
Rivers	rec	L	4.8	4.8	119	95	100							
Rivers	dry	L	4.8	4.8	119	100								
Suisun Bay	wet	L	8.3	8.3	86	95	100							
Suisun Bay	rec	L	8.6	8.6	86	91	100							
Suisun Bay	dry	L	8.7	8.7	86	69	100							
San Pablo Bay	wet	L	7.7	7.7	86	64	99	100						
San Pablo Bay	rec	L	8.6	8.6	86	59	98	100						
San Pablo Bay	dry	L	7.5	7.5	86	59	98	100						
Central Bay	wet	L	3.1	3.1	86	97	100							
Central Bay	rec	L	1.8	1.8	86	91	100							
Central Bay	dry	L	2.8	2.8	86	94	100							
Golden Gate	dry	N	1.6	1.6	86	100								
South Bay	wet	L	5.5	5.5	86	87	100							
South Bay	rec	L	4.2	4.2	86	74	100							
South Bay	dry	L	5.5	5.5	86	90	100							
Lower South Bay	wet	L	9.4	9.4	86	62	99	100						
Lower South Bay	rec	L	10.9	10.9	86	61	99	100						
Lower South Bay	dry	L	9.4	9.4	86	72	100							

Total Chlordanes — % Power at Sample Size of

Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	N	147	147	590	83	100							
San Pablo Bay	wet	L	134	134	590	43	87	98	100					
San Pablo Bay	rec	L	218	218	590	48	92	99	100					
San Pablo Bay	dry	L	109	109	590	43	87	98	100					
Central Bay	wet	N	140	140	590	97	100							
Central Bay	rec	N	105	105	590	99	100							
Central Bay	dry	N	77	77	590	100								
South Bay	wet	N	307	307	590	28	58	80	91	96	98	99	100	
South Bay	rec	N	241	241	590	36	74	93	98	100				
South Bay	dry	N	133	133	590	98	100							

Total DDT's — % Power at Sample Size of

Region	Season	Distribution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Rivers	wet	L	618	531	590	8	10	11	13	14	16	17	18	20
Rivers	rec	L	533	531	590	10	13	17	21	24	27	30	33	36
Rivers	dry	L	463	463	590	11	16	21	25	30	34	38	42	45
San Pablo Bay	wet	L	708	531	590	7	8	9	10	11	12	12	13	14
San Pablo Bay	rec	L	1034	531	590	7	8	9	10	11	11	12	13	13
San Pablo Bay	dry	L	592	531	590	7	8	9	10	10	11	12	13	13
Central Bay	wet	L	233	233	590	34	71	91	97	99	100			
Central Bay	rec	L	274	274	590	24	47	67	80	88	93	96	98	99
Central Bay	dry	L	187	187	590	50	94	100						
South Bay	wet	L	246	246	590	18	33	48	60	70	77	83	88	91
South Bay	rec	L	454	454	590	11	15	20	25	29	33	37	40	44
South Bay	dry	L	150	150	590	49	93	100						

Total Polycyclic Aromatic Hydrocarbons (PAH's) — **% Power at Sample Size of**

(not available: data compilation and calculation error was discovered after the time of analyses)

Total Total PCB's

Region	Season	Distribution	HisMean	μ	WQC	% Power at Sample Size of 2	3	4	5	6	7	8	9	10
Rivers	wet	N	149	149	170	10	14	17	21	24	28	31	34	37
Rivers	rec	N	349	153	170	8	11	13	15	17	18	20	22	24
Rivers	dry	N	225	153	170	8	10	12	14	16	17	19	21	22
San Pablo Bay	wet	L	303	153	170	7	7	8	9	10	10	11	11	12
San Pablo Bay	rec	L	869	153	170	6	7	8	8	9	9	9	10	10
San Pablo Bay	dry	L	534	153	170	7	7	8	9	10	10	11	11	12
Central Bay	wet	L	296	153	170	8	10	12	14	15	17	18	20	21
Central Bay	rec	L	337	153	170	6	7	8	9	9	10	10	11	11
Central Bay	dry	L	366	153	170	8	9	11	12	14	15	16	18	19
South Bay	wet	L	804	153	170	7	8	9	10	10	11	12	13	13
South Bay	rec	L	980	153	170	7	8	9	10	11	12	13	14	15
South Bay	dry	L	704	153	170	8	10	12	14	16	18	19	21	22

Appendix Table 3A.3. Power results for sediment metals and organic contaminants. Empirical data used in these calculations were RMP data measured between 1993 and 1998. Units are in mg/Kg for metals and μg/Kg for organics.

