MAY 1 3 1996

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
NATIONAL OCEAN SERVICE
OFFICE OF OCEAN RESOURCES CONSERVATION AND ASSESSMENT
HAZARDOUS MATERIALS RESPONSE AND ASSESSMENT DIVISION
COASTAL RESOURCES COORDINATION BRANCH
c/o U.S. Environmental Protection Agency (H-8-5)
75 Hawthorne Street
San Francisco, CA 94105-3901

N68311.000753
NAVSTA LONG BEACH
SSIC #5090.3

May 3, 1996

Mr. Mike Radecki
Remedial Project Manager
Southwest Division
Naval Facilities Engineering Command
1220 Pacific Highway
San Diego, CA 92132

SUBJECT: Draft Remedial Investigation (RI) Report Installation Restoration Program For Site 7, Naval Station Long Beach

The U.S. Department of Commerce/National Oceanic and Atmospheric Administration (NOAA) appreciates the opportunity to comment on the subject document. This document was prepared for the Southwest Division, Naval Facilities Engineering Command, by Bechtel National, Inc., San Francisco, California, 22 February 1996.

As co-trustees with the Department of Defense under Subpart G of the National Contingency Plan, § 300.615(a) and CERCLA 104(b)(2), NOAA represents the federal resource trustee interests of the Department of Commerce. NOAA, through the Coastal Resources Coordination program, carries out those trustee responsibilities by providing technical assistance in evaluating the potential impact to NOAA trust resources from releases of hazardous substances.

If you have any questions about these comments, I may be reached at (415) 744-3126.

Sincerely,

Laurie J. Sullivan

Laurie J. Sullivan
Coastal Resources Coordinator

Enclosure

cc: Martin Hausladen/Ned Black,EPA
Bob Hoffman, NMFS
Carol Roberts, FWS
Michael Martin, CFG
Alvaro GutierrezDTSC

National Oceanic and Atmospheric Administration (NOAA) Comments on Draft Remedial Investigation Report for West Basin (Site 7) Naval Station Long Beach, Long Beach, CA, 2/96

General comments

The Navy is to be commended for conducting such a thorough investigation of sediment contamination, sediment toxicity, and contaminant bioaccumulation. The amount and quality of data collected should lead to a good understanding of sediment toxicity and ecological risk in the West Basin.

Although we are pleased with the data collected, NOAA disagrees with the Navy's conclusion that no action is warranted at the West Basin. In our view, the information from the sampling indicates that the sediments in the vicinity of Piers 16, 12, and 7 are highly contaminated with PAHs and are toxic to sensitive benthic invertebrates. PAHs are accumulating in clams near Piers 2, 15, and 6 and in bile of California halibut taken from the basin. Some of the stations with high PAH concentrations had reduced numbers of benthic taxa or crustaceans (for example, stations 10 and 11 near Piers 6 and 7 and station 18 near Pier 3).

Sediments in other portions of the basin around Pier 9 and toward the center of the basin are toxic to echinoderm larvae. We believe this can be best attributed to zinc concentrations in excess of acid volatile sulfide concentrations. Of the primary contaminants of concern, zinc concentrations appear to best explain observed patterns of echinoderm survival. NOAA therefore recommends limited sediment removal where zinc exceeds AVS concentrations and around piers where sediments are toxic and highly contaminated with PAHs. A protective concentration for total PAH and high molecular weight PAH compounds should be derived based on TOC and grain size normalized sediment concentrations. NOAA's preliminary recommendation for a protective sediment concentration would be approximately 10,000 ugtPAH/gOC (normalized to percent fines). NOAA also recommends the use of EPA's sediment quality criteria.

NOAA is also concerned about some of the data analyses performed that led to the Navy's conclusion. Analyses such as comparison to an upper 95% predictive level, and nonconservative definitions of bioassay "hits", and the defining of sediment evaluation zones may all contribute to an

underestimation of risk to marine resources. The reasoning behind some of the analyses are difficult to understand from reading the document. For example, some analyses compare Pier to Basin to Reference sites, while other analyses analyze by zones, which made interpretation difficult for the reader. Also confusing, the term "upper confidence limit" appears to be used interchangeably with the "predictive limit", although it is our understanding that these are different values. A more complete explanation of statistical methods used may help explain these issues. Specific comments regarding the analyses are included in the specific comments below.

