Technical/Regulatory Guidance

Soil Background and Risk Assessment

[DECEMBER 2021]

Prepared by
The Interstate Technology & Regulatory Council
Soil Background and Risk Team

ABOUT ITRC

The Interstate Technology and Regulatory Council (ITRC) is a public-private coalition working to reduce barriers to the use of innovative environmental technologies and approaches so that compliance costs are reduced, and cleanup efficacy is maximized. ITRC produces documents and training that broaden and deepen technical knowledge and expedite quality regulatory decision making while protecting human health and the environment. With private and public sector members from all 50 states and the District of Columbia, ITRC truly provides a national perspective. More information on ITRC is available at www.itrcweb.org. ITRC is a program of the Environmental Research Institute of the States (ERIS), a 501(c)(3) organization incorporated in the District of Columbia and managed by the Environmental Council of the States (ECOS). ECOS is the national, nonprofit, nonpartisan association representing the state and territorial environmental commissioners. Its mission is to serve as a champion for states; to provide a clearinghouse of information for state environmental commissioners; to promote coordination in environmental management; and to articulate state positions on environmental issues to Congress, federal agencies, and the public.

DISCLAIMER

This material was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof and no official endorsement should be inferred.

The information provided in documents, training curricula, and other print or electronic materials created by the Interstate Technology and Regulatory Council ("ITRC" and such materials are referred to as "ITRC Materials") is intended as a general reference to help regulators and others develop a consistent approach to their evaluation, regulatory approval, and deployment of environmental technologies. The information in ITRC Materials was formulated to be reliable and accurate. However, the information is provided "as is" and use of this information is at the users' own risk.

ITRC Materials do not necessarily address all applicable health and safety risks and precautions with respect to particular materials, conditions, or procedures in specific applications of any technology. Consequently, ITRC recommends consulting applicable standards, laws, regulations, suppliers of materials, and safety data sheets for information concerning safety and health risks and precautions and compliance with then-applicable laws and regulations. ITRC, ERIS and ECOS shall not be liable in the event of any conflict between information in ITRC Materials and such laws, regulations, and/or other ordinances. The content in ITRC Materials may be revised or withdrawn at any time without prior notice.

ITRC, ERIS, and ECOS make no representations or warranties, express or implied, with respect to information in ITRC Materials and specifically disclaim all warranties to the fullest extent permitted by law (including, but not limited to, merchantability or fitness for a particular purpose). ITRC, ERIS, and ECOS will not accept liability for damages of any kind that result from acting upon or using this information.

ITRC, ERIS, and ECOS do not endorse or recommend the use of specific technology or technology provider through ITRC Materials. Reference to technologies, products, or services offered by other parties does not constitute a guarantee by ITRC, ERIS, and ECOS of the quality or value of those technologies, products, or services. Information in ITRC Materials is for general reference only; it should not be construed as definitive guidance for any specific site and is not a substitute for consultation with qualified professional advisors.

SBR-# 1

Soil Background and Risk Assessment

[December 2021]

Prepared by
The Interstate Technology & Regulatory Council
Soil Background and Risk Team

50 F Street, NW, Suite 350, Washington, DC 20001

https://sbr-1.itrcweb.org/

ACKNOWLEDGMENTS

The members of the Interstate Technology & Regulatory Council (ITRC) Soil Background and Risk Assessment (SBR) Team wish to acknowledge the individuals, organizations, and agencies that contributed to this guidance.

As part of the broader ITRC effort, the SBR Team effort is funded primarily by the U.S. Department of Energy. Additional funding and support have been provided by the U.S. Department of Defense and the U.S. Environmental Protection Agency.

The SBR team was led by Bonnie Brooks of the Washington Department of Ecology and Claudio Sorrentino of the California Department of Toxic Substances Control.

Team leaders were assisted by program advisors Jim Rocco and Lesley Hay Wilson of Sage Risk Solutions LLC.

Specifically, the team recognizes the efforts and important contributions of the following:

