

DATA REPORT FOR RED BUTTE CREEK HUMAN HEALTH RISK CHARACTERIZATION

PAHs in Sediment/Soil and Water and Macroinvertebrate Metrics, 2014 and 2015

Prepared for

Chevron Energy Technology Company

1200 Smith Street, 31161A

Houston, TX 77002

Prepared by

Integral Consulting Inc.

7030 E. Genesee Street

Suite 108

Fayetteville, NY 13066

June 7, 2016

**DATA REPORT FOR RED BUTTE CREEK HUMAN HEALTH
RISK CHARACTERIZATION**

**PAHs in Sediment/Soil and Water and Macroinvertebrate
Metrics, 2014 and 2015**

Prepared for
Chevron Energy Technology Company
1200 Smith Street, 31161A
Houston, TX 77002

Prepared by
The logo for Integral Consulting Inc. features the word "integral" in a bold, lowercase, sans-serif font. A thin, curved line starts from the bottom of the letter 'i' and sweeps upwards and to the right, ending under the letter 'l'. Below the word "integral", the words "consulting inc." are written in a smaller, lowercase, sans-serif font.
7030 E. Genesee Street
Suite 108
Fayetteville, NY 13066

June 7, 2016

CONTENTS

LIST OF FIGURES	iii
LIST OF TABLES	ix
ACRONYMS AND ABBREVIATIONS.....	x
EXECUTIVE SUMMARY	xi
1 EXPLORATORY DATA ANALYSIS	1-1
1.1 DATA PROCESSING.....	1-2
1.2 BANK AND BED SAMPLES	1-2
1.3 SUMMARY STATISTICS	1-3
1.4 OUTLIER ANALYSIS	1-4
1.5 Q-Q PLOTS.....	1-4
1.6 CORRELATIONS	1-5
1.7 BENTHIC MACROINVERTEBRATE METRICS	1-5
2 UPPER TOLERANCE LIMIT	2-1
3 HYPOTHESIS TESTS.....	3-1
3.1 RED BUTTE CREEK 2011 COMPARED TO RED BUTTE CREEK 2014/15	3-1
3.2 BACKGROUND COMPARED TO RED BUTTE CREEK 2014/15.....	3-1
4 CONCLUSION	4-1
5 REFERENCES.....	5-1

LIST OF FIGURES

- Figure 1. Correlation of Detects in n = 8 Bank and Bed Sediment Samples, 1-methylnaphthalene and 2-methylnaphthalene
- Figure 2. Correlation of Detects in n = 54 Bank and Bed Sediment Samples, Phenanthrene and Pyrene
- Figure A-1. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, 1-methylnaphthalene
- Figure A-2. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, 2-methylnaphthalene
- Figure A-3. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Acenaphthene
- Figure A-4. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Acenaphthylene
- Figure A-5. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Anthracene
- Figure A-6. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Benzo[a]anthracene
- Figure A-7. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Benzo[a]pyrene
- Figure A-8. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Benzo[b]fluoranthene
- Figure A-9. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Benzo[g,h,i]perylene
- Figure A-10. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Benzo[k]fluoranthene
- Figure A-11. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Chrysene
- Figure A-12. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Dibenz[a,h]anthracene
- Figure A-13. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Fluoranthene
- Figure A-14. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Fluorene
- Figure A-15. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Indeno[1,2,3-cd]pyrene
- Figure A-16. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Naphthalene
- Figure A-17. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Phenanthrene
- Figure A-18. Boxplots of 2014/15 Dataset: Bank and Bed Sediment Samples, Pyrene

- Figure B-1. Boxplots of Concentrations in Sediment by Sample Collection Date, 1-methylnaphthalene
- Figure B-2. Boxplots of Concentrations in Sediment by Sample Collection Date, 2-methylnaphthalene
- Figure B-3. Boxplots of Concentrations in Sediment by Sample Collection Date, Acenaphthene
- Figure B-4. Boxplots of Concentrations in Sediment by Sample Collection Date, Acenaphthylene
- Figure B-5. Boxplots of Concentrations in Sediment by Sample Collection Date, Anthracene
- Figure B-6. Boxplots of Concentrations in Sediment by Sample Collection Date, Benzo[a]anthracene
- Figure B-7. Boxplots of Concentrations in Sediment by Sample Collection Date, Benzo[a]pyrene
- Figure B-8. Boxplots of Concentrations in Sediment by Sample Collection Date, Benzo[b]fluoranthene
- Figure B-9. Boxplots of Concentrations in Sediment by Sample Collection Date, Benzo[g,h,i]perylene
- Figure B-10. Boxplots of Concentrations in Sediment by Sample Collection Date, Benzo[k]fluoranthene
- Figure B-11. Boxplots of Concentrations in Sediment by Sample Collection Date, Chrysene
- Figure B-12. Boxplots of Concentrations in Sediment by Sample Collection Date, Dibenz[a,h]anthracene
- Figure B-13. Boxplots of Concentrations in Sediment by Sample Collection Date, Fluoranthene
- Figure B-14. Boxplots of Concentrations in Sediment by Sample Collection Date, Fluorene
- Figure B-15. Boxplots of Concentrations in Sediment by Sample Collection Date, Indeno[1,2,3-cd]pyrene
- Figure B-16. Boxplots of Concentrations in Sediment by Sample Collection Date, Naphthalene
- Figure B-17. Boxplots of Concentrations in Sediment by Sample Collection Date, Phenanthrene
- Figure B-18. Boxplots of Concentrations in Sediment by Sample Collection Date, Pyrene
- Figure C-1. Boxplots of Concentrations in Sediment by Creek Location, 1-methylnaphthalene
- Figure C-2. Boxplots of Concentrations in Sediment by Creek Location, 2-methylnaphthalene
- Figure C-3. Boxplots of Concentrations in Sediment by Creek Location, Acenaphthene
- Figure C-4. Boxplots of Concentrations in Sediment by Creek Location, Acenaphthylene

