Montana Tech Library

Digital Commons @ Montana Tech

Silver Bow Creek/Butte Area Superfund Site

Montana Superfund

Winter 12-2022

COMPARISON OF SIEVED SOIL SAMPLE RESULTS FROM BUTTE-SILVER BOW MULTI-PATHWAY RESIDENTIAL METALS ABATEMENT PROGRAM: 250 µM AND 150 µM SIEVED SAMPLES FOR LEAD AND ARSENIC ENRICHMENT DETERMINATION ATLANTIC RICHFIELD COMPANY BUTTE, MONTANA

Ramboll

Environmental Standards, Inc.

Follow this and additional works at: https://digitalcommons.mtech.edu/superfund_silverbowbutte Part of the Environmental Health and Protection Commons, Environmental Indicators and Impact Assessment Commons, and the Environmental Monitoring Commons Prepared for Atlantic Richfield Company Butte, Montana

Prepared by Ramboll US Consulting, Inc. Seattle, Washington and Environmental Standards, Inc. Valley Forge, PA

Project Number

1690026989

Date December 2022

COMPARISON OF SIEVED SOIL SAMPLE RESULTS FROM BUTTE-SILVER BOW MULTI-PATHWAY RESIDENTIAL METALS ABATEMENT PROGRAM: 250 µM AND 150 µM SIEVED SAMPLES FOR LEAD AND ARSENIC ENRICHMENT DETERMINATION ATLANTIC RICHFIELD COMPANY BUTTE, MONTANA

PREPARED BY

Rosalind Schoof, PhD, DABT

Amanda Bailey, MS

Lester J. Dupes, CEAC

Ramboll US Consulting, Inc. 901 5th Avenue Suite 2820 Seattle, Washington 98164

and

Environmental Standards, Inc. 1140 Valley Forge Road PO Box 810 Valley Forge, PA 19482 CONTENTS

			Page
Executive Summary			1
1.	Introduction		3
	1.1	Background	3
	1.2	Relevant Studies	3
	1.3	Program Description and Oversight	4
	1.4	Current Study Overview and Data Quality Objectives	5
2.	Methodology		6
	2.1	Sample Collection Designations and Characteristics	6
	2.2	Sample Selection	6
	2.3	Analytical Methodologies	7
	2.3	3.1 Initial Sample Preparation and Analysis: 250 μ m aliquot	7
	2.3 2.3 2.3	 3.2 Secondary Sample Preparation and Analysis: 150 µm aliquot 3.3 Quality Control 3.4 Data Validation 	7 8 8
3.	Data Analysis		10
	3.1	Measurement of Enrichment	10
	3.2	Outlier Analysis	10
	3.3	Evaluation of Enrichment by Concentration	10
	3.4	Evaluation of Enrichment by Depth Interval	11
	3.5	Effect of Sieve Size on Action Level Comparisons	11
	3.6	Effect of Sample Size Fractionation	11
	3.7	Statistical Analysis	11
4.	Results		13
	4.1	Outlier Analysis	13
	4.2	Enrichment Summary	14
	4.3	Depth Comparison	14
	4.4	Effect of Sieve Size Relative to Action Levels	15
	4.5	Sample Size Fractionation	15
	4.6	Statistical Analysis	15
5.	Conclusions		17
6.	References		18

DRAFT

TABLES

- Table 1.Summary of Juhasz et al. (2011), Tawinteung et al. (2005), and Karna et al. (2017)
Studies
- Table 2. Samples Identified as Percent Difference Outliers
- Table 3. Lead and Arsenic Enrichment Summary
- Table 4.
 Lead and Arsenic Enrichment by Depth Interval
- Table 5. Statistical Analysis Results

FIGURES

- Figure 1. Lead Percent Difference
- Figure 2. Arsenic Percent Difference
- Figure 3. Lead Percent Difference by Depth Interval
- Figure 4. Arsenic Percent Difference by Depth Interval
- Figure 5. 150 µm vs 250 µm Sieved Lead Result
- Figure 6. 150 µm vs 250 µm Sieved Arsenic Result
- Figure 7. Paired Lead Concentrations in Samples After Sieving to \leq 250 µm and <150 µm
- Figure 8. Paired Arsenic Concentrations in Samples After Sieving to \leq 250 µm and <150 µm

APPENDICES

- Appendix A Sieve Study Data
- Appendix B Data Validation Reports
- Appendix C ProUCL Output Outlier Testing
- Appendix D Reanalysis Summary
- Appendix E Statistical Analysis

EXECUTIVE SUMMARY

During the 2021 execution of the Butte-Silver Bow (BSB) Multi-Pathway Residential Metals Abatement Program (RMAP) Plan (BSB and Atlantic Richfield Company, 2020), 1,553 soil samples were collected, processed, and analyzed for lead and arsenic. Lead and arsenic results were compared to risk-based limits to assess the need for potential remedial action in the vicinity of schools, playgrounds, and sports fields. Consistent with prior sampling programs, the soil samples were air-dried, disaggregated, and sieved to obtain the fraction comprised of particles 250 microns (μ m) or smaller in size, reflecting the fraction of soil that has a greater potential to adhere to children's hands.

The most recent US Environmental Protection Agency (USEPA) guidance (OLEM Directive 9200.1-128; USEPA 2016) recommends sieving to less than (<) 150 μ m. This recommendation is based on "*a growing body of evidence showing that dermally-adhered soil and dust, representative of soil and dust exposure to young children via incidental ingestion, is dominated by particles <150 \mum" (USEPA 2016). The USEPA also cites several studies that suggest lead enrichment occurs in very fine soil fractions (e.g., <63 \mum). There does not appear to be a substantial amount of published data that predict enriched lead concentrations in the <150 \mum fraction compared with the fraction less than or equal to (≤) 250 \mum. Considering this uncertainty, the USEPA previously agreed with the use of the ≤250 \mum fraction for the 2021 sampling program while the particle size enrichment demonstration study described herein was conducted.*

A Study Work Plan was approved by the USEPA in February of 2022 (Environmental Standards and Ramboll 2022). In accordance with the Study Work Plan, a subset (121) of the 1,553 RMAP soil samples sieved to \leq 250 µm and analyzed in 2021 were re-sieved to <150 µm and analyzed for lead and arsenic. Lead and arsenic concentrations in the <150 µm fraction were compared to concentrations in the previously analyzed \leq 250 µm fraction to determine if lead and arsenic enrichment occurs in the <150 µm fraction. Statistical analyses were performed to determine if there was a significant difference between the lead and/or arsenic concentrations in the two size fractions. Percent difference between the two fractions was calculated for lead and arsenic concentrations in each sample pair (i.e., the \leq 250 µm and <150 µm fractions from each sample) as a measure of enrichment. Potential differences in enrichment across concentration ranges and sampled depth intervals were also evaluated.

