N40003.AR.002272
PUERTO RICO NS
5090.3a

LETTER TRANSMITTING U S EPA REGION II COMMENTS ON DRAFT SUMMARY REPORT FOR ENVIRONMENTAL BACKGROUND CONCENTRATIONS OF INORGANIC COMPOUNDS
NAVAL ACTIVITY PUERTO RICO
5/22/2006
U S EPA REGION II

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
REGION 2
290 BROADWAY
NEW YORK, NY 10007-1866

MAY 2 2 2006

CERTIFIED MAIL
RETURN RECEIPT REQUESTED

Mr. Kevin Cloe
Navy Technical Representative
Installation Restoration Section (South)
Environmental Program Branch
Environmental Division,
Atlantic Division (LANTDIV), Code EV23KC
Naval Facilities Engineering Command
6506 Hampton Blvd.
Norfolk, VA 23508-1278

Re: Naval Activity Puerto Rico (former Naval Station Roosevelt Roads) - EPA I.D. Number PRD2170027203, Draft Summary Report for Environmental Background Concentrations of Inorganic Compounds

Dear Mr. Cloe:

The United States Environmental Protection Agency (EPA) Region 2 has completed its review of the Draft Summary Report for Environmental Background Concentrations of Inorganic Compounds, Naval Activity Puerto Rico (the Background Report) submitted on behalf of the Navy by Baker Environmental's letter of April 1, 2006. As part of our review, EPA requested our contractor, Booz Allen Hamilton (Booz Allen), to review the documents.

The Background Report was reviewed to determine if it complies with EPA's Guidance for Comparing Background and Chemical Concentrations I Soil for CERCLA Sites (EPA 540-R-01-003) (CERCLA Guidance). The CERCLA Guidance extensively references the Guidance for Data Quality Assessment - Practical Methods for Data Analysis (EPA QA/G-9) (DQO Guidance). Thus, the Background Report was also reviewed to verify compliance with DQO Guidance, where relevant. The data sets and background statistics presented in the Background Report were reviewed to identify any potential concerns.

Based on Booz Allen's and our own reviews, EPA has identified a number of concerns regarding compliance with the CERCLA and DQO Guidance. Concerns regarding several of the data sets and associated statistic were also identified. These concerns are discussed in the enclosed Technical Review.

To address these concerns, within 45 days of your receipt of this letter, please submit a response to the enclosed comments and/or a revised Background Report addressing those comments.

If you have any questions, please telephone me at (212) 637- 4167.

Sincerely yours,

Timothy R. Gordon

Timothy R. Gordon
Remedial Project Manager
Caribbean Section
RCRA Programs Branch

Enclosure

cc: Ms. Yarissa Martinez, P.R. Environmental Quality Board, w/encl.
Mr. Julio I. Rodriguez Colon, P.R. Environmental Quality Board, w/encl.
Mr. Mark Davidson, U.S. Navy, BRAC Program Management Office SE, w/encl.
Mr. Felix Lopez, U.S. Fish & Wildlife Service, w/encl.
Mr. Mark Kimes, Baker Environmental, w/encl.
Ms. Kathy Rogovin, Booz Allen & Hamilton, w/o encl.

# TECHNICAL REVIEW

# DRAFT SUMMARY REPORT FOR ENVIRONMENTAL BACKGROUND CONCENTRATIONS OF INORGANIC COMPOUNDS APRIL 2006

## NAVAL ACTIVITY PUERTO RICO CEIBA, PUERTO RICO

May 17, 2006
REPA3-2203-084

### General Comment

1. A technical review has been performed on the April 2006 Draft Summary Report for Environmental Background Concentration of Inorganic Compounds (Background Report) at Naval Activity Puerto Rico (NAPR) in Ceiba, Puerto Rico. The Background Report was reviewed to determine if it complies with EPA's Guidance for Comparing Background and Chemical Concentrations I Soil for CERCLA Sites (EPA 540-R-01-003) (CERCLA Guidance). The CERCLA Guidance extensively references the Guidance for Data Quality Assessment - Practical Methods for Data Analysis (EPA QA/G-9) (DQO Guidance). Thus, the Background Report was also reviewed to verify compliance with DQO Guidance, where relevant. The data sets and background statistics presented in the Background Report were reviewed to identify any potential concerns.

