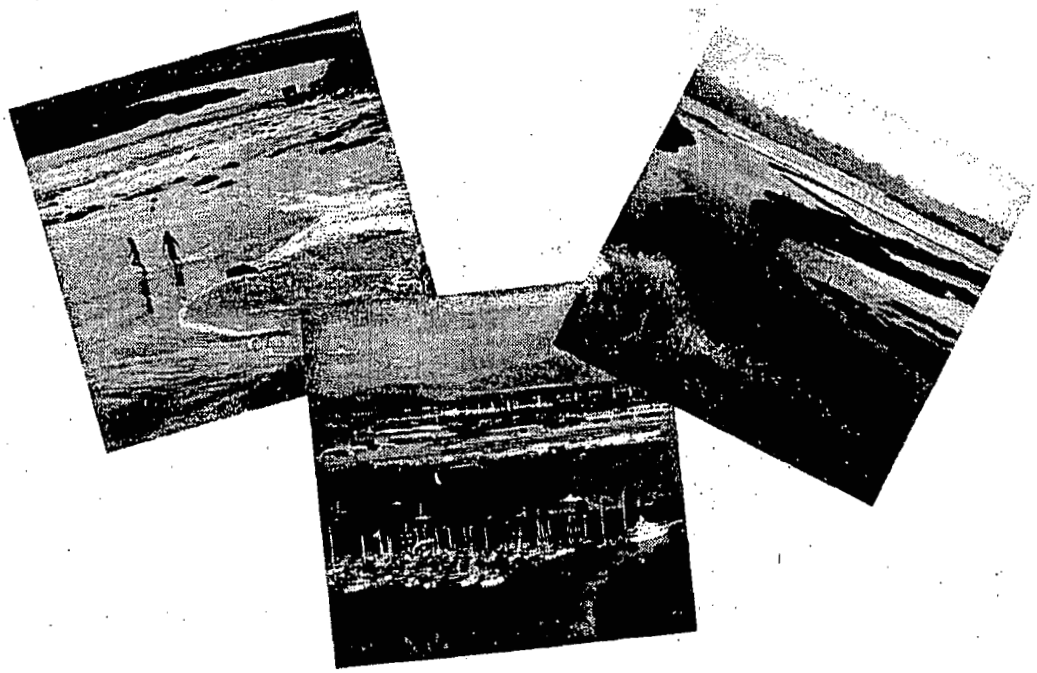

Chemistry, Toxicity and Benthic Community Conditions in Sediments of Selected Southern California Bays and Estuaries

May, 1997

California State Water Resources Control Board

U.S. Environmental Protection Agency

National Oceanic and Atmospheric Administration

California Department of Fish and Game

University of California, Santa Cruz

Moss Landing Marine Laboratories

Columbia Analytical Services

CHEMISTRY, TOXICITY AND BENTHIC COMMUNITY CONDITIONS IN SEDIMENTS OF SELECTED SOUTHERN CALIFORNIA BAYS AND ESTUARIES

May, 1997

California State Water Resources Control Board

U.S. Environmental Protection Agency

National Oceanic and Atmosperic Administration

California Department of Fish and Game

University of California, Santa Cruz

Moss Landing Marine Laboratories

Columbia Analytical Services

Using Cumulative Distribution Frequencies (CDFs) to characterize spatial extent

Cumulative Distribution Frequencies (CDFs) were determined using known areas of each sampling strata normalized to the number of samples per strata. By combining the area represented by each sample with their toxicity designations in a cumulative manner, the CDF's indicated the percentage of total area sampled that was toxic. Sample toxicity was determined from comparisons with laboratory controls as described above; each sample with a mean significantly different from, and less than 80% of, the laboratory control mean was considered "toxic". Calculations used to derive percent areas determined to be toxic are shown on worksheets in Appendix F. CDF's were generated from toxicity tests using *Rhepoxynius* (solid phase) and *Strongylocentrotus* fertilization and larval development in pore water; these were based on 30 random samples. A CDF was also generated from the *Ampelisca abdita* (solid phase) toxicity test based on a smaller subset of 15 random samples. CDF's were used to determine the percentage of area toxic for each toxicity test protocol. A 95% Confidence Interval was calculated for each areal toxicity determination based on EMAP methods.