- Figure C-5. Boxplots of Concentrations in Sediment by Creek Location, Anthracene
- Figure C-6. Boxplots of Concentrations in Sediment by Creek Location, Benzo[a]anthracene
- Figure C-7. Boxplots of Concentrations in Sediment by Creek Location, Benzo[a]pyrene
- Figure C-8. Boxplots of Concentrations in Sediment by Creek Location, Benzo[b]fluoranthene
- Figure C-9. Boxplots of Concentrations in Sediment by Creek Location, Benzo[g,h,i]perylene
- Figure C-10. Boxplots of Concentrations in Sediment by Creek Location, Benzo[k]fluoranthene
- Figure C-11. Boxplots of Concentrations in Sediment by Creek Location, Chrysene
- Figure C-12. Boxplots of Concentrations in Sediment by Creek Location,
Dibenz[a,h]anthracene
- Figure C-13. Boxplots of Concentrations in Sediment by Creek Location, Fluoranthene
- Figure C-14. Boxplots of Concentrations in Sediment by Creek Location, Fluorene
- Figure C-15. Boxplots of Concentrations in Sediment by Creek Location, Indeno[1,2,3-
cd]pyrene
- Figure C-16. Boxplots of Concentrations in Sediment by Creek Location, Naphthalene
- Figure C-17. Boxplots of Concentrations in Sediment by Creek Location, Phenanthrene
- Figure C-18. Boxplots of Concentrations in Sediment by Creek Location, Pyrene
- Figure D-1. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek,
1-methylnaphthalene
- Figure D-2. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek,
2-methylnaphthalene
- Figure D-3. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek,
Acenaphthene
- Figure D-4. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek,
Acenaphthylene
- Figure D-5. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek,
Anthracene
- Figure D-6. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek,
Benzo[a]anthracene
- Figure D-7. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek,
Benzo[a]pyrene
- Figure D-8. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek,
Benzo[b]fluoranthene

- Figure D-9. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek, Benzo[g,h,i]perylene
- Figure D-10. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek, Benzo[k]fluoranthene
- Figure D-11. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek, Chrysene
- Figure D-12. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek, Dibenzo[a,h]anthracene
- Figure D-13. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek, Fluoranthene
- Figure D-14. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek, Fluorene
- Figure D-15. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek, Indeno[1,2,3-cd]pyrene
- Figure D-16. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek, Naphthalene
- Figure D-17. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek, Phenanthrene
- Figure D-18. Q-Q Plots of Sediment Concentrations in Background and Red Butte Creek, Pyrene
- Figure E-1. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, 1-methylnaphthalene
- Figure E-2. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, 2-methylnaphthalene
- Figure E-3. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Acenaphthene
- Figure E-4. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Acenaphthylene
- Figure E-5. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Anthracene
- Figure E-6. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Benzo[a]anthracene
- Figure E-7. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Benzo[a]pyrene

- Figure E-8. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Benzo[b]fluoranthene
- Figure E-9. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Benzo[g,h,i]perylene
- Figure E-10. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Benzo[k]fluoranthene
- Figure E-11. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Chrysene
- Figure E-12. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Dibenz[a,h]anthracene
- Figure E-13. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Fluoranthene
- Figure E-14. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Fluorene
- Figure E-15. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Indeno[1,2,3-cd]pyrene
- Figure E-16. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Naphthalene
- Figure E-17. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Phenanthrene
- Figure E-18. Boxplots of Water Concentrations in Background and Red Butte Creek for 2014/15, Pyrene
- Figure F-1. Boxplots of Macroinvertebrate Metrics for Background and Red Butte Creek - Taxa Richness
- Figure F-2. Boxplots of Macroinvertebrate Metrics for Background and Red Butte Creek - Evenness
- Figure F-3. Boxplots of Macroinvertebrate Metrics for Background and Red Butte Creek - Shannon's Diversity Index
- Figure F-4. Boxplots of Macroinvertebrate Metrics for Background and Red Butte Creek - Hilsenhoff Biotic Index
- Figure F-5. Boxplots of Macroinvertebrate Metrics for Background and Red Butte Creek - Percent Chironomidae (%)
- Figure F-6. Boxplots of Macroinvertebrate Metrics for Background and Red Butte Creek - Percent EPT (%)

