Report on Calculation of 95% UCL Action Levels from USSL Zone 1 Replicates

USEPA FIELDS Group
Chuck Roth
John Canar
2/10/2021

Samples (5 point composites) on 55 Decision Units (DU) in the Zone 1 portion of the USSL Superfund Site were taken to assess the extent of clean-up. Each DU is to be compared to action levels of 8oo mg/kg Lead and 26 mg/kg Arsenic for the top 12 inches of soil. Nine DUs were sampled in triplicate to develop a 95% confidence level for the 800Pb/26As action levels. This report explains the logic and calculations used to develop the 95% CI for the action levels (map of decision units, sample locations, and triplicates in Figure 1).

The data for analysis includes 9 sets of triplicates with varying concentrations of Lead and Arsenic, resulting in 27 results (Table 1). Each of the 9 triplicate sets are independent of each other, whereas the triplicates themselves are highly correlated (same DU with sub-samples about 1 meter apart), which must be accounted for in the analysis. This analysis will use the differences from the mean of each triplicate to provide the variance in the replicates (Table 1, column 4).

The importance of collecting replicates for a subset of Decision Units is to have a method to estimate the variability in all of the DUs. The variability within each DU is assumed to be similar to each other and the replicates representative of the variability associated with the sample design. The sampling variance is associated with the inherent heterogeneity of the contaminant in the soil, the number of subsamples taken, and the methods of sample preparation designed to provide a precise measure of the average concentration in each DU. For this analysis the variance is pooled for each of the triplicate sets and the final standard deviation is used to estimate the 95% Confidence Interval for all of the DUs. The standard formula for estimating the 95% CI calls for the Standard Error (Standard Deviation/SQR RT Sample Size) times the t value (for n degrees of freedom). When calculating the difference from the mean for each replicate the sample size is accounted for in the Standard Error and the formula for the 95% CI becomes the Standard Deviation (of the differences) times the t value.

One of the problems with this (or any other) method for developing the confidence interval is that the variance for each triplicate is likely to be different based on the concentration range of the triplicates. Two approaches work to reduce the effects of different variances, LN transformation and limiting the range of concentrations used in the confidence interval estimate. Since the variability (and the 95% CI) is being determined for specific action levels, the data should be limited to results closer to the action levels. Typically results with concentrations near zero (non-detects) have little or no variability and should be removed (there were no non-detects with these data) and also, possibly, results much higher than the action limits. Each will likely have much lower or greater variability than the results nearer the action levels. This makes intuitive sense as well, since the primary concern is to determine what the variability is when results (of non-replicated samples) are when near the concentration where a decision needs to be made (the action level). In this analysis, outlier tests on the absolute differences were performed to determine which sets of replicates should be removed.

Since an assumption of normality is also required, the resulting difference from the mean of replicates needs to be tested for normality and transformed if needed. The tests for outliers and normality that are available in the ProUCL software were used for this report. The combination of these two procedures is expected to address the parametric assumptions of equal variances and normal distributions.

Outlier Tests

The outlier test on the absolute differences found DU-C6 to have an outlier for Lead and Arsenic and DU-E9 to have an outlier for Arsenic. Although only one of the three replicates in each case was found to be an outlier, all three replicates were removed (since they are dependent on the same calculation of the difference from the mean). The results of the outlier tests are found in Table 5.

Tests for Normality and Lognormality

Tests of normality and lognormality on the difference between means found each dataset to be consistent with a lognormal distribution. The data were then LN transformed for calculating the 95% CI and then back-transformed to determine the appropriate action level based on the 95% CI. The results of the goodness of fit tests are found in Table 4.

