N60138.AR.000156
FISC WILLIAMSBURG
5090.3a

LETTER AND COMMENTS FROM U S EPA REGION III REGARDING DRAFT CALCULATIONS OF BACKGROUND CONCENTRATIONS TECHNICAL MEMORANDUM CHEATHAM ANNEX FISC WILLIAMSBURG VA
05/05/2010
U S EPA REGION III

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
REGION III
1650 Arch Street
Philadelphia, Pennsylvania 19103-2029

Robert Thomson, P.E., R.E.M.
Office of Federal Facility Remediation

Direct Dial (215) 814-3357
Mail Code: 3HS11

Date: May 5, 2010

Mr. Thomas Kowalski
NAVFAC MIDLANT, Code EV3
9742 Maryland Avenue
Building N-26, Room 3208
Norfolk, VA 23511-3095

Re: Naval Weapons Station–Yorktown NPL site, Yorktown, Va.
Naval Supply Center – Cheatham Annex NPL site, Williamsburg, Va
Review of draft *Calculation of Background Concentrations Technical Memorandum*

Dear Mr. Kowalski:

Enclosed, please find the U.S. Environmental Protection Agency's (EPA's) comments pertaining to the review of the U.S. Navy's (Navy's) February, 2010 draft *Calculation of Background Concentrations Technical Memorandum* (TM) for the Naval Weapons Station-Yorktown (NWSY) NPL site, and the Naval Supply Center – Cheatham Annex (CAX) NPL site:

EPA Region 3 has evaluated the appropriateness of statistical methods used by CH2M HILL to evaluate background concentrations at the NWSY NPL site and the CAX NPL site. In order to address the objective of this evaluation, in addition to reviewing the CH2M HILL draft TM, an independent analysis of the background data was performed for selected contaminants at the NWSY and CAX. This review letter therefore summarizes: 1) results of the independent analysis performed by EPA for selected COCs; 2) presents review comments on the statistical background study performed by CH2MHILL; and 3) makes recommendations based upon the statistical analyses performed by EPA.

Establishing Background/Reference Data

Whenever possible, background data sets are collected from pristine unimpacted locations. When using onsite background data to establish site-specific background level concentrations, it is necessary to assure that the selected background/ reference locations are not impacted by onsite activities and/or contamination originating from other industrial activities potentially impacting the site background.

Outliers in Background Data Sets

Elevated outlying observations in a background data set potentially may represent locations impacted by the site activities, especially when background data are collected from locations (e.g., onsite reference area locations chosen at large federal facilities) potentially impacted by the site and/or other industrial activities. In such scenarios (as is the case for CAX and NWSY onsite reference areas), all potential outliers are removed from background data sets (EPA 2000, 2002) before computing decision statistics (e.g., UTLs, UPLs) to estimate site-specific background level concentrations.

Tests for Outliers

Dixon and Rosner tests (EPA 2006) are often used to identify outliers. However the use more effective robust outlier identification procedures (e.g., Tukey's Biweight function – Hoaglin, Mosteller, and Tukey,1983; PROP influence function - Singh, 1993, LMS and MCD methods - Rosseeuuw and Leroy, 1987) is desirable to identify multiple outliers and multiple populations (e.g., several soil associations). In the present case, in order to keep things simple and verify computations performed by CH2M HILL, only classical Dixon and Rosner tests were used.

1.0 Data

Groundwater, surface, and subsurface soil background data were collected from CAX and NWSY sites. These sites consist of four soil assocaition types. For some locations, duplicate samples were collected. In such cases, higher value was retained and used in the statistical analyses perfromed by CH2MHILL to establish background level concentrations. For comaprison sake, the reviewers also used the higher duplicate value, even though the use of the average of duplicates is a preferred method and is commonly used in practice.

However, for background evaluation studies, the use of the lower value of the duplicate results perhaps will be more appropriate resulting in a conservative estimate (95/95UTL) of the background level concentration.

2.0 Constituents studied in this report

The reviewers used EPA software packages, ProUCL 4.00.05 and Scout to perform statistical analyses for selected COCs described below. The notations used in the statistical computation tables are described in ProUCL technical guide. For an example, KM represents Kaplan Meier method used to compute a decision making statistic such as an upper tolerance limit (UTL) based upon data sets with nondetects (NDs). The following constituents have been used to verify the computations described in the Draft CH2M HILL Report (February, 2010).

2.1 Inorganics – surface and subsurface soils

1. Arsenic (data consist of nondetects with multiple detection limits)
2. Lead
3. Manganese
4. Vanadium

2.2 Organics – surface and subsurface soils

1. Benzaldehyde
2. Benzo(a)pyrene
3. Dieldrin
4. Endosulfan sulphate
5. Total Organics

2.3 Groundwater

1. Arsenic, dissolved
2. Mercury, dissolved (data with nondetects at multiple detection limits)
3. Chromium, dissolved (data with nondetects at multiple detection limits)
4. Nitrobenzene (data with nondetects at multiple detection limits)
5. Arsenic
6. Manganese (data with nondetects with multiple detection limits)

3.0 General Procedure to Compute UTLs

In order to compute defensible upper tolerance limits (UTLs) to estimate background level concentrations, outlier analysis was performed for each selected COC. Since outying observation may potentially represent impacted locations, identified outliers are not inlcuded in the computation of UTLs. Outlier analysis was performed using Rosner's test when the number of observations was greater than 25 and Dixon's outlier test when the number was less than 25. For data sets with nondetects (NDs), the outlier tests were performed on data sets obtained using half of the detection limit (DL/2) values for NDs.

