



UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
REGION 1
1 CONGRESS STREET, SUITE 1100
BOSTON, MASSACHUSETTS 02114-2023

N62661.AR 001631
NAVSTA NEWPORT RI
5090 3a

received
5-14-03

JA

May 9, 2003

Franco LaGreca
U.S. Department of the Navy
Naval Facilities Engineering Command
Northern Division
10 Industrial Highway
Code 1823, Mail Stop 82
Lester, PA 19113-2090

Re: Draft Work Plan for Background Soil Investigation for the NUWC Disposal Area, SA-08
at the Naval Station Newport

Dear Mr. LaGreca:

EPA reviewed the draft Work Plan for a background soil investigation at the Naval Undersea Warfare Center (NUWC) Disposal Area (Study Area 08), at Naval Station Newport. The stated purpose of the Work Plan is to describe the process used to select background sampling locations, field activities, analytical parameters, sampling procedures, methods for data analysis and interpretation, and reporting aspects of the investigation. In its review, EPA paid particular attention to the proposed statistical analysis of the background data to be collected according to this Work Plan, as well as to the rationale underlying selection of background sampling locations. This document was reviewed in accordance with EPA's most recent guidance on conducting background investigations and use of background data in site evaluations (USEPA, 2002). Detailed comments are provided in Attachment A.

EPA generally supports the development of a site-specific background investigation instead of relying on more generic, broadly based "background" numbers. The site consists of a storage facility with a fenced, paved area, while the remainder is unimproved, *i.e.* grassy, with some wetlands and a small retention pond. A golf course is adjacent to the site on the north side, and formerly agricultural land lies to the east. Background studies from nearby sites suggest that ambient levels of metals (especially arsenic) and organochlorine pesticides may be elevated in soils from SA-08 relative to other background areas across the state. Therefore, background locations must be selected carefully, in order to determine whether Contaminants of Potential Concern (COPCs) arise because of naturally occurring geochemical processes and geological heterogeneity, or from anthropogenic activities unrelated to on-site releases.

Although this Work Plan specifically addresses the scope and content of a background soil investigation for SA-08, wording in Section 1.1 and elsewhere in the document alludes to the site-to-background comparison that will be performed in a forthcoming SASE. If the collection of site data is affected by the manner in which the background data are obtained, then this discussion needs to be expanded. Will there be a separate Work Plan for the statistical

Toll Free • 1-888-372-7341

Internet Address (URL) • <http://www.epa.gov/region1>

Recycled/Recyclable • Printed with Vegetable Oil Based Inks on Recycled Paper (Minimum 30% Postconsumer)

1522

procedures for site versus background comparisons to be performed in the SASE? The plans should be clearly stated, so that reviewers can determine whether the proposed plan for the background study is appropriate for the intended site sampling.

As stated above, EPA generally supports the development of a site-specific background data. According to EPA guidance (USEPA, 2002), ideally the background soil samples "...should have the same basic characteristics as the site sample (*i.e.*, similar soil depths and soil types)," and the background area should have "...the same physical, chemical, geological, and biological characteristics as the area being investigated..." In practice, this may be difficult or impossible to achieve, without collecting an unrealistically large number of samples. It is apparent from the aerial photographs and statements in the text that as many as three soil types may be represented within the boundaries of SA-08, including fill material. Although a rationale for sample location is clearly stated, limiting sampling to a single soil type may not be adequate for establishing background conditions. Moreover, no provision appears to have been made for characterizing filled or perturbed areas that are unimpacted by site activities (thus constituting "background conditions"). Consensus should be sought among EPA, RIDEM, Navy, and other members of the team with respect to the balance of soil samples from the mapped soil types represented on (or that may influence) the site, including "disturbed" soils.

The work plan does not currently discuss how non-detected results will be treated when comparing the different background data sets. It is likely that pesticide and SVOC analyses for the background data will result in samples with non-detected values. As discussed in EPA (2002) and US Navy (1999), statistical analysis may not be appropriate for data sets containing large numbers of non-detected results using the procedures specified in Section 5.0. For data sets with acceptably low numbers of non-detected values (no more than approximately 40% to 50% non-detects), the work plan should specify the surrogate value to be assigned to the non-detects. Both EPA (2002) and US Navy (1999) provide guidance on selecting appropriate values.