Silver — % Power at Sample Size of

Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers	N	0.11	0.11	1	100								
Suisun Bay	L	0.22	0.22	1	38	78	95	99	100				
San Pablo Bay	N	0.33	0.33	1	70	100							
Central Bay	N	0.33	0.33	1	81	100							
South Bay	N	0.44	0.44	1	92	100							
Lower South Bay	N	0.47	0.47	1	95	100							

Arsenic — % Power at Sample Size of

Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers	L	6.03	6.03	8.2	17	31	44	55	65	73	79	84	88
Suisun Bay	N	8.20	7.38	8.2	6	7	8	9	9	10	10	11	11
San Pablo Bay	N	10.46	7.38	8.2	7	8	10	11	12	12	13	14	15
Central Bay	N	8.86	7.38	8.2	9	11	14	16	18	20	22	24	26
South Bay	N	7.89	7.38	8.2	15	25	35	44	53	60	66	72	76
Lower South Bay	L	6.98	6.98	8.2	20	37	54	67	77	84	89	92	95

Cadmium — % Power at Sample Size of

Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers	L	0.22	0.22	1.2	73	100							
Suisun Bay	L	0.21	0.21	1.2	56	97	100						
San Pablo Bay	L	0.21	0.21	1.2	52	95	100						
Central Bay	L	0.19	0.19	1.2	63	99	100						
South Bay	N	0.16	0.16	1.2	100								
Lower South Bay	L	0.18	0.18	1.2	100								

Chromium — % Power at Sample Size of

Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers	L	80.44	72.90	81	13	20	26	33	39	45	50	55	60
Suisun Bay	L	85.68	72.90	81	9	12	15	18	20	23	25	28	30
San Pablo Bay	L	100.31	72.90	81	9	12	14	17	19	22	24	26	29
Central Bay	L	92.41	72.90	81	9	13	16	19	22	25	27	30	33
South Bay	L	91.44	72.90	81	14	22	30	38	45	52	58	63	68
Lower South Bay	N	104.22	72.90	81	12	18	24	29	34	39	44	49	53

Copper — % Power at Sample Size of

Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers	N	28.55	28.55	34	11	16	21	26	30	35	39	43	46
Suisun Bay	L	31.17	30.60	34	7	7	8	9	10	10	11	11	12
San Pablo Bay	L	39.32	30.60	34	7	8	9	10	10	11	12	13	13
Central Bay	N	39.24	30.60	34	7	7	8	9	10	10	11	12	12
South Bay	N	37.68	30.60	34	16	28	39	50	59	66	73	78	82
Lower South Bay	N	40.29	30.60	34	12	18	23	29	34	39	44	48	53

Mercury — % Power at Sample Size of

Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers	L	0.11	0.11	0.15	9	11	13	16	18	20	22	24	26
Suisun Bay	N	0.19	0.14	0.15	6	7	7	7	8	8	9	9	9
San Pablo Bay	L	0.22	0.14	0.15	6	7	8	9	9	10	11	11	12
Central Bay	L	0.21	0.14	0.15	7	8	9	10	10	11	12	13	13
South Bay	L	0.29	0.14	0.15	11	15	20	24	29	33	36	40	44
Lower South Bay	L	0.33	0.14	0.15	12	19	25	31	37	43	48	53	57

Nickel — % Power at Sample Size of

Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers	N	79.00	18.81	20.9	6	7	8	8	9	9	10	10	10
Suisun Bay	N	85.05	18.81	20.9	6	6	7	7	8	8	8	8	9
San Pablo Bay	L	89.04	18.81	20.9	12	17	23	29	34	39	43	48	52
Central Bay	L	79.40	18.81	20.9	11	17	22	27	32	36	40	45	48
South Bay	L	84.61	18.81	20.9	14	23	32	40	47	54	60	66	70
Lower South Bay	L	95.51	18.81	20.9	20	37	54	67	77	84	89	92	95

Lead — % Power at Sample Size of

Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers	N	12.79	12.79	46.7	95	100							
Suisun Bay	L	15.35	15.35	46.7	30	63	84	94	98	99	100		
San Pablo Bay	L	21.66	21.66	46.7	28	58	80	91	96	98	99	100	
Central Bay	L	22.40	22.40	46.7	36	74	93	98	100				
South Bay	N	23.89	23.89	46.7	74	100							
Lower South Bay	N	25.68	25.68	46.7	43	86	98	100					

Zinc — % Power at Sample Size of

Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers	N	84.39	84.39	150	53	95	100						
Suisun Bay	L	93.73	93.73	150	21	39	56	69	79	86	91	94	96
San Pablo Bay	N	116.90	116.90	150	20	37	53	66	75	83	88	92	94
Central Bay	L	102.62	102.62	150	27	54	75	87	94	97	99	99	100
South Bay	L	114.19	114.19	150	42	85	98	100					
Lower South Bay	L	126.32	126.32	150	21	39	56	69	79	86	91	94	96

