



U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
NATIONAL OCEAN SERVICE
OFFICE OF RESPONSE AND RESTORATION
COASTAL PROTECTION AND RESTORATION DIVISION
c/o U.S. Environmental Protection Agency SFD-8
75 Hawthorne Street
San Francisco, CA 94105-3901

June 29, 2000

Richard G. Mach Jr., P.E.
BRAC Environmental Coordinator
Department of the Navy, Southwest Division
Naval Facilities Engineering Command
1220 Pacific Highway
San Diego, CA 92132-5190

Dear Mr. Mach:

The Department of Commerce, National Oceanic and Atmospheric Administration (NOAA) appreciates the opportunity to comment on the document titled "Parcel F Validation Study, Hunters Point Shipyard." As a Federal Natural Resource Trustee for marine, estuarine, and anadromous species and their habitats, NOAA continues to be concerned regarding the effects of elevated contamination on our trust resources. I appreciate the effort the Navy has expended to bring this project to a resolution.

NOAA's comments are enclosed. If you have any questions or concerns, please telephone me at (415) 744-1893.

Sincerely

Laurie Sullivan
NOAA Coastal Resources Coordinator

cc: Sheryl Lauth, USEPA
Chien Kao, DTSC
Brad Job, RWQCB
Jim Haas, USFWS
Charlie Huang, CDFG
Clarence Callahan, USEPA
Jim Polisini, DTSC/HERD
Don MacDonald, NOAA
Jim Leather, US Navy/SPAWAR
Stacy Curtis, US Navy/SPAWAR

BRAC OFFICE



NOAA Comments on Hunters Point Shipyard Parcel F Validation Study Work Plan

Comments by Don MacDonald and Laurie Sullivan

- Page 1 The last sentence regarding the debris-lined shoreline:
 *"However, if these areas are found to be acting as an ongoing
 source of contamination to intertidal or subtidal sediments
 then they will require evaluation in an FS."* How will it be
 determined that they are an ongoing of source of
 contamination to intertidal or subtidal sediments if they are
 not being sampled?
- Pages 3 & 9 Figures 1-1 and 3-1 are reversed.
- Page 4 For the selected position papers included in Appendix B, are
 there position papers that are not being included? Why?
- Page 5 Section 2.1, first paragraph. Details for the 3 lines of evidence
 continue to be developed through June, 2000.
- Section 2.1, last paragraph. How will exposure-response
 relationships be developed?
- Section 2.2; If any contamination is to be left in place following
 remediation then possible restrictions on future use of the site
 need to be evaluated.
- Page 6 Figure 2-1. The 2nd to last box ("Prepare Validation Study") is
 not clear. Do you mean, prepare validation study report? If so,
 the areas that are excluded also need to be included in the
 report, along with the reasons for exclusion. The "Proceed to
 Feasibility Study" box can then be restated to "Proceed to
 Feasibility Study for areas included in the remedial footprint"
- Page 8 Section 3.1.1. *"All three lines of evidence will be collected at all
 HPS stations and reference sites (although collection of intact
 cores for the SWI test at some of the proposed reference sites
 may not be practicable)."* Why won't it be practicable?
- Page 9 Section 3.1.1, regarding reference sites. (1) In Navy studies
 (Point Molate and Alameda), Paradise Cove has shown
 unacceptable reference toxicity. However, the RWQCB appears
 to have had good performance from these tests. What steps is

the Navy going to take to ensure good reference survival? (2) Alameda Buoy (RMP station BB70) was not used for contributing to the tolerance limit for the BPTCP, according to the BPTCP report.

Page 10

Table 3-1, Step 4, last sentence: *"Subsurface composite samples will be collected in 2-ft increments (i.e., 0-2 ft composite, 2-4 ft composite, etc.)."* Two foot composite samples are fine for evaluating sediments for dredging disposal purposes, but may not be representative of what organisms will be exposed to if the sediments are left in place. For example, if the top two feet are removed for remediation organisms will be exposed to what had been the 2-2.33 ft inclusion of the 2.33-4 ft interval in the evaluation of this sediment may either dilute or increase the apparent concentration within the exposure layer. Also if the top 5 cm sample does not show sufficient contamination to warrant removal but the 0-2 ft sample does, what action will be taken?

Table 3-1, Step 5, Item 1: Clarification is needed on exactly how the ambient ER-M quotient will be determined. Is it going to be a mean of all the reference sites, or is it going to be based on the RWQCB "ambient concentrations" numbers?

Page 12

3rd full paragraph. This paragraph, Table 3-3, and the WOE table handed out on 6/15 do not state clearly how the reference sites will be used for decisions on the amphipod toxicity test. Please clarify this in the response to comments. For example, what is your decision if the reference sites show toxicity compared to the "reference envelope"?

Page 12 & 13

Page 12, Third paragraph: *"In using this approach, the mean percentage survival of E. estuarius for all valid HPS tests will be compared with the lower SWRCB tolerance limit using a one-sample t-test. If survival in a test site bioassay is found to be significantly below this lower limit, then the site will be judged to exhibit greater than ambient (reference) toxicity. Alternatively, if the test site assay is found to be equal to or above the lower limit, the site will be judged nontoxic. For the Validation Study, the Navy proposes to use a p-value of 5 percent. the lower end of the tolerance limit associated with a p-value of 5 percent of E. estuarius is 67.7 percent survival (SWRCB, 1998)."*

Page 13, Table 3-3, Step 5, second paragraph: *"The percent survival of amphipods exposed to sediment from HPS stations*

will be compared with the lower tolerance limit on the 85th percentile of the San Francisco Bay "reference envelope" distribution (67.7% survival). Stations with survival less than 67.7% will be considered toxic to amphipods (positive response), assuming that acceptable survival was observed at reference sites.