In order to demonstrate the influence of outliers on the computations of various statistics, 95/95UTLs were computed in two ways - with and without outliers. Also in order to compute manaeable statistics and estimates, oneway analysis of variance (ANOVA) was performed to determine if concentrations for the COCs in the 4 soil associations differ significantly. If the concentration distributions of a contaminant are comparable for the four soil types, then it is desirable to compute a single UTL to estimate background concentration for all soil associations combined.Statistical tests are also performed to determine: 1) if concentrations of COCs differ significantly in surface and subsurface soils; and 2) if concentrations of COCs found in CAX and NWSY background areas differ significantly.

Background Analysis for Selected Organics and Inorganics in Soils

There are 4 soil associations present in CAX and NWSY sites. The first step is to compare concentrations of the 4 soil associations to see if they can be combined to

compute 95/95UTLs. This was performed using One-way ANOVA (Scout software) for surface and usbsurface soils. The next step is to search for high outliers since those outliers may represent impacted locations resulting in inflated UTL values. Once the outliers were removed, the UTLs were computed for the selected constituents. Depending upon the data distribution (parameteric, nonparametric) and NDs, appropriate 95/95 UTLs were calculated to estimate background level concentrations.

3.1 Background Statistics for CAX Area

3.1.1 CAX surface soils

Classical One-Way ANOVA was per formed to determine if the concentrations of COCs in the 4 soil associations in CAX area are comparable. The results of the ANOVA test are summarized in the Appendix, Table A-1. Based upon ANOVA test statistics, it is concluded that data for the four soil associations can be combined for arsenic, manganese and vanadium. Results of outlier tests performed on the combined data (4 soil associations) are summarized as follows. The outlier test on arsenic data was performed using ½ DL values for NDs.

Table 3.1: Outliers in CAX surface soils

Metals	N	# Outliers	Outlying Value (ID)
Arsenic	40	3	12.6 (BG4-SS01-00) 9.9 (BG1-SS03-00) 8.3 (BG1-SS06-00)
Lead	40	2	136 (BG1-SS03-00) 34.7 (BG1-SS09-00)
Manganese	40	1	435 (BG4-SS10-00)
Vanadium	40	2	50.3 (BG1-SS03-00) 41.2 (BG1-SS09-00)

ANOVA was performed again on lead data after removing the 2 outlying locations shown in the above table. Once the outliers were removed, ANOVA test results (Appendix, Table A-2) suggest that lead concentrations from the four soil associations can also be considered as coming from a single statistical population.

Table 3.2: 95/95UTLs based upon datasets with and without outliers for CAX - surface soil

Metals	N	Max	Raw Statistics		95/95 UTL	
			Mean	SD	No Outliers	With Outliers
Arsenic	37 (17 NDs)	6	1.178 (KM)	1.552 (KM)	4.501 KM	7.921 KM
Lead	38	16	8.963	3.195	15.78 Normal	136 Nonparametric
Manganese	39	304	95.52	74.52	332 HW Gamma	373.4 HW Gamma
Vanadium	38	26.1	9.626	5.848	27.38 Lognormal	36.8 Lognormal

3.1.2 CAX subsurface soils

A similar process was used for subsurface soil data collected from the CAX area. The ANOVA test statistics are given in Appendix, Table A-3. Based upon ANOVA test results, it is concluded that data for the four soil associations can be considered as coming from a single population. The following table summarizes outlier tests results. The outlier test on arsenic data was performed using ½ DL values for NDs.

Table 3.3: Outliers in CAX subsurface soils

Metals	N	# Outliers	Outliers Value (ID)
Arsenic	40	5	18.6 (BG4-SB01-01) 15.9 (BG2-SB01-01) 13.8 (BG4-SB10-01) 12.5 (BG1-SB09-01) 10.4 (BG1-SB03-01)
Lead	40	3	83.7 (BG1-SB08-01) 35.3 (BG1-SB09-01) 35.3 (BG1-SB03-01)

Manganese	40	2	208 (BG3-SB09-01) 195 (BG43-SB01-01)
Vanadium	40	1	53.8 (BG3-SB03-01)

Table 3.4: 95/95UTLs based upon datasets with and without outliers for CAX - subsurface soil

Metals	N	Max	Raw Statistics		95/95 UTL	
			Mean	SD	No Outliers	With Outliers
Arsenic	35 (20 NDs)	6.8	2.73 (KM)	1.202	5.323 KM	12.83 KM
Lead	37	8.3	5.889	1.375	8.831 Normal	83.7 Nonparametric
Manganese	38	154	46.29	39.32	172.5 HW Gamma	210.6 Gamma
Vanadium	39	37.8	17.37	10.15	38.94 Normal	54.98 Gamma

Conclusion: The concentrations of the selected metals found in four soil associations of CAX (for both surface and subsurface) are comparable. The four surface (and subsurface) soil associations can be considered to represent a single population. Therefore, background level concentrations (95/95 UTLs) can be computed using the combined surface (and combined subsurface) soil associations data sets.

3.1.3 CAX surface soil and subsurface soils combined

It is always desirable to evaluate if there are significant differences in contaminants concentrations of surface and subsurface soils. Two sample hypotheses tests were performed to determine if surface and subsurface concentrations for the selected COCs are comparable.

