

## **Evaluation of Background-related DQOs**

### **STATISTICAL METHODOLOGY**

Estimation Uncertainty as Minimum Detectable Difference

Decision Uncertainty

Evaluation of Background-related DQOs Using Estimation and Decision Uncertainty

### **UCL-RELATED DQO EVALUATION**

RCRA Canyon

Remaining Onsite Area

Former Ponds and Pads

### **List of Tables**

Table B6-1 – Background-related DQO Statistics for RCRA Canyon Soils (0-5 feet)

Table B6-2 – Background-related DQO Statistics for Remaining On-site Area Soils  
(0-5 feet)

Table B6-3 – Background-related DQO Statistics for Former Ponds and Pads Soils  
(0-5 feet)

## ATTACHMENT B-6 EVALUATION OF BACKGROUND-RELATED DQOs

The Data Quality Objectives (DQOs) for human health and ecological risk assessment address data adequacy for comparing study area metals concentrations to background. Evaluated in this section, the background-related DQO follows USEPA guidance on background comparisons (USEPA 2002), specifically "Form 1," in which the null hypothesis is that a study area is within background. Recommended statistical errors for this form are alpha (probability of false exceedance of background)  $\leq 20\%$ , and beta (probability of false acceptance within background  $\leq 10\%$ ). Uncertainty in comparing to the background condition is referred to as "estimation uncertainty." Additional methodology is incorporated to identify only those cases when estimation uncertainty is high enough to impact the risk characterization relative to screening levels. This is referred to as "Decision Uncertainty."

Although the concepts of alpha and beta "false acceptance" or "false exceedance," are applicable to other study areas, quantitative assessment of alpha and beta require the assumption that all samples are independent random samples of the same population. Therefore, DQO's are only evaluated in detail for study areas where random sampling was applied, namely RCRA Canyon, Remaining Onsite area, and Former Ponds and Pads.

The statistical methodology and key statistical definitions are presented below followed by the evaluation of UCL-Related DQOs for each of three study areas listed above.

### STATISTICAL METHODOLOGY

#### Estimation Uncertainty as Minimum Detectable Difference

The estimation uncertainty associated with a two-sample test, is quantified as the minimum detectable difference (MDD), i.e., the smallest difference in elevation above background that will cause the test to reject the null (within background) hypothesis. The MDD is calculated as a function of the alpha and beta desired for the test, and various "pooled" statistics calculated from both the background data set and the study area data set. Therefore this MDD describes estimation uncertainty relative to the assumption of the background condition but does not address the importance of this uncertainty relative to toxicity screening levels.

Assessing the MDD or power of a test after it is performed has a circular aspect in that it tells us what we already know most of the time. If the study area is found to exceed, we know with certainty that power was high enough and MDD low enough, we don't need to calculate them. If the study area is not found to exceed, and actually has a much lower sample mean than background sample mean (so that background mean - study area mean  $< \text{MDD}$ ), then power of the one-sided test was also adequate even though the null hypothesis was accepted. However, based on the definition of "background", the occurrence of a study area mean much lower than background is theoretically rare. Therefore the MDD calculation by itself, provides little information, if any, that is additional to the test outcomes themselves. (Note that the "Estimation Uncertainty" column in Tables B6-1 to B6-3 is nearly identically correlated to the answer to the test itself given in the "Study Area Within Background?" column.)

To draw more meaning from the MDD, it must be compared not to the study area or pooled data set but to an independent standard, such as a toxicity screening level. Relating minimum

detectable difference to toxicity screening levels pertains to risk-related decision uncertainty and is described below.

### Decision Uncertainty

Toxicity screening levels are important to consider along side estimation uncertainty. Generally, if estimation uncertainty is high relative to background it may still be well below the screening levels which would indicate that additional sampling is not a priority. However, when estimation uncertainty is close to the screening levels, additional sampling may be beneficial.

To assess decision uncertainty, Tables B6-1 to B6-3 apply a toxicity related standard in two ways, only one of which is used for the final DQO assessment discussed below. The "Actual Decision Uncertainty," identified as "actual" because this is the value used for the DQO assessment, is considered adequately low if the study area sample mean is easily distinguished from the screening level, either much higher than the screening level or much lower. Actual Decision Uncertainty is high if the study area mean is close to the screening level. An older method for determining Decision Uncertainty, identified as "Stated Decision Uncertainty," is also included in these tables. The Stated Decision Uncertainty follows the MDD goal stated and applied in the RI work plan for the random sample size estimation. It is shown for comparison but is not applied in the DQO assessment for this appendix. Although performing correctly most of the time, the Stated Decision Uncertainty was considered to perform inadequately in several cases and therefore was not used in the DQO assessment.