State and Local Government

Naji Akladiss

Maine Department of Environmental Protection

Patrick Boettcher

Delaware Department of Natural Resources and Environmental Control

Willliam Bosan

California Department of Toxic Substances Control

Robin Chinnery

Michigan Department of Environment, Great Lakes and Energy

Weiquan Dong

Nevada Division of Environmental Protection

Erin Endsley

Wisconsin Department of Natural Resources

Brian Faulkner

California Department of Toxic Substances Control

Debi Goodwin

Arizona Dept. of Environmental Quality

Brady Johnson

Idaho Department of Environmental Quality

Margaret Greene

Tennessee Department of Environment and Conservation

Paul Kulpa

Rhode Island Department of Environmental Management

Marsha Lay

Indiana Department of Environmental Management

Gail Lipfert

Maine Department of Environmental Protection

Adam McMahon

Idaho Department of Environmental Quality

Kathryn Morris

Montana Department of Environmental Quality

Norka Paden

Idaho Department of Environmental Quality

Anita Poulsen

Florida Department of Health

Jeff Rackow

Arizona Department of Environmental Quality

Jim Ridenour

New York State Department of Health

Divinia Ries

Michigan Department of Environment, Great Lakes and Energy

Shukla Roy-Semmen

California Department of Toxic Substances Control

Sandra Snyder

South Carolina Department of Health and Environmental Control

Harold Templin

Indiana Department of Environmental Management

Susan Turnblom
Oregon Department of Environmental Quality

Federal Government

Sydney Chan

USEPA

Felicia Barnett

USEPA ORD

Matthew Lambert

USEPA

Cynthia Auld

US Army Corps of Engineers

Amy Rosenstein

US Army Corps of Engineers

Laurel Woodruff

USGS

Academia

Marta Shore

University of Minnesota

Geoff Siemering

University of Wisconsin, Madison

Jennifer Weidhaas

University of Utah

Industry Affiliates

Mark Benotti

NewFields Companies LLC

Gino Bianchi-Mosquera

GSI Environmental, Inc.

Kelly Black
Neptune and Company, Inc.

Doug Blue
ExxonMobil

Mark Bruce
Eurofins

Brian Church
Windward Environmental, LLC

Regan Dahmer
ExxonMobil

Emily Davis
APTIM

Charlie DeWolf
Trihydro Corporation

Joann Dyson
GHD

Allison Geiselbrecht
Floyd|Snider

David Heidlauf
Ramboll

Michael Kierski
Exponent

Alexander Kliminsky
GHD

Jane Parkin Kullmann
Wood PLC

Kanan Patel-Coleman
Burns & McDonnell

Jay Peters
Haley & Aldrich, Inc.

Chrissy Peterson
EHS Support, LLC

Pamela Rodgers
Battelle Memorial Institute

Shahrokh Rouhani
NewFields Companies LLC

Barrie Selcoe
Jacobs

Leyla Shams
NewFields Companies LLC

Anita Singh
Retired, Lockheed Martin

Tamara Sorell
Brown and Caldwell

Scott Stout
NewFields Companies LLC

James Studer
InfraSUR LLC

Anna Sutton
Cardno

Sagar Thakali
AECOM

Karen Thorbjornsen
APTIM

Allen Uhler
NewFields Companies LLC

Ann Verwiel
ToxStrategies, Inc.

Michael Worden
Civil & Environmental Consultants, Inc.

Home Page

While there are already guidance documents regarding soil background, there is not a "*one-stop-shop*" document that provides comprehensive and widely accepted guidance on the state of the science on this topic. This ITRC guidance document fills the gap by providing a consensus-based comprehensive and defensible framework for establishing soil background and using soil background in risk assessments. This guidance also provides references to state, USEPA and other resources, as well as related ITRC documents.

During the initial phases of a site investigation or cleanup project, regulators often use chemical-specific screening values to decide if the presence of chemicals at a site may pose a threat to human health or the environment, or both, and if there is the need for further action. For some chemicals, the screening values are within the range of concentrations found in soils from natural or anthropogenic ambient background. Most regulatory agencies require a response action only when the concentrations of a chemical in soil exceed its background concentrations. Consequently, soil background plays an important role in remedial decisions.

Risk assessors, risk managers, and site investigators, as well as other stakeholders are the target audience of this guidance document.

TABLE OF CONTENTS

1 INTRODUCTION 13

1.1 Audience 13
1.2 Purpose 13
1.3 Use of Background in the Risk Assessment Process 15
1.4 Limitations 15

2 SOIL BACKGROUND DEFINITION 17

2.1 Natural Soil Background 17
2.2 Anthropogenic Ambient Soil Background 18
2.3 Additional Background Definitions 19
2.3.1 Default soil background 19
2.3.2 Site-specific soil background 20
2.3.3 Soil background reference area 21

3 ESTABLISHING SOIL BACKGROUND 22

3.1 Introduction 22
3.1.1 Default soil background 22
3.1.2 Site-specific soil background 23
3.2 Conducting a Soil Background Study 24
3.2.1 Natural background 24
3.2.2 Anthropogenic ambient soil background 26
3.3 Choosing an Area for a Soil Background Study 28
3.3.1 Geologic and geochemical considerations 29
3.3.2 Hydrogeologic conditions 30
3.4 Sampling 30
3.5 Laboratory Analysis 31
3.6 Using an Existing Soil Background Study 31
3.7 Background Dataset Analysis 33
3.8 Establishing Default or Site-Specific Soil Background 34
3.8.1 Using a site-specific background dataset 35
3.8.2 Advanced methods 35
3.9 Extracting Site-Specific Background Dataset from an On-site Dataset 36
3.9.1 Need for background data 37
3.9.2 When to use the background extraction approach 37
3.9.2.1 Sites amenable to the use of the iterative background extraction process 38
3.9.2.2 Assumptions and involvement of the project team 38