Calculated Action Levels

Arsenic concentrations from this sampling ranged from 5 mg/kg to 1,000 mg/kg. Lead concentrations ranged from 190 mg/kg to 14,000 mg/kg. Although about 15% of the DUs were sampled for triplicates only a few DUs provided results near the action limits (26/800 mg/kg), complicating the final replicate dataset. As a result, a wider range above the action levels needed to be used in the analysis. Additionally, there tended to be some large differences within each triplicate further necessitating a data transform. The natural log transformation for Arsenic provides a 95% CI about 25% of the action level of 26 mg/kg which compares well to the results of triplicate samplings at other sites. However, the Lead results may have underestimated the confidence interval as a result of the natural log transformation. The resulting 95% CI for Lead is only 1.6% the action level of 800 mg/kg. This triplicate dataset highlights the importance of collecting a sufficient number of composite subsamples to reduce the measured variability with each DU. When heterogeneity is expected to be high, a higher number of subsamples per composite is needed. Additionally, data processing procedures (collecting equal volumes per subsample, mixing, and drawing representative samples from the composited soil) are often overlooked and should be considered key to obtaining defensible results. Tables 1 and 2 show the calculation process for creating action levels for Lead and Arsenic based on the 95% CI of the triplicate results. A composite sample from the Zone 1 decision units with results less than 19 mg/kg for Arsenic and 787 mg/kg for Lead can be considered to be less than the action levels of 26 mg/kg As and 800 mg/kg Pb with 95% confidence.

Figure 2 and Table 3 show the decisions for each DU based on the above criteria. Thirteen of the 55 DUs are below the action levels for Pb and As. The highest concentration for As in these DUs is 14 mg/kg and 720 mg/kg for Pb. The highest Arsenic concentration remaining (14 mg/kg) is 46% lower than the action limit of 26 mg/kg and the highest Lead concentration remaining (720 mg/kg) is 10% lower than the action limit of 800mg/kg.

Appendix

Table 1. Results from nine decision unit triplicate analyses.

ID	As	AsMean	AsAbsDiff	LnAsAbsDiff	Pb	PbMean	PbAbsDiff	LnPbAbsDiff
DU-A2	9.5	11.17	1.67	0.51	500	530.00	30.00	3.40
DU-A2-TE	11	11.17	0.17	-1.79	520	530.00	10.00	2.30
DU-A2-TS	13	11.17	1.83	0.61	570	530.00	40.00	3.69
DU-A8:	56	48.00	8.00	2.08	2,300	1800.00	500.00	6.21
DU-A8-TE	42	48.00	6.00	1.79	1,700	1800.00	100.00	4.61
DU-A8-TS	46	48.00	2.00	0.69	1,400	1800.00	400.00	5.99
DU-B4	5.0	6.37	1.37	0.31	1,600	1253.33	346.67	5.85
DU-B4-TE	5.9	6.37	0.47	-0.76	1,500	1253.33	246.67	5.51
DU-B4-TS	8.2	6.37	1.83	0.61	660	1253.33	593.33	6.39
DU-C6	180	78.67	101.33	4.62	1,500	4300.00	2800.00	7.94
DU-C6-TE	24	78.67	54.67	4.00	1,400	4300.00	2900.00	7.97
DU-C6-TS	32	78.67	46.67	3.84	10,000	4300.00	5700.00	8.65
DU-D5	15	11.83	3.17	1.15	4,500	2590.00	1910.00	7.55
DU-D5-TE	11	11.83	0.83	-0.18	2,300	2590.00	290.00	5.67
DU-D5-TS	9.5	11.83	2.33	0.85	970	2590.00	1620.00	7.39
DU-E3	11	11.67	0.67	-0.41	5,400	2700.00	2700.00	7.90
DU-E3-TE	14	11.67	2.33	0.85	1,200	2700.00	1500.00	7.31
DU-E3-TS	10	11.67	1.67	0.51	1,500	2700.00	1200.00	7.09
DU-E9	450	673.33	223.33	5.41	10,000	10800.00	800.00	6.68
DU-E9-TE	570	673.33	103.33	4.64	8,400	10800.00	2400.00	7.78
DU-E9-TS	1,000	673.33	326.67	5.79	14,000	10800.00	3200.00	8.07
DU-F1	47	34.00	13.00	2.56	2,000	1723.33	276.67	5.62
DU-F1-TE	37	34.00	3.00	1.10	2,400	1723.33	676.67	6.52
DU-F1-TS	18	34.00	16.00	2.77	770	1723.33	953.33	6.86
DU-F7	13	13.33	0.33	-1.10	610	513.33	96.67	4.57
DU-F7-TE	17	13.33	3.67	1.30	620	513.33	106.67	4.67
DU-F7-TS	10	13.33	3.33	1.20	310	513.33	203.33	5.31

Table 2. Calculations for Action Levels based on 95% Confidence limits of the differences from the mean.