These two descriptions of the proposed approach to evaluate the amphipod toxicity tests are inconsistent. The tolerance limit of 67% calculated by the San Francisco RWQCB "reference envelope" approach is a 95% tolerance limit. If that is the approach, you do not statistically compare your station sample to the lower tolerance limit, you compare the mean without the additional statistics (the statistics were already done in calculating the tolerance limit). In the phrase, "... , the mean percentage survival of *E. estuarius* for all valid HPS tests ... " what does "all" mean? Presumably "all" should be replaced by each."

However, the most recent WOE table presented by the Navy states that an HPS sample will be considered toxic (either high or low positive) if the amphipod result at a station is less than 69.5% of control survival. Please reconcile the latest discussions with the statements on pages 12-13 of the workplan.

Page 12 Last paragraph: With regards to the SWI tests; "... the percentage of normally developed larvae, based on the first 100 individuals observed, will be determined for each test sediment." As per the discussions on the January 25 conference call (see page 17 of Appendix A), this is the less sensitive methodology and the group reached a consensus that the more sensitive methodology should be used.

Page 13 footnote (a). Who is the Agency?

Page 14 Table 3-4, Step 5: "... samples will be considered nontoxic if the percentage of normally developed larvae observed in test sediments is at least 60% of number of normally developed larvae observed in control samples." If the controls have no more than 70% survival, as per Step 3 of Table 3-4, this would mean normal development as low as 42% would be considered toxic. How was the value of 60% derived?

Page 15 top of page: "Additional comparisons may be made using data derived from the alternative counting procedures described by ASTM (1994) and Puget Sound Estuary Program (PSEP) (1995),

if significant toxicity is observed." These alternative counting methods require alternative test procedures than the planned counting method so the tests would have to be done from the beginning to accommodate these alternative methods. The ASTM (1994) citation appears to be incorrect. ASTM Standard E724 (last revision 1998, not 1994) is for bivalve larvae. ASTM Standard E1563 is for echinoderm larvae. The Puget Sound reference is also for bivalve testing, not for echinoderm testing.

- Page 15 Section 3.1.1.3 Bioaccumulation and dose assessment "2 replicates of each sediment sample will be tested".. When were the replicates reduced from 5 to 2 (this issue does not show up in any of the meeting minutes)? What is your justification for doing statistics with an n of 2? The table on page 16 states that you will compare means, but does not say that you will statistically compare means. The latest "weight of evidence" table states that means will be compared by a t-test.
- Page 16 Decision Rule. Please update table with latest agreements.
- Page 17 Regarding Safe Tissue Levels. This will likely require more than a conference call to agree to the safe tissue levels. I recommend that you consider both avian and marine mammal receptors.
- Page 17 TIE. If the TIE results aren't going to be used, why are you doing it, and why are you discussing it in this workplan?
- Page 17, 19 Invertebrate and fish tissue collection. Suggest that you a priori set priorities for taxa for tissue collection, along with a cap on the level of effort you will expend to collect. For fish, the species you are targeting seem appropriate.
- Page 23 Integrated Sample Design. As stated in the meeting of May 1, 2000 (for which there are no meeting minutes in the workplan), I disagree with the stratification scheme, because this essentially weights the sampling towards the areas of apparently lower chemistry (twice as many samples in the A and B areas as in the C and D areas). Using the variance of the ERMQ among areas oversimplifies the problem, because this ignore the fact that there may be different chemicals driving the ERMQ at different stations.
- Page B-3 See previous SWI comments for pages 12, 14 &15 above.
- Page C-1 Bottom of section C.1: *"Given the nature of problems associated with the historical data collected at HPS, relevant*

estimates of variance are available only for the sediment chemistry line of evidence. For this reason, sample size requirements for all lines of evidence are based on the existing sediment chemistry data, and the assumption is made that this number of samples also will provide adequate coverage to understand the toxicity and bioaccumulation potential of these sediments." The variance of the toxicity and bioaccumulation tests are likely to be higher than that of sediment chemistry because they are also influenced by variable bioavailability and organismal variability. Therefore, utilizing the same number of samples for toxicity testing and bioaccumulation as for sediment chemistry would result in less powerful tests for toxicity and bioaccumulation than for sediment chemistry.

- Page C-6 Section C.4. Since ERM quotients were calculated based only on 5 constituents, did you validate the ERM quotients for the historic (thus more complete) data by calculating a "full" ERM quotient?
- Page C-17 First sentence below Table C-4: *"In general, a systematic distribution of sample points within strata was used to provide good spatial coverage and optimize the ability to interpret spatial distribution of results for the three lines of evidence."* By systematically distributing the samples the samples no longer meet the parametric test requirement of randomness, which will reduce the reliability of any statistical inferences.
- Page D-19 Section 4.2.3. Where are you sending your SWI cores?
- Page D-22,E-31 Regarding field collection of invertebrates for bioaccumulation analysis: Organism analyzed should be identified as completely as possible without expending a great deal of energy. They should be identified at least to class and if possible order. Ideally you would want to analyze the same organisms or mix of organisms with each sampling location to reduce the variability between stations due to the variable ability of different organisms for accumulating different contaminants. It is also necessary to be aware that if polychaetes are among the organisms being chemically analyzed, seeing they will not have their guts purged prior to analysis, a large portion of their apparent chemical body burden will actually be contaminants associated with the sediment in their guts.