   The technical review identified a number of concerns regarding compliance with the CERCLA and DQO Guidance. Concerns regarding several of the data sets and associated statistic were also identified. These concerns are presented in the following Specific Comments.

### Specific Comments

1. As indicated in the CERCLA Guidance (page 3-12), it is generally difficult to judge the adequacy of a background data set without first making certain basic decisions regarding the statistical comparison of background and site data. Of particular importance are decisions regarding the desired power and confidence levels of the statistical analysis. These inputs, particularly the desired power of the statistical tests, are closely related to the number of background samples required to achieve the required statistical performance. The adequacy of the number of background samples can only be judged in the context of the specific comparisons being made and the decisions made regarding the desired statistical performance. However, the number of background samples included in some of the data sets raise concerns over the adequacy of these data sets. For example, the weathered bedrock soil background set only contains three samples and the surface

water and sediment background data sets only contain seven to ten samples. These relatively small data sets may limit significantly the statistical performance of any statistical analysis used to compared site data to background. After a more complete review of the implication of these limited background data sets, the Navy may want to consider the collection of additional background data.

2. When discussing the treatment of censored data (non-detects), the Background Report (page 1-6) indicates that for data sets with a frequency of detection (FOD) greater than 50 percent, descriptive statistics were developed using surrogate values for the censored data. This does not appear consistent with the CERCLA Guidance (page 4-7), which indicates that if less than 15 percent of the background samples are non-detects, the distributions of the background sample may be determined using surrogate values. However, if more that 15 percent but less than 50 percent of the measurements in the background sample set are non-detects, the CERCLA Guidance recommends the use of specialized methods for analyzing non-detects and refers the reader to Section 4.7 of the DQO Guidance. The approach that was used in the Background Report to treat background data sets with between 15 and 50 percent non-detects does not appear to conform with the those recommended in the DQO Guidance. NAPR should ensure that the approach used to handle background data sets with between 15 and 50 percent non-detects is consistent with the CERCLA Guidance.

   In addition, the Background Report (page 1-6) indicates that for data sets with a FOD of 50 percent or less, "the data set is truncated such that non-detect and blank results are not considered in the calculation of descriptive statistics." The Background Report further indicates that "although this will reduce the power of the calculated statistics, the use of non-detect or blank results could yield an unacceptably large bias of any calculated statistics." This approach does not appear consistent with CERCLA Guidance. The CERCLA Guidance (page 4-7) indicates that for data sets with more than 50 percent non-detects, "it may not be possible to compare the means of two distributions," and indicates that "an alternative approach is to compare the upper percentiles of two distributions by comparing the proportion of the two populations that is above a fixed level." The DQO Guidance (page 4-50) suggests the use of the Test of Proportions to perform such a comparison. NAPR should ensure that the approach used to handle background data sets with greater than 50 percent non-detects is consistent with CERCLA guidance.

3. When discussing the use of background data sets, the Background Report (page 1-9) indicates that "the use of the upper limit of the means is warranted as an initial step in screening the analytical results for inorganics, consistent with the previous use of background data sets." It is not clear that this approach is consistent with the CERCLA Guidance. The use of this approach to initially screen site data relative to background should be justified based on the CERCLA Guidance.

4. When discussing the analysis of outliers, the Background Report (page 1-7) states that "the discordance test is one of four recommended outlier tests," while referencing Navy guidance. Although the discordance test is referenced, the text does not clearly state how outliers were identified. It should be noted, however, that the DQO guidance (page 4-29) indicates that the discordance test is only suitable for identifying outliers for normally distributed data. The Background Report should clearly identify how each data set was analyzed to identify outliers. The Background Report should also verify that the outlier tests that were used are suitable for the distributions of the data sets tested.

5. The Background Report (page 1-7) indicates that an outlier test was conducted on data sets with a FOD of more than 50 percent. The text further indicates that "in general outliers should not be removed from the data set unless clear evidence shows that they are not based on elements of the population being studied and should not have been included in the data set." As indicated in the tables presenting the results of the background analysis for the individual media, outliers have been identified in a number of the data sets. However, none were removed from the data set because "no errors were found in the sample results." Although these data were not removed from the data set because no errors were found, the outlier tests indicate that these data likely do not belong to the statistical population being studied.