Statistically significant differences were observed between lead and arsenic concentrations in the \leq 250 µm and <150 µm fractions. For both metals, concentrations were higher in the <150 µm fraction. However, consistent with other studies that have compared lead and arsenic concentrations between similar size fractions, the overall magnitude of concentration differences between the two fractions was relatively low for both metals.

The average percent differences observed for lead and arsenic among all samples (excluding outliers) were 8% and 13%, respectively. The greatest enrichment was observed in samples with lower concentrations of both metals, and in samples with very high arsenic concentrations. Among samples with lead and arsenic concentrations indicative of anthropogenic influences, percent differences averaged 6% for lead and 9% to 11.5% (including/excluding the highest concentrations, respectively) for arsenic.

The depth evaluation indicates that lead and arsenic enrichment in RMAP soil samples generally increases with sample depth from the surface to a depth of 12 inches (deeper samples were not included due to small sample size). For both metals, greater enrichment was consistently observed in the 6 to 12 inch depth interval, with the lowest enrichment among samples collected from 0 to 2 inches. The 0 to 2 inch depth interval represents soil most likely to be regularly contacted by people,

DRAFT

and the enrichment in this fraction was very low. Among 0 to 2 inch depth samples with lead and arsenic concentrations indicative of anthropogenic influences, percent differences averaged -2% for lead and 4% for arsenic.

Overall, the results of this study, along with other comparable studies, indicate a small difference in reliance on the \leq 250 µm versus <150 µm sieve size.

1. INTRODUCTION

1.1 Background

Metal concentrations measured in soil samples may vary when samples are sieved to different size fractions. Available data suggest that higher concentrations are often measured in samples sieved to smaller fractions. There are several reasons why this may occur. Sieving to smaller fractions can result in larger surface area to volume ratios, depending on the particle size distribution in a given sample. Clay minerals, organic matter, and iron, manganese, or aluminum oxides may also be retained in the smaller fractions. In such cases, a greater amount of particle surfaces with sorbed or complexed metals can result in higher metal concentrations being measured in soil samples that are sieved to smaller fractions.

In 2016, the US Environmental Protection Agency (USEPA) issued guidance (USEPA OLEM Directive 9200.1-128) recommending that soil samples used for decision-making related to human health risk be sieved to <150 μ m. This guidance was based on studies that show dermally-adhered soil and dust, representative of soil and dust exposure to young children via incidental ingestion, is dominated by particles less than (<) 150 μ m. In addition, the USEPA states that there is a higher likelihood of lead enrichment in smaller particles. However, the USEPA primarily cites studies demonstrating enrichment in very fine particle sizes (e.g., <63 μ m), and no studies that specifically evaluate if lead is enriched in the <150 μ m fraction compared with the <250 μ m fraction. Of the 20 articles cited in USEPA (2016), only two of those articles (Juhasz et al. 2011 and Tawinteung et al. 2005) included data for sieve sizes close to those of interest. Those two studies, along with one additional study published after the USEPA (2016) directive (Karna et al. 2017), are summarized in Section 1.2 and Table 1 of this report.

1.2 Relevant Studies

Juhasz et al. (2011) evaluated 16 soil samples from a range of mining/smelting sites, shooting ranges, incinerators, a gas works and historical fill areas for particle size fractions of <50, <100, <250, and <2,000 µm. The authors did not specifically evaluate percent difference, a measure of enrichment, in lead concentration between the different particle size fractions, but it can be calculated from the information presented. Using the information presented in Juhasz et al. (2011), lead percent difference was calculated for the $<100 \ \mu m$ fraction vs. the <250 µm fraction. Percent differences between these two fractions ranged from 4% to 63% in 15 samples, with one outlier at 209%. Excluding the outlier, the mean enrichment was 19%. In nine samples with lead concentrations between 160 and 2,300 mg/kg (outlier excluded), a range based on the available samples most relevant to legacy mining sites such as Butte, the enrichment for the <100 μ m vs. <250 μ m size fractions ranged from 4% to 25% with a mean of 10%. Extrapolating the data for these nine samples using linear regression predicts similar, slightly lower enrichment for the $<150 \ \mu m$ vs. <250µm size fractions, with an average of 8% (range of 3% to 17%). Based on the data reported by Juhasz et al. (2011), enrichment of lead concentrations in soil samples in the concentration range of interest (i.e., approximately 200 mg/kg to 2,500 mg/kg) is predicted to be less than 10%.

Tawinteung et al. (2005) evaluated lead concentrations in three particle size fractions (<2,000-250 μ m, <250-150 μ m, <150 μ m) in six soil samples collected from three locations with increasing distance from an abandoned battery recycling/secondary smelting factory at depths of 0-15 centimeters (cm; 0 to almost 6 inches) and 15-20 cm (~6 to ~8 inches)

below the ground surface. The authors did not specifically evaluate percent difference in lead concentrations between the different particle size fractions but present the percent mass of composite soil and lead concentrations by size fraction, so percent difference can be calculated between the <250 μ m and <150 μ m size fractions. Using this information, the <250 μ m lead concentration was first calculated using a weighted average of the <250-150 μ m and <150 μ m concentrations (see Table 1 for details), then used to calculate percent difference. Percent differences ranged from 7.5% to 31% for lead in the <150 μ m fraction vs. the <250 μ m fraction. Percent difference estimates decreased with increasing lead concentrations, with percent difference estimates being noticeably higher in the samples with the lowest lead concentrations. Specifically, percent difference was about 30% in the two soil samples with average lead concentrations of less than 30 mg/kg, whereas average percent difference for the other four soil samples (ranging from approximately 200 mg/kg to 2,500 mg/kg) was 9.7%.

