



**TETRA TECH NUS, INC.**

55 Jonspin Road • Wilmington, MA 01887-1020  
Tel 978.658.7899 • Fax 978.658.7870 • www.tetrattech.com

C-NAVY-08-03-1651W

August 28, 2003

Project Number N8856

**received**  
08-29-03

Mr. Franco LaGreca  
Head, New England Restoration Management Branch,  
EFA Northeast, Naval Facilities Engineering Command  
10 Industrial Highway, Mail Stop 82  
Lester, Pennsylvania 19113

Reference: CLEAN Contract No. N62467-D-94-0888  
Contract Task Order No. 0849

Subject Response to Comments, Draft Work Plan for Background Investigations  
Study area 08, NUSC Disposal Area  
Naval Underwater Warfare Center, Middletown Rhode Island

Dear Mr LaGreca:

Enclosed you will find responses to comments on the Draft Work Plan for Background Investigations, prepared for the site referenced above. Comments were provided to the Navy by the US Environmental Protection Agency (USEPA) on May 9, 2003 and by the Rhode Island Department of Environmental Management (RIDEM) on May 21, 2003

This response summary was delayed from our intended release date per RIDEM recommendation, pending their release of a proposed revision to the remediation regulations (DEM-DSR-01-93, August 2003). Following our review of the proposed regulation revision, we do not see reason not to proceed with this background study.

If you have any questions regarding this material, please do not hesitate to contact me

Very truly yours,

A handwritten signature in black ink, appearing to read 'Stephen S. Parker'.

Stephen S. Parker, LSP  
Project Manager

SSP/rp

Attachment

- c R. Machado, NUWC (w/encl - 2)
- K. Finkelstein, NOAA (w/encl - 1)
- K Keckler, U.S. EPA (w/encl. - 3)
- P Kulpa, RIDEM (w/encl. - 4)
- S. McFadden, TAG (w/encl. - 1)
- C Mueller, NSN (w/encl. - 2)
- J Stump, Gannett Fleming (w/encl. - 2)
- J. Trepanowski/G Glenn, TiNUS (w/encl - 1)
- File N8856-3 2 (w/o encl.), N8856-8.0 (w/encl. - 1)

1517

**ATTACHMENT A**  
**Responses to Comments From USEPA**  
**Draft Work Plan For Background Soil Investigation**  
**NUSC Disposal Area, SA 08**  
**Naval Underwater Warfare Center, Middletown Rhode Island**  
**Comments dated May 9, 2003**

**General Comment 1**

*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.*

Response: The Navy concurs with the concerns stated above, and has selected sample stations based on this line of reasoning. Such consideration will also be brought into the field during collection of samples.

**General Comment 2**

*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 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.*

Response: There will not be a separate work plan for discussion of the process for comparison of site data to background data. This information will be expanded on in the revised background work plan.

**General Comment 3**

*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.*

Response: The Navy concurs with the statements above, and hopes that such a consensus can be reached through the work plan revisions and comment response summaries such as this one. As described in the work plan, it is the Navy's intent to evaluate background conditions in the soils adjacent to and upgradient of the site, and compare those conditions to site data. It is the Navy's belief that this can be accomplished within reasonable limitations, and proposes to do so by addressing hydric and non-hydric soils separately, and address each of the two soil types which are represented on site (Se and PmB) separately as well (refer to response to comment 5, below).