3.9.2.3 Treatment of nondetect observations 39
3.9.3 Using the existing off-site background data—highly recommended when available 40
3.9.4 Statistical approach 40
3.9.4.1 Step-by-step summary of the iterative background extraction method 41
3.9.5 Extracting background-level arsenic concentrations from the on-site soil dataset of a Superfund site 42
4 USING SOIL BACKGROUND IN RISK ASSESSMENT 52
4.1 Representative Site Concentration to Compare to a BTV 54
4.1.1.1 Maximum 55
4.1.1.2 95th percentile 55
4.2 Using Default Background 55
4.2.1 Comparing a representative site concentration to a default BTV 56
4.2.2 Test for proportions 56
4.3 Using Site-Specific Background 57
4.3.1 Comparing a representative site concentration to a site-specific BTV 57
4.3.2 Comparison of site concentration dataset to site-specific background concentration dataset 58
4.3.3 Community composition comparison 58
4.3.4 Geochemical evaluations and environmental forensics 59
4.3.4.1 Geochemical evaluations 59
4.3.4.2 Environmental forensics 59
4.3.5 Characterizing risks from background 59
4.3.5.1 Incremental risk analysis of soil background 60
4.3.5.2 Comparing background and site risks 62
4.4 Use of Background for Remedial Goals 66
4.5 Additional Considerations 68
4.5.1 Bioavailability of chemicals in site and background soil 68
4.5.2 Uncertainty and assumptions 69
4.5.2.1 Physical setting 69
4.5.2.2 Methods 70
4.5.3 Professional judgment 71
5 GEOCHEMICAL EVALUATIONS 72
5.1 Geochemistry Is Not Statistics 72
5.2 Uses of Geochemical Evaluations 72

5.2.1 Identify the processes controlling element concentrations 73
5.2.2 Identify contaminated samples 73
5.2.3 Evaluate statistical outliers 73
5.2.4 Confirm or rule out the grouping of candidate background datasets 74
5.3 General Methodology 76
5.4 Nondetects 77
5.5 Key Geochemical Processes 78
5.5.1 Clay minerals 78
5.5.2 Iron oxides 80
5.5.3 Manganese oxides 82
5.5.4 Additional geochemical processes 84
5.6 Extracting Background Data from Existing Data 85
6 USING GEOCHEMICAL EVALUATIONS IN RISK ASSESSMENT 86
6.1 Using Geochemical Evaluations During COPC Selection 86
6.1.1 Example #1: Evaluate anomalous concentrations further 87
6.1.2 Example #2: Samples confirmed to be unrepresentative of background 89
6.2 Using Geochemical Evaluations During Risk Characterization 91
6.2.1 Example #3: Point-by-point spatial evaluation for delineation 92
6.2.2 Example #4: Incorporation into site-specific risk-based cleanup levels or preliminary remediation goals 93
6.3 Considerations 93
7 ENVIRONMENTAL FORENSICS RELATED TO SOIL BACKGROUND 95
7.1 Introduction 95
7.2 Polycyclic Aromatic Hydrocarbons 96
7.2.1 Forensic analysis of PAHs in soils 96
7.2.1.1 Pattern recognition 97
7.2.1.2 Diagnostic ratios 98
7.2.1.3 Spatial and temporal analysis 100
7.2.1.4 Quantitative source apportionment 100
7.3 Total Petroleum Hydrocarbons (TPH) 101
7.3.1.1 Applications and limitations of different types of TPH data 101
7.3.1.2 Characterization of background TPH 102
7.4 Polychlorinated Biphenyls (PCBs) 103
7.4.1 Forensic analysis of PCBs in soil 105
7.4.2 Data screening and treatment of nondetects 106
7.4.3 Forensic PCB compositional pattern analysis techniques 107