More recently, Karna et al. (2017) compared concentrations measured in sieved particle size fractions of <250 μ m, <150 μ m, <75 μ m, and <38 μ m in mining-impacted, smelter-impacted, lead-arsenate pesticide-impacted, and USGS reference material soils, and found that total arsenic and lead concentrations increased with decreasing particle size. Concentrations measured in the smallest particle size fraction (<38 μ m) showed the greatest differences compared to concentrations in the largest fraction (<250 μ m). Comparatively, concentration differences between the <150 μ m and <250 μ m fractions were smaller. Using data presented in the supplemental material of Karna et al. (2017), percent difference between total arsenic and lead concentrations measured in the <150 μ m and <250 μ m fractions was determined. Values ranged from -1% to 11% for lead and -2% to 14% for arsenic. These ranges were the same if all four samples, or only the three falling in the lead concentration range of interest, were included (for this study, lead concentrations within the concentration range of interest ranged from approximately 350 mg/kg to 2,500 mg/kg). The average percent difference in the lead concentration range of interest was 6% for lead.

The three studies that have assessed soil lead concentrations and particle size including sieve sizes similar to those evaluated in this study yield similar conclusions. All three demonstrate lead concentrations generally increase with decreasing particle size, with the greatest differences observed in the finest fraction (as small as <38 μ m) compared to the largest fraction (250 μ m or greater). Comparatively, concentration differences between the <150 μ m and <250 μ m fractions were smaller. Percent difference estimates for lead between the <250 μ m and <150 μ m size fractions in the concentration range of interest (approximately 200 mg/kg to 2,500 mg/kg lead) were generally 10% or less. Karna et al. (2017) was the only study to look at arsenic enrichment between the <150 μ m and <250 μ m fractions reached using data from these studies indicate lead enrichment between the <150 μ m and <250 μ m size fractions is generally 10% or less. However, because there are only three studies for lead and one for arsenic, and the sample size in each study was small, further study of lead and arsenic concentration differences between the <150 μ m and <250 μ m particle size fractions is warranted.

1.3 Program Description and Oversight

The Butte-Silver Bow (BSB) Multi-Pathway Residential Metals Abatement Program (RMAP) Plan (BSB and Atlantic Richfield Company, 2020) was designed to mitigate exposure of residents to sources of lead and arsenic contamination.

Potential releases of metals to the environment may originate from both mining-related (waste rock, tailings, aerial emissions) and non-mining-related sources (e.g., lead-based paint). Potential sources of lead and/or arsenic exposure addressed by the RMAP include lead and arsenic present in soil and dust.

The USEPA Region 8 and the Montana Department of Environmental Quality are responsible for project oversight, review, and approval of all RMAP-generated sampling data and subsequent site-specific remediation plans.

1.4 Current Study Overview and Data Quality Objectives

The Study Work Plan (Environmental Standards and Ramboll 2022) specified two primary data quality objectives:

- Generate data to determine lead and arsenic concentration differences, if any, between 250- μ m sieved and 150- μ m sieved soil sample aliquots to assess the possibility of enrichment in the finer (i.e., less than or equal to [≤] 150 μ m) soil fraction.
- Investigate the potential for lead and arsenic enrichment in the finer soil fraction at various levels of each metal, including naturally-occurring and anthropogenically-influenced concentrations (i.e., does enrichment occur when concentrations are low, moderate, high, or at all levels?).

Soil samples collected as part of the RMAP have historically been sieved to the $\leq 250 \ \mu m$ fraction¹. To address the question of whether there may be meaningful differences in lead and arsenic concentrations between the $\leq 250 \ \mu m$ and $< 150 \ \mu m$ fractions in Butte soils, this study was completed using soil samples collected from nonresidential parcels such as schools, playgrounds, and sports fields as part of the RMAP.² The study evaluated potential enrichment in the $< 150 \ \mu m$ soil fraction using samples with varying lead concentrations, ranging from $< 10 \ mg/kg$ to $> 5,000 \ mg/kg$ in the $\leq 250 \ \mu m$ fraction. While the study was designed to incorporate a broad range of lead concentrations, enrichment among samples with varying levels of arsenic was also assessed. Arsenic concentrations measured in the samples included in this study ranged from $< 6 \ mg/kg$ to $> 700 \ mg/kg$ in the $\leq 250 \ \mu m$ fraction. This study provides additional data characterizing lead and arsenic concentrations and enrichment in the $< 150 \ \mu m$ versus $\leq 250 \ \mu m$ size fractions, and can be used to inform future RMAP soil sampling protocols.

¹ While other studies specify size fractions less than 250 μ m, the fraction passing through the No. 60 sieve used for this study includes particles less than or equal to 250 μ m. Therefore, the fractions included in this and other studies, and USEPA documents, are described slightly differently. ² In addition to the ongoing residential soil sampling, the RMAP has recently included soil sampling of

² In addition to the ongoing residential soil sampling, the RMAP has recently included soil sampling of non-residential parcels (schools, parks, and non-residential daycare centers).

2. METHODOLOGY

Detailed methodology for the sieve study was presented in the Study Work Plan (Environmental Standards and Ramboll 2022) and is summarized here.

2.1 Sample Collection Designations and Characteristics

A total of 20 schools and one park were sampled during the 2021 RMAP sampling events. Sampling was conducted in accordance with the BSB and Atlantic Richfield Company (2021) RMAP Quality Assurance Project Plan (QAPP) (Non-Residential Parcels).

The land use categories sampled included:

- Land Use Category #1 playground areas
- Land Use Category #2 highly accessible areas/barren sports fields
- Land Use Category #3 maintained grass areas/grass sports fields
- Land Use Category #4 low access areas/low maintenance areas/open space
- Land Use Category #5 flower/vegetable gardens

Soil sampling was conducted at multiple depth intervals (0 to 2 inches, 2 to 6 inches, and 6 to 12 inches) to enable assessment of potential health risks under the different land uses, and to obtain data that were comparable to those from previous sampling efforts. Flower/vegetable garden components were sampled at additional depth intervals of 12 to 18 inches and 18 to 24 inches.