Overall, the statistical DQO procedure as proposed in the Background Investigation Work Plan for Disposal Area, SA-08 is consistent with EPA's seven-point DQO process as described in EPA, 2000 and EPA 2002 (*e.g.*, Chapter 3, Section 4). Two background data sets: 20 soil samples and 20 upgradient sediment samples as shown in Figure 3-1 of the Work Plan will be collected. Initially, these 20 sediment samples (upgradient hydric samples) and 20 background soil (non-hydric) samples should be appropriate (according to RIDEM requirements) to characterize the background soil and sediment conditions for the Disposal Area SA-08. Statistical methods as listed in Section 5 (page 5-1) of the Work Plan can be used to evaluate the distributions of the background data sets and to compare the contaminant concentrations of the two background data sets. These methods are consistent with the procedures as described in EPA 2000 Guidance Document.

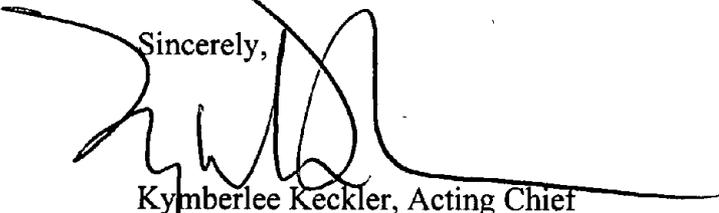
EPA is concerned, however, that some of these procedures (*e.g.*, comparing two background populations on log-scale) may not be implemented properly. I recommend that these procedures be implemented properly once the background and the Site data sets become available. Some

specific observations have been made that are discussed as follows for clarification. Some of these comments will be applicable only when the Site and background data become available. Appropriate evaluation of site data (*e.g.*, comparing a site concentration with respective background Upper Tolerance Limit (UTL)) is dependent upon the adequate statistical analyses of background data.

No mention on the treatment of outliers is included in the proposed Work Plan. Outliers should be identified using appropriate statistical procedures (EPA 2000; EPA 2002). Outliers, when present, distort sample mean and sample variance which in turn will result in inflated 95% Upper Tolerance Limit (UTL) values. Typically, high outliers represent observations that do not fit the general background distribution (EPA, 2002, page 4-6). High outlying values in background data sets may also represent locations that could be impacted by the site activities. Therefore, such outlying background sample locations may not be included in the computation of the background 95% UTL values. Also, while comparing distributions of two populations, in the presence of outliers non-parametric procedures are preferable to parametric procedures, especially those based upon the lognormal assumption (page 4-6, EPA 2002).

I look forward to working with you and the Rhode Island Department of Environmental Management toward the investigation and cleanup of the NUWC Disposal Area. Please do not hesitate to contact me at (617) 918-1385 should you have any questions.

Sincerely,



Kymberlee Keckler, Acting Chief
Federal Facilities Superfund Section

Attachment

cc: Paul Kulpa, RIDEM, Providence, RI
Cornelia Mueller, NETC, Newport, RI
Bart Hoskins, USEPA, Boston, MA
Chau Vu, USEPA, Boston, MA
Jennifer Stump, Gannet Fleming, Harrisburg, PA
Steven Parker, Tetra Tech-NUS, Wilmington, MA

ATTACHMENT A

<u>Page</u>	<u>Comment</u>
p. 1-2, §1-1	The second-to-last paragraph in this section refers to the SASE report that will be produced upon completion of the field investigations for the site. Additional detail regarding plans to conduct the site sampling should be incorporated into the proposed background sampling plan, in order to determine whether the background study will be appropriate for the site-to-background comparison.
p. 2-2, §2-1	The text states that part of the SA-08 site was used as "...a fill area and a storage area for the NUWC" and that "[F]ill activities are evident in air photos taken between 1951 and 1988." Although the sampling to be conducted as part of the forthcoming SASE is not discussed in this Work Plan, is Navy planning to sample this fill? Inasmuch as it may be a soil type that is present on the site, and filled portions may be unaffected by site-related activities, a representative area containing fill material could be included in the background investigation and evaluated separately.
p. 2-5, §2.2.1	The first paragraph under <u>Soil Types</u> states that the predominant soil type on the SA-08 site is Stissing silt loam (also shown on Figure 2-3), which consists of poorly drained soils on glacial upland features in southeastern Rhode Island. However, the delineation shown on Fig. 2-3 suggests that other soil types are found in relatively close proximity to the site. These include: Pittsdown silt loam (PmB, at the northern end of the site), which is a soil that is moderately well drained and deeper than the Stissing silt loam; Mansfield mucky silt loam (Ma, through which the stream entering the site from the south passes), noted for its high water content and association with low-lying wet areas receiving eroded material from surrounding uplands; and the Udorthents soils (UD), disturbed or filled soils underlying the developed area to the west of the site. The rationale for focusing the background soil sampling on the Se type is based on the assumption that, prior to construction of the Navy's facilities to the west, site soils were probably similar to those presently abutting on the north and east sides (<i>i.e.</i> , the Stissing classification). Also, the 1998 SASE Work Plan identified fill areas on site within portions where the soils are classified as Se, so it appears that there is no expectation that any other soil types will be represented. While this may be true, what contingency might the Navy offer to explain possible differences between site and background soil results? In the event that significant differences are observed, how might the Navy discriminate between genuine site-related contamination and merely differences between soil types, Se soil with contributions from surrounding uplands, and/or fill material?