Chlordanes					**% Power at Sample Size of**								
Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers		NA											
Suisun Bay		NA											
San Pablo Bay		NA											
Central Bay	L	0.37	0.37	0.5	10	14	18	22	25	29	32	35	38
South Bay	N	0.51	0.45	0.5	25	49	69	83	90	95	97	99	99
Lower South Bay	L	1.04	0.45	0.5	7	8	9	10	11	12	13	14	15

DDTs					**% Power at Sample Size of**								
Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers	N	1.02	1.02	1.58	18	32	46	58	68	75	81	86	90
Suisun Bay	N	3.71	1.42	1.58	5	6	6	6	6	6	6	7	7
San Pablo Bay	L	2.25	1.42	1.58	6	7	7	8	8	8	9	9	10
Central Bay	N	2.87	1.42	1.58	6	7	8	8	9	9	10	10	10
South Bay	L	2.78	1.42	1.58	7	9	10	12	13	14	15	16	17
Lower South Bay	N	5.35	1.42	1.58	5	6	6	6	6	6	7	7	7

High Molecular Weight Polycyclic Aromatic Hydrocarbons (THPAHs)					**% Power at Sample Size of**								
Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers	L	67.17	67.17	1700	47	91	99	100					
Suisun Bay	N	371.80	371.80	1700	73	100							
San Pablo Bay	L	798.23	798.23	1700	10	15	19	23	27	31	35	38	42
Central Bay	N	2499.73	1530.00	1700	6	6	7	7	7	7	8	8	8
South Bay	L	1867.16	1530.00	1700	8	10	11	13	15	16	18	19	20
Lower South Bay	L	1697.83	1530.00	1700	9	12	15	18	20	23	25	27	30

Low Molecular Weight Polycyclic Aromatic Hydrocarbons (TLPAHs)					**% Power at Sample Size of**								
Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers	L	4.55	4.55	552	64	99	100						
Suisun Bay	N	51.64	51.64	552	100								
San Pablo Bay	L	81.97	81.97	552	26	52	73	86	93	96	98	99	100
Central Bay	L	250.76	250.76	552	12	19	25	32	38	43	48	53	57
South Bay	N	223.92	223.92	552	79	100							
Lower South Bay	L	170.17	170.17	552	58	98	100						

Total Polycyclic Aromatic Hydrocarbons (TPAHs)					**% Power at Sample Size of**								
Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers	L	72.38	72.38	4022	56	97	100						
Suisun Bay	N	422.73	422.73	4022	99	100							
San Pablo Bay	L	890.39	890.39	4022	18	32	46	58	67	75	81	86	90
Central Bay	N	2920.32	2920.32	4022	10	13	17	20	23	26	29	32	35
South Bay	L	2084.10	2084.10	4022	31	65	86	95	98	99	100		
Lower South Bay	L	1881.36	1881.36	4022	45	89	99	100					

Polychlorinated Biphenyls (PCBs)					% Power at Sample Size of								
Region	Distribution	HisMean	μ	ERL	2	3	4	5	6	7	8	9	10
Rivers		NA											
Suisun Bay	L	1.13	1.13	22.7	71	100							
San Pablo Bay	N	2.82	2.82	22.7	90	100							
Central Bay	L	6.20	6.20	22.7	20	38	54	68	77	84	89	93	95
South Bay	L	5.55	5.55	22.7	26	53	74	87	93	97	98	99	100
Lower South Bay	N	7.68	7.68	22.7	56	97	100						

APPENDIX 3B
Power computations & example calculation

The statistical test described in (Figure 3.1 in Chapter 3) is equivalent to a one-tailed, one-sample t-test, which involves comparison of an assumed underlying sample mean with a single constant value that has no underlying variance. Accordingly, we use the power calculations for the one-sample t-test. The power calculations for a chemical constituent require values for the sample size (n, the number of stations), nominal type-1 error rate (α), the assumed underlying sub-region mean (μ), the expected variance of measurements within the sub-region during a survey (σ^2), and the guideline value (G).

For a particular combination of n, α, μ, G, and σ^2, the power can be computed as follows.

1. Compute the noncentrality parameter $\lambda = \sqrt{(\mu - G)^2 / (\sigma^2 / n)}$.

2. Obtain $T_{crit} = t_{n-1,1-\alpha}$, where $t_{n-1,1-\alpha}$ is the $1-\alpha$ quantile of the central Student t distribution with *n-1* degrees of freedom, and α is the nominal type-1 error. Here *n* is the sample size for the sub-region during a survey.