#### **General Comment 4**

*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.*

Response: When there are less than 15 percent non-detects, EPA and Navy guidance cited above allow the option of using 1/2 of non-detects as a proxy value to compute statistics for parametric tests. When there are more than 15 percent non-detects but less than 40 to 50 percent non-detects, other methods are available to convert detection limits into surrogate positive values (e.g., Cohen's method). However, for this project only nonparametric methods and not detection limit adjustments will be used when there are more than 15 percent non-detects. This is the most reliable approach because there are drawbacks with adjustment methods that add uncertainty or bias. First, adjustment methods require that the detection limits must all be the same value (EPA QA-G9, Section 4.7.2.1 and Navy, 2000, Appendix B, Table B-2), which is a technically incorrect assumption because soil samples may have different preparation weights, percent moisture corrections, and dilution factors. In addition, nonparametric methods can be used successfully to avoid the necessity of imputing surrogate values for nondetected measurements (EPA, 2002, Section 4.4, page 4-7). In addition, the underlying validity of parametric tests depends on the reliability of distributional tests, and "tests for the distribution of the data 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 samples" (EPA, 2002, Section 5.3, page 5-3). The nonparametric Gehan's test will be implemented because it can handle multiple detection limits and is statistically equivalent to WRS/Mann Whitney test when there are no non-detects. In addition, the Quantile test will be used for upper rank data subset comparisons that consist of positive values all greater than the highest detection limit in the data sets. Both of these tests are valid given a wide range and number of non-detects.

#### **General Comment 5**

*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.

**Response:** The reviewer should be advised that the Navy intends to add a third set of 20 samples to address the second soil type at the site (PmB). This addition will be described in the revised work plan. In addition, the hydric soil sample stations shown in the Draft Work Plan are intended to be used for comparison of similar hydric soils at the site. No depositional sediments have yet been identified in the background areas, but if enough are found to exist, these may be used for comparison of the depositional sediments on site, using the same statistical testing methods described in the work plan.

#### **General Comment 6**

*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.*

**Response:** Site-to-background comparisons will be performed after data become available and will involve some of the same tests but different p-levels than background-to-background tests. This will be described in more detail in the revised work plan.

#### **General Comment 7**

*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).*

**Response:** More detail regarding handling of outliers will be added to the text. An upper extreme value screening threshold (four times the 75<sup>th</sup> quantile) will be used to spot potential outliers for further examination and testing. Potential outliers will be initially double-checked for calculation or reporting errors and analytical bias and then located on a site map to check for any pattern of localized, anthropogenic contamination that is not representative of the background area as a whole. Confirmed outliers can be eliminated from the background data set based on Dixon's test for normally-distributed data with less than 25 samples or Rosner's test for normally-distributed data with more than 25 samples (EPA QA-G9, Sections 4.4.2 and 4.4.3, pages 4-26 and 4-27).

## SPECIFIC COMMENTS

### Comment 1

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.*

**Response:** Information on the soil and sediment sampling program at the site will be added to the revised background work plan in order to support additional discussions on how the site vs background comparisons will be conducted.

### Comment 2

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.*

**Response:** The site investigation to be performed at the SA-08 site is being performed only to compare contaminant concentrations in soils (including fill) on site to those concentrations in upgradient and background soils near the site. As stated above, additional text will be included in the revised work plan to describe the site investigation data collection effort. While the Navy concurs that fill areas nearby might be used as an additional data set for comparison with the on-site fill, there are too many unknowns regarding the source of fill used for nearby development, and the quality of that fill vs the fill used on the site to try to develop a separate data set characterized as fill.

### Comment 3

p. 2-5, §2.2.1 *The first paragraph under Soil Types 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.e., 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?*

**Response:** The stissing soil type was sought for the reasons given in the work plan. The Navy concurs that there is a second soil type represented on site (PmB) and an additional data set will be developed for background soil for this type. It is anticipated that there will be minor differences in soil chemistry between the Se, and PmB soils, the commentor should be aware that this study is only attempting to determine contaminant concentrations in the background and upgradient soils and compare those concentrations to the site soils and fill. This comparison is being made to determine if the site soils have higher concentrations of contaminants than those in the background and upgradient areas. Variances in soil types and reactive properties in those soil types may cause both false positive and false negative results in the study, but the tests described in the work plan, which will be supplemented by two sets of data for the two soil types, are adequate to account for such limitations.

**Comment 4**

*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.*

**Response:** The statement noting the possibility of combining the data sets was included in the work plan to assure that if appropriate (based on the data analysis), the data sets might be combined. Currently, we speculate that the data sets will remain separate, and would prefer them to remain so. However, such decisions will have to wait until data is collected and evaluated to determine similarities and differences.