- p. 2-7, §2.2.1 In the last paragraph under Soil Types, the Navy acknowledges that differences in soil chemistry may exist even within the Se (Stissing silt loam) soil type. Hydric soils are defined as those in or near stream beds, and non-hydric soils are those at some distance from streams. Thus, even if an area is mapped as the Stissing classification, it is reasonable to expect that differences in water content, redox conditions, type, quantity, and bioavailability of organic matter, physical, chemical, and microbial processes, etc. will affect soil chemistry, and these effects may be significant. Nevertheless, Navy is proposing to sample ‘upgradient’ hydric soil conditions – *i.e.*, in each of the two streams flowing onto the SA-08 site, as flowing surface water may transport sediments onto the site from the upgradient direction – as well as non-hydric soils within the same soil classification. EPA notes that Navy will be subjecting the two groups of data to rigorous statistical analysis (discussed in greater detail in Section 5.0), in order to determine whether the data can be combined into a single data set. Despite the outcome of the statistical assessment, EPA recommends that the two data sets remain separate. Because the site contains some wetlands and a pond, any hydric soils sampled as part of the forthcoming SASE can then be compared to the appropriate background data set; similarly, any non-hydric soils can be compared to the relevant data set as well.
- p. 2-8, §2.2.1 According to the text, Navy is not planning to sample any filled areas, but will restrict sampling to surface soil and stream sediments in upgradient/background areas only. However, portions of the SA-08 site are known to contain fill. What is known about the origin and characteristics of this material? Will it be sampled during the site investigation, inasmuch as it may be “background” in the sense that it is not impacted by site-related activities? Is fill present within the boundaries of the golf course? Are there areas of the golf course known to have been affected by grading, fill, or other perturbations? Please clarify how and when fill samples will be collected and whether these samples will be incorporated into the background soil set.
- p. 2-8, §2.2.1 It appears from the second full paragraph that the null hypothesis to be tested in comparing the two background data sets can be written as $H_0: m_1 = m_2$ where m_1 equals the mean or median of background data set 1 and m_2 equals the mean or median of background data set 2. From Section 2.2.3, it appears that the alpha level selected is $\alpha = 0.05$. Please note that since this is a two-tailed test the critical value will be based on the test criteria of $\alpha/2$. Please confirm that this interpretation is accurate. Also, please note that while this hypothesis and selected alpha value are appropriate for the analysis proposed, these values would not be appropriate for comparing background data to site data as discussed in USEPA (2002). A more comprehensive plan for comparing site to background data will be needed at a some point in this project.

p. 2-9, §2.2.4 Two areas (the polygons labeled “A” and “B” on Figure 2-4) have been identified as “...primary sample areas.” Figure 3-1 shows the same aerial photograph, with the same soil designations and the same polygons, but with sample locations added. EPA notes that the hydric soil sample locations follow the two streams entering the site from the north (the golf course, Area A) and from the south. The latter appears to flow through what may be a wetland surrounding the Mansfield soil type (Ma). Please expand on the rationale for concentrating all non-hydric soil sample locations in the Area A shown on Figure 2-4. In this discussion, please also explain why no samples other than those in the stream will be collected from Area B, from either the Ma soil or from the Se soil surrounding the stream on the south side of the site.

EPA supports the collection of a majority (but not all) of the background samples from the golf course and adjacent agricultural land, because any soil transported into the SA-08 site probably originated in these areas, and the chemistry of soils from these areas is not likely to reflect site impacts. However, it can be expected that routine maintenance activities such as irrigation, fertilizer application, possible herbicide and/or pesticide use, etc. has impacted soil chemistry. Arsenic, iron, and other trace metals are present in soils at Naval Station Newport owing to post-glacial chemical and physical weathering and deposition. In the presence of water, iron (from bedrock sulfides and other minerals, and possibly from fertilizer) oxidizes to form chemically active, hydrous ferric oxide (HFO) surfaces in the overburden. These HFO surfaces consist of positively, negatively, and neutrally charged sites that readily sorb arsenic and other trace metals. However, under some conditions – either a decrease in redox state, and dissolution of the HFO surface, or an increase in aqueous concentrations of competing ligands (*e.g.*, phosphate, nitrate, sulfate, bicarbonate, chloride, *etc.*) – the sorbed constituents may be released back into solution. These elements are then mobile until reaching an environment in which HFO is stable and the concentrations of competing ligands has decreased. According to historical records, the golf course has been maintained since 1922 (p. 2-2), so it is plausible that decades of maintenance have resulted in observable downgradient effects.