Figure F-7. Boxplots of Macroinvertebrate Metrics for Background and Red Butte Creek – EPT/C

LIST OF TABLES

- Table 1. Summary Statistics for 2014/2015 Dataset for Sediment (Combines Bank and Bed Sediment, and All Sampling Dates)
- Table 2. Summary of 95/95 Upper Tolerance Limits for Sediment Collected 2014/2015 in Background Creeks
- Table 3. Summary of Hypothesis Tests for Sediment in Red Butte Creek, Comparing 2011 to 2014/15
- Table 4. Summary of Hypothesis Tests for Sediment, Comparing 2014/15 Red Butte Creek to 2014/15 Background

ACRONYMS AND ABBREVIATIONS

EPA	U.S. Environmental Protection Agency
EPT	Ephemeroptera, Plecoptera, and Trichoptera
FOD	frequency of detection
GOF	goodness-of-fit
HHRA	human health risk assessment
IQR	interquartile range
KM	Kaplan Meier
PAH	polycyclic aromatic hydrocarbon
PQL	practical quantitation limit
RBC	Red Butte Creek
USEPA	United States Environmental Protection Agency
UTL	upper tolerance limit

EXECUTIVE SUMMARY

This data report summarizes an evaluation of polycyclic aromatic hydrocarbon (PAH) concentrations in sediment and surface water from samples collected as part of the human health risk assessment for the Red Butte Creek site. The primary purpose of the data evaluation is to determine if the concentrations collected in sampling events from 2014 and 2015 are similar in Red Butte Creek and neighboring creeks within the watershed that collectively represent background conditions. Sediment data were grouped by location, date of sampling event, and sample type (bank and bed samples) to examine factors that may contribute to heterogeneity and also to verify that assumptions of the statistical analysis are satisfied. Exploratory data analysis and statistical analysis methods were applied using techniques consistent with current U.S. Environmental Protection Agency (EPA) guidance and tools available online.

The analysis supports the overall conclusion that PAH concentrations in Red Butte Creek are similar to concentrations measured in 2011, and consistent with levels present in background areas in 2014 and 2015. PAH concentrations were not detected in the majority of samples collected for both sediment and surface water. Both the FOD and the distribution of concentrations greater than reporting limits (i.e., detects) were the same or lower in Red Butte Creek compared with background locations. The findings for some of the individual PAHs are sensitive to the low FOD. As needed, additional hypothesis testing was applied to detects to further compare the populations and results were interpreted in the full context of the exploratory data analysis.

This data report also summarizes information on macroinvertebrate samples that were collected in Red Butte Creek and the background creeks. Data consisted of various indices or metrics of ecological health and functioning of the aquatic macroinvertebrate community. The findings support the conclusion that the communities are similar across Red Butte Creek and background creek systems.

1 EXPLORATORY DATA ANALYSIS

The sediment dataset consists of 18 polycyclic aromatic hydrocarbons (PAHs) measured in bank and bed sediment samples collected in 2014/15. Samples were collected from various stations throughout Red Butte Creek. All samples were grouped with the exception of the station denoted “Above Amphitheater”, which is an upstream location that was considered a reference area, but excluded from the background data set. The background data were collected from a series of four creeks called City Creek, Emigration Creek, Mill Creek, and Parleys Creek.

For the 2014/15 sediment dataset, samples were collected during four periods – April 2014, October 2014, April 2015, and August 2015. For the 2011 dataset, samples were collected in August 2011.

For the 2014/15 water dataset, samples were collected during the same time periods and in the same creek systems. Samples were also collected from Jordan Creek.

The following questions are addressed in this analysis for each PAH:

1. Is the distribution present in sediment of Red Butte Creek in 2014/15 the same as the distribution in Red Butte Creek in 2011?
2. Is the distribution present in sediment of Red Butte Creek in 2014/15 the same or elevated compared to the distribution in background sediment?
3. Is the distribution present in surface water of Red Butte Creek in 2014/15 the same or elevated compared to the distribution in background creek?

The methods used to address these questions include exploratory data analysis of summary statistics, graphical methods including boxplots and Q-Q plots, screening against background threshold values using upper tolerance limits (UTLs), outlier analysis, correlation analysis, and hypothesis testing. Many of these methods were also used in the evaluation of the 2011 dataset as part of the 2011 Human Health Risk Assessment (HHRA).

The analysis was completed using U.S. Environmental Protection Agency’s (EPA) ProUCL 5.0 software for calculations of UTLs and hypothesis tests (USEPA 2013). Boxplots and Q-Q plots were generated in R version 3.2.3 (R Core Team 2015) using the ggplot2 package (Wickham 2009). The quantile tests were implemented using the WRS Package (Wilcox and Schönbrodt 2014) and the tables were generated with r2excel (Kassambara 2014).