2.2 Sample Selection

A total of 1,553 soil samples were collected from schools, parks and playgrounds within the RMAP Expanded Area during 2021. All RMAP soil samples were air dried, disaggregated, and sieved to obtain the fraction less than or equal to (\leq) 250 microns (µm) prior to analysis for lead and arsenic (see Section 2.3.1 for more detail). Lead concentrations measured in the nonresidential soil samples ranged from 3.8 to 5,220 milligrams per kilogram (mg/kg) and arsenic concentrations ranged from 3.5 to 727 mg/kg. The samples were grouped into lead concentration ranges to help guide sample selection for the sieve study. The lead concentration ranges were selected in consultation with USEPA personnel, with consideration of the BPSOU residential soil action limit for lead (1,200 mg/kg) and the distribution of sample concentrations across the entire data set. The lead concentration ranges, and the number of 2021 nonresidential soil samples with concentrations within each range, are listed below.

- <50 mg/kg: 714 samples
- 50-400 mg/kg: 722 samples
- 400-1,200 mg/kg: 94 samples
- >1,200 mg/kg: 23 samples

All samples with reported lead results >1,200 mg/kg³, and a semi-random selection of samples with concentrations <1,200 mg/kg, were re-sieved to <150 μ m and analyzed for

 $^{^3}$ Two of the 23 samples collected in 2021 that had lead concentrations >1,200 mg/kg were selected for inclusion in the sieve study (see Work Plan Table 1) but were not analyzed for lead and arsenic in the <150 μm fraction due to the laboratory identifying possible mass loss during sample preparation

lead and arsenic in the ${<}150~\mu\text{m}$ fraction. Samples included in the sieve study are listed by concentration range below.

- Less than 50 mg/kg: 18 samples
- 50-400 mg/kg: 41 samples
- 400-1,200 mg/kg: 41 samples
- Greater than 1,200 mg/kg: 21 samples

The semi-random selection of samples in the 400-1,200 mg/kg, 50-400 mg/kg, and <50 mg/kg groups ensured that results from each school and playground and all three depth intervals (0 to 2 inches, 2 to 6 inches, and 6 to 12 inches) were represented in the study. Deeper samples (12 to 18 inches and 18 to 24 inches) were only collected from a limited number of properties where gardens were present. A total of 8 garden samples were randomly selected for this study. Samples that were re-sieved to <150 μ m are listed in the data table provided as Appendix A to this report.

2.3 Analytical Methodologies

2.3.1 Initial Sample Preparation and Analysis: 250 µm aliquot

Sample preparation and analyses were performed in accordance with the USEPA digestion/analytical method specifications, the approved OAPP, and the laboratory's standard operating procedures. Samples were processed (dried and disaggregated using a rolling pin between sheets of paper) at Pace Analytical Laboratories, LLC (Pace) in Green Bay, Wisconsin. Samples were sieved using a No. 60 sieve to obtain the \leq 250 µm fraction. The weights of the >250 μ m and ≤250 μ m fractions were measured and recorded by the laboratory for each soil sample prepared in this manner. The \leq 250 µm fraction was sent under chain of custody to the Pace Minneapolis laboratory. Each sample bag containing the \leq 250 µm fraction was placed flat on a clean laboratory bench top and the air dried and sieved sample was spread-out in the unopened bag to evenly distribute the sample. The sample bags were then gently rolled to further homogenize the sample. This technique provided good sample homogeneity as shaking the sample bag may preferentially cause the finer particles to migrate to the bottom of the bag. After the sample bag was opened, the sample preparation personnel removed a small portion of the sample from several different locations in the bag using a clean metal scoopula and placed a total of 1.0 to 1.1 grams of sample material into a labeled digestion tube. These aliquots were digested for metals analyses according to a modified USEPA Method 3050B. Arsenic and lead concentrations were determined per USEPA Method 6020A (inductively coupled plasma mass spectrometry). The remaining mass of the ≤250 µm sample portion was placed in room temperature archive storage at the Pace Minneapolis laboratory.

2.3.2 Secondary Sample Preparation and Analysis: 150 µm aliquot

The samples designated for re-sieving were removed from laboratory archive storage. The total mass remaining from the initial 250 μ m sieving was calculated by subtracting the weight of the sample aliquot removed for original testing from the weight of the \leq 250 μ m fraction measured prior to analysis. The entire contents of the sample bag were sent back to Pace Green Bay for re-sieving to <150 μ m using a No. 100 sieve; no additional sample

steps (loss during analyst handling and transfer of the sample for digestion). Because these samples (S-S-0015-GA5-3 and S-S-0013-PA1-3) were not analyzed in the sieve study, they are not included in the study sample counts presented in this section.

preparation was performed. The weights of the >150 µm (\leq 250 µm) and <150 µm re-sieved portions were recorded to provide an indication of the sample size fractionation. Once resieved, the <150 µm sample aliquots were returned to Pace Minneapolis for digestion and analysis.

Consistent with the previously analyzed $\leq 250 \ \mu m$ fraction samples, approximately 1 gram of the $<150 \ \mu m$ fraction samples were digested according to a modified USEPA Method 3050B, and arsenic and lead concentrations were determined per USEPA Method 6020A.

2.3.3 Quality Control

Quality assurance/quality control (QA/QC) was performed in accordance with the QAPP (BSB and Atlantic Richfield Company 2021) for the <150 μ m re-sieved sample aliquots using the same procedures as originally performed for the ≤250 μ m samples, as described in the Study Work Plan (Environmental Standards and Ramboll 2022). Laboratory QC samples were analyzed in addition to the calibration samples with each QC batch. Laboratory QC samples are introduced into the measurement process to evaluate laboratory performance and sample measurement bias.

Laboratory blanks, laboratory control samples, analytical duplicates, serial dilutions, and pairs of matrix spike/matrix spike duplicate samples were analyzed in each laboratory QC batch with a minimum frequency of one set for each sample delivery group of a maximum of 20 field samples.

2.3.4 Data Validation

Environmental Standards, Inc. (Environmental Standards) performed Stage 4 analytical quality assurance reviews of the original and re-sieved sample data with guidance from the QAPP (BSB and Atlantic Richfield Company 2021), the *Guidance for Labeling Externally Validated Laboratory Analytical Data for Superfund Use* (USEPA 2009), and the *National Functional Guidelines for Inorganic Superfund Methods Data Review* (National Functional Guidelines; USEPA 2020). The National Functional Guidelines validation guidance documents specifically address analyses performed in accordance with the Contract Laboratory Program analytical methods and are not completely applicable to the type of analyses and analytical protocols performed for the SW-846 methods utilized by the laboratory for these samples.

