TETRA TECH, INC.

May 23, 2022
Nichole Yuen
Department of Toxic Substances Control
700 Heinz Avenue, Suite 200C
Berkeley, California 94710
Sara Ziff
U.S. Environmental Protection Agency, Region 9

75 Hawthorne Street
San Francisco, California 94105

## Via electronic mail

## Subject: Corporation Yard, Triplicates Sample Results Richmond Field Station University of California, Berkeley

Dear Ms. Yuen and Ms. Ziff:

On behalf of the University of California Berkeley, Tetra Tech completed additional triplicates sampling as a follow-up to the removal action conducted at the Corporation Yard in 2017-2018 and presented in the Corporation Yard Data Gaps Sampling Results letter, prepared by Tetra Tech, dated November 22, 2019. The triplicates sampling was completed May 20 and May 21, 2021 following the Corporation Yard, Triplicates Sampling Approach, Revised, prepared by Tetra Tech, dated December 18, 2020.

## BACKGROUND

The purpose of the triplicates sampling was to further determine the mean concentrations of polychlorinated biphenyls (PCB) within the near surface soil ( 0 to 2 inches below ground surface [bgs]) within the Corporation Yard area between Building 120 and the fence line south of Building 185. This area is covered by decision units (DU) designated as DU9 through DU17, as shown on Figure 1.

Initial samples were collected from these DUs for PCB analysis using incremental sampling methodology (ISM) in September 2019, with results shown in Figure 1. Nested triplicate samples (i.e., both field and laboratory subsampling triplicates) were collected at DU11 for quality control (QC) purposes. The laboratory results for DU11 exhibited lower PCB concentrations than several of the other DUs, including one non-detect (ND) result. The combination of two factors: (1) lower relative concentrations; and (2) an ND result, raised the question of whether DU11 results, as measured by the relative standard deviation (RSD), were sufficiently representative of the internal variability of other DUs with significantly different concentrations.

The weighted $-95^{\text {th }}$ percent upper confidence level ( 95 UCL ) is a statistical value representing a conservative upper limit on the estimated PCB concentration for the entire area covered by DU9 through DU17. The ability to accurately represent the internal variability of all the DUs is important when
calculating a weighted-95UCL. The weighted-95UCL is compared to the cleanup goal of 1 milligram per kilogram ( $\mathrm{mg} / \mathrm{kg} \mathrm{)} \mathrm{established} \mathrm{in} \mathrm{the} \mathrm{Removal} \mathrm{Action} \mathrm{Workplan} \mathrm{(RAW)} \mathrm{and} \mathrm{Toxic} \mathrm{Substances} \mathrm{Control}$ Act (TSCA) Risk-Based Approval for the Corporation Yard.

## PROPOSED APPROACH FOR SAMPLING DU9, DU10 AND DU17

The triplicates sampling consisted of nested triplicates from DU9, DU10, and DU17. These DUs were selected for the following reasons:

- DU9, with a 2019 concentration of $0.68 \mathrm{mg} / \mathrm{kg}$, was selected to represent DUs with concentrations approximately one-half to two-thirds of the cleanup goal.
- DU10, with a 2019 concentration of $2.76 \mathrm{mg} / \mathrm{kg}$, was selected because its result was higher than all other DU results and establishing its internal variability is vital to a reliable weighted-95UCL.
- DU17, with a 2019 concentration of $0.92 \mathrm{mg} / \mathrm{kg}$, was selected to represent DUs with a concentration close to the action level.

Due to the heterogeneous nature of soil contamination, an estimate of a DU's actual concentration is the best outcome any soil sampling approach can achieve. ISM sampling approaches have been shown to provide the most reliable estimates of a DU's concentration.

In addition to providing contaminant concentrations, the field triplicate results measure the effectiveness of the ISM sample in capturing PCB contaminant variability within the DU. The field triplicate results inherently include any laboratory variability, since the field triplicates are each analyzed separately by the laboratory.

## SAMPLING METHODOLOGY

ISM was used to collect and analyze soil samples from DU9, DU10, and DU17 to provide a comprehensive and thorough evaluation of chemical concentrations in each DU. ISM involves collecting many small soil masses (called "increments") evenly across the DU and combining them to form a DU field sample.
ISM procedures used were consistent with the proposed sampling approach letter and procedures used previously at Richmond Field Station for PCB sampling.

- The corners and edges of each DU were marked with flags to identify the grids where triplicate increments were collected. The number of grids for DU9, DU10, and DU17 were 84, 85, and 82, respectively, based on the geometry of each DU. The triplicate increments for DU9 and DU10 were placed equidistant in a triangle formation within each grid. DU10 contained two grids which were not square, and therefore the triplicates were spaced equidistantly along a line within the two grids. Different from DU9 and DU10, DU17 is covered with 3 to 6 inches of 3-inch river rocks compacted in place in the majority of the area by heavy truck parking. The river rocks could not be easily removed by hand or shovel and required a small backhoe bucket to uncover the original ground surface. UC Berkeley conducted several test excavations on August 25, 2020 and concluded that three separate excavations per grid produced concerns regarding crosscontamination within triplicate locations, given the amount of movement of the river rocks within a small area. As a result, the triplicates were collected along a single, linear excavation per grid.

This method minimized potential cross contamination between the triplicates by minimizing the movement of river rocks within the DU. Grids and increment triplicate locations for each DU are included on Figures 2, 3, and 4.

- Increments were collected from the top 2 inches of the native surface with a disposable scoop or other disposable sampling equipment. In some areas, the native surface was the current surface cover; however, any gravel was removed prior to collecting the increment. Each increment was approximately 20 grams of soil and the target weight of each ISM sample was approximately 1.5 kilograms.
- Increments from each DU were placed directly into freezer-grade, 1-gallon resealable bags. Each bag was labeled and packed into an insulated cooler with ice packs placed on top of the samples within the cooler. The samples were transported under chain-of custody procedures to Agriculture \& Priority Pollutants Laboratories, Inc. (APPL) in Clovis, California.

Health and safety measures followed the Final Field Sampling Workplan, Appendix B, Health and Safety Plan, prepared by Tetra Tech, dated June 2, 2010. Protocols specific to COVID-19 protections were completed according to the Corporation Yard, Triplicates Sampling Approach, Revised, Attachment B, dated December 18, 2020.

## LABORATORY PROCESSING, SUBSAMPLING AND ANALYSES

Soil samples were processed according to APPL's internal ISM protocol, included in the Corporation Yard, Triplicates Sampling Approach, Revised, Attachment C, dated December 18, 2020. The protocol specifies that the sample is air-dried as necessary, then passed through a 10 -mesh sieve to remove nonsoil material (i.e., rocks or other objects larger than a 2 -millimeter [ mm ] in diameter). The sieved soil is then ground to the consistency of sifted flour and spread into a shallow layer in a pan to form a "slab cake" which is divided into 30 equal-sized grid cells. A 1 -gram increment is taken from each grid cell and the 30 increments are combined to form an analytical sub-sample weighing 30 grams. Each 30 -gram subsample was analyzed for PCBs by U.S. Environmental Protection Agency Method 8082 with 3540C Soxhlet extraction.

One of the field samples from each field triplicate set was sub-sampled and analyzed two additional times (for a total of three sub-sample analyses) to create the laboratory triplicate set. The second sub-sample is collected after the first by taking a second round of increments from the same 30 grid cells. A third round creates the third of the triplicate set. In this way, three independent representative sub-samples are collected.

## ANALYTICAL RESULTS

The field triplicate and laboratory triplicate results from the May 2021 triplicate sampling, as well as results from the previous September 2019 sampling event are provided below.

## Sample Results

| Sample ID | Sampling Result (mg/kg) |  |
| :---: | :---: | :---: |
|  | May 2021 Event | September 2019 Event |
| DU9 R1 | $\mathbf{0 . 4 6}$ | 0.68 |
| DU9 R2 | $\mathbf{0 . 3 9}$ | - |
| DU9 R3A | $\mathbf{0 . 4 7}$ | - |
| DU9 R3B | 0.44 | - |
| DU9 R3C | 0.47 | - |
| DU10 R1 | $\mathbf{1 . 7}$ | 2.76 |
| DU10 R2 | $\mathbf{1 . 7}$ | - |
| DU10 R3A | $\mathbf{1 . 9}$ | - |
| DU10 R3B | 2.4 | - |
| DU10 R3C | 1.9 | - |
| DU17 R1 | $\mathbf{0 . 8 9}$ | 0.92 |
| DU17 R2 | $\mathbf{0 . 8 4}$ | - |
| DU17 R3A | $\mathbf{0 . 8 2}$ | - |
| DU17 R3B | 0.72 | - |
| DU17 R3C | 0.69 | - |

Notes:
Bold results indicate samples that were input into the weighted-95UCL calculator tool for the RSD calculation.

Complete analytical results are included as Attachment A.

## TRIPICATE EVALUATION

QC of field collection procedures in the form of three independent field samples (i.e., field triplicates) assesses the ability of an ISM sample to reliably estimate the DU concentration and quantify inherent soil and contaminant heterogeneity. The field and laboratory sub-sampling triplicates create an ISM "nested triplicate" set from which the amount of variability can be calculated as the RSD to quantify both field heterogeneity and laboratory procedures. These two RSD values measure how much field heterogeneity and laboratory measurement variability contribute to overall data variability. Attachment B provides supporting information regarding the use of RSDs and pooled RSDs in ISM projects.

## Laboratory Triplicate Evaluation

The primary purpose of the laboratory triplicate set is to evaluate the effectiveness of the processing and subsampling protocols for site-specific contaminants and the soil matrix. If the procedures are effective, the three sub-samples should provide results that are close numerically. The closer the agreement among the results, the lower the data variability and RSD for the triplicate set. Variability in the analytical processes of sample extraction, extract cleanup and instrumental measurement is necessarily included in the sub-sampling variability.

Laboratory triplicates were evaluated quantitatively and qualitatively to determine overall data usability. Quantitative evaluation involved calculating an RSD on the three laboratory replicate results as a measure of variability. Qualitative evaluation involved assessing whether laboratory triplicate concentrations agree generally in their concentration ranges (low, moderate, or elevated) and with respect to whether the action level is exceeded. The qualitative evaluation also integrated the concentration variability, where the lower the concentration variability, the higher the data usability. Low concentration variability also offsets any elevated RSDs associated with low detections. The calculated RSDs were considered acceptable, concentration variabilities considered low, and the triplicate concentrations agreed with their concentration ranges. The triplicate results data usability were considered acceptable. The laboratory triplicate evaluation is presented below.

Laboratory Triplicate Evaluation

| Laboratory <br> Triplicate ID | Sample Result (mg/kg) | $\begin{gathered} \hline \text { Quantitative } \\ \text { Evaluation } \\ \text { RSD } \\ \hline \end{gathered}$ | Concentration Variability (mg/kg) | Qualitative <br> Evaluation | Data Usability |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DU9 R3A | 0.47 | 0.038 | 0.03 | Good | Acceptable |
| DU9 R3B | 0.44 |  |  |  |  |
| DU9 R3C | 0.47 |  |  |  |  |
| DU10 R3A | 1.9 | 0.140 | 0.5 | Good | Acceptable |
| DU10 R3B | 2.4 |  |  |  |  |
| DU10 R3C | 1.9 |  |  |  |  |
| DU17 R3A | 0.82 | 0.092 | 0.13 | Good | Acceptable |
| DU17 R3B | 0.72 |  |  |  |  |
| DU17 R3C | 0.69 |  |  |  |  |

## Field Triplicate Evaluation

The field triplicate set provides statistics that, in conjunction with the statistics from the laboratory triplicate set, allow their respective contributions to overall data variability to be determined. Understanding data variability is important because high data variability is detrimental to efficient site investigations and cleanup.

Only one of the laboratory triplicate results was used as the concentration for the parent field sample. By convention and to parallel the data from the other two field triplicate samples, the project used the first sub-sample result reported by the laboratory as the concentration for the parent field sample. The convention of using the first laboratory triplicate result avoided any selection bias to obtain the lowest or highest RSD for the field triplicate set. If the evaluation of the laboratory triplicates had found the data unusable, an alternate method would have been proposed to select the appropriate laboratory triplicate result.

The same evaluation criteria were applied to the field triplicate results as the laboratory triplicate results. The calculated RSDs were considered acceptable, concentration variabilities were considered low, and the triplicate concentrations agreed with their concentration ranges. The triplicate results data usability were considered acceptable. The field triplicate evaluation is presented in below.

## Field Triplicate Evaluation

| Field <br> Triplicate ID | Sample Result (mg/kg) | Quantitative Evaluation RSD | Concentration Variability (mg/kg) | Qualitative Evaluation | Data Usability |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DU9 R1 | 0.46 | 0.099 | 0.06 | Good | Acceptable |
| DU9 R2 | 0.39 |  |  |  |  |
| DU9 R3A | 0.47 |  |  |  |  |
| DU10 R1 | 1.7 | 0.065 | 0.2 | Good | Acceptable |
| DU10 R2 | 1.7 |  |  |  |  |
| DU10 R3A | 1.9 |  |  |  |  |
| DU17 R1 | 0.89 | 0.042 | 0.07 | Good | Acceptable |
| DU17 R2 | 0.84 |  |  |  |  |
| DU17 R3A | 0.82 |  |  |  |  |

## WEIGHTED-95UCL CALCULATION

The ultimate objective of the investigations is to determine the mean and 95UCL concentrations for the large area composed of DU9 through DU17. A weighted mean accounts for the different sizes of the DUs with more influence on the mean concentration from larger DUs than smaller DUs, and better represents the true mean over the entire area.