To facilitate the review using graphical analysis, figures are grouped into batches as follows:

- A. Box and whisker plots for sediment bank and bed samples

- B. Box and whisker plots for sediment data grouped by sample collection date
- C. Box and whisker plots for sediment grouped by creek system
- D. Q-Q plots (lognormal and gamma) for comparison of background to Red Butte Creek 2014/15 and 2011
- E. Box and whisker plots for surface water grouped by creek system
- F. Box and whisker plots for benthic macroinvertebrate metrics

1.1 DATA PROCESSING

The data analysis was applied to the dataset excluding quality control samples (e.g., field duplicates, laboratory duplicates). Nondetects are represented by practical quantitation limits (PQLs) corresponding to analytical method 8270-SIM.

Data from samples collected in Red Butte Creek above the Amphitheater were included in graphical analysis but excluded from statistical tests for consistency with the data aggregation methods applied in the 2011 HHRA.

Outlier screening was conducted, but no outliers were removed.

Concentrations for PAHs in sediment are reported in units of mg/kg and concentrations in water are in units of mg/L, unless specified otherwise.

Data were evaluated to confirm that samples sizes for each PAH matched when grouped by location and sample type (bank or bed).

1.2 BANK AND BED SAMPLES

Boxplots shown in Figure set A illustrate the side-by-side distributions of bank and bed samples for both background and Red Butte Creek datasets. For 11 of 18 PAHs, the distributions for bank and bed samples are similar, and consistent between background and Red Butte Creek. For the remaining 5 PAHs, the median and overall distribution is higher in bank sediment, but the pattern is similar for background and Red Butte Creek.

Based on these observations, the bank and bed samples were combined because the results are not likely to be affected by differences in bank and bed distributions.

1.3 SUMMARY STATISTICS

Table 1 presents the summary statistics for 2014/15 sediment data by PAH and area (background or Red Butte Creek), including frequency of detection (FOD), range of nondetect reporting limits (based on PQLs), range and arithmetic mean of detects, and the median and standard deviation of the full dataset (including all detects and nondetects).

Ratios of sample medians were calculated by dividing the median for background by the median for Red Butte Creek. The ratio is greater than 1 for each PAH, which provides one line of evidence to address the question regarding differences between Red Butte Creek and background.

Ratios of sample standard deviations are also calculated based on the maximum of the ratio of background divided by Red Butte Creek, and vice versa. This metric is used to determine if the assumption of equal variance that underlies parametric and non-parametric hypothesis tests is violated. Ratios greater than approximately 3.0 are indicative of unequal variance, which may introduce uncertainty in the findings of the test. EPA recommends noting such findings, but not attempting to correct for unequal variance through data transformations (USEPA 2007, 2013).

The relatively low FOD in both background and site datasets is noteworthy, as is the observation that the range of nondetects and detects overlap. EPA recommends a minimum of 8 to 10 detects for several statistical tests employed in this analysis (USEPA 2013). Six of 18 PAHs in the background dataset, and 9 of 18 in the Red Butte Creek dataset have fewer than 8 detects for many of the same PAHs. As noted later, this line of evidence also supports the overall finding that PAH concentrations are not elevated above background.

Summary statistics are not tabulated for surface water samples because PAHs were infrequently detected in background and RBC. For the water samples, the PQL for method 8270-SIM is listed as 0.0001 mg/L. A graphical presentation of the data is given in Figure set E and the list of 12 detects is given in the MS Excel workbook attached to this report. Benzo(a)anthracene was detected in 1 of 58 background samples and 0 of 19 Red Butte Creek samples. Three PAHs were detected at 1 of 4 samples collected from the upstream station on Red Butte Creek ("RBC Above Ampitheater"). Eight PAHs were detected in 3 of 19 samples collected from 11 stations in the downstream portion of Red Butte Creek; six of these were from the same location, "RBC Above 1500E", from the sampling event on August 2015. None of these PAHs were detected at this location in April 2014 or April 2015. The remaining two detects were from a sample collected from "RBC Above Sunnyside" in October 2014. This specific location was not sampled during other sampling events in 2014 or 2015.

1.4 OUTLIER ANALYSIS

Table 1 also presents the summary of the outlier analysis applied to the sediment data. The data were initially screened for potential outliers by examining tabular summaries and graphical summaries, including box and whisker plots (Figure sets A, B, and C) and Q-Q plots (Figure set D). Since there were very few water samples with concentrations greater than PQLs, no outlier screening was performed on the water samples.

The premise of outlier screening for this analysis is that it serves as one tool by which to explore the data. The objective is not to “flag and eliminate” spurious high values, but rather to examine the collective set of such high values for potential patterns. For example, a specific sampling event or sample may be impacted by random environmental conditions, sample collection and handling procedures, or laboratory preparation procedures. This screening step serves as a quality control check on the representativeness of the dataset as a whole.

Another screening step that was considered was the interquartile range (IQR) rule, which establishes a threshold above which potential outliers (from the same dataset) are flagged. The calculation involves a multiple, “k”, times IQR added to the median. Common practice is to use k values of 2 or 3, though the choice is somewhat arbitrary and coverage probabilities (i.e., the probability of identifying an outlier if it in fact exists) vary depending on the skewness of the distribution.