Data were examined to determine the usability of the analytical results and compliance relative to the method requirements specified in *Test Methods for Evaluating Solid Waste, Physical/Chemical Methods, SW-846, 3rd Edition* (SW-846) Method 6020A (USEPA 2005). Environmental Standards used professional judgment to determine the usability of the analytical results and compliance relative to the methods utilized by the laboratory. All quality assurance measures stipulated in the QAPP (BSB and Atlantic Richfield Company 2021) were met for the re-sieved sample data with the exception of holding time, and matrix spike/matrix spike duplicate (MS/MSD) recoveries associated with a few samples, as described below. Data validation reports for re-sieved sample data are provided in Appendix B. Data validation reports for the original \leq 250 µm sample data were provided with the reports for each sampled location.

The holding time for arsenic and lead in soil samples is 6 months in accordance with Table 3-2 of SW-846 Chapter 3. Stability of metals may be extended based on USEPA approval. The reported positive results for arsenic and lead in all except four⁴ of the re-sieved samples

⁴ Four sieve study samples (S-S-0009-GA2-SA1-1, -2, -3, and -D-1) collected in November of 2021 were analyzed within the 6 month holding time and therefore not qualified for missed hold time.

should be considered estimated and have been flagged "J" in the qualified electronic data deliverable. The samples were reanalyzed greater than 6 months from sample collection. However, the long-term stability of metals in soil samples (especially those that are dried and sieved) should be considered when evaluating the effects of this qualification on the study data. Additionally, lead and arsenic results for a few samples have been flagged "J+" due to MS/MSD recoveries above the acceptable range. Results for these samples should also be considered estimated and may be biased high.

3. DATA ANALYSIS

Validated analytical data were used to assess relationships between lead and arsenic concentrations measured in soil samples sieved to $\leq 250 \ \mu m$ and $\leq 150 \ \mu m$ and evaluate the effects of sieving to $\leq 150 \ \mu m$ instead of $\leq 250 \ \mu m$. Various analyses completed to evaluate these relationships and influence of sieve size are described in the following sections. Results of the analyses described herein are presented in Section 4.

3.1 Measurement of Enrichment

Percent difference was calculated to assess potential enrichment in the <150 μm fraction compared to the <250 μm fraction for lead and arsenic. Percent difference was calculated for both analytes using the following formula:

$$\left[\left(\frac{<150 \ \mu m \ Result}{\le 250 \ \mu m \ Result}\right) - 1\right] \times 100$$

Percent differences greater than zero occur when the analyte concentration in the <150 µm fraction is greater than the analyte concentration in the ≤250 µm fraction. Therefore, positive percent differences indicate some level of enrichment in the <150 µm fraction relative to the ≤250 µm fraction; a smaller percent difference indicates less enrichment, whereas a larger percent difference indicates more enrichment, in the finer fraction. Percent differences less than zero occur when the analyte concentration in the <150 µm fraction is less than the analyte concentration in the ≤250 µm fraction. Thus, negative percent differences indicate that there is not enrichment in the <150 µm fraction relative to the ≤250 µm fraction.

3.2 Outlier Analysis

Graphical and statistical methods were used to assess potential outliers in the lead and arsenic percent difference datasets. Quantile-quantile (Q-Q) plots were generated using ProUCL[®] software and examined to identify any results appearing notably different from the majority of the data distribution. Following the visual evaluation, the Rosner test was completed using ProUCL[®] to further evaluate the potential presence of outliers. Due to the potential for masking, which can occur when some outliers are obscured by other, more extreme outliers, the tests were run in an iterative fashion by removing any values initially identified as outliers, and then repeating the test to identify any additional outliers. Iterative testing continued until no additional outliers were identified. The Rosner test assumes that the underlying dataset (excluding outliers) is normally distributed. Following outlier removal, goodness of fit testing was therefore completed using ProUCL[®]. Finally, professional judgement was used to interpret the Q-Q plots and Rosner test results for final outlier determinations.

3.3 Evaluation of Enrichment by Concentration

Lead and arsenic data were grouped by concentration ranges, and percent difference was further evaluated across concentration ranges to assess potential differences in enrichment at various (including naturally-occurring and anthropogenically-influenced) levels of the two metals. Concentrations measured in the original ($\leq 250 \ \mu m$) sample analysis were used to group the data.

Lead data were grouped into the concentration ranges used to select samples for the sieve study:

• <50 mg/kg: 18 samples</p>

- 50-400 mg/kg: 41 samples
- 400-1,200 mg/kg: 41 samples
- >1,200 mg/kg: 21 samples

Arsenic data were grouped based on the data distribution (natural sample groupings and percentiles), with consideration of the action level, similar to how the lead concentration ranges were established. The natural background concentration established for the State of Montana (22.5 mg/kg; Hydrometrics, Inc. 2013), also informed the grouping to assess possible differences between arsenic in natural and anthropogenically influenced soils. Initially, arsenic data were grouped into four concentration ranges. Based on the initial analysis, samples in the highest arsenic concentration range (>100 mg/kg) were further grouped into two ranges, 100-250 mg/kg and >250 mg/kg, to examine potential differences above and below the action level.

Arsenic data were grouped into the following concentration ranges:

- <22.5 mg/kg: 11 samples
- 22.5-60 mg/kg: 50 samples
- 60-100 mg/kg: 38 samples
- 100-250 mg/kg: 13 samples
- >250 mg/kg: 9 samples

3.4 Evaluation of Enrichment by Depth Interval

Percent differences for lead and arsenic were also compared between the different sampled depth intervals. Data summary statistics and Q-Q plots were evaluated to assess potential differences in enrichment among samples collected from depth intervals of 0 to 2 inches, 2 to 6 inches, and 6 to 12 inches. Samples collected from depth intervals of 12 to 18 inches and 18 to 24 inches were not included in this evaluation due to low sample counts (only eight samples total from these two depths intervals were included in the study).

3.5 Effect of Sieve Size on Action Level Comparisons

Analyte concentrations in soil samples are compared to action levels (1,200 mg/kg for lead, 250 mg/kg for arsenic) to assess the need for remedial action. For samples with lead and arsenic concentrations measured in the \leq 250 µm fraction below the respective action levels, concentrations measured in the <150 µm fraction were evaluated to assess the effect of using the smaller sieve size in the context of remediation decisions.

