DEPARTMENT OF THE ARMY
OFFICE OF THE DEPUTY CHIEF OF STAFF, G-9
600 ARMY PENTAGON
WASHINGTON, DC 20310-0600

February 22, 2021

Base Realignment and Closure Operations Branch

Mr. Kevin Pierard
Chief, Hazardous Waste Bureau
New Mexico Environment Department
2905 Rodeo Park Drive East, Building 1
Santa Fe, New Mexico 87505-6303

RE: Final Groundwater Background Evaluation, Army's Response to the New Mexico Environment Department Letter of Disapproval dated September 15, 2020. Fort Wingate Depot Activity, McKinley County, New Mexico. EPA# NM6213820974, HWB-FWDA-20-001

Dear Mr. Pierard:

This letter is in reply to the New Mexico Environment Department (NMED) Letter of Disapproval dated September 15, 2020, reference number HWB-FWDA-20-001, Final Groundwater Background Evaluation. The following are Army's response to NMED comments, detailing where each comment was addressed and cross-referencing the numbered NMED comments.

Comments:

<u>SPECIFIC COMMENTS</u>

1. Section 1.3, Study Methodology, lines 36-38, page 2, and lines 4-6, page 3

Permittee Statements: "Alluvial monitoring well TMW28 is located immediately adjacent to the Rio Puerco and groundwater at this location may be influenced by surface water recharge, therefore this well was not included as a background well." and, "BGMW01 was retained as an appropriate background monitoring well because this well is furthest from FWDA operations and is the most hydrogeologically upgradient of the alluvial wells."

NMED Comment: According to Figure 2, *Monitoring Wells Utilized for Statistical Analysis*, well TMW28 is located adjacent to well BGMW01. If well TMW28 is excluded because of the influence from surface water recharge, well BGMW01 located near the Rio Puerco should also be excluded due to the same reason. Provide a more detailed justification for why well BGMW01 is utilized as background evaluation, while well TMW28 is not, in the revised Report.

Army Response, (Section 1.3, page 3 lines 10-22)

Monitoring well BGMW01 was the only alluvial well retained because it is located at one of the most hydraulically upgradient locations (Sundance 2019, Sundance 2017, Sundance 2015) and at the greatest distance from historic FWDA operations. BGMW01 is located approximately 107 feet from a northern branch of the Rio Puerco and at an elevation of 6,690.82 feet mean sea level and hydraulically upgradient of TMW28. Comparatively, TMW28 is located 52 feet from this same northern branch of the Rio Puerco, and at an elevation of 6,688.08 feet mean sea level. According to FWDA personnel, TMW28 is subject to flooding while BGMW01 is not (verbal communication with FWDA staff, 2019). The upgradient location of BGMW01, its distance, from the drainage, and its higher elevation make the groundwater at BGMW01 less likely to be influenced by surface waters. Lastly, the USGS cites TMW28 in its documentation of the Puerco Flow Path: “A flow path originating from the saturated alluvial deposits underlying the South Fork of the Puerco River”. This statement suggests surface water recharge to groundwater at well TMW28 (USGS 2009).

2. Section 1.3, Study Methodology, lines 14-16, page 3

Permittee Statement: “Three wells (BGMW08, BGMW09, BGMW10) located east and hydrogeologically upgradient of the Administration Area were retained as these three wells are furthest from FWDA operations and are the most hydrogeologically upgradient of the bedrock wells.”

NMED Comment: Comment 4 in NMED’s *Approval with Modifications Final Revision 1 Groundwater Periodic Monitoring Report July through December 2018*, dated July 30, 2020 states, “based on the available data, well BGMW08 must not be used as a background monitoring well.” Accordingly, well BGMW08 must not be used as a background well. Revise the Report accordingly.

Army Response:

Pursuant to the Army’s response to the NMED comment #4 from the *Groundwater Periodic Monitoring Report July through December 2018*; it is not BGMW08 that is providing erroneous geochemical data, but TMW02. As presented in the Army’s response (cited above), bedrock groundwater at TMW02 is likely mixing with alluvial groundwater creating erroneous observations. Therefore, the Army has proposed the retention of BGMW08 concurrent with decommissioning TMW02. No changes were made to the revised report.