Null Hypothesis H_0: Data from surface and subsurface soils are comparable and can be combined.
Alternate Hypothesis H_A: Surface and subsurface soils data are different and cannot be combined

Table 3.5: Two Sample Test Results – CAX surface soil versus CAX subsurface soil

Metals	Test	Conclusion	P value ($\alpha = 0.05$)
Arsenic	Gehan	Combine surface and subsurface soils	0.198
Lead	t-Test	Do not combine surface and subsurface soils	0
Manganese	t-Test	Do not combine surface and subsurface soils	0.001
Vanadium	t-Test	Do not combine surface and subsurface soils	0

Arsenic was the only metal where the surface and subsurface soil concentrations are comparable, and the two data sets may be merged to compute a UTL based upon the combined data set. The 95/95UTL for arsenic based upon the combined surface and subsurface soil data sets is given in the following table.

Table 3.6: 95/95UTLs based upon CAX surface and subsurface data combined

Metals	N	Max	Raw Statistics		95/95 UTL
			Mean	SD	
Arsenic	72 (37 NDs)	6.8	1.457 (KM)	1.763	4.947 KM

Conclusion: For most of the CAX COCs evaluated in this report, there are significant differences in contaminant concentrations of surface and subsurface soils, therefore, it is suggested to compute 95/95 UTLs separately for surface and subsurface soils.

3.2 Background Statistics for NWSY Site

The background analysis was performed for NWSY area following the similar procedure as used for CAX area.

3.2.1 NWSY surface soils

Classical ANOVA was performed to determine if the concentrations of the COC in four (4) soil associations are comparable. ANOVA test results are given in Appendix, Table A-4. Based upon the ANOVA test statistics, it is concluded that concentrations of selected

COCs in the 4 soil associations are comparable, and therefore just like the CAX area, 4 soil types can be considered to represent a single statistical populations. Site-wide background level concentrations for surface soils can be computed based upon the combined data set of the 4 soil types. Outlier tests were performed to identify any potentially impacted locations. As before, outlier test on arsenic data set was performed using ½ DL values. The following table summarizes the outlier test results.

Table 3.7: Outliers in NWSY surface soils

Metals	N	# Outliers	Outliers Value (ID)
Arsenic	35	2	63.9 (BGS21) 5.8 (BGS07)
Lead	35	4	43.1 (BGS07) 38.7 (BGS28) 25.3 (BGS22) 22.7 (BGS04)
Manganese	35	2	491 (BGS03) 413 (BGS38)
Vanadium	35	0	-

Table 3.8: 95/95UTLs based upon datasets with and without outliers for NWSY surface soil

Metals	N	Max	Raw Statistics		95/95 UTL	
			Mean	SD	Without Outliers	With Outliers
Arsenic	33	4.2	1.897	0.967	4.528 HW Gamma	63.9 Nonparametric
Lead	31	16.7	11.57	3.08	18.34 Normal	43.1 Nonparametric
Manganese	33	340	99.02	89.83	403.6 HW Gamma	508.1 HW Gamma
Vanadium	35	34.7	15.25	6.808	32.66 HW Gamma	32.66 HW Gamma

From the above table, it is easy to see how outliers inflate the decision statistics (e.g., arsenic) used to estimate background level concentrations.

3.2.2 NWSY subsurface soils

ANOVA test was performed to determine if concentration distributions of COCs in the 4 soil associations are comparable. The ANOVA test results are given in Appendix, Table A-5. Based upon ANOVA test results, it is concluded that concentrations of COCs found in 4 soil associations are comparable, and therefore data of the 4 soil associations can be combined to compute 95/95 UTLs. This means, a single background value can be computed for each of the selected COCs. In order to identify potentially impacted locations present in the site-specific data set, outlier tests were performed. The outlier test on arsenic data was performed using ½ DL values for NDs. The following table summarizes outlier test results.

Table 3.9: Outliers in NWSY subsurface soils

Metals	N	# Outliers	Outliers Value (ID)
Arsenic	13	2	42.7 (BGSB07-11) 26.2 (BGSB04-09)
Lead	13	0	
Manganese	13	3	2940 (BGSB07-11) 353(BGSB04-01) 284(BGSB04-09)
Vanadium	13	0	0

Table 3.10: 95/95UTLs based upon datasets with and without outliers for NWSY subsurface soil

Metals	N	Max	Raw Statistics		95/95 UTL	
			Mean	SD	No Outliers	With Outliers
Arsenic	11	13.8	5.615	4.983	19.64	78.54 Gamma
Lead	13	4.2	11.78	6.778	29.89	29.89 Normal
Manganese	10	94.5	39.88	37.11	325 Gamma	8011 Lognormal*
Vanadium	13	70.3	28.94	21.74	87.01 Normal	87.01 Normal

* It is easy to see how outliers inflate the decision statistics (arsenic, manganese) used to estimate background level concentrations.

Conclusion: The concentrations of the selected metals found in four soil associations of NWSY are (for both surface and subsurface) are comparable. The four surface (and subsurface) soil associations can be considered to represent a single population. Background concentrations (95/95 UTLs) can be computed based upon the combined surface (and subsurface) soil data set.

3.2.3 NWSY surface soil and subsurface soils combined

Just like the CAX site, hypothesis tests were conducted to determine if the concentrations of the COCs in surface and subsurface soils are comparable. The results are summarized in the following table.

Null Hypothesis H_0: Data from surface and subsurface soils are comparable and can be combined.
Alternate Hypothesis H_A: Surface and subsurface soils data are different and cannot be combined.

Table 3.11: Two Sample Test Results – NWSY surface soil versus NWSY subsurface soil

Metals	Test	Conclusion	P value ($\alpha = 0.05$)
Arsenic	WMW	Combine surface and subsurface soils	0.104
Lead	t-Test	Combine surface and subsurface soils	0.885
Manganese	WMW	Do not combine surface and subsurface soils	0.018
Vanadium	WMW	Combine surface and subsurface soils	0.074

Test results for arsenic, lead, and vanadium suggest that surface and subsurface soil data may be combined to compute 95/95UTLs. The following table summarizes UTLs computed using the combined data sets.