**Comment 5**

*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.*

**Response:** Samples will be collected from fill areas at SA-08 (the Site). Fill will be identified in the soil samples collected from the site by an on-site geologist. The Golf course area where background samples are planned are not believed to have been altered or filled in at least the past 80 years. The commentor correctly notes that the Navy intends to compare the background soils to the site soil, even if those site soils are found to be fill. While such a comparison may not be a "pure" study of soil chemistry between two locations, it will meet the purposes of determining if some contaminant concentrations in site soil are higher, lower, or similar to those in the background soils.

#### **Comment 6**

*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 some point in this project.*

**Response:** The commentor has correctly interpreted the text in that alpha levels will be 0.05 for two-way background-versus-background tests, which corresponds to  $\alpha / 2 = 0.025$ . The commentor is correct in that this applies only to background-to-background tests because the alpha to be used for site-versus-background tests is greater than 0.025. The plan to compare site vs background data will be provided in the revised work plan.

#### **Comment 7**

*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.*

**Response:** The Navy concurs with the observations stated above, and it is for just these reasons that the site study and the background study are being undertaken. Conduct of the SASE will need to address contaminant concentrations at the site associated with fill vs contaminants that are ubiquitous, anthropogenic and upgradient of the SA-08 site.

**Comment 8**

*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.*

**Response:** The former is true for a total of 40 samples. The commentor should note that the Navy is planning to add another 20 samples to address the PmB soil type, resulting in a total of 60 samples for this background study.

**Comment 9**

*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  $\alpha$ . The values of two types of error probabilities,  $\alpha$  and  $\beta$  should be chosen according to the recommendations in Chapters 3 and 5 of the EPA Background Guidance document (2002). A higher value of  $\alpha$  ( $=0.1, 0.2$ ) 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  $\beta$  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).*

**Response:** Page 5-15 (EPA, 2002) states that a range of alpha values between 0.05 and 0.2 can be used. The text on page 5-14 states the same thing. Other guidance documents from Navy (2002, 2000, 1999) and EPA (2000) also mention use of a 0.05 alpha value for one-way tests, which is the value selected for this project. An overall false positive rate of greater than 0.05, for example 0.1 or 0.2, would create an unacceptably high risk of unnecessary remediation because of false positives. Because potential risk-driver TAL metals, such as arsenic, are naturally occurring and normally are found in all background samples, it is important to minimize the risk of false positives.

*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.*

*Once the data is received, a retrospective power analysis is not unambiguous – a power analysis still depends on the minimum detectable difference and if the data distribution is nonparametric, the power analysis is uncertain for several nonparametric tests.*

**Response:** A retrospective power analysis can be performed after data are received, but it will be acknowledged that power analysis for nonparametric tests is not reliable especially when the data distribution exhibits a nonparametric shape.

**Comment 10**

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.*

**Response:**     Approximately 22 site surface soil samples, 6 site hydric soil samples, and 3 depositional sediment samples (sediment under water) are planned for the SASE at the SA-08 site. The site investigation will be summarized for the revised background work plan.

**Comment 11**

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.*

**Response:**     Area B sediment samples will be sampled to measure sediment transport from within the defined surface water pathway so there is no need to sample additional Area B soils. Much of Area B is heavily overgrown and wooded, so sampling within Area A is more efficient and is still representative of the Se soil type.

**Comment 12**

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 algaecides? 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)?*

**Response:**     The metals in the ILM04.0 are planned for analysis of samples collected, same as the samples collected from SA-08. Analysis of nitroaromatics and other explosive residuals are not planned, as these compounds would likely only be associated with Navy

activities, and therefore not attributed to anthropogenic or natural conditions. Conversely, analysis of herbicides or algaecides is also not planned for the site or the background areas, as these compounds would likely only be associated with the golf course activities (anthropogenic, but not actionable on Navy property).

**Comment 13**

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.*

**Response:** Outliers will be screened in the background data set to prevent this problem from occurring. In addition, other types of site-versus-background tests may be considered which are not as sensitive to one or two outliers (t-test, Gehan's test, Quantile test).