p.2-12, §2.2.6 This sentence states that the “...site and background data sets are comprised of 20 samples each.” Does this imply that 20 hydric and 20 non-hydric soil samples will be collected from the site (see comment for p. 2-7, §2.2.1), or a total of 20 samples? Please clarify.

p. 2-13 For site versus background comparison, Background Test Form 1 hypothesis has been proposed as suggested by the definition of the error probability α . The values of two types of error probabilities, α and β should be chosen according to the recommendations in Chapters 3 and 5 of the EPA Background Guidance document (2002). A higher value of α (=0.1, 0.2) is preferred (pages 3-7, 3-9 5-

15, EPA, 2002) to an alpha value of 0.05. For Test Form 1 hypothesis, a recommended value of β error is at most 0.1 (page 3-9, EPA 2002). Site sample sizes should be determined using these recommended values of decision errors ($\alpha=0.2$, and $\beta=0.1$ or 0.05).

Also when using Background Test Form 1 hypothesis (as proposed in the Work Plan), a retrospective power analysis (page 3-6, EPA 2002) should be conducted. This is performed to ensure that the test has an adequate power to detect a site with mean contamination that exceeds the background by a specified amount such as the minimum detectable difference.

- pp. 12 & 13 The discussion on limits on decision errors is based upon the assumption that about 20 site samples will be collected. It is also assumed that the standard deviations of the two groups (site vs. background) are equal to compute the approximate sample sizes. These decision error limits will be used when the site mean (or median) will be compared with the background mean (or median). In order to perform such a test, enough site samples need to be collected. This comparison should be done separately for the soil and sediment samples collected from the site. For example, a minimum of 20 hydric (sediment) and 20 non-hydric (soil) samples may have to be collected from the Disposal Area SA-08. The limits on decision errors and sample size determination should be revisited (retrospective power analysis) once the background and site data become available. For example, the determination of number of samples to be collected from the site (separately for soils and sediments) can be performed using the formula given on page 3-4 of EPA 2002 Background Guidance document.
- p. 3-1, §3.2 The first sentence in the second paragraph states that all soil sample locations will be within Areas A and B. As noted above, no soil sample locations appear within Area B on Figure 3-1. Please modify either the figure or the text, as appropriate. It seems likely that the stream could transport sediment from Area B onto the site, so it is EPA's opinion that some of the background soil samples should be located within Area B.
- p. 3-3, §3-2 Samples are to be analyzed for metals, pesticides, and semivolatile organic compounds (SVOCs), according to ILM04.0, OLM03.2, and 8270C (SIM) protocols, respectively (from Table 4-1). Will the entire suite of inorganics (23 metals plus cyanide, listed under ILM04.0) be analyzed? The geochemical behavior of arsenic and numerous other trace metals is controlled, to a large extent, by sorption onto hydrous oxides of iron, manganese, and aluminum. Demonstration of linear correlations between soil iron and other metals is now often accepted as evidence supporting the sorption mechanism. It will be critical to any arguments for a natural origin of arsenic and other trace metals to demonstrate as many correlations as possible.

Will the analyte list also include any herbicides or algicides? Previous discussion in the text (see p. 2-7) indicates that these compounds may also have reached the site from historical agricultural activities in the area, or from the golf course. Please ensure that these compounds are considered as potential analytes. Are there any other classes of compounds that are known or suspected to be present in site soils, such as nitroaromatics (*i.e.*, explosives)?

p. 5-1 The Work Plan states that 95% UTL of background will be calculated for use as a reference value for evaluating site data. Since the site observations will be compared with background 95% UTLs, the Work Plan should explain how the background parametric 95% UTLs will be computed for normally or lognormally distributed data sets. If outliers are present in a background data set, then the use of a 95% UTL based upon a lognormal distribution should be avoided as the presence of outliers may result in an inflated UTL value.

p. 5-1 The Work Plan states that Student's-t test or Satterthwaite t-test will be used on log-transformed data (in case data are lognormally distributed). However, it should be noted that it is not appropriate to use two sample Student's t-test or Satterthwaite t-test on log-transformed data. Equality of means on log-scale does not necessarily imply the equality of means in the original scale. This is because the mean in original scale also depends upon the standard deviation of the log-transformed data that may be different for the two populations. None of the cited EPA guidance documents (*e.g.*, EPA 2000 or EPA 2002) specifically suggests the use of student's t-test or Satterthwaite t-test on log-transformed data. It is preferable to use a two sample nonparametric test (*e.g.*, Mann-Whitney test) rather than using a two sample t-test on log-transformed data sets.