While calculation of a weighted-mean uses a simple equation, the weighted-95UCL calculation is complex and requires a specially designed spreadsheet called the "Combining DUs Calculator." The calculator tool was first designed by Philip Goodrum, Ph.D., a statistician and toxicologist with GSI Environmental, and a contributor to the Interstate Technology and Regulatory Council for 2012 and 2020 ISM guidances. The calculator tool computes the mean and variability from the raw sample data. The original calculator tool was structured to accept ISM field replicate data as the only inputs. The calculator tool was modified in 2020 by Deana Crumbling of Tetra Tech in collaboration with Goodrum to accept DU means, as determined from one or more ISM samples, and RSDs as the inputs. A white paper presenting the calculator tool is presented as Attachment C.

An estimate of the within-DU (i.e., internal) variability is required for each of the DUs to compute a weighted UCL. Field triplicates provide that measure of within-DU variability. Results of the field triplicates (termed "triplicate DUs") can be applied to DUs having only a single ISM sample (termed "singlet DUs"). As a starting point, it is reasonable to assume that within-DU variability (as measured by an RSD) will be similar for DUs that are equivalent from a conceptual site model (CSM) perspective. "CSM-equivalent" DUs have the same contaminant release and transport mechanisms, are in reasonably close proximity, and can be expected to have similar concentrations relative to the action level.

DUs suspected to have a different CSMs and significantly different concentrations should be segregated into their own group or isolated as a lone DU. Each DU group needs its own subset of triplicate DUs; a lone DU will be its own triplicate DU. Note that not all DUs with field triplicates require laboratory triplicates. How well the assumptions behind the CSM-equivalent DU group(s) hypotheses are not known until there is at least one ISM sample from each DU. If the results do not support the hypotheses, DUs may need to be regrouped by concentration and additional triplicate DUs identified for a follow-on
sampling effort. As discussed earlier, this scenario was encountered at DU11 in the September 2019 data set, which prompted this supplemental sampling effort.

## Application of RSDs in the Combining DUs Calculator Tool

The RSDs for each field triplicate set are evaluated. There are two primary options for how RSDs can be applied toward the goal of a weighted UCL which depend on sample results:

Option 1: If only a single RSD is applicable to a particular group of CSM-equivalent DUs, that RSD value can be used to represent the variability of each singlet DU in the group. The Combining DUs Calculator would be populated with each DU's ISM result (for singlet DUs), or the average of a set of field samples (in the case of a triplicate DU), along with that RSD value.

Option 2: If two or more RSDs are applicable to a particular DU group, the RSDs can be pooled to obtain a single "averaged" RSD value to be applied in the calculator tool to singlet DUs. The Excel formula for pooling two RSDs is $<$ Pooled RSD $=\operatorname{sqrt}$ (sumsq(RSD\#1, RSD\#2)/2) $>$. Three RSDs are pooled by the equation $<$ Pooled RSD $=$ sqrt(sumsq(RSD\#1, RSD\#2, RSD\#3)/3)>. DU concentrations are addressed as in Option 1. The pooled RSD value would be entered for the applicable singlet DUs. In Option 2, actual RSDs are normally entered into the calculator tool for triplicate DUs. The average concentration from DUs with field triplicates are applied as the singlet concentrations.

Field triplicate results from DU9, DU10, and DU17 support the use of Option 2, with the averaged RSD of 0.73 applied to the singlet DUs, and actual RSDs applied to DU9, DU10, and DU17. For DU11, the average concentration from the initial field triplicates was applied, consistent with DU9, DU10, and DU17.

## Calculator Tool Output Recommendations

The calculator tool provides a detailed statistical evaluation based on the data inputs presented above. The calculator tool evaluates the different UCL statistical methodologies and provides a recommended UCL. Based on the results and actual or borrowed RSDs, the calculator tool provides two types of overall weighted-95UCL: the Student's-t UCL is used for normal data distributions, and the Chebyshev UCL is used for nonnormal distributions. The calculator tool also recommends which of the two UCL options to use.

## RESULTS AND CONCLUSIONS

The tool concludes that either the Student's-t or Chebychev 95UCL may be appropriate, and recommends using the Student's-t. The resulting weighted-95UCL is $0.62 \mathrm{mg} / \mathrm{kg}$, lower than the $1 \mathrm{mg} / \mathrm{kg}$ cleanup goal for the Corporation Yard, supporting no further evaluation or soil cleanup activities are necessary for the DU9 through DU17areas.

The data collected during this investigation will ultimately be presented with the comprehensive data following completion of all Corporation Yard removal action activities.

If you have any questions or comments regarding this submittal, please call me at (415) 497-9060 or Greg Haet at (510) 812-1541.

Sincerely,


Attachments: Figure 1: Decision Units 9 Through 17
Figure 2: Decision Unit 9 Increment Locations
Figure 3: Decision Unit 10 Increment Locations


Figure 4: Decision Unit 17 Increment Locations
Table 1: Weighted-95UCL Results
Attachment A: Analytical Results
Attachment B: Incremental Sampling Methodology, RSD Calculations and Uses
Attachment C: Combining DUs Calculator Tool White Paper
cc: Greg Haet, UC Berkeley EH\&S
Bill Marsh, Edgcomb Law Group





## TABLE 1

WEIGHTED-95UCL RESULTS

Property/Sample ID: Corporation Yard, Revised Triplicates
Analyte: Total PCBS
Analyte units: milligrams per kilogram
Approach Uses DU singlet data for DUs 12 to 16 and applies the pooled RSD to them. The pooled RSD is derived from the new triplicate sampling/analysis of DU09, DU10, DU 11, and DU17


|  |  |  |  | ata Entry |  | This column |  | Small DU | Actual or | Calculated |  |  |  |  |  |  | 95\% |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Row \# | IDs <br> (of the small DUs) | DU Area ( sq ft ) | Trip or Singlet Concentration | Actual or Borrowed RSD | -- | does not change | Weight for DU | Concentration | predicted replicate SD | SD of increments | calc'd CV <br> for the DU | Adj Factor | adj'd SD of increments | adj'd CV for DU | SE of DU | Student's-t | Chebychev | $\begin{array}{\|c\|} \hline \mathrm{CV} \text { of } \\ \text { Increments } \end{array}$ | 95\% UCL |
| 1 | DU09 | 3291 | 0.44 | 0.099 | -- | 3 | 0.0901 | 0.440 | 0.044 | 0.38 | 0.86 | 1.15 | 0.43 | 0.99 | 0.025 | 0.51 | 0.55 | Low | 0.513 |
| 2 | DU10 | 3608 | 1.77 | 0.065 | -- | 3 | 0.0987 | 1.767 | 0.115 | 1.00 | 0.57 | 1.13 | 1.13 | 0.64 | 0.067 | 1.96 | 2.06 | Low | 1.961 |
| 3 | DU11 | 2946 | 0.077 | 0.073 | -- | 3 | 0.0806 | 0.077 | 0.006 | 0.05 | 0.63 | 1.14 | 0.06 | 0.72 | 0.003 | 0.09 | 0.09 | Low | 0.086 |
| 4 | DU12 | 4363 | 0.61 | 0.073 | -- | 3 | 0.1194 | 0.610 | 0.045 | 0.39 | 0.63 | 1.14 | 0.44 | 0.72 | 0.026 | 0.69 | 0.72 | Low | 0.685 |
| 5 | DU13 | 3313 | 0.06 | 0.073 | -- | 3 | 0.0907 | 0.060 | 0.004 | 0.04 | 0.63 | 1.14 | 0.04 | 0.72 | 0.003 | 0.07 | 0.07 | Low | 0.067 |
| 6 | DU14 | 4697 | 0.2 | 0.073 | -- | 3 | 0.1286 | 0.200 | 0.015 | 0.13 | 0.63 | 1.14 | 0.14 | 0.72 | 0.008 | 0.22 | 0.24 | Low | 0.225 |
| 7 | DU15 | 5208 | 0.47 | 0.073 | -- | 3 | 0.1425 | 0.470 | 0.034 | 0.30 | 0.63 | 1.14 | 0.34 | 0.72 | 0.020 | 0.53 | 0.56 | Low | 0.528 |
| 8 | DU16 | 5135 | 0.84 | 0.073 | -- | 3 | 0.1405 | 0.840 | 0.061 | 0.53 | 0.63 | 1.14 | 0.60 | 0.72 | 0.035 | 0.94 | 0.99 | Low | 0.943 |
| 9 | DU17 | 3977 | 0.85 | 0.042 | -- | 3 | 0.1088 | 0.850 | 0.036 | 0.31 | 0.37 | 1.13 | 0.35 | 0.41 | 0.021 | 0.91 | 0.94 | Low | 0.911 |
| Combine the Weighted DUs: |  | 36538 | -- | -- | -- | 3 | 1.00 | 0.6018 | 0.0171 | 0.1478 | 0.2455 | NA | 0.17 | 0.28 | 0.010 | 0.62 | 0.64 | Low | 0.62 |


| Degrees of freedom by Welch-Satterthwaite approximation | 6.98 |
| :--- | :--- |

Notes
DU Decision unit
SD Standard deviation
SE Standard errror
UCL Upper confidence limit
CV Coefficient of variation (equivalent to "RSD")
RSD Relative standard deviation
Calculation assumes all replicates have the same number of increments

## ATTACHMENT A

ANALYTICAL RESULTS

## Data Validation Package

July 20, 2021

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, California 94612
Attn: Jason Brodersen

Title: Report of Data: Case 96283
Project: 103S582307.07
Dear Mr. Brodersen:
Fifteen soil samples were received May 22, 2021. Written results for the requested analysis are being provided on this July 20, 2021.

Results in this report apply to the samples analyzed in accordance with the chain of custody document. This analytical report must be reproduced in its entirety.

If you have any questions or require further information, please contact your APPL Project Manager, Gregory Salata, gsalata@applinc.com, at your convenience. Thank you for choosing APPL, Inc.

I certify that this data package is in compliance with the terms and conditions of the contract, both technically and for completeness, for other than the conditions detailed above. These test results meet all requirements of NELAC and DoD QSM. Release of the hard copy has been authorized by the Laboratory Manager or his designee, as verified by the following signature.


Loren Portwood, Laboratory Director APPL, Inc.

LP/gs
Enclosure
cc: File
Data Validation Package for103S582307.02ARF 96283TABLE OF CONTENTS
LABORATORY NAME: APPL, Inc.
Case Narrative ..... 3
Sample Management Records ..... 7
Sample Results ..... 13
QC Forms ..... 30
Method 8082A Calibration Data ..... 37
Method 8082A Raw Data ..... 54

## CASE NARRATIVE

# Case Narrative 

ARF: 96283
Project: 103S582307.02

## Sample Receipt Information:

Fifteen soil samples were received May 22,2021 at $2.0^{\circ} \mathrm{C}$. The samples were assigned Analytical Request Form (ARF) 96283. The sample numbers and requested analyses were compared to the chain of custody and e-mail correspondence. No exceptions were encountered.

## Sample Preparation and Analysis:

For the EPA 8082A analysis, the samples were purged according to EPA method 3540.
Percent moisture was determined using ISM02.2, Exhibit D, section 10.0.
Only the portion of the injection log relative to these samples is included. A full sequence $\log$ is available upon request. Measurement uncertainty can be reported upon request.

## Exceptions, Abnormalities and Deviations:

EPA 8082A: None.