Specific Comments:

Page 1-14, section 1.7.4: EPA sediment quality criteria would qualify as ARARs. NOAA sediment quality guidelines would be factors to be considered.

Page 3-5, section 3.2.1.1: The assumption that stations under different piers would represent
the same zone does not appear to be well founded based on the different uses of the piers.

Page 3-5, section 3.2.1.1 (and following sections): The level of confidence in determining the difference between the area maximum and basin median would depend on the variability in concentrations and the minimum difference one would want to detect. Additional information in support of these estimates should be provided.

Page 3-10, section 3.2.6: The reference stations off Cabrillo Beach (40010) were quite a bit more contaminated than the other reference stations. This is unfortunate, but does not render the stations useless as reference areas. Together, the Cabrillo Beach stations and the additional reference stations provide sufficient information to interpret the results of sediment toxicity testing.

Page 3-11, section 3.2.6: The bivalve study used sediments collected from only five basin stations, each of which showed relatively low concentrations of both organic and inorganic substances. None of the stations were representative of moderate hotspot areas where elevated concentrations of PAHs and metals were observed. The study results do not represent the

potential for bioaccumulation in the hotspot areas and probably understate the potential for bioaccumulation in the basin as a whole.

Page 3-15, section 3.4.3.1: Since the sediments beneath piers were sampled by divers using scoops instead of box cores, there may be a difference in the results attributable to sampling technique. Although the sampling technique was designed to minimize disturbance, the scoops most likely created a more disturbed sample than the box core. It may be that the finest sediment escaped from the scoops. This difference in sampling techniques introduces uncertainty into any comparisons between pier and open water stations.

Page 3-19, section 3.4.3.3: The collection of an additional California halibut 17 days after the main sample collection introduces some uncertainty into the results.

Page 3-36, section 3.7: NOAA has several comments regarding the data analysis summary.

1. The comparison of West Basin values to a 95% predictive interval is not clear in this section. From reading the document, it appears that single station values were compared to a 95% predictive interval from reference sites. The rationale for determining whether an individual station is contaminated compared to the 95% interval is not discussed, nor are the analyses for determining significance of a bioassay result to the 95% interval. It should be noted that many of the 95% upper predictive interval numbers for contaminant concentrations in the reference areas exceed the observed contaminant level in the reference area. NOAA does not believe that this kind of comparison is conservative enough to lead to protective decisions. For example, if one looks at Arsenic, only 3 of the West Basin stations exceed the UPL. If one had looked at one standard deviation above the reference mean, 12 stations would exceed the limit. If one had looked at the mean for the reference station, 29 would exceed the limit.

2. The reasoning for the statistical analyses to compare West Basin to reference area are confusing. Sampling design was stratified by areas within the West Basin that are expected to have some similarity in disturbance or contamination (pages 3-4 to 3-6), yet no statistical analyses (other than summary statistics) were done to compare these areas to reference areas. Since much of the textual discussion concerns "pier" versus "basin" stations, it seems odd that these were not analyzed statistically.

3. Section 3 should include a detailed explanation for selecting Sediment Evaluation Zones (SEZ). Since the SEZs are the Navy's basis for deciding future actions within the West Basin, the rationale and decision criteria for selecting SEZs needs to be clearly defined.

4. Although the statistics used to evaluate correlations between contamination and toxicity were appropriate, other methods are also useful to determine correlations and patterns. Since a preponderance of evidence of the biological measures should be used to determine whether sediment is unacceptably affected, sediment contamination should be plotted against different combinations of biological effects indicators. While this is not a sophisticated approach, it can be helpful.

Page 3-38, section 3.7.2: The rationale behind providing the 95th percent confidence interval using a one-tailed distribution at the reference site and using a two-tailed distribution in the West Basin should be provided. Double-check tables and text for possible mixing of terms "predictive limit" and "confidence limit". It is our understanding that these are two different concepts. Also, please identify how this information is used in the risk assessment.

Page 3-38, section 3.7.3: The language used to describe the interpretation of a failure to find a significant difference between two sample groups is not very careful. One can only fail to reject the null hypothesis that the two samples come from the same population.

Page 3-39 and 40, section 3.7.3.4 In the second bullet, please identify how the biological endpoint for each station will be statistically evaluated and significance determined. The explanation given in the text is vague. Second, NOAA disagrees with the definition of a "hit" as being not conservative enough. We believe that if a station meets the relative numerical difference criterion, it should be considered a "hit." Biological justification should be given for selecting the 20 and 30% criteria.