The maximum value was identified as a potential outlier for five PAHs in the background dataset and four PAHs in the Red Butte Creek dataset. These values were noticeably different from the distribution of concentrations illustrated in the Q-Q plots.

Following EPA guidance (USEPA 2013), each outlier was examined for statistical significance using Rosner’s test ($\alpha = 0.05$), which applies for sample sizes greater than 25. Each value was confirmed to be a statistically significant outlier. Given the relatively large sample sizes and the use of nonparametric (rank order) statistics, the results of the analysis summarized in this report are considered robust to these single extreme values. The outliers also provide important context for interpreting the relevance of screening site data against background threshold values.

1.5 Q-Q PLOTS

Q-Q plots were prepared for each dataset using lognormal and gamma distribution assumptions because the background datasets generally fit one or both of these distributions (see Table 2). Three series are shown on each figure – one for background, one for Red Butte Creek 2011, and one for Red Butte Creek 2014/15.

The series can be examined independently to look for potential outliers, as discussed in the Section 1.4, as well as to identify multiple populations. A break in the points or sudden change in slope is indicative of this condition. The series can be examined together to determine if they generally overlap, or if they are shifted to the left or right. If the background distribution is shifted to the left, for example, this means that concentrations tend to be higher at lower quantiles of the overall distribution – so background would be considered elevated relative to Red Butte Creek.

For each PAH distribution, the Q-Q plots support the finding that background concentrations are elevated, or comparable to Red Butte Creek. The exceptions occur in the extreme tails of the distributions for some PAHs, indicating that maximum concentrations are not consistently higher in background compared to RBC.

1.6 CORRELATIONS

Based on the tabular summaries and Q-Q plots (Figure set D), it is apparent that distributions for selected pairs of PAHs are tracking closely together. This is not uncommon for PAHs, and in fact is a property of PAH mixtures used to conduct forensic chemistry evaluations of environmental data. Two correlations are presented graphically based on detected concentrations. Figure 1 shows the strong correlation ($R^2 > 0.99$) for $n = 8$ samples where 1-methylnaphthalene and 2-methylnaphthalene were detected in the combined background and Red Butte Creek datasets. Figure 2 shows a moderate correlation ($R^2 = 0.77$) for $n = 54$ samples where phenanthrene and pyrene were detected. Both plots are on a log-log scale and use a power curve to quantify the fit of the collective data.

Such correlations could be used to extend the data for additional paired concentrations where the result for one PAH is a nondetect and the result for the second, correlated PAH for the same sample is a detect. This method may reduce uncertainty associated with comparisons of distributions that are moderately censored (e.g., FOD equal to 30 to 70%). Such an analysis was beyond the scope of this report.

1.7 BENTHIC MACROINVERTEBRATE METRICS

Macroinvertebrate community structure has commonly been used as an indicator of the condition of an aquatic system (Armitage et al., 1983; Mandaville 2002). The following metrics were calculated based on samples collected twice a year (April and August) on 2013, 2014, and 2015: Taxa Richness, Evenness, Shannon-Wiener Diversity Index, Hilsenhoff Biotic Index, Percent Chironomidae, and Percent Ephemeroptera, Plecoptera, and Trichoptera (EPT). Figure set F presents boxplots that illustrate the distributions of results by creek system, including combined background creeks and Red Butte Creek. The specific sampling dates were: April 2013, August 2013, April 2014, August 2014, April 2015, and August 2015.

The following is a guide to the interpretation of the magnitude of each metric (Mandaville 2002):

- **Taxa Richness (TR):** Taxa Richness equals the total number of taxa represented within the sample. The healthier the community is, the greater the number of taxa found within the community.
- **Evenness:** Relevant abundance of different taxonomic groups, determined by the counts of all organisms collected. Values are distributed between 0 and 1, with values closer to 1 indicating individuals are distributed more evenly (Pielou 1966).
- **Shannon-Wiener Diversity Index (H):** H is a function of the proportion of individuals in a given taxon and the total number of taxa in the community. As the number and distribution of taxa (biotic diversity) within the community increases, the value of H increases.
- **Hilsenhoff Biotic Index:** Provides a measure of family-level tolerance values ranging from 0 (very intolerant) to 10 (highly tolerant). The following scale can be used to evaluate water quality conditions:

Hilsenhoff Biotic Index	Water Quality	Degree of Organic Pollution
0.00 – 3.75	Excellent	Organic pollution unlikely
3.76 – 4.25	Very good	Possible slight organic pollution
4.26 – 5.00	Good	Some organic pollution probable
5.01 – 5.75	Fair	Fairly substantial pollution likely
5.76 – 6.50	Fairly poor	Substantial pollution likely
6.51 – 7.25	Poor	Very substantial pollution likely
7.26 – 10.00	Very poor	Severe organic pollution likely

- **Percent Chironomidae:** The abundance of Chironomidae indicates the system is suitable to more tolerant (less sensitive) organisms. Therefore, higher percentages are an indication of a stressed system.
- **Percent EPT:** The abundance of EPT indicates the system is suitable for less tolerant (more sensitive) organisms. Therefore, higher percentages are an indication of a system considered to be in good biotic condition.
- **EPT/C:** The ratio of EPT/C index is calculated by dividing the sum of the total number of individuals classified as EPT by the total number of individuals classified as Chironomidae. A community considered to be in good biotic condition will display an even distribution among these four groups, while communities with disproportionately high numbers of Chironomidae may indicate environmental stress.