ID	As	AsMean	AsAbsDiff	LnAsAbsDiff
DU-A2	9.5	11.17	1.67	0.51
DU-A2-TE	11	11.17	0.17	-1.79
DU-A2-TS	13	11.17	1.83	0.61
DU-A8:	56	48.00	8.00	2.08
DU-A8-TE	42	48.00	6.00	1.79
DU-A8-TS	46	48.00	2.00	0.69
DU-B4	5.0	6.37	1.37	0.31
DU-B4-TE	5.9	6.37	0.47	-0.76
DU-B4-TS	8.2	6.37	1.83	0.61
DU-D5	15	11.83	3.17	1.15
DU-D5-TE	11	11.83	0.83	-0.18
DU-D5-TS	9.5	11.83	2.33	0.85
DU-E3	11	11.67	0.67	-0.41
DU-E3-TE	14	11.67	2.33	0.85
DU-E3-TS	10	11.67	1.67	0.51
DU-F1	47	34.00	13.00	2.56
DU-F1-TE	37	34.00	3.00	1.10
DU-F1-TS	18	34.00	16.00	2.77
DU-F7	13	13.33	0.33	-1.10
DU-F7-TE	17	13.33	3.67	1.30
DU-F7-TS	10	13.33	3.33	1.20
			StDev	1.13
			t=1.725	1.96
			antilog	7.07
			Action Level	**18.93**

ID	Pb	PbMean	PbAbsDiff	LnPbAbsDiff
DU-A2	500	530.00	30.00	3.40
DU-A2-TE	520	530.00	10.00	2.30
DU-A2-TS	570	530.00	40.00	3.69
DU-A8:	2,300	1800.00	500.00	6.21
DU-A8-TE	1,700	1800.00	100.00	4.61
DU-A8-TS	1,400	1800.00	400.00	5.99
DU-B4	1,600	1253.33	346.67	5.85
DU-B4-TE	1,500	1253.33	246.67	5.51
DU-B4-TS	660	1253.33	593.33	6.39
DU-D5	4,500	2590.00	1910.00	7.55
DU-D5-TE	2,300	2590.00	290.00	5.67
DU-D5-TS	970	2590.00	1620.00	7.39
DU-E3	5,400	2700.00	2700.00	7.90
DU-E3-TE	1,200	2700.00	1500.00	7.31
DU-E3-TS	1,500	2700.00	1200.00	7.09
DU-E9	10,000	10800.00	800.00	6.68
DU-E9-TE	8,400	10800.00	2400.00	7.78
DU-E9-TS	14,000	10800.00	3200.00	8.07
DU-F1	2,000	1723.33	276.67	5.62
DU-F1-TE	2,400	1723.33	676.67	6.52
DU-F1-TS	770	1723.33	953.33	6.86
DU-F7	610	513.33	96.67	4.57
DU-F7-TE	620	513.33	106.67	4.67
DU-F7-TS	310	513.33	203.33	5.31
			St Dev	1.51
			t=1.714	2.59
			antilog	13.32
			Action Level	**786.68**

Table 3. Decisions for each Decision Unit based on the 95% Confidence Interval for the triplicate samples.

Sample ID (DUs)	Triplicates	Arsenic mg/kg	Lead mg/kg
DU-A1		11	300
DU-A2	mean	11.2	530
DU-A3		9.5	410
DU-A4		23	560
DU-A5		69	1,200
DU-A6		64	1,300
DU-A7		19	700
DU-A8	mean	48	1,800
DU-A9		180	8,900
DU-B1		9.3	680
DU-B2		7.3	250
DU-B3		13	1,200
DU-B4	mean	6.4	1,253
DU-B5		24	2,500
DU-B6		87	12,000
DU-B7		97	3,100
DU-B8		20	550
DU-B9		280	4,700
DU-C1		8.5	190
DU-C2		20	1,700
DU-C3		14	720
DU-C4		12	3,300
DU-C5		25	3,400
DU-C6	mean	79	4,300
DU-C7		15	830
DU-C8		28	940
DU-C9		32	770
DU-D1		25	1,000