| SDG | Received | Client ID | APPL ID | Collected DateTime | Matrix | Method | Method Description | Prep DateTime | Analysis DateTime |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 96283 | 05/22/21 | CY-DU17-R1 | BA33121 | 05/20/21 10:00:00 AM | SOIL | EPA 8082A | EPA 8082A ISM SOIL | 06/10/21 4:00:00 PM | 07/07/21 5:59:00 PM |
| 96283 | 05/22/21 | CY-DU17-R1 | BA33121 | 05/20/21 10:00:00 AM | SOIL | CLP MOIST | Moisture | 05/25/21 2:36:00 PM | 05/26/21 11:05:00 AM |
| 96283 | 05/22/21 | CY-DU17-R2 | BA33122 | 05/20/21 10:00:00 AM | SOIL | EPA 8082A | EPA 8082A ISM SOIL | 06/11/21 2:15:00 PM | 07/07/21 6:16:00 PM |
| 96283 | 05/22/21 | CY-DU17-R2 | BA33122 | 05/20/21 10:00:00 AM | SOIL | CLP MOIST | Moisture | 05/25/21 2:36:00 PM | 05/26/21 11:05:00 AM |
| 96283 | 05/22/21 | CY-DU17-R3A | BA33123 | 05/20/21 10:00:00 AM | SOIL | EPA 8082A | EPA 8082A ISM SOIL | 06/10/21 4:00:00 PM | 07/07/21 6:33:00 PM |
| 96283 | 05/22/21 | CY-DU17-R3A | BA33123 | 05/20/21 10:00:00 AM | SOIL | CLP MOIST | Moisture | 05/25/21 2:36:00 PM | 05/26/21 11:05:00 AM |
| 96283 | 05/22/21 | CY-DU10-R1 | BA33124 | 05/20/21 3:00:00 PM | S | EPA 8082A | EPA 8082A ISM SOIL | 06/10/21 4:00:00 PM | 07/07/21 6:49:00 PM |
| 96283 | 05/22/21 | CY | BA33124 | 05/20/21 3:00 | S | CLP MOIST | M | 05/25/21 2:36:00 PM | 05/26/21 11:05:00 AM |
| 96283 | 05/22/21 | CY-DU10-R2 | BA33125 | 05/20/21 3:00:00 PM | SOIL | EPA 8082A | EPA 8082A ISM SOIL | 06/11/21 2:15:00 PM | 07/07/21 7:06:00 PM |
| 96283 | 05/22/21 | CY-DU10-R2 | BA33125 | 05/20/21 3:00:00 PM | SOIL | CLP MOIST | Moisture | 05/25/21 2:36:00 PM | 05/26/21 11:05:00 AM |
| 96283 | 05/22/21 | CY-DU10-R3A | BA33126 | 05/20/21 3:00:00 PM | SOIL | EPA 8082A | EPA 8082A ISM SOIL | 06/10/21 4:00:00 PM | 07/07/21 7:23:00 PM |
| 96283 | 05/22/21 | CY-DU10-R3A | BA33126 | 05/20/21 3:00:00 PM | SOIL | CLP MOIST | Moisture | 05/25/21 2:36:00 PM | 05/26/21 11:05:00 AM |
| 96283 | 05/22/21 | CY-DU9-R1 | BA33127 | 05/21/21 11:15:00 AM | SOIL | EPA 8082A | EPA 8082 | 06/10/21 4:00:00 PM | 07/07/21 7:40:00 PM |
| 96283 | 05/22/21 | CY-DU9-R1 | BA33127 | 05/21/21 11:15:00 AM | SOIL | CLP MOIST | Moisture | 05/25/21 2:36:00 PM | 05/26/21 11:05:00 AM |
| 96283 | 05/22/21 | CY-DU9-R2 | BA33128 | 05/21/21 11:15:00 AM | SOIL | EPA 8082A | EPA 8082A ISM SOIL | 06/10/21 4:00:00 PM | 07/07/21 7:57:00 PM |
| 96283 | 05/22/21 | CY-DU9-R2 | BA33128 | 05/21/21 11:15:00 AM | SOIL | CLP MOIST | Moisture | 05/25/21 3:05:00 PM | 05/26/21 11:13:00 AM |
| 96283 | 05/22/21 | CY-DU9-R3A | BA33129 | 05/21/21 11:15:00 AM | SOIL | EPA 8082A | EPA 8082A ISM SOIL | 06/10/21 4:00:00 PM | 07/07/21 8:14:00 PM |
| 96283 | 05/22/21 | CY-DU9-R3A | BA33129 | 05/21/21 11:15:00 AM | SOIL | CLP MOIST | Moisture | 05/25/21 3:05:00 PM | 05/26/21 11:13:00 AM |
| 96283 | 05/22/21 | CY-DU17-R3B | BA33704 | 05/20/21 10:00:00 AM | SOIL | EPA 8082A | EPA 8082A ISM SOIL | 06/10/21 4:00:00 PM | 07/07/21 9:38:00 PM |
| 96283 | 05/22/21 | CY-DU17-R3B | BA33704 | 05/20/21 10:00:00 AM | SOIL | CLP MOIST | Moisture | 05/25/21 3:39:00 PM | 05/26/21 3:39:00 PM |
| 96283 | 05/22/21 | CY-DU17-R3C | BA33705 | 05/20/21 10:00:00 AM | SOIL | EPA 8082A | EPA 8082A ISM SOIL | 06/10/21 4:00:00 PM | 07/07/21 9:55:00 PM |
| 96283 | 05/22/21 | CY-DU17-R3C | BA33705 | 05/20/21 10:00:00 AM | SOIL | CLP MOIST | Moisture | 05/25/21 3:39:00 PM | 05/26/21 3:39:00 PM |
| 96283 | 05/22/21 | CY-DU10-R3B | BA33706 | 05/20/21 3:00:00 PM | SOIL | EPA 8082A | EPA 8082A ISM SOIL | 06/11/21 2:15:00 PM | 07/07/21 10:12:00 PM |
| 96283 | 05/22/21 | CY-DU10-R3B | BA33706 | 05/20/21 3:00:00 PM | SOIL | CLP MOIST | Moisture | 05/25/21 3:39:00 PM | 05/26/21 3:39:00 PM |
| 96283 | 05/22/21 | CY-DU10-R3C | BA33707 | 05/20/21 3:00:00 PM | SOIL | EPA 8082A | EPA 8082A ISM SOIL | 06/10/21 4:00:00 PM | 07/07/21 10:28:00 PM |
| 96283 | 05/22/21 | CY-DU10-R3C | BA33707 | 05/20/21 3:00:00 PM | SOIL | CLP MOIST | Moisture | 05/25/21 3:39:00 PM | 05/26/21 3:39:00 PM |
| 96283 | 05/22/21 | CY-DU9-R3B | BA33708 | 05/21/21 11:15:00 AM | SOIL | EPA 8082A | EPA 8082A ISM SOIL | 06/10/21 4:00:00 PM | 07/07/21 10:45:00 PM |
| 96283 | 05/22/21 | CY-DU9-R3B | BA33708 | 05/21/21 11:15:00 AM | SOIL | CLP MOIST | Moisture | 05/25/21 3:39:00 PM | 05/26/21 3:39:00 PM |
| 96283 | 05/22/21 | CY-DU9-R3C | BA33709 | 05/21/21 11:15:00 AM | SOIL | EPA 8082A | EPA 8082A ISM SOIL | 06/10/21 4:00:00 PM | 07/07/21 11:02:00 PM |
| 96283 | 05/22/21 | CY-DU9-R3C | BA33709 | 05/21/21 11:15:00 AM | SOIL | CLP MOIST | Moisture | 05/25/21 3:39:00 PM | 05/26/21 3:39:00 PM |

Page 1
5 of 60

# APPL Inc. <br> Abbreviations and Flags 

| FLAG | DESCRIPTION |
| :---: | :---: |
| \# | Recovery or RPD outside control limits |
| * | Recovery or RPD outside control limits |
| B | Analyte detected in associated method blank |
| C1 | Reason for correction: wrote incorrect response |
| C2 | Reason for correction: calculated incorrectly |
| C3 | Reason for correction: needs to be rechecked |
| C4 | Reason for correction: data not usable |
| DO | Diluted out |
| E | Exceeds linear range |
| F | Estimated value |
| G1 | Includes a wide range of hydrocarbons which does not match our gasoline standard |
| G10 | Includes a match to hydrocarbon profiles within the range of mineral spirits |
| G11 | Includes a match to hydrocarbon profiles within the range of JP-4 |
| G12 | Pattern does not match the gasoline standard; the carbon range for this sample is consistent with JP8 |
| G13 | Closely resembles the hydrocarbon profile of aviation gasoline |
| G14 | Analyte concentration may be biased due to carry over |
| G2 | Closely resembles the boiling point hydrocarbon profile consistent with weathered gasoline |
| G3 | Includes higher boiling hydrocarbons |
| G4 | Includes dominant peak(s) not indicative of petroleum hydrocarbons |
| G5 | Is mainly dominant peak(s) not indicative of petroleum hydrocarbons |
| G6 | Contains recognizable contaminant peak(s) which has been removed from quantitation |
| G7 | Is mainly a match to hydrocarbons within the range of gasoline |
| G8 | Closely resembles the boiling point hydrocarbon profile consistent with weathered gasoline |
| G9 | Includes hydrocarbons within the range of kerosene |
| J | Estimated value |
| M | Matrix effect |
| MI1 | Manual integration: integration does not follow baseline |
| MI2 | Manual integration: non-target peak interference |
| MI3 | Manual integration: to split a peak that was integrated as one peak by the computer. |
| MI4 | Manual integration: to integrate a split peak |
| MI5 | Manual integration: the whole peak or part of the peak was not integrated |
| MI6 | Manual integration: computer integrated wrong peak |
| MI7 | Manual integration: other - (See case narrative) |
| MDL | Method detection limit |
| ND | Not detected |
| NT | Non-target |
| Q | Acceptance criteria not met |
| T1I | Includes wide range of hydrocarbons not indicative of diesel |
| T1 M | Is mainly wide range of hydrocarbons not necessarily indicative of diesel |
| T2 I | Includes lower boiling hydrocarbons, e.g. mineral spirits, kerosene, stoddard solvent, white gas |
| T2 M | Is mainly lower boiling hydrocarbons, e.g. mineral spirits, kerosene, stoddard solvent, white gas |
| T3 I | Includes higher boiling hydrocarbons, e.g. asphaltene, waste oil, motor oil, or weathered diesel fuel |
| T3 M | Is mainly higher boiling hydrocarbons, e.g. asphaltene, waste oil, motor oil, or weathered diesel fuel |
| T4 I | Includes dominant peak(s) not indicative of hydrocarbons |
| T4 M | Is mainly dominant peak(s) not indicative of hydrocarbons |
| T5 | Contains recognizable contaminant peak(s) which has been removed from quantitation |
| T6 | Is mainly a match to hydrocarbons within range of diesel fuel |
| T7 | Closely resembles the boiling point hydrocarbon profile consistent with diesel fuel |
| T8 | Includes a match to hydrocarbon profiles within range of diesel and kerosene fuel |
| T9 I | Includes non-diesel hydrocarbons within boiling point range of diesel fuel |
| T9 M | Is mainly non-diesel hydrocarbons within boiling point range of diesel fuel |
| U | Not detected |
| Y | Percent difference between primary and confirmation column $>40 \%$ |

# SAMPLE MANAGEMENT RECORDS 

CHAIN OF CUSTODY,
ARF, CRF, AND CLIENT COMMUNICATION


## Comments:

AN: 'U' Prints MDL report, DVP3.
Login to Jason.Brodersen@tetratech.com
Dry, sieve, and grind prior to PCB MIS (30 subsamples)
Prep using Soxhlet
FR: PDF to Jason.Broderen@tetratech.com
EDD: Excel to Jason.Brodersen@tetratech.com


| 7. CY-DU9-R1 | $\begin{aligned} & \hline \text { BA33127S } \quad 05 / 21 / 21 \quad 11: 15 \\ & \|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\mid \end{aligned}$ | \$82ADOD51SMI, \$PCBS, MOIST |
| :---: | :---: | :---: |
| 8. CY-DU9-R2 | $\begin{aligned} & \text { BA3128S } \quad 05 / 21 / 21 \quad 11: 15 \\ & \|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\mid \end{aligned}$ | \$82ADOD51SMI, \$PCBS, MOIST |
| 9. CY-DU9-R3A | $\begin{aligned} & \text { BA33129S } 05 / 21 / 21 \quad 11: 15 \\ & \|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\mid \end{aligned}$ | \$82ADOD51SMI, \$PCBS, MOIST |
| 10. CY-DU17-R3B | $\begin{aligned} & \text { BA33704S } \quad \text { 05/20021 } 10: 00 \\ & \|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\mid \end{aligned}$ | \$82ADOD51SMI, \$PCBS, MOIST |
| 11. CY-DU17-R3C |  | \$82ADOD51SMI, \$PCBS, MOIST |
| 12. CY-DU10-R3B | BA33706S $05 / 20 / 21 \quad 15: 00$ $\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\mid$ | \$82ADOD51SMI, \$PCBS, MOIST |
| 13. CY-DU10-R3C | $\begin{aligned} & \text { BA3707S } 05 / 20 / 21 \text { 15:00 } \\ & \text { \|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\| } \end{aligned}$ | \$82ADOD51SMI, \$PCBS, MOIST |
| 14. CY-DU9-R3B | BA3308S $05 / 21 / 21 \quad 11: 15$ $\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\mid$ | \$82ADOD51SMI, \$PCBS, MOIST |
| 15. CY-DU9-R3C | $\begin{aligned} & \text { BA3379S } 05 / 21 / 21 \quad 11: 15 \\ & \|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\|\mid \end{aligned}$ | \$82ADOD51SMI, \$PCBS, MOIST |


| Sample | Container Type | Count | p |
| :---: | :---: | :---: | :---: |
| BA33121 | 2332 oz Jar | 1 | NA |
| BA33122 | 23 320z Jar | 1 | NA |
| BA33123 | 2332 oz Jar | 1 | NA |
| BA33124 | 2332 oz Jar | 1 | NA |
| BA33125 | 2332 oz Jar | 1 | NA |
| BA33126 | ${ }^{23} 32 \mathrm{oz} \mathrm{Jar}$ | 1 | NA |
| BA33127 | 2332 oz Jar | 1 | NA |
| BA33128 | $23320 z \mathrm{Jar}$ | 1 | NA |
| BA33129 | 2332 oz Jar | 1 | NA |
| BA33704 | 2332 oz Jar | 1 | NA |
| BA33705 | ${ }^{23} 32 \mathrm{oz} \mathrm{Jar}$ | 1 | NA |
| BA33706 | 2332 oz Jar | 1 | NA |
| BA33707 | 2332 zz Jar | 1 | NA |
| BA33708 | 2332 oz Jar | 1 | NA |
| BA33709 | 2332 oz Jar | 1 | NA |

Sample Container Type
Count p


1) Project: RFS Corp Yard $103 S 582307.02$

Date Received:
2) Coolers:

Number of Coolers:
1
3) YES Were custody seals present and intact?
How many? $\quad 1$
1 Name/Date on seal? SEE BELOW
4) YES Was there a shipping slip? Carrier name: FEDEX
Type of packing in cooler:

| X bubble wrap | popcorn |
| :--- | :--- |
| $\overline{\mathrm{X}}$ wet ice | foam |
| dry ice | no ice |

$\qquad$ gel ice
6)

YES Were cooler temperatures acceptable?
7)
8)

Serial number of calibrated thermometer used: R5 CF:+0.0${ }^{\circ} \mathrm{C}$
Cooler temp(s): In ${ }^{\circ} \mathrm{C}$. Thermometer Temp / Corrected Temp
1: 2.0/2.0
7: $\qquad$
2 :
8: $\qquad$
3:
4: $\qquad$
5: 11: $\qquad$
6: $\qquad$

Chain of custody:
9) YES Was a chain of custody received?
10) YES Were the custody papers complete/signed in the appropriate places?