The reference envelope approach to distinguish the most toxic samples

The second objective of this study was to assist in the identification of "toxic hotspots", where adverse biological impacts are observed in areas with localized concentrations of pollutants. Identification of problem sites is an essential step in prioritizing efforts to improve sediment and water quality through regulation and remediation programs. An efficient use of funds requires that efforts be focused on localized areas that are significantly more toxic than optimal ambient conditions that presumably exist in the greater portion of the Southern California bays, estuaries, and coastal lagoons. In this study, we have employed a "reference envelope" statistical approach (Smith, 1995) to identify samples that exhibit significantly greater toxicity than expected in the area as a whole.

The reference envelope approach uses data from "reference sites" to characterize the response expected from sites in the absence of localized pollution. Using data from the reference site population, a tolerance limit was calculated for comparison with data from test sites. Samples with toxicity values greater than the tolerance limit were considered toxic relative to the optimal ambient condition of the area studied.

This relative standard established using reference sites was conceptually different from what might be termed the absolute standard of test organism response in laboratory controls. Rather than

comparing sample data to control data using t-tests, with laboratory replication used to characterize the variance component (as in the "t-test-control approach" described above), the reference envelope approach compares sample data against a percentile of the reference population of data values, using variation among reference sites as the variance component. The reference envelope variance component, therefore, includes variation among laboratory replicates, among field replicates, among sites, and among sampling events.

The reference stations were assumed to be a random sample from an underlying population of reference locations that served as a standard for what we considered relatively non-impacted conditions. The toxicity measured at different reference locations will vary due to the different local conditions that can affect the toxicity results. In order to determine whether sediments from a test location were toxic, the bioassay results for the test locations were compared with the bioassay results from the population of reference locations.

If it is assumed that the bioassay results from the population of reference locations were normally distributed, then we could get an idea of the probability that the test sediment was from the underlying reference station distribution. For example, if the result for a test sediment was at the first percentile of the underlying reference location distribution (in the direction of toxicity), then there would be approximately a 1% chance that the test sediment was from the distribution of reference locations.

The toxicity level at the first percentile of the reference distribution was not known because the number of samples from the underlying distribution were limited. Therefore, the location of the first percentile could only be estimated. If this value was estimated a large number of times using different random samples from the reference distribution, a non-central t distribution of estimates would be obtained, with the distribution mode at the actual first percentile (**Figure 3**). This figure shows that for this distribution of estimates, about one half of the time the estimate from the sample will be above the actual first percentile. Ideally, it would be preferable to identify an estimated toxicity value that would cover the actual first percentile for a large percentage of the estimates (say 95% of the time). This value can be obtained from the left tail of the distribution of estimates where 5% of the estimates are less than the chosen value. We define p as the percentile of interest, and alpha as the acceptable error probability associated with an estimate of the pth percentile. Thus, in this example, p=1 and alpha = .05.