7.5 Polychlorinated Dibenzo-p-Dioxins and Dibenzofurans (PCDD/F) 110
7.5.1 Data screening and treatment of nondetects 113
7.5.2 Forensic data analysis of PCDD/F in soil 113
7.6 Perfluoroalkyl Substances (PFAS) 117
7.6.1 PFAS uses 117
7.6.2 Forensic PFAS compositional pattern analysis techniques 118
7.7 Remote Sensing 119
7.7.1 Overview 119
7.7.2 Items to consider 119
7.7.3 Application 120
8 CONCEPTUAL SITE MODEL AND DATA QUALITY OBJECTIVES 121
8.1 Conceptual Site Model 121
8.2 Data Quality Objectives 122
8.2.1 Quality assurance project plan 123
8.2.2 Sampling and analysis plan 123
9 SAMPLING 125
9.1 Background Reference Areas 125
9.1.1 Natural soil background 126
9.1.2 Anthropogenic ambient soil background 126
9.1.3 Geological and other considerations 127
9.2 Sample Depth 128
9.3 Sample Size 128
9.3.1 Sample size tools 129
9.3.1.1 Visual Sample Plan software 129
9.3.1.2 ProUCL 129
9.4 Sample Methods 130
9.4.1 Discrete samples 130
9.4.1.1 Advantages 130
9.4.1.2 Disadvantages 131
9.4.2 Composite samples 131
9.4.2.1 Advantages 131
9.4.2.2 Disadvantages 131
9.4.3 Incremental sampling methodology (ISM) 132
9.4.3.1 Advantages 132
9.4.3.2 Disadvantages 132

9.5 Sampling Design ... 133
9.5.1 Judgmental (targeted or biased) sampling design ... 133
9.5.2 Probability-based (statistical) sampling design ... 134
9.6 Sample Collection Methods ... 134
9.7 Sample Handling ... 135
10 ANALYTICAL METHODS ... 138
10.1 Introduction ... 138
10.2 Obtaining Reliable Analytical Data ... 139
10.2.1 Data quality ... 139
10.2.2 Test method bias and precision ... 139
10.2.3 Laboratory quality system and analyte group accreditation ... 139
10.2.4 Data validation ... 140
10.3 Analytical Limits ... 140
10.4 Sample Preparation ... 143
10.5 Analytical Test Methods ... 151
11 STATISTICS ... 157
11.1 Data Requirements ... 157
11.1.1 Precautions ... 157
11.1.2 Appropriate use of statistics ... 157
11.1.3 Minimum sample size ... 158
11.1.4 Data selection ... 158
11.1.5 Target population ... 159
11.1.6 Underlying assumptions ... 159
11.2 Data Distribution ... 160
11.2.1 Parametric and nonparametric statistical methods ... 161
11.2.2 Uncertainties ... 162
11.3 Treatment of Nondetects ... 163
11.3.1 Substitution methods ... 163
11.3.2 Kaplan-Meier method ... 164
11.3.3 Regression on order statistics (ROS) ... 165
11.3.4 Maximum likelihood estimation (MLE) ... 167
11.3.5 Summary of nondetect treatment ... 168
11.4 Graphical displays ... 168
11.4.1 Quantile-quantile plot (Q-Q plot) ... 168
11.4.2 Histogram ... 170
11.4.3 Box plot ... 172

11.4.4 Probability plots 173
11.4.5 Percentile plots 174
11.5 Outliers 175
11.5.1 Graphical plots 175
11.5.2 Statistical tests 176
11.6 Confidence Interval Limit, Coefficient, and Limit 178
11.7 Statistical Values Used to Represent Background 179
11.7.1 Upper percentile 180
11.7.2 Upper confidence limit 181
11.7.3 Upper prediction limit 181
11.7.4 Upper tolerance limit (UTL) 181
11.7.5 Upper simultaneous limit (USL) 182
11.8 Statistical Tests to Compare Site and Background Datasets 182
11.9 Statistical Software 186
11.9.1 USEPA's ProUCL 188
11.9.2 Microsoft Excel 190
11.9.3 R 190
11.9.4 Python 191
11.9.5 Integrated statistical programs 191
12 REGULATORY FRAMEWORK FROM STATE SURVEY 193
12.1 Description of State Survey 193
12.2 Overview of State Survey Results 193
12.3 State Survey Results 194
13 EXISTING GUIDANCE AND STUDIES 199
14 CASE STUDIES 200
14.1 Minnesota Pollution Control Agency (MPCA) Soil Background Case Study 200
14.1.1 Existing data 200
14.1.2 USGS and USEPA analytical methods 202
14.1.3 USEPA method reanalysis 203
14.1.4 Statistics 204
14.1.5 Results 205
14.2 Former Firearms Training Range Soil Background Case Study 207
14.2.1 Development of a site-specific arsenic cleanup goal 207
14.2.2 Recommendations 209
14.3 Region 4 RARE Urban Background Study 210
14.3.1 Sample Design 210