Page 3-43: A discussion of the rationale behind collecting AVS concentrations should also be provided.

Page 4-2, section 4.1.1: Even though zinc and butyltin compounds were found in some samples at less than 5 times the blank concentrations, because these are site-related contaminants, these results should not have been discarded.

The fact that these uncommon compounds were found in blanks should indicate that lab or field procedures were not as clean as necessary. The uncertainty introduced by these findings should be discussed.

Page 4-3, section 4.1.3: The fact that only two animals were found in one of the samples from reference station 40010.3 indicates that it may not be an appropriate reference station.

Page 4-4, section 4.2.1: The visual observations of the sediments provides additional information in support of NOAA's conclusions that sediments under the piers are highly contaminated with PAHs.

Page 4-9 and 10, section 4.3.2.1. The polychaete growth data should be rejected as not meeting the performance criteria.

Page 4-19, section 4.4.1.2: The concentration of sulfides in the reference sediments should also be discussed here.

Page 4-24, section 4.4.2.1: The statements about the concentrations of TBT in California halibut in the text are confusing and inconsistent with the results in Table 4-24. According to the table, TBT in halibut from the basin is higher than in the reference area fish. Table 4-25 does not support the statement in the text that white croaker from the reference station contained more zinc than those from the West Basin. According to the table, the fish from the basin contained higher zinc concentrations than those from the reference areas.

Page 4-25, section 4.4.2.2: The fact that every compound in bile was higher in the fish taken from the basin than from the reference area indicates that the PAH compounds in the basin are available to fish.

Page 4-25, section 4.4.2.3: Since pyrene was not detected in clams from the reference areas, it is a logical conclusion that it is present and bioavailable in the basin. How can there be a statistical analysis between reference and West Basin stations when pyrene was not detected in reference stations (so there would be no variance in reference station data)?

Page 4-26, section 4.5.1.1: The discussion of the toxicity data from pier stations indicates that the pier stations are clearly toxic.

Page 4-28, section 4.5.2.2: Station 8 near the marina also had a reduced number of taxa. Stations near Piers 10, 11, and 18 contained no or only one species of crustacean, which are among the most sensitive of benthic invertebrates.

Page 4-32 and 33, section 4.6.1. Although there is a discussion of criteria used in establishing SEZs, there is no information on what the final selection criteria for drawing the lines. For example, why weren't F and E combined to be on group? We also note that since many contaminants were not included in the analysis (all LPAHs, some of the HPAHs, for example), you may be disregarding some of the very compounds that make one station more similar to another. We also believe that results of toxicity tests and benthic community similarity should have been used as selecting factors for defining SEZs. Because of these problems, we believe that analyses by SEZ group are flawed, and should be reconsidered.

The discussion of the selection of SEZs certainly needs to be addressed in the uncertainty analysis.

Page 4-37, section 4.6.5: When viewed in combination with the other data, the benthic community at pier stations should be classified as stressed, instead of "semi-healthy". The criteria for classifying a community as "healthy", "semi-healthy" or "stressed" needs to be discussed.

Page 4-38, section 4.7.2.1: In light of the confusing discussion regarding zinc concentrations in fish on page 4-24, the data should be reexamined to consider including zinc as a contaminant of concern for fish.

Page 4-40, section 4.8. Reasons for rejecting SEZs are not clear. Reference is made to detections that do not exceed the "reference station value". Presumably the reference station value is the 95% prediction limit, not the mean. If so, that should be made clear.

Figure 6-1: The pathways should indicate how the fish and shellfish are exposed.

Page 6-2, section 6.1.2: The conceptual model should include a discussion of the toxic mechanisms of the contaminants of concern. If the contaminants cannot be evaluated individually, they should be discussed in groups. The discussion should include the fact that many of the contaminants are directly

toxic to sensitive invertebrates and some of the contaminants bioaccumulate and are indirectly toxic, causing reproductive problems in fish and wildlife.

Page 6-8, section 6.1.3.1: The assessment endpoints for the ecological risk assessment should be more specific. For example, an assessment endpoint of "the health and vitality of the .... communities" is somewhat vague. The benthic community should be protected against direct toxicity. The fish and mammals should be protected against both direct and indirect toxicity.

Page 6-8, section 6.1.3.2: It is customary to describe the measurement endpoints in terms of testable statements, generally exceedances of reference station values or exceedances of toxic thresholds.