The distribution for each metric for Red Butte Creek generally overlaps with the combined background distribution. The median for Red Butte Creek is slightly greater for Taxa Richness, Evenness, Shannon-Weiner Diversity Index, and Percent EPT, lower for Percent Chironomidae, and higher for the EPT/C ratio – all indications that conditions for biota in Red Butte Creek are the same or more favorable than that of background.

For the Hilsenhoff Biotic Index, 11 of 21 (52%) sampling events in Red Butte Creek yielded index values classified as *very good* or *excellent* water quality, 5 of 21 (24%) are classified as *good*, and the remaining 5 of 21 (24%) are classified as *fair*. This distribution matches almost exactly the classifications for the combined background creek dataset (n= 38): very good or excellent (55%), good (23%), fair (15%), and fairly poor (7%).

Collectively, the findings from the analysis of the biotic indices support the conclusion that the communities are similar across Red Butte Creek and background creek systems.

2 UPPER TOLERANCE LIMIT

The UTL is considered an upper bound concentration in the background dataset which, if exceeded in the site dataset, may indicate that the distributions represent different populations. It is considered a “background threshold value”, and an initial line of evidence to be further explored using two-sample hypothesis tests.

The UTL can be thought of as an upper confidence limit on an upper percentile of the sample results. The standard approach is to select the 95th percentile and the 95 percent upper confidence limit, which is denoted the “95/95 UTL”. Parametric and nonparametric methods can be implemented with ProUCL 5. For nonparametric methods, a minimum of 59 observations is needed for the sample maximum to be able to provide the desired confidence level and coverage the distribution. All of the background datasets for this analysis consist of 104 observations.

Table 2 provides the summary statistics and results of the UTL calculations. Each distribution shape was explored with goodness-of-fit (GOF) tests using ProUCL as well as by visual inspection of the Q-Q plots for lognormal and gamma distributions. The selection criteria, also summarized in the table footnotes, are as follows:

- If the GOF supports a normal distribution, with or without additional support for lognormal and gamma distributions, the UTL is based on the normal UTL equation using Kaplan Meier (KM) parameter estimates. EPA recommends KM parameter estimates for left censored data, such as with each PAH dataset in this analysis (USEPA 2013).
- If the GOF supports a gamma and lognormal distribution, the UTL is based on the gamma UTL using KM with Wilson Hilferty estimations.
- If the GOF supports a lognormal distribution, the UTL is based on the lognormal UTL using KM parameter estimates.
- If the GOF does not support normal, gamma, or lognormal, the UTL is based on the rank ordered value ($n = 102$ of 104).

An insufficient number of detects ($n < 8$) are present to calculate a UTL for 6 of 18 PAHs. For 8 of the remaining 12 PAHs, the UTL is exceeded by the maximum of the Red Butte Creek dataset. It is noteworthy that for two of these PAHs – benzo(a)pyrene and phenanthrene – the maximum values differ from the 95/95 UTL by less than 5%. In addition, for three of these PAHs – benzo(a)anthracene, benzo(b)fluoranthene, and chrysene – the maximum values in the

Red Butte Creek datasets are statistical outliers, and the second highest value is less than the 95/95 UCL.

Finally, it is important to note that, for each PAH background dataset, the maximum concentration exceeds the 95/95 UTL calculated from the same dataset (see Table 2). In other words, if the very same results were obtained from RBC, one or more exceedances of the 95/95 UTL would have been noted in every case. This finding further underscores the concept that the utility of a UTL is as a screening tool, not as a conclusive line of evidence regarding the comparability of the background and site distributions.

3 HYPOTHESIS TESTS

Given that the PAH datasets are all left-censored with multiple PQL reporting limits, the Gehan test was selected to compare distributions from two samples, consistent with EPA guidance (USEPA 2013). The Gehan test examines the differences in the sample medians. Results are summarized in Table 3 for the RBC 2011 and RBC 2014/15 datasets, and Table 4 for the background and RBC datasets from 2014/15.

It is important to note that one-sided two-sample hypothesis tests require a definition for the null hypothesis. Often this assumption determines the outcome of the test because large differences in sample statistics are required to reject the null hypothesis. For this reason, a background evaluation should not hinge on a single test of statistical significance, but rather include hypothesis tests as a line of evidence together with exploratory data analysis procedures.

3.1 RED BUTTE CREEK 2011 COMPARED TO RED BUTTE CREEK 2014/15

Table 3 summarizes the hypothesis test results for comparisons of Red Butte Creek 2011 and Red Butte Creek 2014/15 datasets. Sufficient sample samples (i.e., $n > 8$ detects in both groups) are available to perform this analysis for four PAHs: chrysene, fluoranthene, phenanthrene, and pyrene. The ratio of standard deviations are all less than 3, so the assumption of homogeneity of the variances of the two distributions is not violated.