3.6 Effect of Sample Size Fractionation

Percent differences were compared with consideration of sample size fractionation to examine if there was any apparent relationship between the distribution of sample mass between the \leq 250 µm and <150 µm fractions and enrichment.

3.7 Statistical Analysis

In addition to percent difference calculations, hypothesis testing was used to evaluate potential differences in metals concentrations between soil samples sieved to <150 μ m and <250 μ m. Analytical results were compared to determine if there is a statistically significant difference in lead and/or arsenic concentrations between the two particle size fractions. Each sample was sieved, analyzed, then re-sieved and analyzed to measure concentrations in the two different particle size fractions; the results therefore represent paired datasets that are

not independent of one another. The goal of the statistical analysis was to identify differences based on particle size and not influenced by other factors such as location. To meet this goal, a paired t-test was used for the comparisons.

Parametric hypothesis tests are appropriate for comparing datasets that follow a normal distribution, while nonparametric methods do not require datasets to meet the assumption of normality. To select the appropriate test for this analysis, goodness-of-fit testing was therefore completed using the USEPA's ProUCL[®] software prior to hypothesis testing. Both the Lilliefors and Shapiro-Wilk tests were used to assess normality for each of the four analytical datasets (lead / arsenic concentrations in \leq 250 µm and <150 µm fractions). None of the datasets were normally distributed.

Paired analytical results for the \leq 250 µm and <150 µm sieved fractions of each sample were statistically compared using the Wilcoxon Rank Sum (WRS) test, which is the nonparametric equivalent of a paired t-test, using R version 4.1.0. The statistical analysis tested the null hypothesis that lead or arsenic concentrations in the <150 µm fraction are equal to the lead or arsenic concentrations in the <150 µm fraction are equal to the lead or arsenic concentrations in the <150 µm fraction are equal to the lead or arsenic concentrations in the <150 µm fraction are equal to the lead or arsenic concentrations in the <150 µm fraction are equal to the lead or arsenic concentrations in the <250 µm fraction, using a two-sided test. Alpha (a) was set equal to 0.05; if the p-value was less than 0.05, the null hypothesis was rejected. If the null hypothesis was rejected, i.e., the two-sided test indicated that the two populations are not equal, then a one-sided test was used to test the null hypothesis that analyte concentrations in the <250 µm fraction are less than or equal to concentrations in the <150 µm fraction.

4. **RESULTS**

The results of the analyses described in Section 3 are presented here. The outlier analysis results are presented first. Unless otherwise specified, other analyses were completed using datasets with outliers removed, and results described in subsequent sections exclude outliers.

4.1 Outlier Analysis

As described in Section 3.2, both a graphical evaluation and formal outlier testing were used to assess the presence of potential outliers in the lead and arsenic percent difference datasets. The Q-Q plots for lead and arsenic percent difference, shown on Figures 1 and 2 respectively, indicated the presence of five potential outliers for lead, and one potential outlier for arsenic. Because these results appeared notably separate from the remainder of their respective data distributions, they were further evaluated using the Rosner test as described in Section 3.2. The Rosner test confirmed that these values were statistical outliers. Following removal of the five lead and one arsenic percent difference outliers, the Rosner test was repeated and confirmed that no additional outliers were present. The lead percent difference dataset was determined to be normally distributed following outlier removal, based on both normality tests in ProUCL® (Lilliefors and Shapiro-Wilk). For arsenic, percent difference values were not normally distributed either before or after outlier removal. However, based on the visual evaluation and confirmation through the statistical outlier test, one value in the percent difference dataset for arsenic was concluded to be an outlier, and no other potential outliers were identified. Percent difference values identified as outliers are described below and summarized in Table 2. ProUCL[®] output is provided in Appendix C.

For arsenic, a single percent difference value of 210% was identified as an outlier. Three negative percent difference results and two positive results were identified as outliers for lead. The negative outliers ranged from -92% to -95%, meaning that very little of the lead mass in these samples remained in the finer (<150 μ m) fraction after re-sieving. Although they stand out when compared to the majority of the dataset, these results don't appear to represent errors.

The positive results identified as lead percent difference outliers were 394% and 754%. These results indicate that the amount of lead present in the finer fraction after re-sieving was substantially greater (close to 400% or more) compared to the lead mass present in the original sample (\leq 250 µm fraction prior to re-sieving). Based on the total amount of lead that could have been present in each sample (estimated based on sample masses and concentrations measured in the \leq 250 µm fraction), the lead concentrations measured in the <150 µm fractions of these two samples do not appear plausible, suggesting that potential measurement or other errors may have occurred. A review of the data packages did not identify any obvious discrepancies or potential explanation.

Because these results did not appear plausible and could not be explained, samples S-S-0012-G1-4 and S-S-0014-HA2-1 were selected for reanalysis. The same methods used for the original sample analysis were applied to new aliquots sampled from the remaining <150 μ m re-sieved mass for the two reanalyzed samples. Because all of the mass that remained after the original sample analysis was re-sieved to <150 μ m, new aliquots representing the original <250 μ m fraction could not be sampled. Instead, aliquots were sampled from the remaining mass that did not pass through the No. 100 sieve, representing particle sizes <250 μ m, but >150 μ m. Because this size fraction is not directly comparable to the <250 μ m fraction, evaluation of the reanalysis results focused on the <150 μ m fraction. Original

and reanalysis results are summarized in Appendix D; for completeness, results are presented for both reanalyzed fractions.

The reanalysis results for the <150 µm fraction were similar to the original results for this fraction, and generally confirmed the percent differences calculated using the original paired concentrations. The reanalysis did not provide any further clarification regarding the extreme differences observed between lead concentrations in the \leq 250 µm and <150 µm fractions of these two outlier samples. Therefore, analyses described below (results) and in Section 3 of this report (methods) were completed using original sample results. Because the two samples selected for reanalysis were identified as outliers for lead percent difference, they were excluded from the lead data analyses along with other lead outliers as discussed below. The arsenic percent differences observed in the two samples selected for reanalysis were consistent with other arsenic results (i.e., not identified as outliers), thus the original results for these two samples were retained for further arsenic data analyses.

The five lead outliers shown in Table 2 were removed from the lead dataset, and the single arsenic outlier (Table 2) was removed from the arsenic dataset, before the additional data analyses described in Section 3 (methods) and the subsections below (results) were completed.