**Comment 14**

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.*

**Response:** Given an appropriate match of distributional shapes, there is no particular reason why the comparison of means of logs is not as appropriate as other nonparametric statistical contrasts. If the mean of logs differs between two populations that both follow a lognormal distribution, then this information is a valid alternative indicator of population dissimilarity even though such dissimilarity is obviously defined by different attributes than a comparison of arithmetic means, and the latter of which is not possible in this situation. An overall conservatism will be maintained because all valid parametric and nonparametric tests will be considered, including the nonparametric tests recommended by the commentor. Therefore, if any valid test indicates two populations are dissimilar, then the overall conclusion is that they are not the same and one is greater than the other in the context of the property from which the contrast was drawn.

*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.*

**Response:** This situation will not present a problem because the log-transformed test is not the only test being used. As stated previously, a nonparametric test will also be performed, so if either nonparametric or parametric tests indicate a difference then the conclusion will to accept that a difference exists.

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.

**Response:** While a log-transformed test tends to be in between a test of means versus medians, technically it is still a parametric test and is by no means mathematically equivalent to a median test, which is a nonparametric test.

*It should also be noted that the decision errors ( $\alpha$  and  $\beta$ ) 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.*

**Response:** Decision errors cannot be evaluated precisely until after data are collected, as stated in the text. Distributional shape (normal, lognormal, or nonparametric) cannot be predicted in advance. Once data are collected, then the distributional shape, standard deviation, and background mean or mean of logs can be used where appropriate to evaluate decision error tolerance. The mean of log differences corresponds to a back-transformed mean ratio, which is a statistically valid interpretation with lognormal data. There is no compelling reason that the metric for discussing and evaluating decision errors has to be in terms of differences rather than ratios.

**Comment 15**

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.*

**Response:** The commentator's interpretation is correct. This will be emphasized in the text.

**Comment 16**

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 95<sup>th</sup> 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?).*

**Response:** The passage should read "Because there are two separate..."

**Comment 17**

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).*

**Response:** Outliers will be evaluated using the test appropriate for the distributional shape encountered and the number of samples involved. Guidance recommends that normally distributed data be tested using Dixon's test for  $N < 25$  and Rosner's test for  $N > 25$ . The tests will be listed. In addition to the recommended outlier tests, a qualitative evaluation against an extreme value screening threshold of 4 times the 75<sup>th</sup> percentile will be employed to designate values for further investigation into possible causes of abnormal concentration values.

**Comment 18**

*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.

**Response:** The UTL is useful for identifying hot spot areas needing further investigation or to determine the areal extent of previously confirmed elevated contamination areas. The UTL will not be used for global evaluation across all site samples but instead will be used in conjunction with other statistical tests. This will be emphasized.

**Comment 19**

*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).

**Response:** This has already been stated and will be further emphasized.

**Comment 20**

*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.

**Response:** This will be further emphasized.

## 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.

**ATTACHMENT B**  
**Responses to Comments From RIDEM**  
**Draft Work Plan For Background Soil Investigation**  
**NUSC Disposal Area, SA 08**  
**Naval Underwater Warfare Center, Middletown Rhode Island**  
**Comments Received May 21, 2003**

**General Comment:**

*The navy has elected to perform a background study prior to conducting a remedial investigation of the site. Typically a remedial investigation is performed prior to, or at the same time as the background study. This sequence is normally done as site specific information is used to guide the background study, (i.e. soil types, where samples will be taken, etc.). In addition, performing the background study prior to the collection of site specific information may bring into question the conclusions of the study and may result in the need to perform another study. Finally, and fundamentally, the site-specific information is used to determine if an extensive background study, similar to the one proposed, is even needed.*

*In addition to the above considerations be aware that the Office of Waste Management is currently reevaluating the arsenic standard and back ground study requirements. In light of these considerations the Office of Waste Management suggest that the Navy postpone conducting the background study at least until after site-specific information has been collected.*

**Response:**

The Navy guidance and policy documents cited in the work plan encourage the development of background data as described in the work plan. In fact, it is recommended that background soil information be developed as early in the process as possible. Data from the site investigation will likely be received and evaluated before the background sampling begins, however, planning for the background study is necessary in the meantime.