### Evaluation of Background-Related DQO Using Estimation and Decision Uncertainty

The evaluation rule considering both types of uncertainty is relatively simple:

*Cases are considered to adequately meet background-related DQOs if either or both estimation uncertainty or decision uncertainty is low.*

*A re-statement of this rule is that cases are considered to inadequately meet DQOs only when both estimation and decision uncertainty are high.*

The rationale for concluding the DQO to be met even if only one type of accuracy is adequate is as follows.

If estimation uncertainty is high (see column labeled "Estimation Uncertainty") yet decision uncertainty (column labeled "Decision Uncertainty") is low, then the conclusion of "within background" may be uncertain however the study area and background distributions are well below the screening level. DQO is adequately met because the estimation uncertainty, regardless of how high, does not impact the decision.

Alternatively, if decision uncertainty is high (i.e., the estimation interval spans the SL) but estimation uncertainty is low (the study area is demonstrated to be clearly higher than background or clearly lower than background), then additional sample data will provide little or no benefit and the DQO is adequately met. This case arises when the true study area mean is either less than, but very close, to the background mean or well above it, and both are close to the screening levels

In short, it is only in cases where decision uncertainty is high and estimation uncertainty is high that the beta-type error can be reduced (with additional sampling) while at the same time maintaining alpha at 0.05.

## **UCL-RELATED DQO EVALUATION**

Compounds which do not meet the UCL-Related DQOs, as stated above, are identified with underline below for each study area. Additional detail about these cases as well as example cases which do meet DQOs is also provided.

### ***RCRA Canyon***

For RCRA Canyon, the background-related DQO was met for all metals except for chromium, selenium, and tin (Table B6-1). The DQO was not strictly met for tin, however the case of tin is considered below and determined to meet general standards of sample number adequacy

The two-sample t-tests of background condition, for both chromium and selenium, have minimum detectable differences (at the stated error rates, alpha = 0.2, beta = 0.1) that could be improved (narrowed) with additional sampling, potentially to the extent needed to detect a difference over background. Additionally the lowest risk-based screening levels for both metals are well below background so exceedance of the background analysis could potentially drive COPC selection. Selenium has the additional uncertainty of having substantial censoring due to "below detection limit" data.

The background sample size for tin (10) reflects the removal of 7 RI Phase II results for tin, all of which were non-detect below the relatively low detection limit of 2.8 mg/kg. These data were outliers compared to both the RI Phase 1 samples and the site-wide study are data set for tin. The two-sample t-test of background condition for tin has a minimum detectable difference (at the stated error rates, alpha = 0.2, beta = 0.1) that could be improved (narrowed) with additional sampling, potentially to the extent needed to detect a difference over background. Additionally the lowest risk-based screening levels for tin is near background so exceedance of the background condition could potentially drive COPC selection. As discussed in Attachment A-1 of Appendix A, the distribution of tin across the site, down to all sample depths, appeared as a single slightly skewed lognormal distribution, which though shifted somewhat from the background sample distribution, had all characteristics of being ambient rather than contamination-related. Therefore, the background condition for tin was characterized using a site-wide ambient UTL (N > 500) rather than only the 10 RI-Phase I samples applied in this table.

### ***Remaining On-site Area***

For Remaining Onsite Soils, the background-related DQO was considered met for all metals (Table B6-2) except tin. The DQO was not strictly met for tin, however background characterization for tin applied the site-wide data set (> 500 samples) and development of the background concentration based on ambient distribution reasoning, as noted directly above and detailed in the attachment A1 of the background analysis Appendix A.

**Former Ponds and Pads**

For Former Ponds and Pads soils, the background-related DQO was met for all metals except for nickel, tin, and zinc (Table B6-3). The DQO criteria were not strictly met for tin, however background characterization for tin applied the site-wide data set (> 500 samples) and development of the background concentration based on ambient distribution reasoning, as noted directly above and detailed in the attachment A1 of the background analysis Appendix A.

For zinc, the study area mean is relatively close to the background mean, and therefore additional sampling has only a moderate chance of reducing the minimum detectable difference enough to detect an exceedance above background. However due to the proximity of the screening level to the study area mean, the DQO for zinc is not met.

For nickel, the difference between the study area and background means for is larger and closer to what the current minimum detectable difference is. Therefore, additional sampling may change the result of the two-sample test for nickel. Additionally, the screening level is similar to the background mean and so the exceedance of background could potentially drive the COPC selection.