3. Section 1.3, Study Methodology, lines 22-27, page 3

Permittee Statement: “The single alluvial well BGMW01, which is at the most upgradient location for the alluvial aquifer in this portion of FWDA, has sufficient data points (15 monitoring events collected across 7 years) to conduct a statistical analysis of alluvial data. The three bedrock background wells BGMW08, BGMW09, BGMW10 each had 3 monitoring events (collected across 1-1/2 years); therefore, the data were pooled to derive a sufficient number of samples for a statistical analysis of bedrock groundwater data (USEPA 2009).”

NMED Comment: Discuss whether compounds originating from anthropogenic sources (e.g., explosives, volatile and semi-volatile compounds) have historically been detected from the groundwater samples collected from wells BGMW01, BGMW09, and BGMW10. If such compounds have been detected, these wells are not appropriate for background evaluation. In addition, compare the concentrations of nitrate, nitrite, and perchlorate detected in the selected background groundwater monitoring wells relative to the other upgradient wells. Based on the comparison, justify the use of these wells for background evaluation for nitrate and perchlorate in the revised Report.

Army Response, (Section 1.3, page 3, lines 33-39, page 4, and page 5 lines 1-18)

There is no evidence to suggest that historic contaminating operations were performed at or near background locations BGMW01, BGMW08, BGMW09 or at BGMW10, or that historic operations influence groundwater quality at these locations. However, the hydraulically upgradient locations of these wells does not preclude the potential for detections of anthropogenic compounds.

To determine whether or not BGMW01, BGMW08, BGMW09 and BGMW10 should be included or excluded due to the presence anthropogenic constituents, a review of groundwater analytical results for anthropogenic compounds (explosives, volatile semi-volatile compounds, polychlorinated biphenyls, herbicides and pesticides) from BGMW01, BGMW08, BGMW09, and BGMW10 was performed.

It was determined that there were no detections of anthropogenic compounds in samples collected from BGMW01 and BGMW09. A single detection of one constituent (methyl acetate) out of all of the compounds in these analysis suites was reported from BGMW10. Upon further review of the laboratory chromatogram and mass spectrum for this sample result, it was determined that although the target analyte identification criteria (mass spectra ion relative intensities and relative retention times) were nominally met, the analyte identification was still questionable. The secondary ion at m/z 43 maximized 1.4 seconds after the primary ion at m/z 74, the tertiary ion at m/z 59 was not present (albeit possibly due to the low concentration), and there were a number of ion masses that appeared in the sample spectrum at approximately the same relative intensities as m/z 43. Further, m/z 43 is a common ion fragment mass and shows up commonly in the mass spectra of lighter analytes, such as methyl acetate. This result would have been more appropriately reported as non-detect at the sample limit of quantitation (LOQ). This single result is not considered valid, as it should have been reported as non-detect (U) at the LOQ. In addition, at BGMW08, low level detections of one herbicide (dinoseb) and two VOCs (benzene and toluene) were reported at estimated concentrations below the limit of detection (LOD) from groundwater samples collected during the 2018 groundwater sampling events. Herbicides have not been reported in bedrock wells BGMW09 or BGMW10, and it has since been recommended to discontinue analyzing for herbicides in these wells during future groundwater monitoring events, as agreed upon in the Army response to the Approval with Modifications letter HWB-FWDA-17-008 dated January 8, 2018. For this reason, the single report of one herbicide (dinoseb) is considered insignificant. The two VOCs (benzene and toluene) are also anthropogenic compounds. Based upon the low estimated concentrations of these two compounds and the hydraulic upgradient location of BGMW08, the estimated detection of the two VOCs (benzene and toluene) are not considered significant to this background evaluation.

The USEPA has recognized, in several documents, the potential for anthropogenic compounds to be present at background monitoring locations (USEPA 2018). The issues for background monitoring points is not the lack of anthropogenic compounds, but the lack of hydrogeologic influence from the site-specific contaminating source(s). Therefore, detections of anthropogenic compounds, if any, do not preclude the use of these wells as background monitoring points, as these detections are representative of local or regional conditions. For example, the concentrations of nitrate and perchlorate detected in the alluvial background wells are orders of magnitude less than the highest concentrations observed in the downgradient wells.