4.2 Enrichment Summary

Percent differences for lead and arsenic are summarized across all samples and by concentration ranges in Table 3. Percent differences for lead ranged from -58% to 84% across all samples, with an overall average of 8%. Greater enrichment was observed among samples with the lowest lead concentrations (<50 mg/kg), with an average percent difference of 21%. Among samples with higher concentrations that are more likely to be associated with anthropogenic, rather than naturally occurring, sources of lead, overall enrichment was lower with an average percent difference of 6%.

Arsenic showed similar results to lead, with an overall average percent difference of 13% and values ranging from -84% to 113%. Similar to lead, greater enrichment was observed in samples with lower arsenic concentrations. The average percent difference among samples with arsenic concentrations below the Montana background concentration of 22.5 mg/kg was 23%. Among samples with arsenic concentrations in the ranges more likely to be associated with anthropogenic sources, enrichment was lower except among the highest concentrations. The average percent difference for samples with arsenic concentrations between background and the action level (i.e., 22.5 to 250 mg/kg) was 9%. In contrast, samples with arsenic concentrations above the action level of 250 mg/kg showed the greatest enrichment, with an average percent difference of 38% across these nine samples.

4.3 Depth Comparison

A comparison of enrichment between soil samples collected from different depth intervals is provided in Table 4 and on Figures 3 and 4 for lead and arsenic, respectively. As shown on the Q-Q plots, enrichment was generally greatest in soil samples collected from depths of 6 to 12 inches, and lowest in samples collected from the 0 to 2 inch depth interval, for both lead and arsenic. This observation is consistent with the average percent differences shown in Table 4 for each depth interval. While this pattern was generally consistent for both metals across the percent difference distributions for the three depth intervals evaluated, some deviation was observed for both metals at the high end of the distributions. Specifically, the highest overall percent difference values for lead and arsenic corresponded to samples collected from the 2 to 6 inch depth interval. Two samples collected from 0 to 2 inches also

had some of the highest percent differences for arsenic. However, overall, lead and arsenic enrichment appears to increase with depth.

The 0 to 2 inch depth interval represents soil most likely to be regularly contacted by people, and the enrichment in this fraction was very low. For samples with lead concentrations >50 mg/kg in the \leq 250 µm fraction, average concentrations in the <150 µm fraction were actually lower (enrichment = -2%). For samples with arsenic concentrations >22.5 mg/kg in the \leq 250 µm fraction, average concentrations in the <150 µm fraction were 4% higher. When all samples were included enrichment was still low for the 0 to 2 inch depth interval (0.67% for lead and 6% for arsenic).

4.4 Effect of Sieve Size Relative to Action Levels

For samples with lead and arsenic concentrations measured in the $\leq 250 \ \mu m$ fraction below their respective action levels, concentrations measured in the $< 150 \ \mu m$ fraction were evaluated to assess the effect of using the smaller sieve size in the context of remediation decisions. Figures 5 and 6 show the effects of changing sieve sizes for lead and arsenic, respectively. For both metals, the outcomes of action level comparisons were minimally affected by the change in sieve size. Details are described below.

For lead, 95 of the 116 samples sieved to \leq 250 µm that were not identified as outliers had concentrations <1,200 mg/kg. Of these, only eight samples had concentrations >1,200 mg/kg measured in the <150 µm fraction. Of those eight samples, two had arsenic concentrations measured above the action level in the \leq 250 µm fraction, meaning that overall remedial action decision outcomes would be affected by sieving to a finer fraction only 6% of the time (i.e., six of 95 samples).

Effects of sieving to the finer fraction were negligible for arsenic; of the 111 non-outlier samples with concentrations <250 mg/kg measured in the <250 μ m fraction, only two had concentrations ≥250 mg/kg measured in the <150 μ m fraction (one equal to and one >250 mg/kg). Both samples also had lead concentrations above the action level in one or both fractions, thus overall conclusions would not change based on use of the smaller sieve size for arsenic analysis.

4.5 Sample Size Fractionation

A comparison of the measured weights of each re-sieved sample fraction comprised of particles measuring <150 μ m versus those measuring between >150 and 250 μ m showed that overall, the finer fraction made up a greater proportion of sample mass. About 80% of samples contained a greater proportion of mass in the finer fraction. The percentage of total mass (calculated as the sum of weights of re-sieved fractions) comprised of particles <150 μ m ranged from 27% to 82% with an average of 60%. Despite the consistent trend of finer particles comprising more of the overall sample mass, preliminary evaluations showed no apparent relationship between sample size fractionation and enrichment. Therefore, the potential influence of sample size fractionation was not further evaluated.

4.6 Statistical Analysis

The results of hypothesis tests comparing lead and arsenic concentrations in paired samples sieved to $\leq 250 \mu m$, and then to $< 150 \mu m$, are shown in Table 5 with additional supporting information provided in Appendix E. Graphical comparisons of lead and arsenic concentrations in the paired sample fractions are shown on Figures 7 and 8, respectively.

Results of the two-tailed tests indicate that for both analytes there is a statistically significant difference between concentrations measured in each fraction. Specifically, the p-values

(0.002 for lead and 0.00001 for arsenic) were less than the specified value of a (0.05). Thus, the null hypothesis that lead or arsenic concentrations in the <150 μ m fraction are equal to the lead or arsenic concentrations in the <250 μ m fraction was rejected, indicating that concentrations are not equal between the two fractions.

Since the null hypothesis of the two-tailed WRS test was rejected for both metals, additional one-tailed tests were used to determine which fraction contained greater concentrations of each metal. The null hypothesis that lead or arsenic concentrations in the \leq 250 µm fraction are less than or equal to concentrations in the <150 µm fraction was accepted based on p-values greater than a (0.999 for lead and 1 for arsenic). These results indicate that the <150 µm fraction contains greater concentrations of both metals compared to the \leq 250 µm fraction.

5. CONCLUSIONS

This study compared concentrations of lead and arsenic measured in soil samples analyzed after sieving to $\leq 250 \mu$ m, then re-sieving to $< 150 \mu$ m, to evaluate potential enrichment in the $< 150 \mu$ m fraction relative to the $\leq 250 \mu$ m fraction. While hypothesis testing showed a statistically significant difference between lead and arsenic concentrations between the two fractions, with higher concentrations in the finer fraction, overall enrichment is relatively low. Among samples with concentrations indicating a likelihood of anthropogenic influence, an average enrichment of 6% for lead, and 9% for arsenic⁵, was measured. In the 0 to 2 inch depth interval most likely to be regularly contacted by people, among samples with lead and arsenic concentrations indicative of anthropogenic influences, percent differences averaged -2% for lead and 4% for arsenic.