**1. Section 1.1 Specific Investigation Objectives;  
Page 1-1, Whole Section.**

*This section of the work plan includes a discussion of the Site Remediation Regulations requirements for a background investigation. The discussion implies that if there are elevated levels of contaminants at the site and at neighboring areas remediation is not required. Presence of contamination on the site and or the neighboring area does not negate the need to address the contamination. Therefore, please remove this discussion from the work plan, as these statements concerning the application of the regulations are incorrect.*

**Response:**

The section in question has been reviewed, and the implication cited in the comment above does not appear to be apparent. The text states cites the passage from RIDEM regulations allowing sampling background areas to "distinguish concentrations related to the contaminated site from... (those not related)". It also states the intention to compare the data from the site against the data from the background areas to identify site contaminants from non-site contaminants. The section does not provide any statements relating to remediation, or addressing (or not addressing) the contamination based on the findings of the study. The Installation Restoration

Program and CERCLA mandate cleanup based on risk and risk based cleanup goals for each site determined through accepted practices and risk management, in concert with state and federal regulations.

**2. Section 1.2 Project Deliverables; Page 1-3.**

*Please add the following to the list of deliverables:*

*Map depicting the concentrations of arsenic and lead found in the various sampling locations.*

*Summary table of analytical data, as well as, a separate summary table for arsenic and lead.*

*Map and summary table as described above for any other analyte, which the Navy intends to request a background exemption.*

*Table describing the depth and geologic characteristic of each sample. Summary tables of statistical tests employed including sample results and statistical critical values (as an illustration if a test for normality is conducted the table should include the test result for the sample data as well as the critical or cut off value from the test which will determine whether the sample is normal).*

**Response:**

If background exemptions are sought as a part of the CERCLA studies to be performed at this site, the appropriate support will be provided for those exemptions at the appropriate time, including tables of data as described above. As it appears that RIDEM is particularly interested in arsenic and lead, and as noted in the general comment above regarding the reevaluation of the arsenic standard, the requested maps showing arsenic and lead concentrations will be provided as a deliverable item for this study. Additional maps will be provided if appropriate based on the data reported. For each statistical test, critical values, calculated values, p-values, and parameters used in the calculation (for example, mean, standard deviation, number of data points) will be provided.

**3. Section 2.2.1, Soil Types; Page 2-7, Paragraph 3.**

*The work plan proposes conducting separate background studies on the hydric and non hydric soils. The background samples for the hydric soils will be collected from the streambeds. Although not stated it is assumed that the samples will be collected from soils immediately adjacent to the stream and not from any soils submerged in stream water. Please confirm.*

**Response:**

Samples are to be collected from hydric soils that are on the edges of streams. Hydric soils are not the same as depositional sediment (silts and fine materials under water). The background hydric soils data are to be collected to compare to data from hydric soils found on site, and if depositional sediments are found in the background areas, those sediments will be collected and compared to depositional sediments collected from the site. This will be clarified in the revised work plan.

**4. Section 2.2.1, Soil Types; Page 2-7, Paragraph 3.**

*The work plan notes that the soil adjacent to the stream beds in the same soil classification will be hydric, as opposed to the non-hydric soils located further away from the stream beds. Accordingly,*

*two background studies will be performed at the site, one for hydric and the other for non-hydric soils. These studies will entail the collection of twenty background soils samples for each soil type. The site does not lie in the flood plain of a large river. In fact the streams entering the site are small, and in some locations they can be jumped across. Further, disposal activities have resulted in nearly vertical slopes along sections of the stream, and overall the wetlands at the site itself is limited. Therefore, the hydric soils at the site may not be significant enough to warrant a separate, intensive, background investigation, such as the one proposed in the work plan. Without site-specific information demonstrating the need to perform a separate background assessment, the Office of Waste Management does not concur with the proposed background study for hydric soils.*

**Response:**

The separate data set for hydric soils was deemed necessary because hydric soil, non hydric soil, and even depositional sediment at the site will be evaluated separately if possible and compared against different risk screening criteria in the SI report. Because there are streams actively carrying sediment particles to the site from upgradient properties, and because the activities at these nearby properties may involve chemical uses that are not conducted at the site in question, it is important to identify those contaminants as a part of the site investigation to meet the objectives stated in the response to the general comment and comment no. 1, above.