- From April 2012 to October 2018, 327 samples were collected from the alluvial aquifer downgradient wells for analysis of nitrite. The alluvial aquifer downgradient nitrite detections ranged from 0.063 to 4.7 mg/L. Background concentrations of nitrite were not reported for the alluvial background well for this study.

- From April 2012 to October 2018, 327 samples were collected from the alluvial aquifer downgradient wells for nitrate. The alluvial aquifer downgradient nitrate detections ranged from 0.044 to 97 mg/L; and in the single background alluvial well nitrate detections ranged from 0.0970 to 1.90 mg/L.

- From April 2012 to October 2018, 262 samples were collected from the alluvial aquifer downgradient wells for perchlorate. The alluvial aquifer downgradient perchlorate detections ranged from 0.0000041 to 0.0015 mg/L; and in the single background alluvial well perchlorate was non-detect (LOQ 0.00005 - 0.0002)

The concentration of nitrate and perchlorate detected in the bedrock background wells are orders of magnitude less than the highest concentrations observed in the downgradient wells.

- From April 2012 – October 2018, 179 samples were collected from the bedrock aquifer downgradient wells with a nitrite range of 0.053 to 1.2 mg/L. Background concentrations of nitrite were not reported for the bedrock background wells for this study.

- From April 2012 – October 2018, 179 samples were collected from the bedrock aquifer downgradient wells for nitrate. The bedrock aquifer downgradient nitrate detections ranged from 0.046 to 50 mg/L; with a single background bedrock detection of 0.0870 mg/L.

- April 2012 – October 2018, 194 samples collected in the bedrock aquifer downgradient wells for perchlorate. The bedrock aquifer downgradient perchlorate detections ranged from 0.0000063 to 4.4 mg/L; with a bedrock background detections range of 0.0000057 to 0.00000950 mg/L.

4. Section 1.3, Study Methodology, lines 36-39, page 3
Permittee Statement: "The downgradient wells were selected as those wells which are hydrogeologically downgradient from the background wells and completed within the same aquifer as the background well(s). Downgradient well observations are tested against the BTVs to identify groundwater quality changes."

NMED Comments: Table 1, *Monitoring Wells Utilized for Statistical Analysis* lists 23 alluvial downgradient wells where the analytical results are to be compared. However, many of these wells are located in the areas where the water leak from well 69 is likely affecting the groundwater quality and flow direction in the alluvial aquifer. Once the well is abandoned, the flow direction and water quality may change. Comment 8 in NMED's *Disapproval Final Groundwater Periodic Monitoring Report July through December 2018*, dated January 30, 2020, states, "Figure 4-1, *Northern Area Alluvial Groundwater Contour Map July 2018* and Figure 4-2, *Northern Area Alluvial Groundwater Contour Map October 2018*, indicate that the groundwater flow direction in the alluvium is predominantly southwest and west.

However, groundwater leakage from well 69 may have been affecting the Facility's natural groundwater flow direction. According to the figures, the groundwater flow direction south of the Administration Area (e.g., areas around the TNT Leaching Beds) is north to northwest. Presumably, the natural alluvial groundwater flow direction is consistent with local topography toward the Rio Puerco; therefore, a northerly and northwesterly groundwater flow direction is more likely. The areas in the south of the Administration Areas may be less affected by the well 69 leakage and more representative of natural groundwater flow direction. The flow direction may significantly change once the leakage is repaired." Accordingly, the data collected from the wells located south of the Rio Puerco are not appropriate for comparison. Wells BGMW02, BGMW03, MW23, and MW24 are located north of the Rio Puerco and these wells are located downgradient of well BGMW01. The groundwater quality and flow direction at these well locations are less likely to be affected by the water leak from well 69. Remove all downgradient wells located south of the Rio Puerco, if appropriate, and include wells located north of the Rio Puerco for data comparison; revise the Report accordingly.

Army Response, (Section 1.3, page 6 lines 3 -24)
Background wells are used to calculate BTVs, with downgradient monitoring points only used to determine significance levels for the UPLs. Analytical results from the downgradient wells were not used to calculate BTVs and were not used for comparative purposes; therefore, any potential influence of the Rio Puerco and/or from well 69 is not pertinent to this statistical analysis.