Sample ID (DUs)	Triplicates	Arsenic mg/kg	Lead mg/kg
DU-D2		13	950
DU-D3		52	2,900
DU-D4		9.1	2,100
DU-D5	mean	11.8	2,590
DU-D6		20	1,900
DU-D7		42	1,800
DU-D8		70	2,300
DU-D9		34	1,100
DU-E1		12	900
DU-E2		30	12,000
DU-E3	mean	11.7	2,700
DU-E4		22	5,700
DU-E5		43	7,300
DU-E6		4.9	240
DU-E7		9.1	650
DU-E8		19	840
DU-E9	mean	673	10,800
DU-F0		8.0	450 B
DU-F1	mean	34	1,723
DU-F2		18	900 B
DU-F3		26	1,300 B
DU-F4		25	1,300 B
DU-F5		6.1	200 B
DU-F6		7.7	240 B
DU-F7	mean	13.3	513
DU-F8		33	880 B
DU-F9		46	740 B

13 No Action	10 Based on Pb
5 Based on As	27 Based on Both

Table 4. Tests for Goodness of Fit for Normal and Lognormal Distributions.

Pb Absolute Value for the Difference from the Mean
With 1 Replicates Removed as an Outlier

Total Number of Observations	24
Minimum	10
Maximum	3200
Mean	841.7
Median	450
SD	919.3
Skewness	1.307
Coefficient of Variation	1.092

Shapiro Wilk Normal GOF Test	
Shapiro Wilk Test Statistic	0.824
5% Shapiro Wilk Critical Value	0.916
Data Not Normal at 5% Significance Level	

Lilliefors GOF Test	
Lilliefors Test Statistic	0.196
5% Lilliefors Critical Value	0.177
Data Not Normal at 5% Significance Level	

Shapiro Wilk Lognormal GOF Test	
Shapiro Wilk Test Statistic	0.954
5% Shapiro Wilk Critical Value	0.916
Data appear Lognormal at 5% Significance Level	

Lilliefors Lognormal GOF Test	
Lilliefors Test Statistic	0.0916
5% Lilliefors Critical Value	0.177
Data appear Lognormal at 5% Significance Level	

As Absolute Value for the Difference from the Mean
With 12 Replicates Removed as an Outlier

Total Number of Observations	21
Minimum	0.167
Maximum	16
Mean	3.508
Median	2
SD	4.127
Skewness	2.148
Coefficient of Variation	1.177

Shapiro Wilk Normal GOF Test	
Shapiro Wilk Test Statistic	0.708
5% Shapiro Wilk Critical Value	0.908
Data Not Normal at 5% Significance Level	

Lilliefors GOF Test	
Lilliefors Test Statistic	0.294
5% Lilliefors Critical Value	0.188
Data Not Normal at 5% Significance Level	

Shapiro Wilk Lognormal GOF Test	
Shapiro Wilk Test Statistic	0.973
5% Shapiro Wilk Critical Value	0.908
Data appear Lognormal at 5% Significance Level	

Lilliefors Lognormal GOF Test	
Lilliefors Test Statistic	0.149
5% Lilliefors Critical Value	0.188
Data appear Lognormal at 5% Significance Level	

Table 5 Tests for Outliers.

STEP 1 for Arsenic Results

Rosner's Outlier Test for AsAbsDiff (n>=25)
Number of observatiosn = 27
Observation Value 326.7 is a Potential Outlier

For 5% Significance Level, 326.7 is a Potential Outlier
For 1% Significance Level, 326.7 is a Potential Outlier
Triplicate DU-E9 for Arsenic to be removed from dataset

STEP 2

Dixon's Outlier Test for AsAbsDiff (n<25)
Number of Observations = 24
5% critical value: 0.413
1% critical value: 0.497
Observation Value 101.3 is a Potential Outlier
Test Statistic: 0.542

For 5% significance level, 101.3 is an outlier.