## Sample Labels:

11) YES Were all sample labels complete (sample ID, date/time of sampling, etc.)?
12) YES Did all container labels agree with custody papers?

## Sample Containers:

13) YES Were all containers sealed in separate bags?
14) YES Did all containers arrive in good condition:(unbroken, no leakage, no cracked/broken lids)?
15) YES Were correct containers and preservatives used for the tests indicated?
16) YES Was a sufficient amount of sample sent for tests indicated?
17) NA Were bubbles present in volatile samples?

If yes, the following were received with air bubbles:
Larger than a pea:
Smaller than a pea:
Preservation Hold time:
18) Yes Was a sufficient amount of holding time remaining to analyze the samples?
19) NA Was the pH taken of all non-VOA preserved samples and written on the sample container?
20) NA Was the pH of acid preserved non-VOA samples < 2?
21) NA Was the pH of the "basic" preserved samples for Cyanide $>12$, Sulfide $>9$, Hexchrom $>9$ ?
22) NO Were unpreserved VOA Vials received for VOA Dept analysis?
23) NA If "yes", are the unpreserved VOA vials noted in the ADD TEST FIELD on the ARF? pH strip lot number:


## SAMPLE RESULTS

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU17-R1
Sample Collection Date: 05/20/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707011
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU17-R2
Sample Collection Date: 05/20/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707012
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU17-R3A
Sample Collection Date: 05/20/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707013
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU10-R1
Sample Collection Date: 05/20/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707014
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU10-R2
Sample Collection Date: 05/20/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707015
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU10-R3A
Sample Collection Date: 05/20/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707016
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU9-R1
Sample Collection Date: 05/21/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707017
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU9-R2
Sample Collection Date: 05/21/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707018
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU9-R3A
Sample Collection Date: 05/21/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707019
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU17-R3B
Sample Collection Date: 05/20/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707024
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU17-R3C
Sample Collection Date: 05/20/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707025
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU10-R3B
Sample Collection Date: 05/20/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707026
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU10-R3C
Sample Collection Date: 05/20/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707027
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU9-R3B
Sample Collection Date: 05/21/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707028
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

## EPA 8082A ISM SOIL

Tetra Tech, Inc.
1999 Harrison St., Suite 500
Oakland, CA 94612
Attn: Jason Brodersen
Project: RFS Corp Yard 103S582307.02
Sample ID: CY-DU9-R3C
Sample Collection Date: 05/21/21

APPL Inc.
908 North Temperance Avenue
Clovis, CA 93611

| Method | Analyte | Result | RL | MDL | Units | Extraction <br> Date |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |  |  |
| Analysis |  |  |  |  |  |  |
| Date |  |  |  |  |  |  |

Quant Method: PCB0629.M
Run \#: 0707029
Instrument: Lucy
Sequence: 210707
Dilution Factor: 1
Initials: BTI
Printed: 07/20/21 11:02:00 AM APPL-F1-SC-NoMC-REG MDLs

| Wetlab Results | Tetra Tech, Inc. |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  | 1999 Harrison St., Suite 500 |  |  |  |
| ARF: 96283 | Oakland, CA 94612 |  |  |  |
| APPL Inc. |  |  |  |  |
| 908 North Temperance Avenue Clovis, CA 93611 | Attn: Jason Brodersen |  |  |  |
|  |  |  |  |  |
| Method Analyte | Result | RL MDL Units | Prep Date Analysis Date |  |
| APPL ID: BA33121 -Client Sample ID: CY-DU17-R1 | -Sample Collection Date: 05/20/21 |  | Project: RFS Corp Yard $103 \mathrm{S582}$ |  |
| CLP MOIST MOISTURE | 7.8 | 2.0 | \% 05/25/21 | 05/26/21 |
| APPL ID: BA33122 -Client Sample ID: CY-DU17-R2 |  | -Sample Collection Date: 05/20/21 | Project: RFS Corp Ya | ard 103S582 |
| CLP MOIST MOISTURE | 7.1 | 2.0 | \% 05/25/21 | 05/26/21 |
| APPL ID: BA33123 -Client Sample ID: CY-DU17-R3 |  | -Sample Collection Date: 05/20/21 | Project: RFS Corp Ya | ard 103 S 582 |
| CLP MOIST MOISTURE | 6.4 | 2.0 | \% 05/25/21 | 05/26/21 |
| APPL ID: BA33124 -Client Sample ID: CY-DU10-R1 |  | -Sample Collection Date: 05/20/21 | Project: RFS Corp Ya | ard 103S582 |
| CLP MOIST MOISTURE | 3.5 | 2.0 | \% 05/25/21 | 05/26/21 |
| APPL ID: BA33125 -Client Sample ID: CY-DU10-R2 |  | -Sample Collection Date: 05/20/21 | Project: RFS Corp Ya | ard 103 S 582 |
| CLP MOIST MOISTURE | 3.4 | 2.0 | \% 05/25/21 | 05/26/21 |
| APPL ID: BA33126 -Client Sample ID: CY-DU10-R3 |  | -Sample Collection Date: 05/20/21 | Project: RFS Corp Ya | ard 103S582 |
| CLP MOIST MOISTURE | 3.1 | 2.0 | \% 05/25/21 | 05/26/21 |
| APPL ID: BA33127 -Client Sample ID: CY-DU9-R1 |  | -Sample Collection Date: 05/21/21 | Project: RFS Corp Ya | ard 103 S 582 |
| CLP MOIST MOISTURE | 4.6 | 2.0 | \% 05/25/21 | 05/26/21 |
| APPL ID: BA33128 -Client Sample ID: CY-DU9-R2 |  | -Sample Collection Date: 05/21/21 | Project: RFS Corp Ya | ard 103 S 582 |
| CLP MOIST MOISTURE | 6.6 | 2.0 | \% 05/25/21 | 05/26/21 |
| APPL ID: BA33129 -Client Sample ID: CY-DU9-R3 |  | -Sample Collection Date: 05/21/21 | Project: RFS Corp Ya | ard 103S582 |
| CLP MOIST MOISTURE | 5.4 | 2.0 | \% 05/25/21 | 05/26/21 |

## ATTACHMENT B

## INCREMENTAL SAMPLING METHODOLOGY RSD CALCULATIONS AND USES

# Incremental Sampling Methodology RSD Calculations and Uses 

### 1.0 Introduction to the RSD

A relative standard deviation (RSD) is a measure of data variability, which serves as a reverse measure of precision. In other words, "precision" (how closely replicate values agree) is measured by determining how much disagreement there is. The less disagreement, the better the precision. Therefore, the larger the RSD value, the larger the variability, and the worse the precision. A low RSD value reflects low data variability, which implies good precision.

### 2.0 Understanding Data Variability

Concentration data are necessarily produced by some measurement technique (such as an analytical chemistry method). Although overall data precision is affected by several factors, the most fundamental factor is the ability of the measurement technique to produce the "same" result from repeated measurements of the "same" sample.

### 2.1 Zero Variability Equals Perfect Precision

As an example, assume a $1-\mathrm{mL}$ vial of liquid is completed analyzed by injecting $100.1-\mathrm{mL}$ aliquots into an ideal analytical instrument 10 times. The 10 ideal results are

$$
20.0,20.0,20.0,20.0,20.0,20.0,20.0,20.0,20.0 \text {, and } 20.0 \text { (i.e., all are } 20.0 \mathrm{ppm} \text { ) Data Set \#1 }
$$

The statistics for the above data set are a mean of 20.0 ppm with a variability of 0 . Since the vial was "exhaustively" measured (i.e., the entire vial was analyzed), we can be confident the true concentration is 20.0 ppm . Variability must intuitively be 0 because there is no variation at all. Thus, this ideal instrument shows perfect measurement precision.

### 2.1 Low Variability Equals Good Precision

Perfect analytical instruments do not exist in the real world. If another 1-mL vial of the exact same liquid were measured in the real world, an accurate method with good precision might give 10 results that look like this:
19.8, 20.1, 19.5, 20.4, 20.3, 19.5, 19.9, 20.2, 20.0, and 20.3.

Data Set \#2
The mean of this exhaustive data set is also 20.0 ppm; however, most individual results vary from the true 20.0 ppm concentration. The variations are small, so intuitively the variability is evaluated as low and precision was good. The statistic commonly used to measure variability is the standard deviation (SD). The SD for Data Set \#1 was 0 ppm , and the SD for Data Set \#2 is 0.3 ppm . More details about how to calculate the SD and its statistical relevance are easily found on the Internet.

### 2.3 High Variability Equals Poor Precision

The good precision shown in Data Set \#2 can be contrasted with the poorer precision of the following data set:

$$
24.6,17.8,15.4,20.4,14.2,28.7,19.9,22.5,17.8 \text {, and } 19.0
$$

Data Set \#3
The mean of this data set is also 20.0 ppm , but individual results often vary quite a bit from 20.0. Intuitively the variability of Data Set \#3 is larger than that of Data Set \#2: the SD of Data Set \#3 is 4.3 ppm, over 10 times larger than the SD of 0.3 for Data Set \#2.

### 3.0 Relative Standard Deviation (RSD)

The RSD is useful because it "normalizes" the SD for easier comparisons among different data sets. As the name suggests, the magnitude of variability is assessed "relative to" something else, which is the mean. The RSD is calculated by dividing the standard deviation by the mean. The RSD for Data Set \#2 is $0.3 / 20.0=0.016$. For Data Set \#3, the RSD is $4.3 / 20.0=0.216$. For convenience, RSDs are often communicated as a percent: $1.6 \%$ for Data Set \#2, and $21.6 \%$ for Data Set \#3. The magnitude of variability in Data Set \#2 is very small in proportion to the mean, whereas Data Set \#3's variability is $1 / 5$ th of the mean's magnitude.

### 3.1 Low RSDs and High Concentrations

When the means of data sets are very similar, there is little advantage to using the RSD rather than the SD to compare the degree of variability among data sets. The RSD becomes more useful when comparisons involve data sets where the concentrations differ. Consider Data Set \#4 from the complete analysis of another 1 mL vial, but one with a different concentration from the previous examples.

$$
\text { 145.4, 150.8, 134.2, 167.4, 184.2, 158.7, 139.3, 156.0, 140.8, } 166.7 \quad \text { Data Set \#4 }
$$

Because this data set is from an exhaustive analysis, the arithmetic mean of 154.4 ppm for the 10 analyses must be the vial's true mean. The SD of the data set is 15.4 ppm -almost four times larger than the SD for Data Set \#3 (which was 4.3 ppm ).

Which data set is more variable?
If judged by the SD, Data Set \#4 is more variable since its SD is greater. However, if the degree of variability is considered relative to concentration, Data Set \#4 has the lower variability, since 15.4/154.4 = $10.0 \%$ RSD, compared to Data Set \#3's RSD of 21.6\%.

### 3.2 Calculating the RSD in Excel

The RSD for triplicate subsampling sets is calculated in Excel using the following notation:
=STDEV.S(C3:C5)/AVERAGE(C3:C5)

The \%RSD is simply the RSD decimal value times 100.
Note that Excel's "sample" standard deviation (STDEV.S) function is used, not the "population" standard deviation (STDEV.P). The same formula can be used to calculate the RSD directly for duplicate sets. Note also that RSD is unitless because both SD and mean have the same units and division cancels out the units.

### 3.2.1 Relationship between RSD and Relative Percent Difference (RPD)

When measuring the different between duplicates, most practitioners are accustomed to using the relative percent difference (RPD). As suggested by the name, the RPD is a measure of the difference (as found by subtraction) between two values and expressed relative to (i.e., divided by) the average of the two values. Like RSD, RPD has no units.

Although both RPD and RSD measure variability, it is important to realize they are different "units." For the same amount of variability, the RPD value is higher than the RSD value. For example, for the twonumber data set $(10,15)$, the RPD is $40.0 \%$ and the RSD is $28.3 \%$.

An analogy is the English vs. metric measures of distance: 1 inch is the same magnitude of length as 2.54 cm . Just as length in cm can be converted to inches units by dividing by 2.54, RPD and RSD are mathematically related by the square root of 2 . An RPD of 35 can be converted into RSD units by dividing 35 by sqrt(2) to give an RSD of $25 \%$ (i.e., $35 \%$ RPD/sqrt(2)) $=25 \%$ RSD).

### 3.2.2 Use the Same Measure of Variability to Avoid Confusion

ISM QC practices usually measure variability using three repeat measures (i.e., triplicates) as a more robust measure of variability than just two (i.e., duplicates). Because an RPD involves subtraction, it cannot be calculated on more than two replicates. That is one reason why ISM projects use RSDs to measure and express data variability. Another advantage is that the RSD is much easier to integrate into calculations of other important statistics, such as the upper confidence limit (UCL), than the RPD is. To avoid confusion, it is best to choose just one (either RSD or RPD) to use throughout a project or report so all measures of variability are in the same "units." If necessary, RSD values can be converted to RPD values by multiplying by sqrt(2).

### 3.3 RSDs at Very Low Concentrations

As mentioned above, the larger the RSD value, the larger the variability and the poorer the precision. However, the usefulness of this relationship breaks down when concentrations are very low. Since dividing by a small number increases the quotient, the RSD value will be larger at lower concentrations, even if the absolute difference among the replicate values is very small.