It is the number of downgradient wells that are used to determine the individual test significance level per UPL - not the well location (i.e., wells located south of the Rio Puerco) and not the groundwater analytical results at the downgradient location. The number of downgradient wells is used as input to satisfy the site-wide false positive rate and produce a test significance level. It is only the count of downgradient wells that is required, not the actual location of the wells. This is because the requirement to establish the site-wide false positive rate is of a statistical nature, not of a hydrogeological nature. The Unified Guidance methodology was followed to set individual test significance levels such that their sum over all potential tests in a year does not exceed the recommended site wide false positive rate (SWFPR) of 10 percent (see Section 6, page 6-9, USEPA Unified Guidance 2009).

While groundwater mixing certainly occurs in the Administration Area, and surface water from the Rio Puerco infiltrates to the alluvial aquifer, these inputs to the hydrogeologic regime do not preclude comparison of BTVs calculated from the background monitoring wells to be used as comparative values at downgradient locations. The Army based its rationale on the selection of these 23 downgradient wells per USEPA Federal Register (80 FR 21399), which states "*Because hydrogeologic conditions vary so widely from one site to another, the rule does not prescribe the exact number, location and depth of monitoring wells needed to achieve the general performance standard*." Table 1 identifies the background and downgradient wells by aquifer.

5. Section 2.1, Outlier Test Results – Background Wells, lines 25-26, page 6
Permittee Statement: "For constituents that had NDs, the NDs were removed prior to testing for outliers."

NMED Comment: The ProUCL Technical Guide (Guide) suggests that relevant statistics be computed using data sets with outliers and without outliers for comparison. This extra step helps to see the potential influence of outlier(s) on the various decision-making statistics and helps in making informed decisions about the disposition of outliers. It is important to identify high outliers representing contaminated locations and hot spots. The occurrence of non-detect (ND) observations and other low values is quite common in data sets, especially when the data are collected from a background or a reference area. For the purpose of the identification of high outliers, ProUCL indicates that the ND values may be replaced by their respective detection limits (DLs) or half of the DLs or in some cases, may be ignored (e.g. high reporting limits are associated with NDs) in the outlier tests. Except for the identification of high outlying observations, the outlier test statistics computed with NDs or without NDs are not used in any of the estimation and decision-making processes. Therefore, for the purpose of testing for high outliers, it does not matter how the ND observations are treated. As such, eliminating the NDs from the datasets prior to testing for outliers may be an acceptable approach. Clarify if the NDs had high reporting limits and whether inclusion of the ND values resulted in any changes in identified outliers in the revised Report.

Army Response, (Section 2.1, page 8 Lines 26 – 39 to page 9 Lines 1 - 19)
Based upon review of the groundwater analytical results utilized for this background evaluation, there were no elevated laboratory detection limits. For those constituents that had NDs, the NDs were removed prior to conducting the Dixon's outlier tests. The NDs that were excluded for Dixon's outlier tests did not have high reporting limits. NDs were not included as part of Dixon's outlier tests for the following reasons:

- While ProUCL states to include the NDs at the limit or ½ limit value, this inclusion is misleading. Researchers in the last 20 years have expounded on not using this substitution method, as it produces bias in the estimates. For example, we have found that when using ND values with ProUCL, the NDs themselves became outliers.

- ProUCL chose to use the variant of the Dixon's outlier test, which assumes an underlying normal distribution except for the outlier, and that there is only one outlier in the sample. The choice of the normal distribution is unexpected, as most of the technical guidance emphasizes the gamma and lognormal distributions as appropriate distributions to model groundwater constituent concentrations. Even the lognormal distribution comes with caveats. Including data samples with NDs and then using the simple substitution method could further distort the distribution from an assumed normal one.

- Dixon's outlier test was conducted with what is available in ProUCL, but multiple lines of evidence were applied using scatter plots and box and whiskers plots, which included using all detect and non-detect values combined with scientific and historic knowledge of the area to determine if the highest value should be removed or kept.

While there are other methods to test for outliers in the presence of NDs in very small samples, the Army's approach is in compliance with the Unified Guidance (USEPA 2009). Due to the small sample sizes, the Army supplemented the statistical outlier tests with visual means using scatter plots and box and whisker plots. When background sample sizes accumulate to 20 or more samples, a sensitivity test may be done to compare outlier test results with or without NDs.