In each case, the Gehan test supports the conclusion that the distributions are the same because the difference in the medians of the background and RBC data are not statistically significant (p -values are much greater than 0.05).¹

3.2 BACKGROUND COMPARED TO RED BUTTE CREEK 2014/15

The sample size is sufficient with at least 8 detects in background and RBC datasets from 2014/15 to apply hypothesis tests for 9 of 18 datasets. As shown in Table 4, the null hypothesis that RBC is less than or equal to background is rejected for one PAH – indeno(1,2,3-cd)pyrene². For this PAH, even though the sample median for background is greater than that of RBC for the full dataset of detects and nondetects (with a ratio of sample medians equal to 1.08), this

¹ The null hypothesis is based on a two-sided test to give equal chance of concluding that the distribution from 2014/15 is different (higher or lower) than that of 2011.

² The null hypothesis is a one-sided test that the RBC median is less than or equal to the background median (called Form 1 by USEPA 2013). The Form 2 test was not evaluated given that the relatively low frequency of detects for these PAH datasets reduces the information content required to achieve a desired error rate of just 5 percent.

difference is not large enough to be considered statistically significant³. It should be noted that this PAH has the lowest FODs (16 percent for background, 29 percent for RBC) of the 9 PAHs evaluated, meaning that the summary statistics (median and SD) are dominated by the PQL values used to represent nondetects. Not only is the sample median higher for background, but the sample maximum for background is also greater by approximately a factor of two (see Table 3) and the overall distribution of detects (illustrated by the Q-Q plot for background in Figure D-15) is greater for the background dataset compared with the RBC. Indeed, when the hypothesis test is rerun on detects only, the distribution for RBC is considerably lower than that of background ($p=0.47$; see Table 4). Therefore, collectively, the summary statistics, graphical presentation of the data, and hypothesis testing suggest that the concentrations in the region of the detects supports the assumption that indeno(1,2,3-cd)pyrene concentrations are not elevated in RBC compared to background.

It is noteworthy that for each PAH, the sample median for the background dataset is higher than that of the Red Butte Creek dataset. The ratio of standard deviations for the candidate PAHs with sufficient samples sizes of detects are all less than 3.

Upper tail tests, such as the Quantile test, are no longer supported in ProUCL 5.0 and 5.1 due to lack of resources and because some users previously misapplied the tests to datasets with sample sizes greater than 18. EPA discontinued this functionality because exact critical values for the larger samples sizes have not yet been developed. Results for Quantile tests implemented in R are shown in Table 4 for comparison purposes - in the event that these test were run on earlier datasets from Red Butte Creek.

³ The p -value for the Gehan Test for indeno(1,2,3-cd)pyrene is 0.04; a value of 0.05 or more is needed to conclude the assumption that the median for Red Butte Creek is less than or equal to background is supported by the data.

4 CONCLUSION

Using the same methods as the 2011 HHRA, this report presents multiple lines of evidence to address the questions outlined at the start of the report

- Is the distribution present in sediment of Red Butte Creek in 2014/15 the same as the distribution in Red Butte Creek in 2011?

The median concentrations were comparable for each, though slightly lower for 2014/15, perhaps due to improvements in analytical detection limits. The four PAHs that had sufficient detects to evaluate with hypothesis tests confirmed that the PAH concentrations are comparable between the two time periods.

- Is the distribution present in sediment of Red Butte Creek in 2014/15 the same or elevated compared to the distribution in background sediment?

The following observations are noted for each PAH:

- **1-methylnaphthalene and 2-methylnaphthalene** (highly correlated): very low FOD in both background (8 percent) and RBC (10 percent); Q-Q plots overlap (exhibiting similar concentrations across the range); maximum concentration in background appears to be a statistical outlier.
- **Acenaphthene, acenaphthylene, anthracene, dibenz(a)anthracene, fluorene, and naphthalene:** All have very low FOD in background and Red Butte Creek datasets, and exhibit similar magnitude and distributions of detects.
- **Benzo(a)anthracene:** Slightly greater FOD in Red Butte Creek (41% compared to 35%); Q-Q plots overlap (exhibiting similar concentrations across the range); just the maximum value exceeds the 95/95 UTL and this value is a statistical outlier; ratio of medians is 1.16 (higher in background).
- **Benzo(a)pyrene:** Slightly greater FOD in Red Butte Creek (43% compared to 37%); Q-Q plots overlap (exhibiting similar concentrations across the range); just the maximum value (0.166 mg/kg) exceeds the 95/95 UTL (0.163), a difference of 2%; ratio of medians is 1.14 (higher in background).
- **Benzo(b)fluoranthene:** Greater FOD in Red Butte Creek (43% compared to 24%); Q-Q plot for background is shifted to the left (exhibiting higher concentrations for background across the range); just the maximum value (0.263 mg/kg) exceeds the 95/95 UTL (0.149) and this value is a statistical outlier; ratio of medians is 1.14 (higher in background).
- **Benzo(g,h,i)perylene:** Slightly greater FOD in Red Butte Creek (31% compared to 24%); Q-Q plot for background is shifted to the left (exhibiting higher concentrations for

background across the range); maximum value does not exceed the 95/95 UTL; ratio of medians is 1.17 (higher in background).