The conclusions of this study are consistent with those of other studies that have assessed soil lead concentration and particle size and included sieve sizes similar to those evaluated herein. An independent evaluation of data from those studies (Juhasz et al. 2011; Tawinteung et al. 2005; Karna et al. 2017) resulted in percent difference estimates for lead between the <250 µm and <150 µm size fractions, within the concentration range of interest relative to this study (approximately 200 mg/kg to 2,500 mg/kg lead), that were generally 10% or less. One of these studies (Karna et al. 2017) also evaluated arsenic and found concentration differences between fractions similar to those observed in this study. Independent calculations using arsenic data from Karna et al. (2017) yield an average percent difference of 8%. It should be noted that arsenic concentrations in all four of those samples were quite high relative to the current study.

Differences in lead and arsenic concentrations measured in RMAP soil samples after sieving to ≤ 250 µm versus < 150 µm are consistent with typical variability observed among analytical soil sample data for metals. As an example, the acceptable criteria for laboratory and field duplicate sample precision under the RMAP are 20% and 35% relative percent difference, respectively (BSB and Atlantic Richfield Company 2021). Additionally, the equation used to calculate relative percent difference results in lower values compared to the percent difference calculations used for this study. In the context of data quality criteria, the percent differences observed between lead and arsenic concentrations in the ≤ 250 µm and < 150 µm fractions are relatively low.

In conclusion, the results of this and other comparable studies indicate a small difference in reliance on the \leq 250 µm versus <150 µm sieve size.

⁵ Excludes nine samples with very high (>250 mg/kg) arsenic concentrations; the average enrichment among all samples with arsenic concentrations greater than background is 11.5%.

6. **REFERENCES**

- Butte Silver Bow County (BSB) and Atlantic Richfield Company, 2021. Residential Metals Abatement Program (RMAP) Quality Assurance Project Plan (QAPP) (Non-Residential Parcels). June.
- BSB and Atlantic Richfield Company, 2020. Revised Final Multi-Pathway Residential Metals Abatement Program (RMAP) Plan. Priority Soils Operable Unit Silver Bow Creek/Butte Area, National Priorities List Site, Butte, Montana. Butte-Silver Bow County and Atlantic Richfield Company, November 2020.
- Environmental Standards, Inc. (Environmental Standards) and Ramboll. 2022. Study Work Plan, Comparison of Sieved Soil Sample Results at Lead Site for Assessment of Incidental Ingestion: 250 µm and 150 µm Sieved Samples for Lead Enrichment Determination. February 10.
- Hydrometrics, Inc. 2013. Project Report, Background Concentrations of Inorganic Constituents in Montana Surface Soils. Prepared for Montana Department of Environmental Quality, in collaboration with Energy Laboratories, Inc. September.
- Juhasz AL, Weber J, Smith E. 2011. Impact of Soil Particle Size and Bioaccessibility on Children and Adult Lead Exposure in Peri-Urban Contaminated Soils. Journal of Hazardous Materials, 186: 1870-1879.
- Karna RR, Noerpel M, Betts AR, Scheckel KG. 2017. Lead and Arsenic Bioaccessibility and Speciation as a Function of Soil Particle Size. Journal of Environmental Quality, 46(6): 1225-1235. DOI 10.2134/jeq2016.10.0387.
- Pace Analytical. 2021. Standard Operating Procedure "Soil Sieve" ENV-SOP-GBAY-0164. April.
- Tawinteung N, Parkpian P, DeLaune RD, Jugsujinda A. 2005. Evaluation of Extraction Procedures for Removing Lead from Contaminated Soil. Journal of Environmental Science and Health, 40:2, 385-407.
- US Environmental Protection Agency (USEPA). 2020. Contract Laboratory Program National Functional Guidelines for Inorganic Superfund Data Review. November.
- USEPA. 2016. Recommendations for Sieving Soil and Dust Samples at Lead Sites for Assessment of Incidental Ingestion. OLEM Directive 9200.1-128. Assessment and Remediation Division, Office of Superfund Remediation and Technology Innovation. July.
- USEPA. 2009. Guidance for Labeling Externally Validated Laboratory Analytical Data for Superfund Use. OSWER No. 9200.1-85, EPA 540-R-08-005, 13. January.
- USEPA. 2005. Test Methods for Evaluating Solid Waste Physical/Chemical Methods, SW-846, Third Edition. 2005 Compendium Method 3050B and Method 6020A.

TABLES

- Table 1. Summary of Juhasz et al. (2011), Tawinteung et al. (2005), and Karna et al.(2017) Studies
- Table 2. Samples Identified as Percent Difference Outliers
- Table 3. Lead and Arsenic Enrichment Summary
- Table 4. Lead and Arsenic Enrichment by Depth Interval
- **Table 5. Statistical Analysis Results**

FIGURES

- Figure 1. Lead Percent Difference
- Figure 2. Arsenic Percent Difference
- Figure 3. Lead Percent Difference by Depth Interval
- Figure 4. Arsenic Percent Difference by Depth Interval
- Figure 5. 150 µm vs 250 µm Sieved Lead Result
- Figure 6. 150 µm vs 250 µm Sieved Arsenic Result
- Figure 7. Paired Lead Concentrations in Samples After Sieving to \leq 250 µm and <150 µm
- Figure 8. Paired Arsenic Concentrations in Samples After Sieving to ≤250 µm and <150 µm

APPENDICES

- DRAFT
- Appendix A Sieve Study Data
- Appendix B Data Validation Reports
- Appendix C ProUCL Output Outlier Testing
- Appendix D Reanalysis Summary
- Appendix E Statistical Analysis

DRAFT

APPENDIX A SIEVE STUDY DATA

DRAFT

APPENDIX B DATA VALIDATION REPORTS

Sieve Study Report December 2022

APPENDIX C PROUCL OUTPUT – OUTLIER TESTING

DRAFT

APPENDIX D REANALYSIS SUMMARY

DRAFT

APPENDIX E STATISTICAL ANALYSIS