14.3.2 Results 210
14.3.3 Uses of Methods and Data 211
14.4 Geochemical Evaluation Case Study—Statistical Outlier Is an Uncontaminated Soil Sample 213
14.5 Geochemical Evaluation Case Study—Statistical Outlier Is a Contaminated Soil Sample 217
14.6 Geochemical Evaluation Case Study–Contaminated Soil Sample Is Not a Statistical Outlier 220
14.7 Environmental Forensics Case Study—PAHs from Leaked Petroleum Versus Contaminated Fill 224

FRAMEWORKS

Framework 1. Establishing Default Soil Background

Framework 2. Establishing Site-Specific Soil Background

Framework 3 Using Soil Background in Risk Assessment

APPENDICES

Appendix A. Upper Limits Used to Estimate Background Threshold Values

Appendix B. Index Plots

Appendix C. Additional Sources of Information for PAHs in Soil

Appendix D. Team Contacts

Appendix E. Glossary

Appendix F. Acronyms

LIST OF TABLES

Table 3-1. Tarone-Ware test results comparing arsenic in surface and subsurface soils 43
Table 3-2. Calculation of BTV estimates for arsenic in surface soils 49
Table 7-1. Diagnostic ratios used to assess PAH sources 99
Table 7-2. Composition of Aroclors by percent level of chlorination 105
Table 7-3. Most commonly reported PCDD and PCDF congeners 112

Table 9-1. Soil sample collection methods 136
Table 10-1. Sample preparation 147
Table 10-2. Analytical methods 152
Table 11-1. Common distributions of environmental data 161
Table 11-2. Outlier tests 177
Table 11-3. Summary of values used to represent BTV 180
Table 11-4. Relationship between number of samples tested and the resulting UCL for an upper 5th percentile of the background population 181
Table 11-5. Assumptions, advantages, and disadvantages of common two-sample tests 183
Table 11-6. Statistical software commonly used in computations involving background concentrations in risk assessments 187
Table 14-1. USEPA analytical method results 206
Table 14-2. Lead summary statistics 211

LIST OF FIGURES

Figure 3-1. Q-Q plots comparing arsenic in surface and subsurface soils. Using the pooled redataset consisting of arsenic concentrations of the existing background and various AOCs; a horizontal line is displayed at the largest detection limit. 44
Figure 3-2. Index plot comparing arsenic in surface and subsurface soils. Using the pooled dataset consisting of arsenic concentrations of the existing background and various AOCs; a horizontal line is displayed at the largest detection limit. 44
Figure 3-3. Normal Q-Q plot of detected arsenic concentrations in the pooled dataset consisting of the existing background (bk) and AOCs surface soil data. 45
Figure 3-4. Normal Q-Q plot of detected arsenic concentrations <33 mg/kg in the pooled dataset consisting of the existing background (bk) and AOCs surface soil dataset. 46
Figure 3-5. Normal Q-Q plot of detected arsenic concentrations ≤18.5 mg/kg in the pooled dataset consisting of the existing background (bk) and AOCs surface soil dataset. 47
Figure 3-6. Normal Q-Q plot of detected arsenic concentrations ≤ 15.1 mg/kg in the pooled dataset consisting of the existing background (bk) and AOCs surface soil data. 47
Figure 3-7. Normal Q-Q plot based upon the extracted background data (*bk-extrct*) for arsenic in surface soil with concentrations ≤ 15.1 mg/kg (detects and nondetects). 48
Figure 3-8. Index plot of arsenic in surface soil comparing AOCs data with extracted arsenic background data (*bk-extrct*) and BTV estimates: 95-95 UTL = 12.1 and 95 USL = 15.1 mg/kg. 50
Figure 3-9. Post plot of arsenic in surface soil showing locations exhibiting background level concentrations (green), intermediate arsenic concentrations (yellow lying between 12.1 mg/kg and 15.1 mg/kg) and concentrations potentially representing impacted site locations (red). 51
Figure 5-1. Cobalt vs. manganese in background surface and subsurface soil samples. 75
Figure 5-2. Cobalt vs. Co/Mn ratios in background surface and subsurface soil samples. 76
Figure 5-3. Barium vs. aluminum in background soil samples. 79
Figure 5-4. Barium vs. Ba/Al ratios in background soil samples. 80
Figure 5-5. Vanadium vs. iron in background soil samples. 81
Figure 5-6. Vanadium vs. V/Fe ratios in background soil samples. 82
Figure 5-7. Cobalt vs. manganese in background soil samples. 83
Figure 5-8. Cobalt vs. Co/Mn ratios in background soil samples. 84