For 1% significance level, 101.3 is an outlier.
Triplicate DU-C6 for Arsenic to be removed from dataset

STEP 3

Dixon's Outlier Test for AsAbsDiff (n<25)
Number of Observations = 21
5% critical value: 0.44
1% critical value: 0.524
Observation Value 16 is a Potential Outlier
Test Statistic: 0.515

For 5% significance level, 16 is an outlier.
For 1% significance level, 16 is not an outlier.
No other Triplicates to be removed from dataset

STEP 1 for Lead Results

Rosner's Outlier Test for AsAbsDiff (n>=25)
Number of observations = 27
Observation Value 5700 is a Potential Outlier

For 5% Significance Level, 5700 is a Potential Outlier

For 1% Significance 5700 is a Potential Outlier
Triplicate DU-C6 for Lead to be removed from dataset

STEP 2

Dixon's Outlier Test for AsAbsDiff (n<25)
Number of Observations = 24
5% critical value: 0.413
1% critical value: 0.497
Observation Value 3200 is a Potential Outlier
Test Statistic: 0.253

For 5% significance level, 3200 is not an outlier.
For 1% significance level, 3200 is not an outlier.
No other Triplicates to be removed from dataset

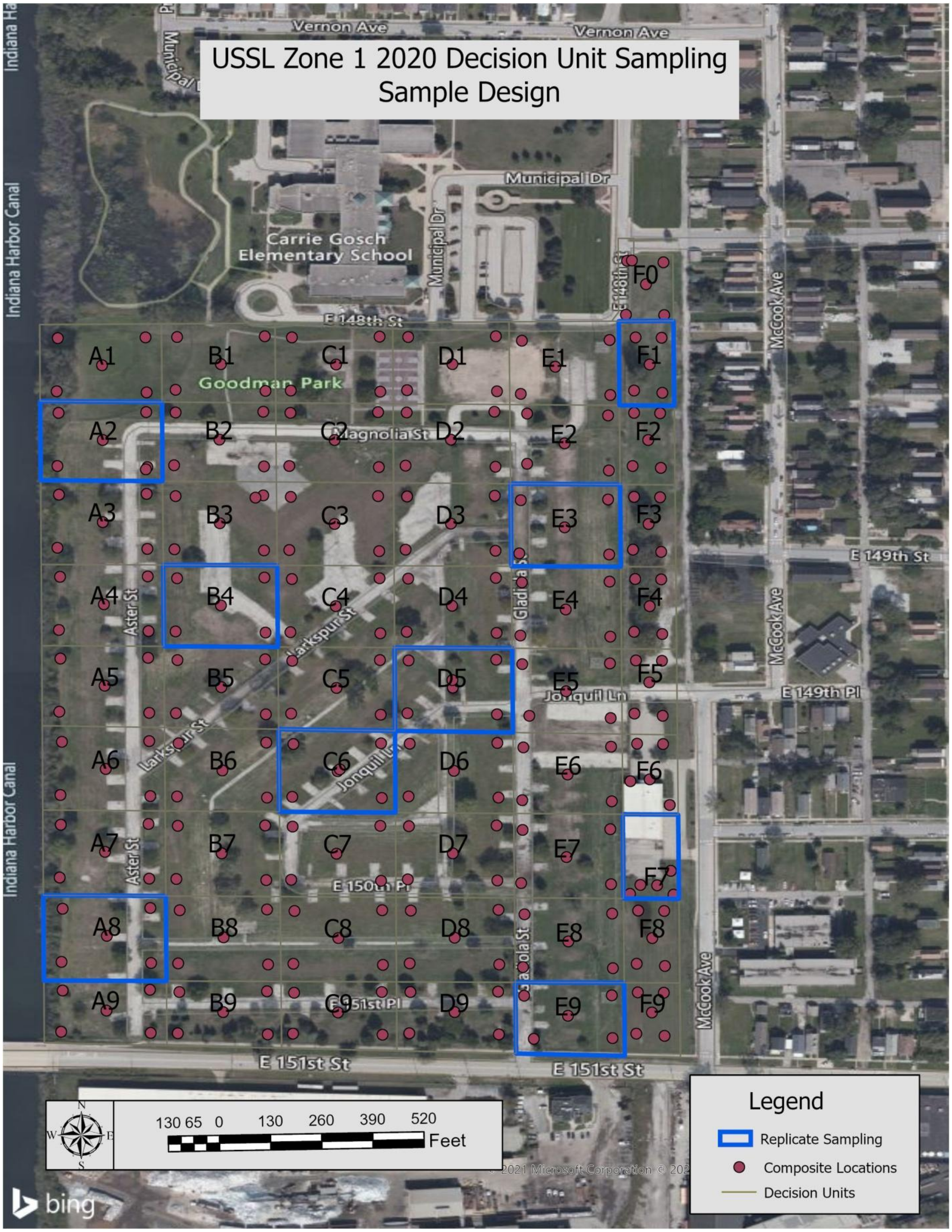

Figure 1. Decision Units and Sample Locations.