3. Obtain the type-2 error β = the probability that an observation from a *noncentral* t distribution (with *n-1* degrees of freedom and noncentrality parameter λ) is less than or equal to T_{crit}.

4. Power = $1 - \beta$.

The underlying sub-region mean (μ) and within-subregion variance (σ^2) for a survey are not known, so are based on estimates from the available historical data. To compute the historical estimates for a sub-region, the mean and variance for each *survey* are first computed. Following this, the means of the survey means and variances are computed to give the values for μ and σ^2 respectively. This approach estimates the average within-survey variance for the subregion, and gives equal weight to each survey in the estimate. The within-survey variance for the sub-region is appropriate since the statistical test will be repeated separately for each survey in each sub-region.

It should be noted that after repeated samplings within a sub-region, a pooled estimate of σ^2 could be computed if it can be shown that the within-region variance does not vary significantly from year to year. This would allow for a t-test with more than n-1 degrees of freedom and consequently higher power than a test without a pooled variance estimate.

Example Computations

Assumed underlying mean (μ) = 3.3
Assumed underlying within-region variance for a survey (σ^2) = 0.65

WQC guideline (G) = 5.0

Sample size within the sub-region during a survey (n) = 3

Type-1 error rate of the t-test (α) = 0.05

The non-centrality parameter (λ) = $\sqrt{(3.3-5.0)^2/(0.65/3)}$ = 3.65

$T_{crit} = t_{n-1,1-\alpha} = t_{2,.95} = 2.92$

If using SAS, you would use function TINV(0.95,2) to compute this.

Type-2 error (β) = .25386.

If using SAS, you would use function PROBT(2.92,2,3.65) to compute this.

Power = 1 - β = 0.74614

Addendum
New Power Analyses for the Lower South Bay

Since the completion of the Re-design Work Group process, new site-specific aquatic life water quality objectives for dissolved copper and nickel were adopted by the State of California in 2003 and approved by the U.S. EPA for Lower South San Francisco Bay (south of the Dumbarton Bridge). The dissolved copper objective changed from 4.8 µg/L to 10.8 µg/L acute (exposure for one hour) and from 3.1 µg/L to 6.9 µg/L chronic (exposure for four days). The dissolved nickel objectives changed from 74 µg/L to 62.4 µg/L acute and from 8.2 µg/L to 11.9 µg/L chronic. Table 1 shows power results for that region for the new objectives. With these new objectives, only 2 or 3 samples are needed to achieve greater than 80% power.

Table 1. Power Results using the new Lower South Bay site-specific objectives for copper and nickel. Refer to appendix table 3A for a description of table elements.

Dissolved Copper — **% Power at Sample Size of**

Region	Season	Distri-bution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Lower South Bay	dry	L	3.6	3.6	6.9	91	100							
Lower South Bay	rec	L	3.6	3.6	6.9	63	99	100						
Lower South Bay	wet	L	2.3	2.3	6.9	100								

Dissolved Nickel — **% Power at Sample Size of**

Region	Season	Distri-bution	HisMean	μ	WQC	2	3	4	5	6	7	8	9	10
Lower South Bay	dry	L	3.6	3.6	11.9	72	100							
Lower South Bay	rec	L	3.1	3.1	11.9	89	100							
Lower South Bay	wet	L	3.0	3.0	11.9	90	100							

Because the Re-design Work Group used the power analyses for dissolved copper as guidance in determining the number of water samples to allocate into each region, it may be warranted to reallocate the number of samples collected in the Lower South Bay (currently six samples are collected annually). The Regional Board and RMP staff considered making these adjustments in 2003 but decided that it would be best to re-evaluate statistical power against various effects thresholds with data from the new random sample design, which will provide better spatial coverage of the Estuary for estimating contaminant concentrations within each region.

Adding Additional Historical Sites for Water Monitoring

The Re-design Work Group originally felt that since estuarine water bodies do not maintain site-specific characteristics between sampling periods only the two Rivers stations and the Golden Gate background stations should be maintained as historic sites. However, the Regional Water Board NPDES permit department uses contaminant data from three historic RMP sites in their permitting process (Sacramento River (BG20), Yerba Buena Island (BC10) in the Central Bay, and Dumbarton Bridge (BA30) in the South Bay regions). Therefore two historic sites (not originally in the new sampling design) were added back into the monitoring program in 2003 (BC10 and BA30) and two randomly allocated samples were dropped (one each from the South Bay and Lower South Bay regions) in order to maintain program costs.