Page 6-10, section 6.2.2: The statements regarding California halibut should be reexamined based on the data in Table 4-24.

Page 6-12, section 6.3.2: While the classification of stations into zones allows statistical comparisons with the reference area (if one accepts the assumptions that the zones represent the same "population" of sediments), it is not very helpful to examine the contamination data and toxicity data on a zonal basis. It seems more logical to examine stations individually, correlating sediment contamination with toxicity. The correlation analysis discussed on page 6-13 attempted to do this but did not use normalized concentrations, which would have alleviated some problems with the data. For example, plotting zinc/AVS against echinoderm survival indicates that this is a reasonable predictor (see attached figure).

Page 6-13, section 6.3.2.1: In this case, sulfides are not a convincing explanation of toxicity. However, since sulfides and PAHs are correlated with each other, these factors cannot be separated.

Page 6-17, section 6.3.2.2: The discussion of risk to the benthic community is not consistent with the data presented in the previous sections of the report. The statement that "The laboratory analyses performed as part of this RI....indicated that the benthic community is healthy and not at risk from possible chemical effects" contradicts the data demonstrating that the sediments beneath the piers is highly contaminated, toxic to invertebrates, and contains components which are bioaccumulating in clams. The fact that there is a "semi-healthy" benthic community beneath the piers does not indicate that the benthic community is not at risk. Significant toxicity in laboratory tests indicates that the benthic community IS at risk. Since there

were no reference stations beneath uncontaminated piers, one cannot determine what the composition of the benthic community would be in the absence of such high levels of PAH contamination. NOAA concludes that the benthic community beneath the piers is at risk from PAH contamination.

The statement that no discernible relationship between "between-pier" and "beneath pier" benthic infaunal composition indicates that there is no influence of beneath pier sediments to between pier sediments is unclear. Influence would be better indicated by chemical contamination than by infaunal community. EPA's sediment quality criteria should be used as ARARs.

Page 6-18, first paragraph: The correlation between HPAH compounds and infaunal indices should be explained.

Page 6-18, third paragraph: It was our understanding that sulfides and PAHs are correlated with each other, therefore, these factors cannot be separated.

Page 6-19, section 6.3.2.3, first paragraph: The statements regarding California halibut should be reexamined based on the data in Table 4-24.

Page 6-19, fifth paragraph: Possible sources of tributyltin in croaker tissue should be discussed since the fish are expected to be year round residents in the basin. The number of fish captured and their external appearance does not provide evidence that no detrimental effects of contamination are occurring.

Page 6-21, section 6.3.2.3: The fish bile analysis indicates that the fish have been recently exposed to PAH contamination. This finding is consistent with the results of sediment chemical analysis and bioaccumulation in clams. The data should not be ignored. The number of fish captured and their external appearance does not provide evidence that no detrimental effects of contamination are occurring.

Page 6-22, section 6.3.2.3: Correlations between sediment contamination and clam bioaccumulation may be weak due to the very small numbers of samples available for comparison. The fact that pyrene is accumulating in clam tissue is consistent with findings that fish are metabolizing PAHs, and that sediments high in PAHs are toxic to benthic invertebrates. This information indicates that PAHs are a problem in the West Basin.

Page 6-23, section 6.3.3.1: The lack of an obvious correlation between metal concentrations and toxicity may be due to examining the data by zone. It appears that the echinoderm survival may be reduced by zinc concentrations in excess of AVS concentrations (see attached figure). The significant toxicity at stations where zinc/AVS ratios are less than one may be attributable to other contaminants (for example, stations 33 and 45).

Page 6-24, section 6.3.3.4: Zone H is not discussed. The fact that many of the samples that were toxic contained concentrations of PAHs that were below the maximum at non-toxic stations does not mean that PAHs are not the cause of the observed toxicity. Concentrations should be normalized to TOC concentrations and possibly grain size, since these parameters varied considerably. There may be other factors which prevent PAHs from being bioavailable at the stations where no toxicity was observed.

Section 6.4.2: The zone approach may not be the most useful in interpreting causes of sediment toxicity throughout the basin. There are portions of zones A, B, C, and F that may require cleanup based on zinc concentrations in excess of AVS. However, the entire zones would not require action to meet these objectives. NOAA disagrees with the Navy's assessment of the lack of risk in zones G and H. Based on a preponderance of the evidence from toxicity testing and bioaccumulation studies, limited action is warranted to reduce the risk. The observations of the benthic community under the piers do not explain away the toxicity and bioaccumulation data. Many sections refer to contaminants "not detected" in particular SEZs (see, for example, SEZG). In fact, these contaminants are detected, in many cases well above the ER-L, but do not exceed the 95 % predictive level.