Figure 6-1. Scatter plot of Trace Element A vs. Major Element B shows anomalously high concentrations of Trace Element A in specific samples. 88
Figure 6-2. Ratio plot of Trace Element A vs. Trace Element A/Major Element B. 89
Figure 6-3. Scatter plot of Trace Element A vs. aluminum. 91
Figure 6-4. Ratio plot of Trace Element A vs. Trace Element A/aluminum. 91
Figure 7-1. PAH histograms for materials that can contribute to anthropogenic ambient soil background. (A) road asphalt and (B) urban dust (NIST SRM1649a). Dark blue bars represent USEPA priority pollutant PAHs; light blue bars represent alkylated PAHs often used in forensic assessments; red lines depict (A) bell-shaped and (B) skewed homologue profiles (see (Stout et al. 2015)[369] for additional details). 98
Figure 7-2. Principal component factor score plot. PC1 vs. PC2 (left) showing overall trends among PAHs in 350 sediment samples. Most samples appear as mixtures of weathered creosote and urban (anthropogenic ambient) soil background. PAH profiles for representative samples (A-C) are shown (right). 101
Figure 7-3. General structure for polychlorinated biphenyls. The molecule can contain one to ten chlorine atoms (x), resulting in 209 possible polychlorinated congeners. 103
Figure 7-4. The PCB congener distribution for commercial Aroclors. Congener number (and number of chlorines per molecule) and LOC increase from left to right. 104
Figure 7-5. Candidate background soil samples organized by Aroclor type and concentration. Sixty-five percent of the samples have Aroclor concentrations less than 7 µg/kg and contained either Aroclor 1242 or Aroclor 1248. Higher concentration (>12 µg/kg) soil samples contained higher molecular weight Aroclor 1254 and Aroclor 1260, suggesting potential point source impacts. 108
Figure 7-6. Low PCB concentration background soil samples have a lower LOC composition than suspected impacted site soil samples. 109
Figure 7-7. Generalized structures for polychlorinated dibenzo-p-dioxins and dibenzofurans. The number of chlorine atoms (x) can range from one to eight. 110
Figure 7-8. Examples of combustion and chemically produced PCDD/F congener distributions. 114
Figure 7-9. Sum-standardized star plots showing comparison of PCDD/F pattern for off-site background soil samples with the average PCDD/F pattern for impacted soils. The plot on the left shows the PCDD/F pattern for off-site background soil samples that compare favorably with the pattern for USEPA mean urban background soil, while the plot on the right shows that the average PCDD/F pattern for impacted soils from the related industrial site exhibits a clearly different PCDD/F pattern compared to background. 117
Figure 10-1. Relationship between various analytical limits. 141
Figure 11-1. Examples of assumed nondetect distributions and their corresponding substitution values. 164
Figure 11-2. ROS application example. 166
Figure 11-3. Example Q-Q plot—Similar site concentration and background concentration. 169
Figure 11-4. Example Q-Q plot—Site concentrations diverging from background concentrations. 170
Figure 11-5. Example histogram—Lognormal distributed data versus Q-Q frequency plot. 171
Figure 11-6. Example histogram—Normal distributed data versus Q-Q frequency plot. 171
Figure 11-7. Illustration of box plot and key characteristics. 172

Figure 11-8. Example comparison of the range of metal concentrations in soils at two sites.... 173
Figure 11-9. Example percentile plot.. 174
Figure 11-10. Outlier in a Q-Q plot. 176
Figure 11-11. Statistical tests for comparison of two datasets.. 186
Figure 11-12. Example USEPA ProUCL interface. 189
Figure 11-13. Example USEPA ProUCL box plots.. 190
Figure 12-1. Definitions for natural and/or anthropogenic ambient soil background. 195
Figure 12-2. Recommend using natural and/or anthropogenic ambient soil background values in risk assessment.. 196
Figure 12-3. Established statewide or regional background soil values.................................. 197
Figure 12-4. Studies to establish statewide or regional soil background with or without agency approval/endorsement.. 198
Figure 12-5. Allow geochemical evaluations to support risk assessment. 198
Figure 14-1. Probability plot of 30 samples.. 207
Figure 14-2. Box and whisker plot of confirmation data.. 208
Figure 14-3. Box and whisker plots for lead... 211
Figure 14-4. Scatter plot of aluminum versus iron for background subsurface soil.................. 214
Figure 14-5. Scatter plot of vanadium versus iron for background subsurface soil. 215
Figure 14-6. Ratio plot of vanadium versus V/Fe for background subsurface soil. 216
Figure 14-7. Scatter plot of aluminum versus iron for background subsurface soil.................. 217
Figure 14-8. Scatter plot of arsenic versus iron for background subsurface soil....................... 218
Figure 14-9. Ratio plot of arsenic versus As/Fe for background subsurface soil...................... 219
Figure 14-10. Scatter plot of aluminum versus iron for background surface soil. 221
Figure 14-11. Scatter plot of lead versus aluminum for background surface soil. 222
Figure 14-12. Ratio plot of lead versus Pb/Al for background surface soil. 223
Figure 14-13. Principal component analysis (PCA) factor score plot. 225