The issue of elimination of values with high reporting limits is not of concern, as the eliminated values did not have high reporting limits. Only detected values were analyzed when running Dixon's outlier test. Using multiple lines of evidence, the statistical outlier tests were substantiated with visual means using scatter plots and box and whiskers plots, which used all detect and non-detect values along with scientific and historic knowledge of the area to determine if the highest value should be removed or retained. In accordance with the Unified Guidance (USEPA 2009), HDR's review by a statistician and hydrogeologist determined that the values exhibited unexceptional ranges from a geological perspective for the constituents in the area.

6. Section 2.1, Outlier Test Results – Background Wells, line 40, page 6, and lines 1-2, page 7
Permittee Statement: "Given the variable nature of groundwater samples and the small sample sizes, the statistical outliers should not be removed from the data set at this time for purposes of determining background concentrations."

NMED Comment: The Guide states that potential outliers be manually identified where they may be present in a data set before proceeding with the computation of the various background threshold value (BTV) estimates and that upper limits computed by including a few low probability high outliers tend to represent locations with those elevated concentrations rather than representing the main dominant population. The Guide further suggests that relevant statistics be computed using data sets with outliers and without outliers for comparison, to see the potential influence of outlier(s) on the various decision-making statistics about the disposition of outliers. Based on the recommendations in ProUCL, inclusion of the outliers may or may not be appropriate as inclusion of the outliers may not represent the main population of background data and could adversely set a higher compliance level than is representative of background. In looking at the box and whisker plots presented in Figures 4A through 4D, some of the data appear to be outliers. In the revised Report, provide additional justification for the inclusion or removal of the identified outliers in the background datasets.

Army Response, (Section 2.1, page 9, Lines 38-39, page 10, and page 11 lines 1-9)
The justification for inclusion of the apparent outliers is based on a review of range of the concentrations in light of the small sample sizes. The review was done in four steps:

1. Identify which constituent-well pairs had statistical outliers based on ProUCL's Dixon's test.
2. Study range of data in scatter plots from the same constituent-well pairs identified in Step 1. Note if the highest values were a magnitude or less in difference from the other values.
3. Study range of data from box and whisker plots for all constituent-well pairs, including those identified in Step 1. Note if the highest values were a magnitude or less in difference from the other values.
4. Check if human activity or error occurred on the dates of the apparent statistical outliers. Since no reason could be found, the values were included in the development of the BTVs.

Outlier tests, whether they are based on Dixon's or use graphical means (e.g., box and whisker plots), are less definitive for very small datasets (less than 20 observations) as we have gathered from the alluvial and bedrock wells. What appears to be an extreme value or outlier in the small dataset is most likely from a portion of the background distribution that has yet to be sampled (See Unified Guidance (2009), page 5-5). Outliers flagged by statistical tests or by means of scatter plots (Figures 3a – 3c) or box and whisker plots (figures 4A – 4D) were evaluated for anomalous ranges and the values are deemed reasonable and expected. The Army does not have other technical information, knowledge, or basis at this time with a very small dataset to remove the highlighted values. Lastly, the values in figures 4A and 4D appear to be outliers because they are plots of all constituents and some of the constituents had statistical outliers based on Dixon's tests as listed in Table 4. Based upon review of the box and whisker plots for ranges within one order of magnitude or less, it was determined the values were not due to data errors, discrepancy or other non-background populations.

The Unified Guidance (USEPA 2009) recommends not removing statistically identified outliers unless some basis for a likely error or discrepancy can be identified, or they are of high magnitude compared to other concentrations. The Unified Guidance (USEPA 2009) recognizes that statistical outlier tests should be done on datasets; however, the decision to drop them rests on whether the values are in error or of very high magnitude (USEPA Unified Guidance page 5-5). Some statisticians have discussed dropping outliers as a rule; however, this is not substantiated with very small sample sets. Small sample sets are not fully representative of the populations from which they are drawn, as they only have a partial picture of the underlying distribution of groundwater concentrations. What appears to be an extreme value or outlier in the small dataset is most likely from a portion of the background distribution that has yet to be sampled (Unified Guidance 2009). Nevertheless, the outlier analysis for this work was not performed solely by the software; the model output was reviewed by a degreed and practicing statistician and hydrogeologist to ensure the outlier evaluation was appropriate for this specific site evaluation.