Table 3.12: 95/95UTLs based upon NWSY surface and subsurface data combined

Metals	N	Max	Raw Statistics		95/95 UTL
			Mean	SD	
Arsenic	44	13.8	2.826	3.02	13.8 Nonparametric
Lead	44	25.5	11.63	4.41	20.86 Normal
Vanadium	48	70.3	18.96	13.86	64.95

Conclusion: For CAX area, it is concluded that there are significant differences in concentrations of selected COCs found surface and surface soils. Since only four metals were evaluated for NWSY area, and manganese concentrations are found to be significantly different in surface and subsurface soils. In order to keep things simple and comparable for the two background areas, it is suggested that 95/95 UTLs be computed separately for surface and subsurface soils. Often in practice, separate background values are computed for surface and subsurface soils.

3.3 Background Statistics for CAX and NWSY Background Areas all soils combined

Next, hypothesis tests were conducted to determine if concentrations of COCs in surface and subsurface from CAX are comparable to corresponding surface and subsurface concentrations found in NWSY. The detailed test results for each of the metals are given in Appendix, Tables A-6 to A-9. The main results are summarized in the following table.

Null Hypothesis H_0: Data from CAX and NWSY can be considered coming from a single population
Alternate Hypothesis H_A: Data from CAX and NWSY represent different populations

Table 3.13: Two Sample Test Results – CAX soils versus NWSY soils – (surface and subsurface combined)

Metals	Test	Result	P value ($\alpha = 0.05$)
Arsenic	WMW	Do not combine CAX and NWSY soils	0.000
Lead	WMW	Do not combine CAX and NWSY soils	0.000
Manganese	WMW	Combine CAX and NWSY soils	0.504
Vanadium	WMW	Do not combine CAX and NWSY soils	0.0074

Table 3.14: NWSY/CAX Statistics – Using all background data

Metals	N	Max	Raw Statistics		95/95 UTL
			Mean	SD	
Manganese	121	340	76.26	72.05	259.5

For most of the COCs considered in this report, it is noted that concentrations of those COCs found in CAX and NWSY are not comaprable. Therefore, it is suggested that estimates (95/95UTLs) of background concentrations should be computed separately for the two background sites: CAX and NWSY.

Conclusion: Concentrtaions for the selected COCs in the two background areas: CAX and NWSY are significantlly different. Therefore, separate background level concentrations should be computed (95/95 UTLs) for the two background areas.

Error in Computation: While reviewing the CH2M HILL Report, some discrepancies were idenified. For an example, location BGSB04-09 (353) represents an outlier for manganese (also found in the CH2M report) for soil association 3 and not from soil association 1 (incorrectly included in calcluations for soil asscoiation type 1). It appears that this outlier has been incorrectly included in the calculation of manganese UTL of 244.8 for soil association 1.

It is suggested that CH2M HILL double check its calculations to assure that all statistics are correctly computed.

4.0 Background Analysis for Organics in Soils

CAX background data sets consist mostly of NDs for the four soil associations. Hence the data from the four soil associations for both surface and subsurface were combined to compute UTLs after testing for outliers. No organics background data are available for the NWSYsite. Due to lack of detected results, 95/95 UTLs are computed based upon the combined surface and subsurface data collected from the CAX site. Several (6) outliers were identified for total organics as summarized in Table 4.1.

Table 4.1: Outliers for Total Organics - Surface and Subsurface Data Combined

Chemical	N	# Outliers	Outliers Value (ID)
Total Organics	40	6	133000 (BG1-SB03-01) 55100 (BG1-SB09-01) 52360(BG1-SB06-01) 40740 (BG1-SB05-01) 34600 (BG1-SB04-01) 28660 (BG1-SB01-01)

The 95/95UTLs for selected COCs are summarized in Table 4.2.

Table 4.2: Statistics for Organic Compounds - Surface and Subsurface Data Combined

Organics	# Detects	# NDs	Max (detected)	Raw Statistics (using detects)		95/95 UTL	
				Mean	SD	EPA*	CH2M HILL
Benzaldehyde	4	75	100	76	27.72	123.1 KM UTL*	123 KM UTL
Benzo(a)pyrene	1	78	4.2	-	-	-	-

Dieldrin	7	72	4.5	3	0.927	3.282 KM UTL	3.54 KM UTL
Endosulfan Sulphate	7	72	3.4	2.6	0.735	3.516 KM UTL	3.54 KM UTL
Total Organics	34	0	10600	5315	2240	11314 HW Gamma	14900 Gamma

***EPA = EPA Las Vegas**

For Benzaldehyde, using the procedure used by CH2M HILL, the reviewers also get the same value for 95/95UTL. This gives us confidence in the computations made by CH2M HILL. However, based upon the sample size requirements summarized in ProUCL Tech Guide (Chapter 1), 95/95 UTL of 123.1 for Benzaldehyde (4detects and 75 NDs) deemed as an unreliable estimate of the background concentration.

For total organics, EPA identified 6 statistical outliers, and CH2M reported only 5 outliers (first 5 in Table 4.1 above). Inclusion of the last outlier, 28660 is the cause of a higher value, 14900 for 95/95UTL.