SOIL BACKGROUND AND RISK ASSESSMENT

1 INTRODUCTION

Soil background is an important factor to consider when conducting human health and ecological risk assessments. Regulators use numerical screening values specific to contaminants commonly found in soil to evaluate potential risks to human receptors and the environment. Screening values are derived using toxicity data, exposure parameters, and chemical-specific parameters. For some chemicals, the screening value may be within the range of concentrations found in soils from natural or anthropogenic ambient background. Because most regulatory agencies do not require action to be taken if the concentrations of a chemical in soil represent background concentrations, soil background plays an important role in remedial decisions.

Although some state and federal agencies and other entities have guidance documents regarding soil background (for example, (USEPA 1992)[486], (USEPA 2002)[154], (USEPA 2002)[131]), there is not one comprehensive and widely accepted guidance document that summarizes the state of the science on this topic. This ITRC guidance document is intended to fill the gap by providing a comprehensive defensible framework for establishing soil background and using soil background in risk assessments.

1.1 Audience

The primary audience for this guidance document is individual risk assessors, risk managers, and site investigators, which may include federal, state, tribal, and various local agency employees; contractors to these agencies; and potentially liable parties and their consultants. Generally, those involved in developing plans for remedial investigations that include the collection of data for the purpose of risk assessment and background evaluation would benefit from this guidance.

Additional audiences that may find this guidance useful include regulatory agency management and other stakeholders, which could include members of the public and other interested parties. This guidance could be a tool for providing stakeholders a better understanding of how soil background may be applied in a risk assessment.

1.2 Purpose

This document is designed to provide comprehensive guidance regarding the establishment and use of soil background values in risk assessment. It focuses on the process of establishing defensible background concentrations of naturally occurring (for example, metals) or anthropogenic substances (for example, polycyclic aromatic hydrocarbons, polychlorinated biphenyls, dioxins/furans, and per- and polyfluoroalkyl substances) to be used for risk assessment at contaminated sites. The following topics are included in this guidance:

- Soil background definition (Section 2): Provides definitions of natural and anthropogenic ambient soil background, default, and site-specific background, as used throughout this document.

- Establishing default and site-specific soil background (Section 3): Provides information and best practices for establishing soil background by conducting a soil background study or using existing data.

- Using soil background in risk assessment (Section 4): Explains how soil background may be applied in risk assessment.

- Geochemical evaluations (Section 5): Explains the purpose of geochemical evaluations and how they are performed.

- Using geochemical evaluations in risk assessment (Section 6): Explains how a geochemical evaluation can be used as an additional line of evidence to determine whether site concentrations reflect background.

- Environmental forensics related to soil background (Section 7): Explains the purpose of an environmental forensics evaluation and how it is performed.

- Conceptual site model and data quality objectives (Section 8): Discusses how conceptual site model and data quality objectives processes should be considered in developing soil background.

- Sampling (Section 9): Provides recommendations for best practices when designing a soil background sampling program.

- Analytical methods (Section 10): Provides recommendations for best practices when analyzing background soil samples to develop a background soil dataset.

- Statistics used for soil background (Section 11): Identifies statistical methods and best practices for evaluating background soil data.

- Regulatory framework (Section 12): Establishes baseline understanding of level of development by states of background definitions, default background values, and guidance for using background in risk assessment as developed from an ITRC-administered survey to states.

- Existing guidance and studies (Section 13): Provides references to background values and associated guidance developed by states and the U.S. Environmental Protection Agency (USEPA), and background studies from non-state sources.

- Case studies (Section 14): Provides examples of establishing background and using background in risk assessment.

- Frameworks: depict the process generally used (Frameworks).

1.3 Use of Background in the Risk Assessment Process

Background concentrations of chemicals are a factor that should be considered in the risk assessment process, particularly when chemical concentrations at a cleanup site may originate from releases attributable to the site, as well as other sources, including natural sources and anthropogenic ambient sources not attributable to the specific site under investigation. This is especially important when chemical concentrations exceed risk-based values (which are typically numerical chemical concentrations that are protective for specified levels of health risk). Additionally, if soil background levels are elevated or potentially present health risks, this information can be evaluated and presented in the risk assessment process so that the public can be informed (USEPA 2002)[131]. The comparison of site and background concentrations informs risk management decisions and aids in understanding the magnitude and spatial patterns of chemicals of potential concern (COPC) in site media.