Page 6-30, section 6.5: This section addresses only a few of the uncertainties found in this risk assessment, and generally lists only uncertainties that would overestimate risk. A more balanced discussion of the uncertainties should be provided.

Page 7-3, section 7-3: NOAA disagrees with the Navy's conclusions that no action is required.

Table 4-10. RPD exceedances need to be underlined

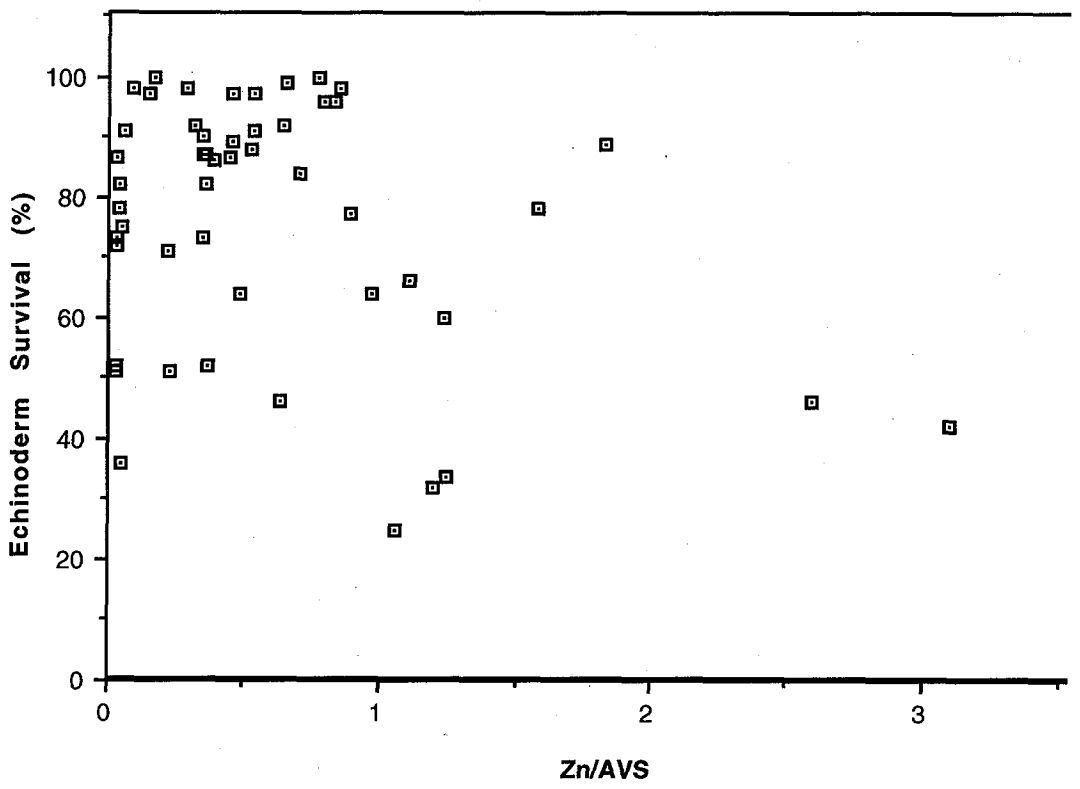
Table 4-11. Polychaete growth is mislabeled as a percent.

Tables 4-18, 19, 20. Summary statistics given by group of Pier, Basin and Reference. Why no statistical analyses by these groups for chemistry, infauna, and toxicity tests?

Tables 4-30 - 4-32. Statistical significance of individual bioassays needs to be explained. It is our understanding that this refers to the comparison to the lower predictive limit from reference areas, but it is not clear where the statistical significance comes in. Comparison to the lower predictive limit is not conservative enough to make protective decisions regarding the site.

Appendix M. Laboratory toxicity test data sheets should be included, not just the summary report.

Appendix S. These statistical analyses need to be double-checked. ANOVAs for amphipod survival and echinoderm development are missing. Others appear to have been rank-transformed, then analyzed with a standard ANOVA rather than the non-parametric version. Percent data should be arcsin transformed.



Zn/AVS and Echinoderm Survival in Sediments of West Basin