Dealing with nondetects: It is not easy to compute reliable decision statistics based upon data sets consisting of only a few detected values (e.g., 1 or 2). In a data set consisting mostly of NDs (e.g., >95%), a 95/95 UTL may also be considered a ND value. One can use nonparametric KM method to compute a 95/95UTL when number of detected observations is at least 4-5 (more are preferred for higher confidence) and percentage of NDs is less than 95%. However, the uncertainty associated with such a decision statistic (95/95 UTL) remains high. In such situations, one may want to use the largest or next largest detected value (provided they are not outliers) as an estimate of the background level concentration. However, it is the recommendation of EPA that organic background concentrations be eliminated given the high uncertainty surrounding the calculations involving a large number of nondetects.

5.0 Background Analysis for Groundwater

Following, a similar procedure as used for metals in soils, background evaluations were performed for groundwater data for the selected COCs identified by EPA. Results of outlier analysis performed on groundwater data are summarized in the following table.

Table 5.1: Outliers in groundwater

Chemical	N	# Outliers	Outliers Value (ID)
Arsenic, dissolved	32	3	6.5 (YBKG-GW09A-YE-0809) 2.5 (YBKG-GW08A-CC-0809) 2 (YBKL-GW01P-CC-0809)
Chromium, dissolved	32	1	27.8(YBKG-GW09A-YE-0809)
Arsenic	32	6	17.8 (YBKL-GW01P-CC-0809) 17.2 (YBKL-GW01-CC-0809) 15.4 (YBKG-GW01A-YE-0809) 8.9 (YBKG-GW07-CC-0809) 7.8 (YBKG-GW09A-YE-0809) 4.3 (YBKG-GW08A-CC-0809)
Manganese	32	4	279 (YBKG-GW01A-YE-0809) 126 (YBKG-GW10-CC-0809) 104 (YBKG-GW07-CC-0809)

Table 5.2: 95/95 UTLs with and without Outliers

Chemical	N (no outliers)	Max	Raw Statistics		UTLs		
			Mean	SD	EPA (no outliers)	EPA (with outliers)	CH2M
Arsenic, dissolved	29	1.6	0.614	0.329	1.552 HW Gamma	3.18 Lognormal	1.37 Gamma
Mercury, dissolved	32 (29 NDs)	0.14	0.104	0.0279	0.165 KM		0.118 KM
Chromium, dissolved	31 (19 NDs)	5.8	2.05	1.399	5.123 KM	13.1 Lognormal	6.04 KM
Nitrobenzene	32 (30 NDs)	9.4	2.613	1.183	5.199 KM		5.64 KM
Arsenic	26	2.5	0.903	0.552	2.518 HW Gamma	17.8 Nonparametric	2.21 Gamma
Manganese	28 (4 NDs)	13.4	29.54	12.97	58.66 KM	151.6 KM	57.9 Normal

Conclusion

For mercury, dissolved in groundwater, there are only 3 detects and they are all lower than the 29 nondetects with detection limit of 0.2. Three detected values for mercury are: 0.072, 0.1 and 0.14. The rest of the observations were nondetects with a detection limit of 0.2. As mentioned before, it is not possible to compute a reliable decision statistic (e.g., 95/95UTL) based upon such data sets with a few detects. In such situations, one may use the maximum or the next to the maximum detected value (provided that value is not an outlier) as an estimate of the background threshold concentration. As mentioned previously, EPA recommends that background concentrations calculated with a large number of nondetects be eliminated given the high uncertainty. For mercury, the below maybe acceptable to EPA, given that mercury is not identified as a COC for the Operable Unit. This should be discussed with the project team, as the statistics are highly suspect.

In the present case, for mercury one may use 0.1 (second largest detected value) as an estimate of the background level concentration.

For Nitrobenzene, there were only 2 detected values: 2.4 and 9.4, and rest of the 30 results are nondetects with detection limit of 0.25. It appears the detected value 9.4 may represent an outlier. Its cause should be determined as it may be a release. In such situation, the background concentration may be estimated by the next largest value which is = 2.613 or just a nondetect value. Alternatively, since most of the values for nitrobenzene are NDs, it can be determined that nitrobenzene is not present in background groundwater.

6.0 Summary and Conclusions

- For most of the COCs evaluated in this report, significant differences were found in contaminant concentrations of surface and subsurface soils (for both CAX and NWSY), therefore, it is suggested to compute 95/95 UTLs separately for surface and subsurface soils.

- The concentrations of the selected metals found in the four soil associations of CAX area (for both surface and subsurface individually) are comparable. Therefore, four surface (and subsurface) soil associations can be considered to represent a single surface (and subsurface) population for CAX area. Background concentrations (95/95 UTLs) can be computed based upon the combined surface soil data set, and combined subsurface soil data set. A similar conclusion is derived for the four soil associations of NWSY area.

  This is contrary to the CH2M Report which states that only soil associations 3 and 4 could be combined.

- Concentrations for the selected COCs in the two background areas: CAX and NWSY are significantlly different. Therefore, separate background level concentrations should be computed for the two background areas.

- As mentioned before, it is not possible to compute a reliable decision statistic (e.g., 95/95UTL) based upon data sets consisting mostly of nondetect results (e.g., mercury and nitrobenzene in groundwater). EPA recommends that background concentrations calculated with a large number of nondetects be eliminated given the high uncertainty. Mercury maybe an exception as discussed above, based upon input from the project team.

- Overall, 95/95 UTL computations made by CH2M HILL appear to be correct except in cases when the majority of data are nondetects.

- Since some errors were identified in the calculations performed, it is suggested that CH2M HILL double check its calculations to assure that all statistics are correctly computed.

- In some cases as identified in this report, it is noted that CH2M HILL included outliers in their compuations resulting in inflating values of UTLs. These outlier concentrations may have come from areas that have potentially been impacted by site operations. Elevated observations (outliers) from these areas may represent releases. Therefore, outliers should not be included in the computation of 95/95UTLs.