The example data sets at right illustrate this effect. Although the absolute difference between the values in the two sets is the same ( $4-3=1$ and $10-9=1$ ), their RSDs are quite different. An absolute difference of 1 may be insignificant to decisions and well within the range of simple analytical variability. But the relative difference causes the RSD to be significantly higher for the lower concentration set. This fact is important when setting limits on RSDs for QC purposes: any limits placed on RSDs must consider the concentrations likely to be obtained. A QC limit that may be achievable at higher concentrations may be mathematically unachievable at lower concentrations.
Comparison of RSDs for Data Sets
of Lower vs. Higher Concentrations

| 2 sets of duplicate results |  |  |  |
| ---: | :---: | :---: | :---: |
|  | Set 1 | Set 2 |  |
| Value 1 | 4 | 10 |  |
| Value 2 | 3 | 9 |  |
| RSD | 2 | $20.2 \%$ | $7.4 \%$ |

### 3.3.1 RSD Limits at Low Concentrations

If RSD limits need to be set for analytes that are typically reported at low concentrations (such as PCBs, dioxins, dioxin TEQ, and pesticides), the absolute magnitude of differences (in terms of concentration) that might be expected due to analytical variability alone should be considered before setting an RSD limit. Specifically, what analytical precision can be expected when sample concentrations are near the detection or reporting limits. An analytical chemist familiar with the method and soil data may be able to determine what degree of absolute precision can be expected at different concentration levels. Mathematically feasible RSDs can then be predicted.

### 3.3.2 Use UCL Calculations to Determine Acceptable RSD For a Given DU Mean Concentration

The best approach is to determine the amount of absolute variability that is acceptable or reasonable from a decision-making purpose. When UCLs are the basis for decisions, limits on field replicate RSDs can be expressed in the context of limits on decision uncertainty. As an expression of variability, RSD interacts with concentration to determine whether the UCL exceeds a screening level. The closer the concentration is to the action level, the less the RSD can be before it causes the UCL to exceed the screening level.

A calculator can be set up in Excel so that trial and error entering numbers for triplicate results can find the maximum variability possible so that the UCL will not exceed the action level at a given mean concentration. Refer to the table at right for examples:

- When the mean of triplicate field samples is 0.90 ppm , the RSD cannot be higher than $6.1 \%$

|  | trplicate ave = | 0.90 | 0.80 | 0.70 |
| :---: | :---: | :---: | :---: | :---: |
|  | ISM Rep 1 | 0.85 | 0.71 | 0.60 |
|  | ISM Rep 2 | 0.90 | 0.80 | 0.65 |
|  | ISM Rep 3 | 0.96 | 0.90 | 0.84 |
|  | Mean $=$ | 0.90 | $0.80{ }^{\prime}$ | 0.70 |
|  | Std Dev $=$ | 0.06 | $0.10^{\prime}$ | 0.13 |
|  | $\mathrm{n}=$ | 3 | 3 | 3 |
|  | \%RSD = | 6.1 | 11.8 | 18.2 |
|  | 1 -sided 95\% t-UCL $=$ | 1.00 | 0.96 | 0.91 |
| 1-sided Chebyshev 95\% UCL $=$ |  | 1.04 | 1.04 | 1.02 | (since at that RSD the 1 -sided $95 \%$ t-UCL is exactly 1.0 ).

- If the mean concentration is 0.8 ppm , the highest allowable RSD is $11.8 \%$.
- When there are 75 increments per DU and the field triplicate RSD exceeds $14.5 \%$, the UCL must switch to the Chebyshev, and the maximum RSD is $18.2 \%$ if the DU mean is 0.70 ppm .


### 4.0 Pooling of RSDs

When there are two or more RSDs with relevant concentrations, those RSDs may be "pooled" to produce the equivalent of an "average" RSD. Combining multiple RSDs from equivalent SUs with similar concentrations produces a more representative RSD to apply to singlet SUs (i.e., SUs with only a single field sample). The equation used to pool two RSDs in Excel is:
Pooled RSD = sqrt(sumsq(RSD\#1, RSD\#2)/2).

The equation for pooling three RSDs is:
Pooled RSD = sqrt(sumsq(RSD\#1, RSD\#2, RSD\#3)/3).

These equations are applicable where

- All replicate sets have the same n from which the RSD is calculated (e.g., $\mathrm{n}=2$ for all SU field sampling replicates),
- The concentrations of the SUs being pooled are similar enough so that the DUs can be considered CSM-equivalent, and
- The standard deviation (SD) values for the SUs are similar (so that the DUs can be considered CSM-equivalent).


### 4.1 Pooled RSD as Parallel to Pooled Standard Deviation (SD)

The concept of pooling RSDs to obtain an "average" of two or more RSDs mirrors the pooling of SDs which is a standard statistical calculation. The following text is copied from the statistical website called "Statistics How To" (accessed 3/3/2020)
(https://www.statisticshowto.datasciencecentral.com/pooled-standard-deviation/).

## What is a Pooled Standard Deviation?

The Pooled Standard Deviation is a weighted average of standard deviations for two or more groups. The individual standard deviations are averaged, with more "weight" given to larger sample sizes.

Once the pooled standard deviation has been calculated, $S D_{\text {pooled }}$ is used in place of $S D_{1}$ and $S D_{2}$ in the formula for standard error. Along with an updated degrees of freedom formula ( $d f=n_{1}+n_{2}-2$ ), the idea is that you would be able to get a better model for the sampling distribution of the sample mean.

Pooled standard deviations are used in many areas in statistics, including: effect size calculations, t-tests, and ANOVAs. They are also used in lab-based sciences like biology and chemistry, where they can be an indication for repeatability of an experiment.

## How to Calculate the Pooled Standard Deviation

Cohen (1988) offers a couple of options for calculating the pooled standard deviation. The simplest is:

$$
S D_{\text {poobled }}=\sqrt{\frac{\left(S D_{1}^{2}+S D_{2}^{2}\right)}{2}}
$$

Where:

- $\mathrm{SD}_{1}=$ standard deviation for group 1
- $\mathrm{SD}_{1}=$ standard deviation for group 2
(as long as both data sets have the same $n$ )

RSDs can be pooled In a similar way:
Pooled RSD = sqrt(sumsq(RSD\#1, RSD\#2)/2),

Remembering that RSDs must be converted to relative variance by squaring before averaging can be performed. After the average relative variance is obtained, its square root is taken to get back to standard deviation.

### 4.2 Derivation of the Pooled RSD Equation from Ingersoll's Work

As described below, Ingersoll explained the usefulness of partitioning variability when assessing data quality and identifying corrective actions (Ingersoll, 2001). He used an equation to derive an averaged variabilty term on page 44 of his paper:

$$
s_{r}=\left(\Sigma\left(d_{r}^{2}\right) / 2 k\right)^{1 / 2}
$$

This formula uses rather esoteric notation but is translated this way:

$$
\text { Pooled RSD } \left.=\operatorname{sqrt}\left\{\left[\text { the sum of the (relative differences }{ }^{2}\right)\right] /\left(2^{*} \mathrm{k}\right)\right\},
$$

where k is the number of replicate sets whose variability is being pooled.
Notice that Ingersoll uses the "relative difference ( $\mathrm{d}_{\mathrm{r}}$ )" in his equation to accommodate traditional laboratory duplicate sets whose variability was calculated using the RPD (relative percent difference). "Relative difference" is the same as RPD except that it is expressed as a decimal rather than as the percent. As discussed earlier, RPD and RSD are related by the sqrt(2) in that RPD/sqrt(2) = \%RSD. Therefore, the term "dr/sqrt(2)" equates to RSD (expressed as a decimal).
Isolating the $\mathrm{d}_{r} / \mathrm{sqrt}(2)$ term in $\mathbf{s}_{\mathbf{r}}=\left(\Sigma\left(\mathbf{d}_{\mathbf{r}}{ }^{2}\right) / 2 \mathbf{k}\right)^{1 / 2}$ and using the plus sign rather than the summation sign arranges Ingersoll's equation to mirror the pooled SD equation: Pooled RSD $\left.=\operatorname{sqrt[}\left[\left(d_{r_{1}}{ }^{2} / 2 k\right)+d_{r_{2}}{ }^{2} / 2 k\right)\right]$, where $\mathrm{k}=2$ when two RSD are being pooled.

The equation now becomes Pooled RSD $\left.=\operatorname{sqrt}\left[\left(\mathrm{drl}^{2} / 2 * 2\right)+\mathrm{d}_{\mathrm{r} 2}{ }^{2} / 2^{*} 2\right)\right]$.
Eqn. 1

Since $d_{r} /$ sqrt(2) $=$ RSD, squaring both sides gives $d_{r}^{2} / 2=$ RSD $^{2}$. Substituting $R^{2} D^{2}$ for $\mathrm{dr}^{2} / 2$ in Eqn 1 gives Pooled RSD $=\operatorname{sqrt}\left(\mathrm{RSD}_{1}{ }^{2} / 2+\mathrm{RSD}_{2}{ }^{2} / 2\right)$, which can be rewritten as Pooled $\mathrm{RSD}=\operatorname{sqrt}\left[\left(\mathrm{RSD}_{1}{ }^{2}+\right.\right.$ $\left.\left.\mathrm{RSD}_{2}{ }^{2}\right) / 2\right]$, or generalized as

$$
\begin{equation*}
\text { Pooled RSD }=\operatorname{sqrt}\left[\left(\mathrm{RSD}_{1}^{2}+\ldots+\mathrm{RSD}_{k}^{2}\right) / \mathrm{k}\right] . \tag{Eqn. 2}
\end{equation*}
$$

## 5.0 "Measurement Method" Includes More than Just the Laboratory Analysis

A "measurement method" encompasses all activities involved in generating a data result. For soil data, the measurement method includes activities on both the sampling side and analytical side. The sampling side includes sampling design (how many samples from what locations), sample collection, and sample processing. The analytical side includes subsampling of the processed sample, analytical preparation of the subsample to produce an extract, possibly cleanup of the extract, and instrumental analysis of the extract. Each step offers an opportunity for data variability to creep in and magnify. Data variability attributable to each of these steps can be represented as a "nested configuration." QC checks are targeted to each step in order to separate out the variability contributed by each step (Ingersoll, 2001 and 2006).

### 5.1 Assessing Data Variability and Its Components

Conceptual Model for Data Uncertainty Steps Introducing Variability and QC to Measure


The consensus of experts is that activities on the sampling side contribute much more heavily to data uncertainty than does the analytical side: "It has been estimated that up to 90 percent of all environmental measurement variability can be attributed to the sampling process." (Homsher, 1991). ISM address this by routinely using QC replication to measure variability stemming from field sampling and sample handling. Field triplicates serve as co-located field samples (USEPA, 2006). Since they are at the "top of the chain," they measure all components, but allow estimation of variability attributable to the field sampling design. The difference among field replicates is calculated as a total RSD.

If DU-level RSDs are elevated, the reasons for elevation are explored by calculating the RSDs for QC replicates for nested steps. For example, variability attributable to the processing/subsampling procedures will be measured using subsampling triplicates. Laboratory control sample (LCS) data can be used to evaluate analytical variability, and surrogate recovery QC data can be used to extend the evaluation of analytical variability to include matrix-specific sample preparation effects. If corrective action is required to reduce data variability, these variability evaluations can help pinpoint which corrective actions would be most effective for reducing data uncertainty, as reflected by the width of confidence intervals around individual sample results.

### 5.2 Measuring Variability Components

As illustrated by the above nested variability diagram, field sample precision is dependent on analytical precision. If the analytical process is not precise, different results can be expected even if field replicate samples were to have identical concentrations. For that reason, subsampling variability (which
incorporates all components of analytical variability "downstream" of sample processing) must be known to have adequate precision. If subsampling variability is not measured, it cannot be known whether field sample variability stems from field heterogeneity or poor sample processing. If QC checks show good analytical precision, but field sample replicates are too variable, the problem is isolated to some aspect of the field sample collection design.

### 5.2.1 Key Variability Components and Their QC Measures

QC checks should be structured to provide QC data that quantifies the major components of data variability.

- Laboratory control samples (LCS) can quantify analytical variability that is independent of the sample matrix. Other QC measures, such as surrogate compounds, might be used for this purpose if LCS data are not available. Such measures of analytical variability incorporate the extraction step, as well as extract cleanup and concentration steps, through to the last step which is instrumental determination/quantitation,
- Subsampling QC triplicates quantify the combined variabilities from sample processing through analysis.
- Field sample triplicates add in the variability from sampling design and field collection procedures so that all contributions to data variability are encompassed.


### 5.2.2 Literature References for the Mathematical Strategy Used in this Project

The above relationships can be expressed mathematically using the RSDs from replicate QC. This process is described in technical detail in various documents:

- Ingersoll, 2001: Environmental Analytical Measurement Uncertainty Estimation: Nested Hierarchical Approach, available through the Defense Technical Information Center (https://discover.dtic.mil/); also see References for the direct URL
- Ingersoll, 2003: Standard Operating Procedure Estimation of Analytical Measurement Uncertainty
- Ingersoll's uncertainty calculation strategy was referenced in the Department Of Defense Quality Systems Manual for Environmental Laboratories (Prepared by the DoD Environmental Data Quality Workgroup, Final Version 3, January 2006), page 61.