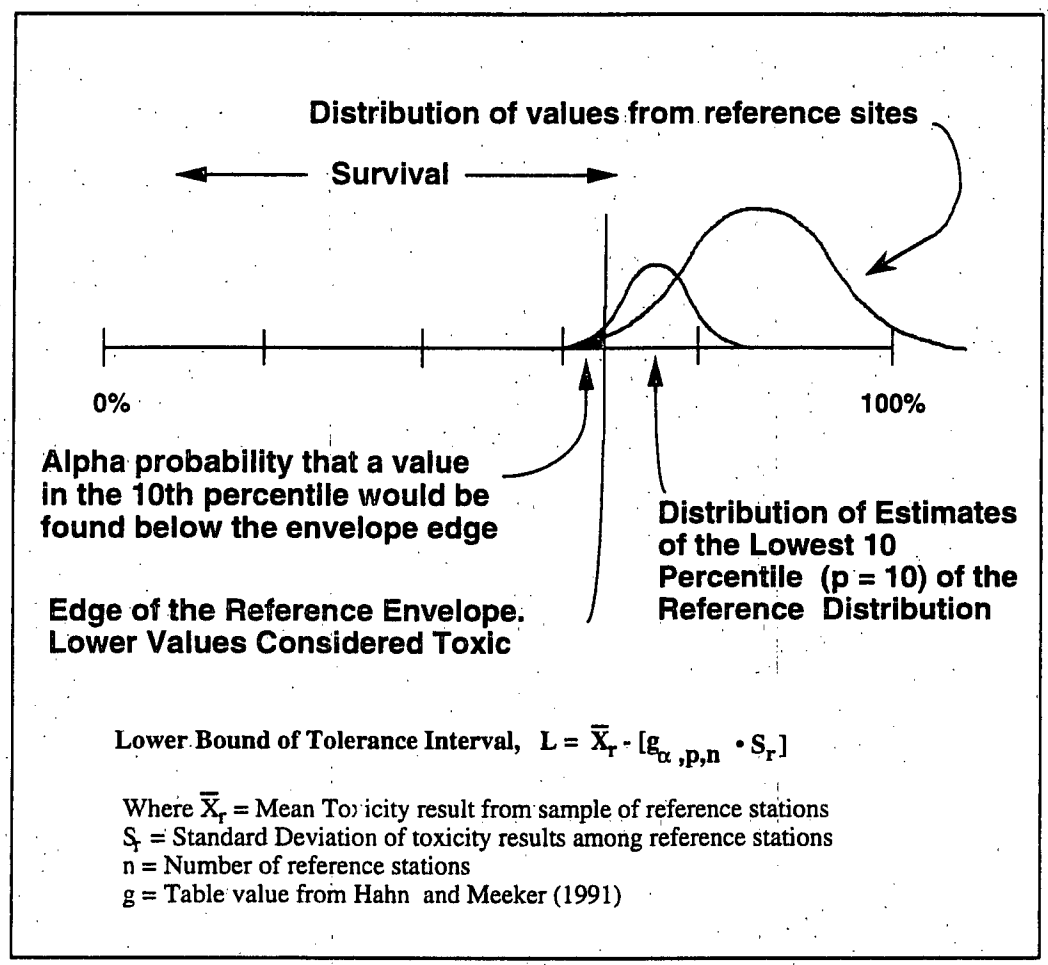

Figure 3. Schematic illustration of the method for determining the lower tolerance interval bound (edge of the reference envelope) to determine sample toxicity relative to a percentile of the reference site distribution.

The toxicity level that will cover the pth percentile 1 minus alpha proportion of the time can be computed as the lower bound (L) of a tolerance interval (Vardeman 1992) as follows:

$$L = X_r - [\, g_{a,p,n} * S_r \,]$$

where X_r is the mean of the sample of reference stations, S_r is the standard deviation of the toxicity results among the reference stations, and n is the number of reference stations. The g values, for the given alpha, p, and n values, can be obtained from tables in Hahn and Meeker (1991) or Gilbert (1987). S contains the within- and between- location variability expected among reference locations. If the reference stations are sampled at different times, then S will also incorporate between-time variability. L is called the "edge of the reference envelope" because it represents a cutoff toxicity level we will use to distinguish toxic from non-toxic sediments. The value used for p will depend on the level of certainty needed for a particular regulatory situation. In this study we chose p values equal to 1 and 10%, to distinguish the most toxic samples, that is, the samples that we are 95% certain are the most toxic 1 and 10% relative to the reference conditions defined below.

Reference station selection for use in developing reference envelope

Reference stations were selected to represent optimal ambient conditions available in the Southern California bays and estuaries sampled, based on available chemistry and benthic community data. Toxicity data were not used in the selection process. Stations were selected if both of the following criteria were met: 1) the benthic communities appeared relatively undisturbed (based on indices described in the benthic community analysis section), and 2) sediment chemical concentrations were below Effects Range Median (ERM) levels (Long *et al.*, 1995) and Probable Effects levels (PELs; McDonald, 1994). Among all stations, both randomly and non-randomly selected, a total of 43 samples were analyzed for toxicity, chemistry and benthic ecology in this study. After screening these 43 samples, six stations were selected as reference stations. Five stations were selected as baseline or reference stations from the results of P450 RGS analyses, as these produced low values of 1.7 to 2.5 µg of benzo(a)pyrene equivalents per g dry weight. It should be noted these stations were not selected prior to the initiation of the study, but were selected after all of the analyses for the study were completed.