7.0 References

Hoaglin, D.C., Mosteller, F., and Tukey, J.W. 1983. *Understanding Robust and Exploratory Data Analysis*. John Wiley, New York.

Rousseeuw, P.J., and Leroy, A.M. 1987, *Robust Regression and Outlier Detection*, John Wiley and Sons, NY.

Singh, A. 1993, *Omnibus Robust Procedures for Assessment of Multivariate Normality and Detection of Multivariate Outliers, In Multivariate Environmental Statistics*, Elsevier Science Publishers, Patil G.P. and Rao, C.R., Editors, 1993, pp. 445-488.

Singh, A., Maichle, R.W. and S. Lee 2006. On the Computation of a 95% Upper Confidence Limit of the Unknown Population Mean Based Upon Data Sets with Below Detection Limit Observations. EPA 2006, EPA/600/R-06/022.

U.S. Environmental Protection Agency (EPA) U.S. Nuclear Regulatory Commission, *et al.* 2000. *Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM). Revision 1. EPA 402-R-97-016*. Available at http://www.epa.gov/radiation/marssim/ or from http://bookstore.gpo.gov/index.html (GPO Stock Number for Revision 1 is 052-020-00814-1).

U.S. Environmental Protection Agency (EPA). 2002. *Guidance for Comparing Background and Chemical Concentrations in Soil for CERCLA Sites*. EPA 540-R-01-003-OSWER 9285.7-41. September, 2002.

U.S. Environmental Protection Agency (EPA). 2010. *ProUCL 4.00.05 Software for Calculating Upper Confidence Limits (UCLs)*, Office of Research and Development, April 2010. http://www.epa.gov/esd/tsc/TSC_form.htm

U.S. Environmental Protection Agency (EPA). 2010. *ProUCL 4.00.05 Technical Guide (Draft)* EPA/600/R-07/041, Office of Research and Development, April 2010.

U.S. Environmental Protection Agency (EPA). 2009. *Scout 2008 – A Robust Statistical Package*, Office of Research and Development, February 2009. http://www.epa.gov/esd/databases/scout/abstract.htm#Scout2008v101

Appendix

Table A-1 Classical One-Way ANOVA for the 4 metals in surface soil of the CAX Background Area

Arsenic

Group	Obs	Mean	SD	Variance
ssi-bk1	10	3.953	3.382	11.44
ssi-bk2	10	1.286	1.06	1.124
ssi-bk3	10	0.908	0.447	0.2
ssi-bk4	10	2.423	3.853	14.84
Grand Statistics (All data)	40	2.143	2.794	7.809

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	56.13	3	18.71	2.712	0.0593
Within Groups	248.4	36	6.9		
Total	304.5	39			

Pooled Standard Deviation	2.627
R-Sq	0.184
Adj (R-Sq)	0.116

Lead

Group	Obs	Mean	SD	Variance
ssi-bk1	10	23.98	40.46	1637
ssi-bk2	10	8.5	1.57	2.464
ssi-bk3	10	9.07	2.161	4.669
ssi-bk4	10	9.58	3.631	13.18
Grand Statistics (All data)	40	12.78	20.63	425.5

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	1678	3	559.2	1.35	0.274
Within Groups	14917	36	414.4		
Total	16595	39			

Pooled Standard Deviation	20.36
R-Sq	0.101
Adj (R-Sq)	0.0262

Manganese

Group	Obs	Mean	SD	Variance
ssi-bk1	10	42.94	35.02	1227
ssi-bk2	10	128.2	62.29	3880
ssi-bk3	10	112.7	86.69	7515
ssi-bk4	10	132.2	132.5	17548
Grand Statistics (All data)	40	104	91.06	8292

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	51850	3	17283	2.291	0.0947
Within Groups	271533	36	7543		
Total	323383	39			

Pooled Standard Deviation	86.85
R-Sq	0.16
Adj (R-Sq)	0.0904

Vanadium

Group	Obs	Mean	SD	Variance
ssi-bk1	10	18.78	16	256
ssi-bk2	10	8.49	2.938	8.634
ssi-bk3	10	9.21	5.532	30.6
ssi-bk4	10	9.25	6.711	45.04
Grand Statistics (All data)	40	11.43	9.853	97.08

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	723.5	3	241.2	2.835	0.0518
Within Groups	3063	36	85.08		
Total	3786	39			

Pooled Standard Deviation	9.224
R-Sq	0.191
Adj (R-Sq)	0.124

Table A-2 Classical One-Way ANOVA for Lead in surface soil of the CAX Background Area after removing the 2 outliers

Lead

Group	Obs	Mean	SD	Variance
ssi-bk1	8	8.638	5.183	26.87
ssi-bk2	10	8.5	1.57	2.464
ssi-bk3	10	9.07	2.161	4.669
ssi-bk4	10	9.58	3.631	13.18
Grand Statistics (All data)	38	8.963	3.195	10.21

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	6.913	3	2.304	0.211	0.888
Within Groups	370.9	34	10.91		
Total	377.8	37			

Pooled Standard Deviation	3.303
R-Sq	0.0183
Adj (R-Sq)	-0.0683

Table A-3 Classical One-Way ANOVA for the 4 metals in subsurface soil of the CAX Background Area

Arsenic

Group	Obs	Mean	SD	Variance
sbi-bk1	10	4.427	4.083	16.67
sbi-bk2	10	4.82	4.222	17.83
sbi-bk3	10	1.173	0.444	0.197
sbi-bk4	10	4.526	6.297	39.65
Grand Statistics (All data)	40	3.737	4.407	19.42