7. Section 2.1, Outlier Test Results – Background Wells, lines 3-4, page 7
Permittee Statement: "As new data become available, outlier test results may change and earlier observations thought to be outliers may no longer be outliers."

NMED Comment: It is agreed that the background dataset should also be re-evaluated with each sampling round. In future periodic groundwater monitoring reports, provide a section that includes a discussion of the background level re-evaluation.

Army Response, Report not updated
Per the Permit, the GPMRs are to evaluate groundwater concentrations and trends. Interpretation of data from the GPMRs and background wells will be conducted after the remedy is in-place.

8. Section 2.3, Background Trends, lines 8-10, page 8, and Section 2.4, Statistical Comparison of Alluvial and Bedrock Background Wells, lines 24-27, page 9
Permittee Statements: "Although statistical trends were identified for these three constituents [dissolved nickel, total arsenic and total chromium], these trends were not consistent among alluvial and bedrock monitoring wells (Table 6)." and, "While the sample sizes for the alluvial and bedrock background wells are relatively small, the ANOVA test results and the apparent differences of the distributions based on the side- by-side box and whisker plots between the alluvial and bedrock background wells suggest that the BTVs should be specific to an aquifer and should not be pooled."

NMED Comment: Because statistically significant trends were not observed in the background well regression analyses for dissolved nickel, total arsenic, and total chromium, it is unlikely that many of the underlying conditions needed to pool the background data would be met. The results of the comparison of the alluvial and bedrock data indicate that there are apparent differences in distributions within the two aquifers. Thus, NMED agrees that the background data for the alluvial and bedrock aquifers should not be pooled. No response required.

Army Response,
Comment noted, no action required

9. Section 2.4, Statistical Comparison of Alluvial and Bedrock Background Wells, lines 36-38, page 8
Permittee Statement: "As many groundwater constituents follow lognormal or lognormal- like distributions, the raw data were transformed by taking their natural logarithms."

NMED Comment: The Kruskal-Wallis test does not require log-transformation of the data; however, as log transformation will have no effect at all on the value of the Kruskal-Wallis statistic, there are no issues regarding the use of log-transformed data. No response required.

Army Response,
Comment noted, no action required

10. Section 4, Findings, lines 18-21, page 14
Permittee Statement: "Samples collected from FWDA often have high turbidity which can affect observed analytical results (Sundance, 2019). Well turbidity can introduce excess naturally occurring trace elements into the samples and result in elevated, inconsistent, and incomparable metals results within and between wells."

NMED Comment: The groundwater samples with high turbidity readings must be removed from data evaluation to be used for statistical analyses because inclusion of such data may adversely set a higher compliance level than is representative of background. Revise the Report, as necessary.

Army Response, Report Not updated
The Groundwater Periodic Monitoring Reports from Spring 2009 to Spring 2012 show a collection of 449 samples, with 27% of samples having turbidity greater than100 NTU. In Fall 2019, 69% of samples had turbidity greater than 100 NTU.

As these high turbidity levels reflect conditions at the site, all background samples were retained to establish representative conditions for ground water samples. Removing a large number of samples where turbidity was high would leave fewer data points and less representative points to establish the BTVs.

11. Table 9, Background Threshold Values for Monitoring Constituents, page 44
NMED Comment: The BTVs for the alluvial and bedrock nitrogen concentrations are indicated as 1.90 and 0.0870 mg/L, respectively. Nitrogen was not included in the analytical suite for groundwater samples. The referenced constituent is likely nitrate. Correct the typographical error in the revised Report.

Army Response,
Acknowledged. Table 10 of the revised report was corrected.

12. Appendix A, Statistical Summary Reports, Upper Prediction Limits for Monitoring Constituents in the Alluvial Aquifer, Footnote Number 5. Permittee Statement: "To optimize the process for estimating a reference background concentration when there are multiple forms of the [upper prediction limit (UPL)] formulae for data sets with assumed distributional properties, the averaging of the results over the multiple methods is done. The formulae are comparable and each one offers advantages and also has inherent errors in their methods. The action of averaging these estimates produces a pooled estimate closest to the unknown common truth based on how this error is perceived."