### 5.3 Mathematical Strategy for Partitioning Variability

Ingersoll's calculation strategy and equations can be adapted and simplified to meet specific projects needs. For the purpose of this project, the simplified equation is based on the key variability components listed in 5.2.1. The mathematical strategy is shown in the graphic below:

```
    Total variability = Field variability + Subsampling variability
Subsampling variability = Within-sample variability + Analytical variability
                        equivalent to:
                            (Total %RSD)}\mp@subsup{)}{}{2}=(\mathrm{ Field %RSD)}\mp@subsup{}{}{2}+(\mathrm{ Subsampling %RSD)}\mp@subsup{}{}{2
                        rearrange equation to solve for field component:
Eqn 1: Field %RSD = sqrt[(Total %RSD)}\mp@subsup{}{}{2}-(\mathrm{ (Subsampling %RSD)}\mp@subsup{}{}{2}
    ===========================================================
        If subsampling RSD too high, the analytical portion should be
Eqn 2: subtracted to evaluate the within-sample RSD:
    Within-sample %RSD = sqrt[(Subsampling %RSD)}\mp@subsup{}{}{2}-(\mathrm{ Analytical %RSD)}\mp@subsup{}{}{2}
```

The overall mathematical strategy can be summarized as:

The total variability (as measured by DU field replicates) is the sum of two major components: field heterogeneity and subsampling/analysis variability (as measured by subsampling triplicates).

Although field heterogeneity cannot be measured directly, it can be partitioned (i.e., separated or isolated) out of the total, This is done by subtracting the subsampling/analytical variability from the total variability per Equation 1. The remainder is attributed to the field component. In this way components that cannot be measured directly can be isolated from an aggregated variability.

Similarly the portion of variability attributable solely to sample processing and slab cake subsampling can be backed out from the analytical variability (Eqn 2 above). It is the remainder after the analytical variability is subtracted from the total within-sample variability (i.e., subsampling/analytical component measured by the subsampling triplicates).

Remember that RSDs cannot be added or subtracted directly since they are values obtained as square roots. The RSDs must be squared first to obtain relative variances (USEPA, 2002, p. 197). The relative variances are used to carry out the math, then the square root of the equation's output is taken to convert back to an RSD value.

### 5.3.1 Indications for Corrective Action

If the field component is identified as an issue, corrective action would be to evaluate field sample collection procedures such as the number of increments per field sample, the efficiency and consistency of sample collection tools and potential losses during transfer of increment masses into the sample container. Identified deficiencies would be corrected and samples recollected if necessary.

If the subsampling component is identified as an issue, corrective action is to re-evaluate sample processing and subsampling procedures. More rigorous sample processing may be needed (such as milling of the field sample, rather than sieving alone). Slabcake subsampling procedures (number of increments, the tools used) may need revision. It is also important to check the laboratory reports to make sure that the designated subsampling mass ( 30 grams) was actually used. Deficiencies should be corrected. If necessary, samples may need to be reprocessed or re-subsampled and reanalyzed.

If the pure analytical component is higher than desired, corrective action involves close review of the data validation reports and laboratory packages for any problems identified by laboratory QC such as
surrogate recoveries, blanks, calibration issues and clerical errors. Reanalysis of processed samples may be needed.

### 5.3.2 Dioxin Project Example of Partitioning to Identify Corrective Action

An example of this partitioning strategy and corrective action is provided by a past dioxin TEQ project to which the author provided assistance. The project consisted of two nearby areas (referred to here as "Site A" and (Site B") which were sampled at different times. The results of QC variability partitioning are summarized in the table below:

| Replication QC Summary <br> for Dioxin Site | Total RSD <br> $(\%)$ | Field RSD <br> $(\%)$ | Within-sample <br> RSD (\%) | Analytical RSD <br> $(\%)$ |
| :--- | :---: | :---: | :---: | :---: |
| Site A Data (Oct 2011) | Subsampling done in Lab from single 100-g sample jar |  |  |  |
| QC calculations | 31 | 18 | 25 | 2.5 |
| Site B Data (Jun 2012) | $15-\mathrm{g}$ analytical subsampling done in field-1 analysis/jar |  |  |  |
| QC calculations | 16 | 15 | 5.4 | 2.7 |

Site A was sampled in 2011. Its QC data showed a higher RSD for sample processing (25\%) than for field variability (18\%), which is unusual and a cause for concern. Fortunately, TEQ concentrations were well below the screening level ( 51 ppt ) so that the added data uncertainty caused by the total RSD of $31 \%$ did not compromise decision-making, and reanalysis of samples was not needed.

A follow-up query to the laboratory revealed that their subsampling method involved simple scooping off the top of the sample in its jar. Although the field sample had been correctly subsampled by field staff, 100-gram jars were filled for the laboratory. Therefore, the laboratory was scooping off the top of stratified/segregated soil samples, which introduced the data variability that was isolated to the subsampling step. The lab was unwilling to implement slabcake subsampling for the project's future work at Site B.

When Site B was sampled in 2012, corrective action was implemented by sending no more sample than the analytical mass ( 15 grams) used by the laboratory, with the instructions that the complete mass inside the jar was to be analyzed. Site B's QC results and variability partitioning showed great improvement: the processing RSD was down to $5.4 \%$ (from Site A's $25 \%$ ), which was only twice the LCS RSD ( $2.7 \%$ ) and much less than the field RSD (15\%). This reduction in subsampling variability brought total data variability down to $16 \%$ from the previous $31 \%$ RSD observed in Site A's sampling effort.

## References

Homsher, M.T.; F. Haeberer; P.J. Marsden; R.K. Mitchum; D. Neptune; and J. Warren. 1991. "Performance Based Criteria, A Panel Discussion" in Environmental Lab. Part I in the October/November 1991 issue; Part II in the December/January 1991/92 issue. Both parts combined in this online document: https://clu-in.org/download/char/dataquality/perfbased.pdf

Ingersoll, W.S. 2001. Environmental Analytical Measurement Uncertainty Estimation: Nested Hierarchical Approach, Defense Technical Information Center \# ADA396946. https://cluin.org/conf/tio/xrf 082108/cd/Ingersoll-uncertainty-calculator-2006/Ingersoll-paper.pdf and https://apps.dtic.mil/dtic/tr/fulltext/u2/a397300.pdf

Ingersoll, W.S. 2006. Estimation of Uncertainty Based on QC Data: Nested Hierarchical Approach, in Proceedings of the 22nd Annual National Environmental Monitoring Conference, pp 367-389, Aug 2831. http://www.nemc.us/docs/2006/2006 nemc.pdf

Myers, Jeffry C. 1997. Geostatistical Error Management: Quantifying Uncertainty for Environmental Sampling and Mapping. Van Nostrand Reinhold (Wiley), New York. ISBN-13: 978-0471285564.

USEPA. 2006. Applicability of Superfund Data Categories to the Removal Program. Quality Assurance Technical Bulletin No. 2 of 3. Office of Solid Waste and Emergency Response. OSWER 9360.4-21FS. July. See Notes 1 and 3 for Table 1. https://www.epa.gov/sites/production/files/2015-10/documents/dcgatib final for sig2.pdf

USEPA. 2002. RCRA Waste Sampling Draft Technical Guidance. EPA530-D-02-002. Office of Solid
Waste. August. https://www.epa.gov/sites/production/files/2015-10/documents/rwsdtg 0.pdf

## ATTACHMENT C

COMBINING DUS CALCULATOR TOOL

WHITE PAPER

# ISM Calculator (v. 3.0, August 2020) for Calculating 95\% UCL with ISM Data 

## Instructions


**This calculator updates v. 2 (Sept 2017, prepared by D. Crumbling). Specific updates applied here are noted in RED font.

| Item | Note |
| :---: | :---: |
| 1 | This calculator can be used to calculate a 95\% UCL using ISM data fom a single DU, or from multiple DUs. If multiple DUs are included, the calculator will also yield a result for the weighted avearge. Weighting factors are applied to the summary statistics and reflect the relative proportion by area, volume, or depth interval (as specified by the User). Equations that incorporate the weighting factors are given in the notes in "Caculator-example" in this spreadsheet. |
| 2 | Two types of $95 \%$ UCLs are calculated. The $95 \%$ Student's $t$ UCL is recommended for data sets with low variability within DUs, while the $95 \%$ Chebyshev UCL is recommended for data sets with high variability within decision units (ITRC 2020). Section 3.2 of Incremental Sampling Methodology (ITRC 2020) discusses the performance metrics for the UCL that can be used to determine which UCL method may be more likely to achieve the study objectives. There are also considerations based on the size of the data set that affect which UCL is recommended. |
| 3 | The calculator recommends the $95 \%$ Student's $t$ UCL for data sets where the coefficent of variation (CV) of increments is low (defined as CV < 1.5), and the 95\% Chebyshev UCL for data sets where the CV is medium (defined as $1.5<\mathrm{CV}<3$ ) or high (defined as CV $>3$ ). Because the CV of increments is typically not available from standard sampling programs, this calculator estimates the CV based on summary statistics of replicates. Lognormal distrubutions may also be investigated for their suitability. If the data distribution is lognormal, the geometric standard deviation (GSD) may be used as an alternate measure of dispersion to evaluate the possibility of using the $95 \%$ Student's $t$ UCL. The calculator does not include these calculations; see ITRC 2020 for guidance. |
| 4 | In July 2017, the original calculator was updated by its creator (Dr. Phil Goodrum, statistician with the ITRC ISM team, pegoodrum@gsi-net.com) to correct an underestimation bias in the calculation of the within-DU/between-increment coefficient of variation (CV). This is important because this CV is used to determine which UCL (the Student's $t$ or the Chebyshev) is appropriate. If the CV of the underlying increment population is underestimated, the lower UCL (the Student's $t$ ) could be indicated by the calculator when the higher UCL (the Chebyshev) is actually more appropriate. <br> Statistical simulations discovered the slight underestimation of the between-increment CV that occurs when it is calculated using the usual formula for standard error (SE), which obtains the between-increment standard deviation (SD) by multiplying the between-replicate SD by the square root of the number of increments used in each ISM sample replicate (see Column N formula). <br> Statistical simulations also supplied adjustments factors that can be used to remove the underestimation bias from the calculated CV. The value of the adjustment factor (shown in Column P) varies depending on the number of replicate ISM samples in the DU and the value of the calculated between-increment CV. The adjusted between-increment/within-DU CV is then used to determine which UCL the calculator will recommend to the user. <br> A 10-page white paper discusses the need for the withn-DU CV adjustment and derivation of the adjustment equations. This paper is provided on the "CV Adjustment" sheet. The calculator applies adjustment factors for DUs with replicates ranging from $r=3$ to 6 . Consult $P$. Goodrum if conditions fall outside this range. |

## ISM Calculator (v. 3.0, August 2020) for Calculating 95\% UCL with ISM Data

## Instructions

 **This calculator updates v. 2 (Sept 2017, prepared by D. Crumbling). Specific updates applied here are noted in RED font.

| Item | Note |
| :---: | :---: |
| 5 | The method used to estimate the coefficient of variation of the individual and combined DUs is an approximation. In addition, implementation of the WelchSatterthwaite approximation in the calculation of the $95 \%$ Student's $t$ UCL is recommended in cases where only moderate departures from normality are observed. The ITRC ISM Workgroup has not thoroughly tested the performance of this approach under conditions where distributions are highly skewed and weighting factors are very different. |


|  | Inputs <br> The green-shaded cells in the calculator indicate cells that the User should provide input for: |
| :---: | :---: |
| 6 | Select whether the sampling units (SU) are based on area, volume or depth-interval - this field (a drop-down menu) should be set to the value that represents what the SUs are based on (usually this will be area). Note that the "depth-interval" value assumes that the area for each depth-interval is the same. |
| 7 | Number of increments per replicate sample - this is the number of increments present in each replicate sample. For data sets where the number of increments is not the same for all samples, it is recommended that a statistician be consulted. |
| 8 | DU Name - this field is optional, but provides a place for the User to add a descriptive name for each DU. |
| 9 | DU Area/Volume or Depth-Interval - this field is required. The name of the column will adjust to reflect the choice made in the drop-down menu described above. The User should enter the quantitated value (in any consistent units) for each DU. For example, if area is the measure, the cells in column C should be populated with the DU area in sq.ft. or sq.m. (the actual units do not matter, as long as all measures are in the same units). If depth-interval is the measure, the thickness of the interval should be entered, such as 6 inches, 0.5 ft or 0.5 m . |
| 10 | Replicate concentration - these fields are required. Although there are spaces for up to 10 DUs and 6 replicates per DU, not all of the cells must have entries. At a minimum, three replicates should be entered for each DU (as a reminder, the text for each DU will remain red until three replicates have been entered for that DU). It is not required to have the same number of replicates for each DU. At least two DUs should be entered into the calculator. |

## Outputs for each component DU

The blue-shaded cells in the calculator show outputs for the individual DUs that the ITRC guidance recommends for evaluating combined DUs:

12 Weight for each DU - the fraction of the total combined DU that corresponds to each individual DU.
13 Arithmetic mean for each DU - the sum of the concentrations for all replicate samples in the DU divided by the number of samples in the DU.
14 SD of replicates for each DU - the standard deviation (SD) of the ISM replicate samples for the specific DU.
 calculated by multiplying the SD for the replicates by the square root of the total number of increments within each replicate.