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	88.46	3	29.49	1.586	0.21
Within Groups	669.1	36	18.59		
Total	757.5	39			

Pooled Standard Deviation	4.311
R-Sq	0.117
Adj (R-Sq)	0.0432

Lead

Group	Obs	Mean	SD	Variance
sbi-bk1	10	17.9	25.37	643.6
sbi-bk2	10	6.86	1.251	1.565
sbi-bk3	10	5.2	1.169	1.367
sbi-bk4	10	6.06	1.274	1.623
Grand Statistics (All data)	40	9.005	13.3	177

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	1069	3	356.2	2.198	0.105
Within Groups	5834	36	162		
Total	6902	39			

Pooled Standard Deviation	12.73
R-Sq	0.155
Adj (R-Sq)	0.0844

Manganese

Group	Obs	Mean	SD	Variance
sbi-bk1	10	24	25.8	665.5
sbi-bk2	10	65.76	37.65	1418
sbi-bk3	10	46.68	58.87	3466
sbi-bk4	10	79.78	62.85	3950
Grand Statistics (All data)	40	54.06	51.41	2642

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	17565	3	5855	2.465	0.0779
Within Groups	85492	36	2375		
Total	103057	39			

Pooled Standard Deviation	48.73
R-Sq	0.17
Adj (R-Sq)	0.101

Vanadium

Group	Obs	Mean	SD	Variance
sbi-bk1	10	18.09	16.09	258.8
sbi-bk2	10	22.44	8.491	72.1
sbi-bk3	10	13.42	9.647	93.07
sbi-bk4	10	19.16	10.43	108.7
Grand Statistics (All data)	40	18.28	11.56	133.6

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	417.4	3	139.1	1.045	0.385
Within Groups	4794	36	133.2		
Total	5211	39			

Pooled Standard Deviation	11.54
R-Sq	0.0801
Adj (R-Sq)	0.00343

Table A-4 Classical One-Way ANOVA for the 4 metals in surface soil of the NWSY Background Area

Arsenic

Group	Obs	Mean	SD	Variance
1	3	2.233	1.106	1.223
2	12	2.101	1.227	1.505
3	13	2.054	1.296	1.679
4	7	10.53	23.55	554.6
Grand Statistics (All data)	35	3.78	10.52	110.7

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	398.4	3	132.8	1.223	0.318
Within Groups	3367	31	108.6		
Total	3765	34			

Pooled Standard Deviation	10.42
R-Sq	0.106
Adj (R-Sq)	0.0193

Lead

Group	Obs	Mean	SD	Variance
1	3	15.63	7.537	56.81
2	12	10.42	3.325	11.05
3	13	16.88	11.6	134.5
4	7	13.87	1.161	1.349
Grand Statistics (All data)	35	13.96	7.911	62.58

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	270.3	3	90.1	1.504	0.233
Within Groups	1857	31	59.91		
Total	2128	34			

Pooled Standard Deviation	7.74
R-Sq	0.127
Adj (R-Sq)	0.0426

Manganese

Group	Obs	Mean	SD	Variance
1	3	184	126.4	15988
2	12	162	158.6	25146
3	13	82.67	89.5	8011
4	7	85.91	72.86	5309
Grand Statistics (All data)	35	119.2	120.8	14595

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	59658	3	19886	1.412	0.258
Within Groups	436571	31	14083		
Total	496229	34			

Pooled Standard Deviation	118.7
R-Sq	0.12
Adj (R-Sq)	0.0351

Vanadium

Group	Obs	Mean	SD	Variance
1	3	16.43	6.679	44.6
2	12	13.28	4.607	21.22
3	13	15.79	8.378	70.19
4	7	17.13	7.43	55.21
Grand Statistics (All data)	35	15.25	6.808	46.34

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	79.53	3	26.51	0.549	0.652
Within Groups	1496	31	48.26		
Total	1576	34			

Pooled Standard Deviation	6.947
R-Sq	0.0505
Adj (R-Sq)	-0.0414

Table A-5 Classical One-Way ANOVA for the 4 metals in subsurface soil of the NWSY Background Area

Arsenic

Group	Obs	Mean	SD	Variance
3	3	12.83	12.63	159.5
2	5	6.14	4.37	19.09
4	5	12.29	17.83	317.8
Grand Statistics (All data)	13	10.05	12.22	149.3

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	124.8	2	62.41	0.374	0.697
Within Groups	1667	10	166.7		
Total	1792	12			

Pooled Standard Deviation	12.91
R-Sq	0.0697
Adj (R-Sq)	-0.116

Lead

Group	Obs	Mean	SD	Variance
3	3	8.767	4.486	20.12
2	5	15.74	7.963	63.4
4	5	9.64	5.678	32.24
Grand Statistics (All data)	13	11.78	6.778	45.95

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	128.5	2	64.27	1.52	0.265
Within Groups	422.8	10	42.28		
Total	551.4	12			

Pooled Standard Deviation	6.502
R-Sq	0.233
Adj (R-Sq)	0.0798

Manganese

Group	Obs	Mean	SD	Variance
3	3	243.8	133.8	17916
2	5	40.5	38.2	1459
4	5	608.4	1304	1699605
Grand Statistics (All data)	13	305.8	799	638437

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	821153	2	410576	0.6	0.567
Within Groups	6840087	10	684009		
Total	7661240	12			

Pooled Standard Deviation	827
R-Sq	0.107
Adj (R-Sq)	-0.0714

Vanadium

Group	Obs	Mean	SD	Variance
3	3	32.83	23.02	529.7
2	5	33.06	17.85	318.5
4	5	22.48	27.39	750
Grand Statistics (All data)	13	28.94	21.74	472.7