NMED Comment: Pooling of UPLs is not typical and appears to be inappropriate; no technical guidance and/or justification has been found to support pooling UPLs (averaging various UPLs) to derive a single UPL. Further, the distinct tests have differing equations and different underlying assumptions. Revise the determination of the UPL to be based on the best fit or provide additional technical justification and references to support pooling UPLs in the revised Report.

Army Response, Appendix A, Footnote 5

To optimize the process for estimating a reference background concentration when there are multiple forms of the UPL equations for data sets with assumed distributional properties, the averaging of the results over the multiple methods or equations under the recommended distribution assumption is done. The equations are comparable, and each one offers advantages but also has inherent errors in its methods. The action of averaging these estimates produces a pooled estimated closest to the unknown common truth based on how this error is perceived.

The first step in this process is to select one parametric distribution or failing results from the goodness of fit tests (or having 50% or more NDs), selecting nonparametric methods. The ProUCL authors have recognized that the GOF tests at a given significance level can indicate multiple distributions as the best or discernable fit. We see examples of all three parametric distributions passing the GOF tests, two of the three passing, or just one. The very small datasets are unreliable to have confirmation as to the best fit, though results are indicative of a distribution from a type of parametric distribution. For this reason, the latest version of ProUCL no longer recommends a best fit. When multiple distributions are noted as fitting a distribution, we prioritize the selection of the 'recommended' distribution in this order: Gamma, lognormal (provided the standard deviation of the log-transformed data is less than 1), and normal. The ProUCL authors highly recommend gamma over lognormal, even if both pass the GOF tests as the lognormal can have extremely long tails and may produce unrealistic BTVs. (e.g., page 38 of ProUCL Version 5.1 Technical Guide).

Once a parametric distribution has been assigned, and depending on the censorship level of the sample, one or multiple equations are available to estimate the UPL. If there is more than one equation, we average the outputs. The averaging of multiple equations is a method drawn from both meta-analysis (combining outputs from the experts) and machine learning (combining outputs from different algorithms trying to predict the same value).

The work contained in the Unified Guidance and ProUCL are based on information as of the late nineties or early 2000's. Nuances as to how to interpret UPL results using very small datasets are not addressed in depth (e.g., the issue of overfitting a dataset due to small sample size), leaving the ambiguity to be resolved by the practitioner. We have resolved such ambiguity by averaging the different equations within a distribution class as supported by state of the practice for statistical and machine learning analyses. The averaging or pooling is a means to negate the biases from the different equations under very small sample conditions to produce a UPL that is more representative than any one estimate produced by a particular equation. It is noted that the averaging approach in this study will have minimal impact on the recommended UPLs, as the approach specific methods within a parametric class produce very similar UPL estimates.

13. Appendix A, Statistical Summary Reports, Upper Prediction Limits for Monitoring Constituents in the Alluvial Aquifer

NMED Comment: In some cases, the data follow a discernable distribution as determined by ProUCL (e.g., gamma or lognormal); however, a non-parametric test was selected over a distribution-based test. In most cases, the resulting UPL is equivalent to the maximum detected background value and/or is conservative (lower than distribution based UPLs).
Therefore, the use of these UPLs is acceptable. However, provide the rationale for selecting a non-distribution based UPL in the revised Report.

Army Response, report not updated
In section 2.2, the rationale for using nonparametric methods is presented. As well, please see page 15-25 of the Unified Guidance, which also uses this method when there are more than 50% NDs in a sample.

If you have questions or require further information, please contact me at George.h.cushman.civ@mail.mil, 703-455-3234 (Temporary Home Office, preferred) or 703-608-2245 (Mobile).

Sincerely,

George H. Cushman IV

George H. Cushman IV
BRAC Environmental Coordinator
Fort Wingate Depot Activity
BRAC Operations Branch
Environmental Division

Enclosures
CF:

Dave Cobrain, NMED, HWB
Ben Wear NMED, HWB
Michiya Suzuki, NMED, HWB
Lucas McKinney, U.S. EPA Region 6
Ian Thomas, BRACD
Steven Smith, USACE
Michael Falcone, USACE
Saqib Khan, USACE
Admin Record, NM
Admin Record, Ohio