**5. Section 2.2.4, Definition of Study Boundaries;  
Page 2-9.**

*This section of the work plan includes a discussion of the site and the different areas where background samples may be collected. The work plan notes that the site and the proposed background areas were used for agricultural purposes, golf course, etc. The Navy notes that pesticides, herbicides and other agricultural chemicals were commonly and consistently used at these sites. Be advised that it is inappropriate to collect background samples from release areas. Therefore, all of the proposed background areas are inappropriate and the Office of Waste Management does not concur with the proposed locations and will not accept or review any reports based upon samples taken in these areas. The work plan should focus on non-release areas, that is, areas where pesticide, herbicides, etc were not used. The criteria of collecting samples in non release areas was employed in the background studies performed at other sites on the base. Accordingly, the work plan should be modified and alternate sampling areas should be proposed.*

**Response:**

As evident on Figure 2-2 of the background work plan, the entire area was once agricultural, and therefore it is not possible to locate an area near the site where historical use of pesticides or herbicides has assuredly not occurred. Uses of pesticides or herbicides in accordance with manufacturers instructions are typically not considered releases, and thus should qualify as background locations. While it is not possible to determine if the use and application of any chemical in the past has occurred in accordance with manufacturers instructions, the data set is large enough to show whether a release event has occurred, and the report will provide site comparisons only as appropriate.

**6. Section 3.2, Soil Sampling; Page 3-1, Whole Section.**

*The proposed soil sampling locations are not acceptable. Please submit alternative sampling areas for review. Be advised that background samples should not be collected from release areas.*

**Response:**

The reviewer is asked to refer to the response to comment no. 5 above. There is no known history of releases at the proposed locations. The reviewer should be aware that an additional set of background samples will be added to address the PmB soil type, and these samples will be added within the north and west sections of the golf course shown in Figure 3-1.

**7. Section 3.0, Field Sampling Plan; Page 3-1, Whole Section.**

*The work plan should include a section on regulatory notification. The work plan should specify that the regulatory agencies will receive a schedule for field activities and will be notified one week prior to the start of the sampling effort. In addition, when possible, a twenty-four hour notification should be given for any cancellation of field activities.*

**Response:**

The Navy concurs, and this information will be provided, following completion of the final work plan, and prior to initiation of the field investigation.

**8. Section 5.0, Data Analysis and Statistical Testing; Page 5-1, Whole Section.**

*This section of the work plan discusses the statistical test that will be used to evaluate the data. Although not stated it is assumed that this evaluation will include results for standard statistical test. These test include, but are not limited to, the mean (geometric/arithmetic), median, mode, variance, range, minimum, maximum standard deviation, interquartile range, percentiles, variation, sum, count confidence level skewness, and kurtosis. All of this information should be presented in table format as appropriate. In addition the sample results for a particular contaminant that the Navy is performing a background assessment on, will be depicted in tables in ascending order. The Office of Waste Management recommends placing the above statistical data below the ascending order values.*

**Response:**

Complete supporting information associated with each statistical test will be provided in tables containing results for each test. Other descriptive information that is not directly used in quantitative statistical tests but which is useful to describe the data will be provided as well (mean, minimum, maximum, 25<sup>th</sup>, 75<sup>th</sup>, and 95<sup>th</sup> percentiles of both positives and non-detects, and frequencies of detection). The mean of logarithms (i.e., geometric mean) and standard deviation of logarithms will be provided wherever the distributional shape matches lognormal. Sample results will be presented in ascending order in an appendix to aid in identification of outliers and document limitations of statistical tests related to the range and number of non-detects.