Classical One-Way Analysis of Variance Table

Source	SS	DOF	MS	V.R.(F Stat)	P-Value
Between Groups	339	2	169.5	0.318	0.735
Within Groups	5334	10	533.4		
Total	5673	12			

Pooled Standard Deviation	23.09
R-Sq	0.0598
Adj (R-Sq)	-0.128

Table A-6 Gehan's Test for Arsenic in CAX and NWSY Background Areas

Sample 1 Data: Arsenic(cax)
Sample 2 Data: Arsenic(nwsy)

Raw Statistics

	Sample 1	Sample 2
Number of Valid Samples	72	44
Number of Non-Detect Data	37	0
Number of Detect Data	35	44
Minimum Non-Detect	0.23	N/A
Maximum Non-Detect	3.1	N/A
Percent Non detects	51.39%	0.00%
Minimum Detected	0.094	0.36
Maximum Detected	6.8	13.8
Mean of Detected Data	2.709	2.826
Median of Detected Data	2.4	1.65
SD of Detected Data	1.816	3.02

Sample 1 vs Sample 2 Gehan Test

H0: Mu of Sample 2 = Mu of Sample 1

Gehan z Test Value	3.453
Lower Critical z (0.025)	-1.96
Upper Critical z (0.975)	1.96
P-Value	5.5409E-4

Conclusion with Alpha = 0.05
Reject H0, Conclude Sample 2 <> Sample 1
P-Value < alpha (0.05)

Table A-7 WMW Test for Lead in CAX and NWSY Background Areas

Sample 2 Data: Lead(nwsy)
Sample 1 Data: Lead(cax)

Raw Statistics

	Sample 1	Sample 2
Number of Valid Samples	75	44
Number of Distinct Samples	47	38
Minimum	1.1	4.2
Maximum	16	25.5
Mean	7.447	11.63
Median	6.8	11.4
SD	2.901	4.41
SE of Mean	0.335	0.665
W-Stat Rank Scores	3515	3626
Expected W Scores (H0)	4500	2640
SD of W Scores (H0)	181.7	181.7

Wilcoxon-Mann-Whitney (WMW) Test

H0: Mean/Median of Sample 2 = Mean/Median of Sample 1

Sample 2 Rank Sum W-Stat	3626
WMW U-Stat	2636
WMW Expected U	1650
WMW SD of U	181.7
Standardized WMW U-Stat	5.422
Lower Critical Value (0.025)	-1.96
Upper Critical Value (0.975)	1.96
Approximate P-Value	5.8855E-8

Approximate P-Value < alpha (0.05)

Conclusion with Alpha = 0.05
Reject H0, Conclude Sample 2 <> Sample 1

Table A-8 WMW Test for Manganese in CAX and NWSY Background Areas

Sample 2 Data: Manganese(nwsy)

Sample 1 Data: Manganese(cax)

Raw Statistics

	Sample 1	Sample 2
Number of Valid Samples	77	43
Number of Distinct Samples	74	43
Minimum	2.9	3.5
Maximum	304	340
Mean	71.23	85.27
Median	51.1	68.9
SD	64.37	84.16
SE of Mean	7.335	12.83
W-Stat Rank Scores	4573	2688
Expected W Scores (H0)	4659	2602
SD of W Scores (H0)	182.7	182.7

Wilcoxon-Mann-Whitney (WMW) Test

H0: Mean/Median of Sample 2 = Mean/Median of Sample 1

Sample 2 Rank Sum W-Stat	2688
WMW U-Stat	1742
WMW Expected U	1656
WMW SD of U	182.7
Standardized WMW U-Stat	0.468
Lower Critical Value (0.025)	-1.96
Upper Critical Value (0.975)	1.96
Approximate P-Value	0.64

Approximate P-Value >= alpha (0.05)

Conclusion with Alpha = 0.05

Do Not Reject H0, Conclude Sample 2 = Sample 1

Table A-9 WMW Test for Vanadium in CAX and NWSY Background Areas

Sample 2 Data: Vanadium(nwsy)
Sample 1 Data: Vanadium(cax)

Raw Statistics

	Sample 1	Sample 2
Number of Valid Samples	77	48
Number of Distinct Samples	71	44
Minimum	2.1	1.2
Maximum	37.8	70.3
Mean	13.55	18.96
Median	10.9	14.9
SD	9.13	13.86
SE of Mean	1.041	2
W-Stat Rank Scores	4324	3552
Expected W Scores (H0)	4851	3024
SD of W Scores (H0)	197	197

Wilcoxon-Mann-Whitney (WMW) Test

H0: Mean/Median of Sample 2 = Mean/Median of Sample 1

Sample 2 Rank Sum W-Stat	3552
WMW U-Stat	2376
WMW Expected U	1848
WMW SD of U	197
Standardized WMW U-Stat	2.675
Lower Critical Value (0.025)	-1.96
Upper Critical Value (0.975)	1.96
Approximate P-Value	0.00747

Approximate P-Value < alpha (0.05)

Conclusion with Alpha = 0.05
Reject H0, Conclude Sample 2 <> Sample 1

This concludes EPA's review of the Navy's February, 2010 draft *Calculation of Background Concentrations TM* for the NWSY NPL site, and the CAX NPL site. If you have any questions, please feel free to call me at (215) 814-3357,

Sincerely,

Robert Thomson, P.E., R.E.M.
Federal Facility Remediation (3HS11)

Cc: Wade Smith (VaDEQ, Richmond)
Dawn Ioven (USEPA, 3HS41)