Example: Let the mean of the log-transformed data of the two background groups be roughly equal to 5 with standard deviation of one group (*e.g.*, soils) be equal to 2 and of the other group (*e.g.*, sediments) be equal to 3. The means on the log-scale are both equal to 5, but the mean of background soil in original scale = 1096.63 and mean of the upgradient sediments = 13359.73. Obviously, the two means in the original scale are significantly different.

Please note that by using statistical tests on log-transformed data, one is testing the equality of the medians and not of the means of the two populations. Often the sample size determination formula is based upon the relative or actual differences of the means (*e.g.*, page 3-4, EPA 2002) and not of the medians.

It should also be noted that the decision errors (α and β) and sample size computed to detect a specified mean difference in the original scale do not transform to mean difference (*e.g.*, $\log(a-b)$ is not equal to $\log(a)-\log(b)$), decision errors and sample size in the log-scale (Singh, Singh, and Engelhardt, 1999).

Sample size should be computed using the mean difference (actual or relative) in the original scale.

- p. 5-1 The Work Plan states that because of the potential differences in the chemistries of upgradient sediment samples and background soil samples, the two background data sets may not be combined. In this case, two sets of background 95% UTLs should be computed. Site sediment samples should be compared with the upgradient sediment 95% UTL, and the site soil samples should be compared with the background 95% UTL based upon soil samples.
- p. 5-1, §5.0 The last sentence of the second paragraph is written as follows: “Because the two separate soil chemistries, it is anticipated that the datasets for the upgradient sediments and background soils will remain separated through the evaluation process.” Is a word or two missing from the first part, or does this refer to the use of quartiles and 95th percentile differences (in the previous sentence) to distinguish two different populations? Please clarify the wording here (*i.e.*, is “separate” used here as an adjective or a verb?).
- p. 5-1, §5.0 The list of statistical analyses that will be applied to the background data sets appears to be reasonably complete, and lists tests that determine population distributions, equality of variances, similarity of rank distributions, and similarity of frequency of detections. What test(s) will be used to identify statistical outliers, *e.g.*, Rosner’s test (Gilbert, 1987) or other appropriate approach? Please review this list, and add the appropriate test(s).
- p. 5-1, §5.0 The final paragraph of this page states that a 95 percent UTL of background will be calculated and used to in some way “evaluate” background. As discussed in the US Navy (1999) and mentioned on page 5-2 of this document, comparisons to UTLs are predicted to result in exceedances even when sample population characteristics are similar unless very few site samples are collected. Therefore, it is unclear what type of evaluation of site data will be performed using the 95% UTLs of the background data. All statistical or qualitative comparisons of background to site data should be clearly developed in a work plan separate from the present document or included in a separate section in this document.
- p. 5-2, §5.0 The first paragraph indicates that nonparametric tests are not recommended by Navy guidance for comparisons between data sets with normally distributed data. It should be noted that if data are found to be non-normally or non-lognormally distributed, then nonparametric tests will be needed in order to compare medians as per USEPA (2002).

p. 5-2 The Work Plan states that UTL test will not be used as a stand-alone test (Navy, 1999). All alternative tests (e.g., t-test, Mann-Whitney test) that will be used to compare site and background concentrations should be listed in the Work Plan.

REFERENCES

Gilbert, R. O., 1987, *Statistical Methods for Environmental Pollution Monitoring*. John Wiley & Sons, New York. 320 pp.

Singh, A. K., Singh, A., and Engelhardt, M. (1999). *Some Practical Aspects of Sample Size and Power Computations for Estimating the Mean of Positively Skewed Distributions in Environmental Applications*. EPA/600/S-99/006.

US EPA, 2000. *Guidance for Data Quality Assessment*. EPA QA/G-9. EPA/600/R-96/084. July 2000.

USEPA, 2002, *Guidance for Comparing Background and Chemical Concentrations in Soil for CERCLA Sites*. EPA 540-R-01-003, OSWER 9285.7-41, September 2002.

US Navy, 1999. *Handbook for Statistical Analysis of Environmental Background Data*. Southwest Division, Naval Facilities Engineering Command. September 1999.