**9. Section 5.0, Data Analysis and Statistical Testing;  
Page 5-1, Whole Section.**

*This section of the work plan notes that the Shapiro Wilks Test will be used to assess normality. There are a number of statistical tests that may be used to assess normality. Should the Shapiro Wilks be inappropriate other tests may be employed.*

**Response:**

The Navy concurs, and the data will be evaluated for appropriateness of the tests to be performed.

**10. Section 5.0, Data Analysis and Statistical Testing; Page 5-11, Whole Section.**

*This section of the work plan lists the different tests that will be performed to analyze the background data set. Prior to performing these analysis tests for outliers should be performed on the data sets. This step is necessary as it may affect which sample locations are used in the background analysis.*

**Response:**

More detail regarding handling of outliers will be added to the text. An upper extreme value screening threshold (four times the 75<sup>th</sup> quantile) will be used to spot potential outliers for further examination and testing. Potential outliers will be initially double-checked for calculation or reporting errors and analytical bias and then located on a site map to check for any pattern of localized, anthropogenic contamination that is not representative of the background area as a whole. Confirmed outliers can be eliminated from the background data set based on Dixon's test for normally-distributed data with less than 25 samples or Rosner's test for normally-distributed data with more than 25 samples (EPA QA-G9, Sections 4.4.2 and 4.4.3, pages 4-26 and 4-27).

**11. Section 5.0, Data Analysis and Statistical Testing; Page 5-1, Whole Section.**

*This section of the work plan lists the different tests that will be applied to the data sets. Please be advised that each test has limitations as to whether the test can be applied to normal, log normal, or non-normal data. Prior to performing the particular test the Navy will have to demonstrate that the test is appropriate for the particular data set.*

**Response:**

The Navy concurs, and the data will be evaluated for appropriateness of the tests to be performed.

**12. Section 5.0, Data Analysis and Statistical Testing; Page 5-1, Whole Section.**

*The work plan notes that a statistical test will be performed to determine whether the hydric and nonhydric background data sets will be combined. The plan has also included a lengthy discussion concerning the differences in the soil types and where the soils would be found, (i.e. hydric soils are in the vicinity of the streams, etc, non hydric soils are upland). Unless there is evidence that flooding*

at the site has resulted in hydric soils being deposited on the nonhydric portions of the site, there is no reason to combine the two data sets.

**Response:**

The Navy concurs, and believes that the data sets will likely remain separated. However, the statements are correct in that the data sets could be combined if the data shows great similarity and little variation. It is unlikely that it could be conclusively shown that hydric soils are moved in significant quantity on top of non-hydric soils in the streambed that is as small as this one is. Evaluation of the air photos shows that the stream channel locations vary little over time, and periodic flooding is not anticipated to be a significant sediment transport mechanism to upland portions of the site.

**13. Section 5.0, Data Analysis and Statistical Testing; Page 5-1, Last Paragraph.**

*This section of the work plan states that the 95 % UTL will be used to determine the background concentration. It is premature to state whether the 95 % UTL will be employed as the background concentration. The value employed will be based upon the data. Accordingly, the work plan should note that 95 % UTL, the 80 %, the mean etc., may be used as a reference value for existing site data.*

**Response:**

It is agreed that the 95% UTL may not always be applicable, specifically if the distributional shape is not normal or lognormal. In addition, the UTL test is not generally valid as a stand-alone background test to judge whether any remedial action is needed at a site. Statistical guidance (US Navy, 2000) acknowledges that the UTL test can produce an unacceptably high false positive rate in cases where the site population is really no different from the background population. Therefore, UTL exceedances should be confirmed by additional statistical tests in which the false positive rate is controlled to less than a 5 percent error rate. In particular, an "elevated" concentration for a metal should be indicated only if there is found to be either an overall difference between the entire populations of site and background sample results (the t-test, the Mann-Whitney test, or Gehan's Test); hot spots at multiple locations (the upper ranks test); or if no other tests are conclusive, an elevated frequency of detection in site versus background (the test of proportions or Fisher's Exact Test).