- **Benzo(k)fluoranthene:** Slightly greater FOD in Red Butte Creek (12% compared to 8%); Red Butte Creek has the same distribution compared to Red Butte Creek upstream; Q-Q plots overlap (exhibiting similar concentrations across the range); 2 of 42 values for Red Butte Creek exceed the 95/95 UTL; ratio of medians is 1.05 (higher in background).
- **Chrysene:** Greater FOD in Red Butte Creek (41% compared to 34%); pattern of higher bank compared to bed is similar in background and Red Butte Creek; Red Butte Creek has the same distribution compared to Red Butte Creek upstream; Q-Q plots overlap (exhibiting similar concentrations across the range); just the maximum value (0.244 mg/kg) exceeds the 95/95 UTL (0.138) and this value is a statistical outlier; ratio of medians is 1.16 (higher in background).
- **Fluoranthene:** Greater FOD in Red Butte Creek (41% compared to 34%); pattern of higher bank compared to bed is similar in background and Red Butte Creek; Red Butte Creek has the same distribution compared to Red Butte Creek upstream; Q-Q plots overlap (exhibiting similar concentrations across the range); just the maximum value (0.244 mg/kg) exceeds the 95/95 UTL (0.138) and this value is a statistical outlier; ratio of medians is 1.16 (higher in background).
- **Indeno(1,2,3-cd)pyrene:** Greater FOD in Red Butte Creek (29% compared to 16%); pattern of higher bank compared to bed occurs in Red Butte Creek but not background; Red Butte Creek has the same distribution compared to Red Butte Creek upstream; Q-Q plot for background is shifted to the left (exhibiting higher concentrations for background across the range); 2 of 42 value results for Red Butte Creek exceed the 95/95 UTL; hypothesis test of difference in medians of full datasets indicates Red Butte Creek is elevated, but the ratio of sample medians is 1.08 (higher in background) and the hypothesis test for detects only demonstrates that RBC is less than or equal to background ($p=0.47$).
- **Phenanthrene:** Same FOD in Red Butte Creek and background (36%); pattern of higher bank compared to bed is similar in background and Red Butte Creek; Red Butte Creek has the same distribution compared to Red Butte Creek upstream; Q-Q plots overlap (exhibiting similar concentrations across the range); just the maximum value (0.237 mg/kg) exceeds the 95/95 UTL (0.231) and this difference is 3%; ratio of medians is 1.19 (higher in background).
- **Pyrene:** Slightly lower FOD in Red Butte Creek (50% compared to 51%); pattern of higher bank compared to bed is similar in background and Red Butte Creek; Red Butte Creek has the same distribution compared to Red Butte Creek upstream; Q-Q plots overlap (exhibiting similar concentrations across the range); maximum value for RBC does not exceed the 95/95 UTL; ratio of medians is 1.06 (higher in background).

The following summarizes findings regarding the following question regarding PAHs in surface water:

- Is the distribution present in surface water of Red Butte Creek in 2014/15 the same or elevated compared to the distribution in background creek?

Very few detects are present in the PAH water dataset. These detects were present in both background and Red Butte Creek surface water, supporting the conclusion that the conditions are similar in background creeks and Red Butte Creek.

In addition, the benthic macroinvertebrate community indices support the conclusion that the communities are similar across Red Butte Creek and background creek systems.

5 REFERENCES

Armitage, P.D., D. Moss, J.F. Wright, and M.T. Furse. 1983. The performance of a new biological water quality score system based on macroinvertebrates over a wide range of unpolluted running-water sites. *Water Res.* 17: 333-47.

Kassambara, A. 2014. r2excel: Read, write and format Excel files in R. R package version 1.0.0. www.sthda.com.

Mandaville, S.M., 2002. Benthic macroinvertebrates in freshwater – taxa tolerance values, metrics, and protocols, Project H - 1. Nova Scotia: Soil & Water Conservation Society of Metro Halifax.

Pielou, E.C., 1966. The measurement of diversity in different types of biological collections, *J. Theoret. Biol.*, 13: 131-144.

R Core Team. 2015. R: A language and environment for statistical computing. Available at: www.R-project.org. R Foundation for Statistical Computing, Vienna, Austria.

USEPA. 2007. Performance of statistical tests for site versus background soil comparisons when distributional assumptions are not met. Technology Support Center Issue. EPA/600/R-07/020. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC. March.

USEPA. 2013. ProUCL Version 5.0.00 technical guide. Statistical software for environmental applications for data sets with and without nondetect observations. EPA/600/R-07/041. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC. September.

Wickham, H. 2009. *ggplot2: Elegant graphics for data analysis*. Springer-Verlag, New York, NY.

Wilcox, R.R., and F.D. Schönbrodt. 2014. The WRS package for robust statistics in R (version 0.24). Retrieved from <http://r-forge.r-project.org/projects/wrs/>.