   When discussing outliers, the CERCLA guidance (page 4-6) indicates that:

   > The use of nonparametric hypothesis tests for background comparisons greatly reduces the sensitivity of test results to the presence of outliers. Parametric tests based on the lognormal distribution may yield results that are extremely sensitive to the presence of one of more outliers.

   The CERCLA Guidance (page 5-6) further indicates that:

   > If the data sets contain outliers or non-detect values, an additional level of uncertainty is faced when conducting parametric tests. Since most environmental data sets do contain outliers and non-detect values, it is unlikely that the current widespread use of parametric tests is justified, given that these tests may be adversely affected by outliers and by assumptions made for handling non-detect values.

   Thus, the retention of the outliers in the background data sets will likely require that nonparametric tests be used when comparing these sets with site data, although distributional tests may identify the populations as normal or lognormal.

6. The Background Report (page 1-7) indicates that the **Shapiro-Wilk's W-test** was performed on all data sets with frequencies of detection over 50 percent. The text further indicates that "the W test is a 'goodness-of-fit' test considered to be effective for determining whether a data set can be described as 'normally' or lognormally distributed

for sample sets with 50 or fewer samples." This statement is in agreement with the test of normality presented in the CERCLA Guidance (page 4-2). However, the CERCLA Guidance (page 5-3) adds further qualifications to the use of the W-test for determining normality by also indicating that:

> Tests for the distribution of the data (such as the Shapiro-Wilk test for normality) often fail if there are insufficient data, if the data contain multiple populations, or if there is a high proportion of non-detects in the sample. Test for normality lack statistical power for small sample sizes. In this context, "small" may be defined roughly as less than 20 samples, either on site or in background areas. .... Therefore, for small sample sizes or when the distribution cannot be determined, non parametric tests should be used to avoid incorrectly assuming the data are normally distributed when there is not enough information to test this assumption.

Many of the background data sets presented in the Background Report have less that 20 samples. Thus, it does not appear appropriate to use the results of the W-test to identify normally or lognormally distributed populations for purposes of later recommending the use of parametric over nonparametric tests. For those sample populations with less than 20 samples that are found to be normally or lognormally distributed using the W-test, the **Background Report should either remove their designations as normally or lognormally distributed or clearly identify these designations as qualified based on sample size.**

7. The data sets used to establish background for groundwater do not appear to include multiple measurements from the same background well over the period of a year or more. Consequently, these data sets may not adequately include any temporal variability inherent in background groundwater quality, such as that introduced by seasonal effects. NAPR should demonstrate that data sets used to establish groundwater background adequately represent seasonal and other temporal effects.

8. Background for groundwater has been established without any apparent regard for the geologic strata from which the groundwater samples were derived. Frequently, groundwater quality is influenced by geochemical differences between the various geologic materials through which groundwater passes. NAPR should demonstrate that it is not necessary to establish separate groundwater backgrounds for each of the various strata present at the former Roosevelt Roads site. Otherwise, a separate groundwater background should be established for each geologic strata in which groundwater is present and in which contamination is present at the facility.

9. Table 5.3 indicates that the mean copper concentration in the data set used to establish background for estuarine wetland surface water is 12.2 micrograms per liter (µg/L). This is nearly three times the chronic marine ambient water quality criteria for copper (3.1 µg/L). NAPR should provide further discussion and/or analysis that demonstrate the suitability of the data set proposed for establishing background for estuarine wetland surface water.

10. Based on a comparison to EPA's National Coastal Assessment (NCA) data, concentrations of cadmium and selenium in NAPR's estuarine background sediment samples appear to be somewhat greater than typical background levels observed in Puerto Rico. For example, the mean cadmium concentration reported in Table 5-9 is 0.527 milligrams per kilogram (mg/kg), while only two of 43 samples in the NCA data set had detected cadmium concentrations of 0.5 mg/kg or greater. (NCA data were obtained in June 2005 from John Macauley of EPA's Environmental Effects Research Laboratory in Gulf Breeze, Florida). NAPR should discuss possible reasons for elevated cadmium and selenium concentrations in the background estuarine sediment samples and provide adequate justification for the continued use of the data set proposed for establishing background for cadmium and selenium in estuarine sediments.