## ISM Calculator (v. 3.0, August 2020) for Calculating 95\% UCL with ISM Data

## Instructions

 **This calculator updates v. 2 (Sept 2017, prepared by D. Crumbling). Specific updates applied here are noted in RED font.

| Item | Note |
| :---: | :---: |
| 16 | Calc'd CV of DU for each DU - the CV for the specific DU as determined by simple calculation. The font in this column is gray to reflect the underestimation bias of the simple CV calculation (see discussion in Item 4 above). |
| 17 | Adjustment Factor - the factor applied to the calculated CV to remove the underestimation bias. This is the update that was applied in version 2 of this calculator. |
| 18 | Adjusted CV of the DU for each DU - the adjusted within-DU CV used to determine which UCL the calculator recommends. |
| 19 | SE for each DU - the standard error (SE) for the specific DU. The standard error is the SD divided by the square root of the number of samples for the DU. |
| 20 | Student's-t 95\% UCL - the 95\% UCL for the Student's t method. See ITRC 2020, Section 3.2.4.2 for the equation. |
| 21 | Chebyshev 95\% UCL - the 95\% UCL for the Chebyshev method. See ITRC 2020, Section 3.2.4.2 for the equation. |


|  | Outputs for the Combined DUs <br> The purple-shaded cells show outputs for the combined DU, weighted by area, volume, or depth. |
| :---: | :---: |
| 22 | DU Area - the total area for the combined DU. |
| 23 | Total number of replicates - this includes the original sample as well as all replicates. |
| 24 | Weight - the total weight. This should always be 1; if it is not then there is an error in the calculations. |
| 25 | Mean - the weighted mean for the combined DU. |
| 26 | $\underline{S D}$ - the weighted SD for the combined DU. |
| 27 | SE - the weighted SE for the combined DU. |
| 28 | CV - the CV for the combined DU. |
| 29 | Student's-t 95\% UCL - the weighted 95\% UCL for the Student's t method. Relevant equations are given in the notes in the sheet "Caculator-example" in this workbook. |
| 30 | Chebyshev 95\% UCL - the weighted 95\% UCL for the Chebyshev method. Relevant equations are given in the notes in the sheet "Calculator-example" in this workbook. |
| 31 | Degrees of freedom by Welch-Satterthwaite approximation - used to calculate the Student's-t 95\% UCL. Note that the Welch-Satterthwaite approximation relaxes the assumption of equal variance of the DUs, but still carries the assumption of normality. |
| 32 | Recommended UCL - the recommendation is based on the dispersion of the individual DUs; see Section 3.2.4.2 of ITRC 2020 for a more detailed explanation of the recommendations. |

## Notes


 assistance with calculating weighted $95 \%$ UCLs for data sets with non-detect results.

## ISM Calculator (v. 3.0, August 2020) for Calculating 95\% UCL with ISM Data

## Instructions

This spreadsheet is a tool for use in accordance with the Interstate Technology \& Regulatory Council (ITRC) document: Incremental Sampling Methodology (ISM) Update (ITRC 2020). **This calculator updates v. 2 (Sept 2017, prepared by D. Crumbling). Specific updates applied here are noted in RED font.

| Item | Note |
| :---: | :---: |
| References |  |
| 34 | The statistical method of combining and weighting DUs to produce a combined mean and UCL from incremental samples is equivalent to the statistical method for processing discrete data from "stratified sampling" as described in the following EPA statistical guidance (see Section 3.2.1.3, Stratified Random Sampling): <br> EPA QA/G-9S Data Quality Assessment: Statistical Methods for Practitioners, EPA/240/B-06/003, February 2006 |
| 35 | ITRC. 2020. Technical and Regulatory Guidance, Incremental Sampling Methodology Update. October. |
| 36 | EPA. 2010. EPA Basic KM TEQ and ISM UCL Calculator. Available online at: https://www.epa.gov/sites/production/files/2015- |
| 37 | EPA. 2015. ProUCL Version 5.1 Technical Guide. Office of Research and Development, Washington, DC. EPA/600/R-07/041. Prepared by Singh, A. and A.K. Singh. October. Available online at: https://www.epa.gov/land-research/proucl-version-5100-documentation-downloads |
| Referencing this calculator |  |
| 38 | ITRC (Interstate Technology \& Regulatory Council). 2020. Incremental Sampling Methodology. ISM-2. Washington, D.C.: ITRC Incremental Sampling Methodology Team. Calculator for the Weighted 95\% UCL for a Combined DU from Several Smaller DUs, version 3 (August 2020) webpage URL: not known at this time - refer to ITRC's ISM web page |

# WHITE PAPER - BACKCALCULATION OF STANDARD DEVIATION OF UNDERLYING DISTRIBUTION OF INCREMENTS USING INCREMENTAL SAMPLING METHODS 

This white paper presents statistical concepts related to estimating parameters of skewed distributions using results from a dataset generated from Incremental Sampling Methods (ISM). The final section provides recommendations on conditions when the ISM summary statistics and corresponding backcalculation of the underlying distribution match conditions specified in the guidance provided in the ITRC 2012 report ${ }^{1}$. Practitioners may find this useful in the following contexts:

- Selection of 95UCL calculation method (i.e., Student's $t$ or Chebyshev) to apply to ISM data, based on ITRC's 2012 guidance on ISM regarding the coefficient of variation (CV) of the distribution of increments; and
- Sampling design considerations - relationship between number of increments or replicates and the width of confidence intervals for the mean.


## Key Terms and Concepts

The following terms and concepts are used in this white paper:

- Underlying distribution - this is the distribution of all possible increments collected from a decision unit (DU). The distribution is contingent on the sample mass of each increment (discussed below). With ISM, a set of increments (e.g., $\mathrm{n}=30$ ) is combined into a single field replicate and further prepared in a laboratory such that a final analytical result is representative of the average concentration; individual analytical results for each individual increment are typically not available, so we do not have a direct measure of the variance (or standard deviation) of the underlying distribution. The central question then becomes - can we estimate that variance using sample statistics from the ISM dataset?
- Sample support - refers to the sample mass of increments. There is no single "true" underlying distribution of chemical concentrations in soil within a DU; rather, the underlying distribution depends, in part, on the choice of sample support. Two different sampling designs (applied to the same DU), each with different sample support, may have different underlying distributions. This partly explains why the underlying distribution for discrete sampling events (with smaller

[^0]sample support) is not the same as the underlying distribution of an ISM sampling event. Note that the sample size (i.e., number of increments) does not influence the underlying distribution, but it does influence the distribution of the sample means, which is what each ISM replicate result measures.

- Population parameter - the parameters of a population are typically not known. Using a random sampling design, we can obtain unbiased estimates of the population parameters as well as confidence intervals for those estimates. The population mean $(\mu)$ and population standard deviation $(\sigma)$ are estimated by the sample mean ( $\bar{x}$ ) and sample standard deviation (s).
- Negative bias - a sample summary statistic such as the arithmetic standard deviation, s , is said to be an unbiased estimate of the population parameter, $\sigma$, if $s=$ $\sigma$ on average (i.e., across repeated sampling events), or equivalently $|\mathrm{s}-\sigma|=0$ on average. The estimator has negative bias if it tends to underestimate the population parameter, on average.
- Coefficient of variation (CV) - refers to the standard deviation divided by the arithmetic mean (i.e., $C V=\sigma / \mu$ ). Because the $C V$ is unitless (i.e., the units for $\sigma$ and $\mu$ are the same and, therefore, cancel each other out), it is a convenient metric for comparing the spread of multiple variables. The CV may be estimated from the summary statistics $s$ and $\bar{x}$ (i.e., $C V=s / \bar{x}$ ). We can calculate CV for the underlying distribution and/or for the distribution of means. They will not yield the same result because $s$ of ISM replicates is less than $s$ of increments. We typically do not have a direct measure of $s$ for increments.
- Distribution of means - Assume we repeated a sampling event many times such that each event involves collecting $n$ increments (of equal mass) from the same DU. Each event would yield a separate estimate of the sample mean, $\bar{x}$. The distribution of the sample means can be said to have an arithmetic mean $\mu$ and standard deviation $\sigma / \sqrt{n}$, where $\sigma$ is the standard deviation of the underlying distribution of increments. Note that the mean of the distribution of sample means, which we'll denote as $\mu_{\bar{x}}$, is equal to the mean of the underlying distribution, $\mu$. However, the standard deviation of the means, $\sigma_{\bar{x}}$, is synonymous with the standard error of the mean. The values $\sigma_{\bar{x}}$ and $\sigma$ are related, but not the same, is as discussed below.
- Standard error (SE) - refers to the "error" or expected variability in the sample means, and is one of the factors that informs the confidence interval for the mean when $\bar{x}$ is used to estimate $\mu$. The SE can be thought of as the standard deviation of the distribution of arithmetic means, sometimes denoted as $s_{\bar{x}}$. Since we typically do not repeat a sampling event many times to obtain multiple estimates of $\mu$, it is common to calculate the SE from the sample standard deviation, $s$, and sample size,
$n$, using the equation $\mathrm{SE}=s_{\bar{x}}=s / \sqrt{n}$. We sometimes say that the standard deviation of ISM replicates provides a direct measure of the SE of the mean of the underlying distribution because each ISM replicate result is a point on the distribution of sample means.
- Central Limit Theorem (CLT) - as $n$ increases, the distribution of sample means tends to become normal, even if the underlying distribution is non-normal. However, the skewness of the underlying distribution directly affects the symmetry of the distribution of means. It takes larger sample sizes to achieve approximate normality of the distribution of means as the asymmetry of the underlying distribution increases. As discussed in the 2012 ITRC document (Section 4.2.2.1 and Section 4.3), the performance of the Student's $t$ UCL calculation using ISM replicate results varied as a function of the CV for the underlying lognormal distribution, whose population parameters were specified as part of a simulation study.


## Simulation Study

A new simulation study was conducted to explore the following question:

- If a DU is sampled using ISM, yielding $r$ replicate results, each a composite of $n$ increments, can we estimate the standard deviation of the underlying distribution, $s$, by rearranging the standard error equation to solve for $s$ ?

$$
\begin{gathered}
s_{\bar{x}}=s / \sqrt{n} \\
s=s_{\bar{x}} \times \sqrt{n}
\end{gathered}
$$

Assume the summary statistics for the ISM replicates yields a standard deviation of replicates, $s_{\bar{x}}$, and a grand mean of the set of $\bar{x}$, which we'll denote $\bar{X}$.
The underlying distribution from which increments are randomly selected is defined as a lognormal distribution with $\mu=1$ and $\sigma=1,2$, or 3 (equivalent to, $C V=1,2$, or 3 ).

The simulation was repeated 5,000 times for each of the three underlying lognormal distributions ( $\mathrm{CV}=1,2$, and 3), and for a range of replicates ( $\mathrm{r}=3,4,5,6$, and 10 ). For any single ISM dataset, we expect some difference between the calculated $s$ and the true underlying distribution parameter $\sigma$. However, the calculation of $s$ is considered unbiased if $s=\sigma$ on average (i.e., across all 5,000 results), or equivalently $|s-\sigma|=0$ on average.

## Results

Table 1 provides a comprehensive summary of the negative bias in the use of backcalculated $s$ for simulations (5,000 iterations) based on the following combinations of ISM sampling scenarios:

- underlying distributions defined as lognormal with $C V=1,2$, and 3;
- range of increments ( $n=20,30$, and 40 ); and
- range of replicates ( $\mathrm{r}=3,4,5,6$, and 10 ).

Table 1. Negative bias in the estimate of the standard deviation of the underlying lognormal distribution with $\mu=1$ and $\sigma=1,2$, and 3 (i.e., $C V=1,2$, and 3 ), expressed as an absolute difference and a percent difference.


## Notes:

$\mathrm{CV}=$ coefficient of variation $=\sigma / \mu$
${ }^{\text {a }}$ percent difference $=100 \% \times(s-\sigma) / \sigma$

In general, there is a negative bias in the estimate of $\sigma$ across all simulations. In terms of the ISM sampling design options, the bias is more sensitive to the choice of ISM replicates than ISM increments. For example, when $\sigma=2$ and $n=30$, the calculated $s$ underestimates $\sigma$ by $20 \%$ for $\mathrm{r}=3,13 \%$ for $\mathrm{r}=5$, and $8 \%$ for $\mathrm{r}=10$; however, changing the number of increments ( $\mathrm{n}=20,30$, and 40), results in variation in $s$ (i.e., all underestimates of $\sigma$ are within $5 \%$ ) across
all of the replicate scenarios. Similar examples are provided in Figures 1 and 2 discussed below.

Figure 1 shows the frequency distribution of the difference between the calculated standard deviation, $s$, and the population parameter, $\sigma$, for $C V=1$ (i.e., underlying lognormal distribution with $\mu=\sigma=1$ ). The distribution of ( $s-\sigma$ ) is right-skewed with an arithmetic mean of $-0.13,-0.07$, and -0.03 for $\mathrm{r}=3, \mathrm{r}=5$, and $\mathrm{r}=10$, respectively. Increasing the number of replicates effectively reduces the skew of $(s-\sigma)$, resulting in lower bias in the estimate $s$.

Figure 2 shows the frequency distribution of $(s-\sigma)$ for $C V=2, r=3$ replicates, and increments ranging from $\mathrm{n}=20$ to 40 . The distribution of $(\mathrm{s}-\sigma)$ is right-skewed with an arithmetic mean of $-0.44,-0.39$, and -0.37 for $n=20, n=30$, and $n=40$, respectively. Increasing the number of increments has a relatively minor effect on the skew of ( $s-\sigma$ ), resulting in little change in the bias in the estimate of $s$.