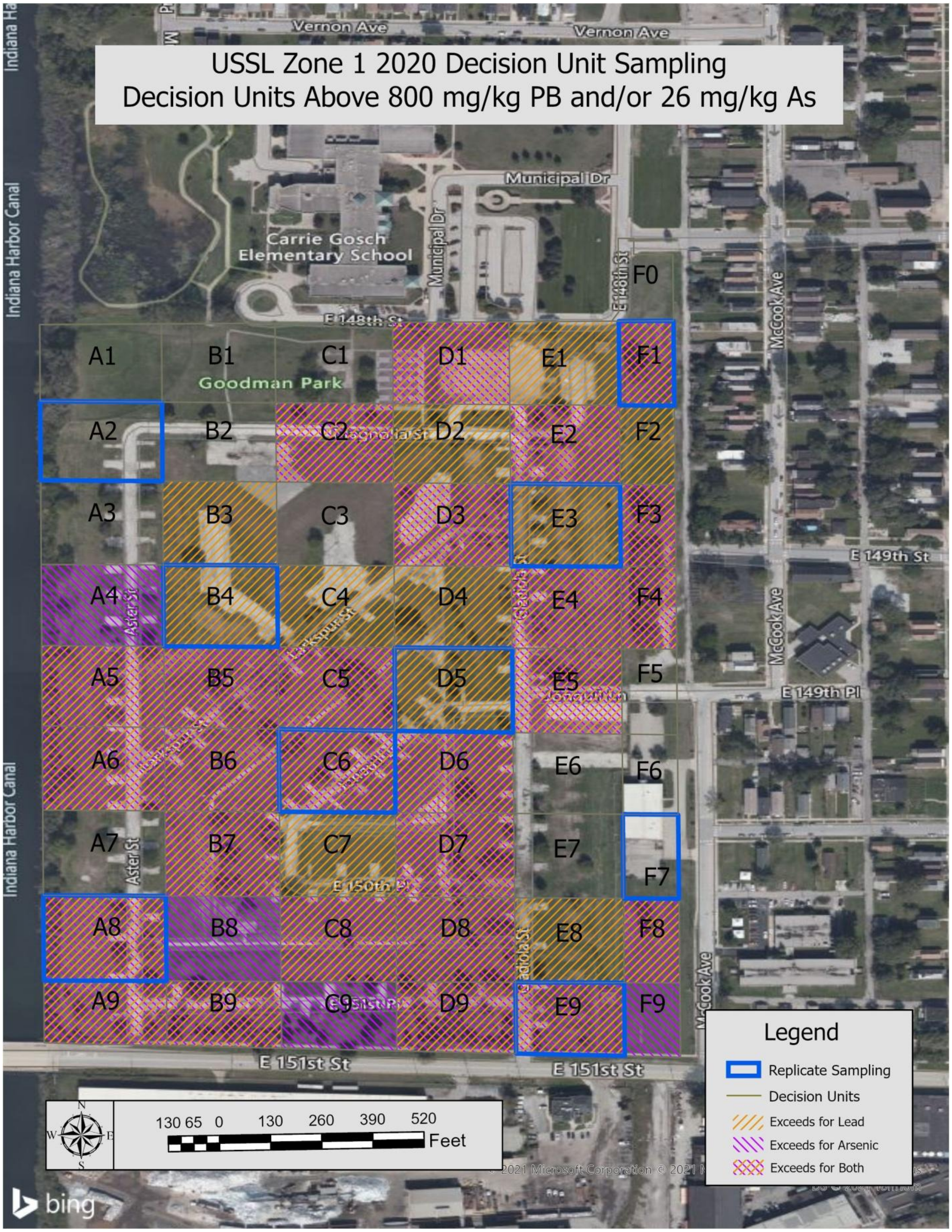

Figure 2. Decision units requiring remediation based on 95% Confidence Intervals of the replicate samples.