Figure 3 provides the simulation results for $\mathrm{n}=30$ increments and lognormal distributions with $\mathrm{CV}=1,2$, and 3 . The best fit lines show the relationship between the number of replicates and the percent underestimate of $\sigma$. Based on these results, we can determine the adjustment factor (i.e., multiplier) that must be applied to the estimate of $s$ in order to more closely approximate $\sigma$. For example, for $\mathrm{r}=3$ ISM replicates collected from an underlying distribution of $\log (1,1)$, the backcalculated $s$ will underestimate $\sigma$ by $13 \%$, on average so we expect $s=1 \mathrm{x}(1-0.13)=0.87$. The multiplier needed to adjust this result is $1 / 0.87=1.151$. This approach to calculating a multiplier can be extended to any lognormal distribution by applying the multiplier to the estimate of the CV. Figure 4 shows the full set of multipliers needed to adjust backcalculated CV , given different numbers of replicates.

ISM introduces two sample sizes - the number of increments ( n ) and number of replicates (r). The influence of $r$ on the backcalculation of CV stems from the reduction in the error on the standard deviation of replicates, $s_{\bar{x}}$, afforded by increasing the number of replicates used in the calculation. Just as the confidence interval in the mean is tighter with increasing $r$, so too is the standard deviation. The central limit theorem confirms that the normality of the distribution of $s_{\bar{x}}$ also improves with larger sample sizes. This has the effect of "centering" and thereby reducing the bias in the mean $s_{\bar{x}}$ across the 5,000 iterations as $r$ increases.

Page 6
ISM Backcalculation of Standard Deviation
P. Goodnom, Integral Consulting Inc.

A


B


C


Figure 1. Simulation results for $\log (\mu=1, \sigma=1)$, based on 5,000 trials of an ISM sampling design with $\mathrm{n}=30$ increments and $\mathrm{r}=3$ (Panel A), $\mathrm{r}=5$ (Panel B), and $\mathrm{r}=10$ (Panel C) replicates.

A


B

c


Figure 2. Simulation results for $\log (\mu=1, \sigma=2)$, based on 5,000 trials of an ISM sampling design with $\mathrm{r}=3$ replicates and $\mathrm{n}=20$ (Panel A), $\mathrm{n}=30$ (Panel B), and $\mathrm{n}=40$ (Panel C) increments.


Figure 3. Simulation results with lognormal distributions with $\mathrm{CV}=1,2$, and 3. Each replicate is generated based on the arithmetic mean of $n=30$ increments. Points represent the percent difference between $s$ and $\sigma$ for $r=3,4,5,6$ and 10 . The backcalculated estimate of the standard deviation of the lognormal distribution, $s$, is calculated from the standard deviation of the replicates, $s_{\bar{x}}$, multiplied by $\sqrt{30}$.


Figure 4. Simulations of backcalculated CV and the adjustment factors required to compensate for the negative bias. The multiplier depends on the number of replicates. Equations for best-fit polynomials are given and each fits the points with an r-square of nearly 1.0. See Table 2.

So all the pieces are in place to address the question of how to use ISM results to yield an unbiased estimate of the CV for an underlying lognormal distribution. We know that the degree of negative bias depends on the ISM sampling design - particularly the number of replicates. We can use the $2^{\text {nd }}$ order polynomial equations given in Figure 4 to solve for the adjustment factor to apply to a backcalculated CV. Table 2 provides an example for the specific cases where the backcalculated CVs all yield an adjustment factor to accurately estimate CV =1.5 and 3.0. These are two of the triggers identified by ITRC (2012, Table 4-4Likelihood that UCL method achieves coverage) for selecting a UCL calculation method. If the CV of the underlying distribution is less than 1.5 , then both Student's $t$ and Chebyshev can be expected to yield the desired $95 \%$ coverage. If the CV is greater than 1.5 , the Chebyshev is recommended and is expected to yield the desired coverage up to about $\mathrm{CV}=3.0$. The examples in Table 2 are specifically based on simulations with $\mathrm{n}=30$ increments per replicate, but as noted previously, the variability in the negative bias is low for ISM sampling designs with $\mathrm{n}=20$ to 40 increments.

Table 2. Adjustments (multipliers) to apply to calculated CVs in order to compensate for negative bias in the backcalculation using ISM results ( $\mathrm{r}=3$ to $6, \mathrm{n}=30$ increments).

| Coefficients of $2^{\text {nd }}$ Order <br> Polynomial ${ }^{\text {a }}$ | $r=3$ |  | $\mathrm{r}=4$ |  | $\mathrm{r}=5$ |  | r=6 |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathrm{x}^{2}$ | 0.0645 |  | 0.0856 |  | 0.0747 |  | 0.0657 |  |
| x | -0.0327 |  | -0.1253 |  | -0.1094 |  | -0.1019 |  |
| Intercept | 1.1306 |  | 1.147 |  | 1.1156 |  | 1.1017 |  |
| Backcalculated (biased) CVs | 1.26 | 2.19 | 1.33 | 2.29 | 1.36 | 2.36 | 1.38 | 2.42 |
| x Adjustment factor ${ }^{\text {b }}$ | 1.19 | 1.37 | 1.13 | 1.31 | 1.10 | 1.27 | 1.09 | 1.24 |
| Corrected CV | 1.5 | 3.0 | 1.5 | 3.0 | 1.5 | 3.0 | 1.5 | 3.0 |

Notes:
a See Figure 4 for polynomials fit to results for each replicate scenario.
${ }^{\text {b }}$ Adjustment factors are multipliers applied to the backcalculated CV. CV is the " $x$ " in the polynomial equations. For example, for $\mathrm{r}=3$, if a backcalculated CV is 1.26 , we solve for the "corrected CV " by first calculating the adjustment factor specific to this result: $0.0645 x^{2}-0.0327 x+1.1306=0.0645(1.26)^{2}-0.0327(1.26)+1.1306$ $=1.19$. The product of the biased $\mathrm{CV} \times$ adjustment factor is $1.26 \times 1.19=1.5$.

## Recommendations

The Student's t UCL will likely achieve the desired 95\% coverage if the CV of the replicates and equivalent (biased) backcalculated CV satisfy the following conditions (determined by multiplying by $\sqrt{30}$ ):

- if $\mathrm{r}=3$
- CV of ISM replicates $<0.23$
- Equivalent (biased) CV of underlying distribution $<1.26$
- if $\mathrm{r}=4$
- CV of ISM replicates $<0.24$
- Equivalent (biased) CV of underlying distribution $<1.33$
- if $\mathrm{r}=5$
- CV of ISM replicates $<0.25$
- Equivalent (biased) CV of underlying distribution $<1.36$
- if $\mathrm{r}=6$
- CV of ISM replicates $<0.26$
- Equivalent (biased) CV of underlying distribution $<1.38$

The Chebyshev UCL will likely achieve the desired 95\% coverage for the following conditions:

- if $\mathrm{r}=3$
- CV of ISM replicates $<0.40$
- Equivalent (biased) CV of underlying distribution $<2.19$
- if $\mathrm{r}=4$
- CV of ISM replicates $<0.42$
- Equivalent (biased) CV of underlying distribution $<2.29$
- if $\mathrm{r}=5$
- CV of ISM replicates $<0.43$
- Equivalent (biased) CV of underlying distribution $<2.36$
- if $\mathrm{r}=6$
- CV of ISM replicates $<0.44$
- Equivalent (biased) CV of underlying distribution $<2.42$

Any questions regarding this white paper can be directed to Philip Goodrum, Ph.D., DABT at pgoodrum@integral-corp.com, phone 315-396-6655.

Note: P. Goodrum's contact information as of August 2019 is: pegoodrum@gsi-net.com, ph: 315-396-6655.

Calculation of Weighted 95\% UCLs for a Combined Decision Unit (DU) from Several Smaller DUs Having Replicate Incremental Samples
Enter information in green highlighted cells. See the "Instructions" tab for detailed instructions.

| Project ID: |
| :---: |
| Property/Sample ID: |
| Date of calculations: |
| Calculator completed by: |
| Analyte: |
| Analyte units: |
| DU metric units: |
| Notes: |

Click in green cell below to select from drop-down menu DU size metric: area, volume, or depth interval:
$\qquad$

Note: Assumes all replicates have the same number of increments Number of increments per replicate:

| Row \# | IDs/Names of the Smaller DUs | DU Area () | Replicate field sample concentrations |  |  |  |  |  | Number of Replicates | Weight | Arithmetic Mean | SD of Replicates | calc'd SD of Increments | calc'd CV for the DU | Adj Factor | adj'd SD of Increments | adj'd CV for DU | $\begin{aligned} & \text { SE } \\ & \text { of DU } \end{aligned}$ | 95\% UCL |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | Rep 1 | Rep 2 | Rep 3 | Rep 4 | Rep 5 | Rep 6 |  |  |  |  |  |  |  |  |  |  | Student's-t | Chebychev | CV of Increments | 95\% UCL |
| 1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  | \#N/A | \#VALUE! | \#VALUE! |  |  |  | \#VALUE! | \#VALUE! |
| 2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  | \#N/A | \#VALUE! | \#VALUE! |  |  |  | \#VALUE! | \#VALUE! |
| 3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  | \#N/A | \#VALUE! | \#VALUE! |  |  |  | \#VALUE! | \#VALUE! |
| 4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  | \#N/A | \#VALUE! | \#VALUE! |  |  |  | \#VALUE! | \#VALUE! |
| 5 |  |  |  |  |  |  |  |  |  |  |  |  |  |  | \#N/A | \#VALUE! | \#VALUE! |  |  |  | \#VALUE! | \#VALUE! |
| 6 |  |  |  |  |  |  |  |  |  |  |  |  |  |  | \#N/A | \#VALUE! | \#VALUE! |  |  |  | \#VALUE! | \#VALUE! |
| 7 |  |  |  |  |  |  |  |  |  |  |  |  |  |  | \#N/A | \#VALUE! | \#VALUE! |  |  |  | \#VALUE! | \#VALUE! |
| 8 |  |  |  |  |  |  |  |  |  |  |  |  |  |  | \#N/A | \#VALUE! | \#VALUE! |  |  |  | \#VALUE! | \#VALUE! |
| 9 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 10 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | Sum: |  | -- | -- | -- | -- | -- | -- |  |  |  |  |  |  | NA |  | \#VALUE! |  |  |  | \#VALUE! | \#VALUE! |
| df by Welch-Satterthwaite approximation:\| |  |  |  |  |  |  |  |  | Recommended UCL: \#VALUE! 0 |  |  |  | \#VALUE! |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  | \#VALUE! |  |  |  |  |  |  |  |  |  |

set
=adjusted df = degrees of freedom SD
CV = coefficient of variation RSD $=$ relative standard deviation sts.

Calculation of Weighted 95\% UCLs for a Combined Decision Unit (DU) from Several Smaller DUs Having Replicate Incremental Samples
Enter information in green highlighted cells. See the "Instructions" tab for detailed instructions.

| Project ID: | ABC Oil Field |
| ---: | :--- | :--- |
| Property/Sample ID: | OU 1 and OU 2 |
| Date of calculations: | August 23, 2020 |
| Calculator completed by: | Jane Doe |
| Analyte: | Pb |
| Analyte units: | $\mathrm{mg} / \mathrm{kg}$ |
| DU metric units: | acres |
| Notes: | Example |

Click in green cell below to select from drop-down menu
DU size metric: area, volume, or depth interval:

| Row \# | IDs/Names of the Smaller DUs | DU Area (acres) | Replicate field sample concentrations |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | Rep 1 | Rep 2 | Rep 3 | Rep 4 | Rep 5 | Rep 6 |
| 1 | DU-1 | 2.0 | 10 | 12 | 14 |  |  |  |
| 2 | DU-2 | 2.5 | 10 | 12 | 14 | 19 |  |  |
| 3 | DU-3 | 3.0 | 10 | 13 | 15 | 16 | 18 |  |
| 4 | DU-4 | 1.2 | 12 | 14 | 15 |  |  |  |
| 5 | DU-5 | 1.0 | 13 | 15 | 14 | 15 |  |  |
| 6 | DU-6 | 1.3 | 16 | 25 | 28 |  |  |  |
| 7 | DU-7 | 3.6 | 29 | 30 | 34 | 43 |  |  |
| 8 | DU-8 | 1.2 | 16 | 18 | 19 | 23 | 45 |  |
| 9 |  |  |  |  |  |  |  |  |
| 10 |  |  |  |  |  |  |  |  |
| Sum: |  | 16 | -- | -- | -- | -- | -- | -- |
| df by Welch-Satterthwaite approximation: 8.7 |  |  |  |  |  |  |  |  |

— 8.7

| Note: Assumes all replicates have the same number of increments |
| :--- |
| Number of increments per replicate: |

Number of increments per replicate: 30

*Student's t UCL is acceptable if adj'd CV for DU is "Low" (e.g., CV $\leq 1.5$ ). The User should consult the instructions for additional guidance on which $95 \%$ UCL is recommended for specific data
Notes
adj'd = adjusted df $=$ degrees of freedom sets.
$\begin{array}{llll}\text { calc'd }=\text { calculated } & \text { df } & =\text { degrees of freedom } & \text { SD } \\ & \text { DU } & =\text { decision unit } & \text { SE }\end{array}$
$\begin{array}{llll}\text { CV }=\text { coefficient of variation } \quad \text { RSD } & =\text { relative standard deviation } & 95 \% \text { UCL } & =95 \% \text { upper confidence limit for arithmetic mean }\end{array}$


[^0]:    ${ }^{1}$ ITRC (Interstate Technology \& Regulatory Council). Incremental Sampling Methodology. ISM-1. Washington, DC: Interstate Technology \& Regulatory Council, Incremental Sampling Methodology Team. 2012. Available at: www.itrcweb.org.