A risk assessment is conducted to characterize the current and potential threats to human health and the environment that may be posed by contaminants present at a site. Soil background concentrations are used at two distinct points in the risk assessment process.

- Soil background concentrations can be used in the risk characterization to select COPC and to distinguish between risk contributions from site-related releases and background conditions. To select COPC, site soil concentrations can be compared to default or site-specific background soil concentrations. Many regulatory frameworks allow chemicals that do not exceed soil background concentrations to be excluded as COPC because it is not reasonable to expect cleanups to achieve concentrations less than soil background. A comparison of risk contributions from site and background concentrations may help risk managers make decisions concerning appropriate remedial actions, including evaluating potential risks that can be reduced or controlled by remedial actions and what risks will likely remain due to soil background concentrations.

- Soil background concentrations can be used to establish remedial goals. When background concentrations are greater than risk-based remedial goals, background values are often used to establish remedial goals, because it is not reasonable to expect cleanups to achieve concentrations less than soil background.

A more detailed discussion of the use of soil background in risk assessments is found in Section 4.

1.4 Limitations

This guidance document is intended to describe scientifically sound methods for establishing technically defensible soil background values and appropriate ways to use soil background in risk assessment. It is not intended to provide specific soil background values for chemicals; rather, it is intended to describe appropriate methods and approaches to establish soil background. The focus of this document is on soils, not sediments. This guidance document is intended to inform and support risk-based decision-making and complement existing guidance documents. However, the regulatory framework and the policies of the agency with jurisdiction over the site should be reviewed to ensure compliance with applicable guidance.

Additionally, this guidance is not intended to provide in-depth details regarding sampling and laboratory analysis methods, statistics, geochemistry, or forensics. Rather, it is intended to provide useful information covering these topics to help understand the value and methods to evaluate when establishing soil background and using soil background in risk assessment. Similarly, this guidance is not intended to be a comprehensive and detailed textbook on statistics, geochemistry, or forensics; rather, it is intended to describe statistical, geochemical, and forensic approaches that may be helpful for determining soil background and using soil background in risk assessment.

References appropriate to these topics, including ITRC guidance documents, are cited throughout the document.

2 SOIL BACKGROUND DEFINITION

The soil background definitions provided below were developed for purposes of this guidance document only. These soil background definitions are not intended to capture all of the existing definitions that regulatory agencies and other entities use. They are intended to provide sufficient context to inform and support the reader.

Soil is fragmented particulate material consisting of discrete rock and mineral particles less than 2.0 millimeters in size and varying amounts of organic matter (humus) (ITRC 2020)[431]. At some locations, the particulate material may have formed in place from the physical and chemical erosion of the underlying bedrock, been deposited by geological processes, or be associated with material that was transported to the location.

Soil background includes both natural soil background and anthropogenic ambient soil background. These terms, as well as other terms, are described in detail below. Soil background concentrations vary depending on many factors, including local geology and physical and chemical properties of the soil. The relationship between the soil background concentration to the physical and chemical properties of the soil is described in Section 5. There may be heterogeneity both laterally and vertically and over small and large areas. The background concentrations of a chemical are more accurately described as a range. However, for practical purposes, regulatory agencies often use one specific value to represent soil background. It is important to consider the many factors that can influence soil background concentrations of a chemical and their variability when evaluating whether observed concentrations are a result of a release or natural or anthropogenic ambient soil background.

Soil background can be derived and presented as a range of values or can be expressed as a single numerical value. The decision on whether to present background as a concentration range, a population, or as a single value is based on the project- or program-specific needs. For example, states often use one single value to represent background across an entire state or region of a state to determine whether concentrations for a site reflect soil background. In contrast, responsible parties often compare (using a statistical test) an entire dataset from a site that is being evaluated to an entire dataset of a soil background reference site to determine whether site chemical concentrations reflect only soil background or if they also reflect soil contamination. Understanding the applicable state regulations, study objectives, and how the soil background will be used is critical to determining which approach is most appropriate.

2.1 Natural Soil Background

For purposes of this document, natural soil background is defined as the concentration of a substance, or family of closely related substances (for example, similar element species or similar compounds), present in soil due to geological characteristics, natural processes, or releases from nonanthropogenic sources (for example, wildfires, volcanic activity). Natural soil background does not include releases from local anthropogenic sources, releases from distant anthropogenic sources of persistent chemicals, or other anthropogenic sources of ubiquitous or widespread contamination.

Natural soil background has been defined in multiple federal and state guidance documents. For instance, the USEPA defines natural (soil) background at Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) sites as substances present in the environment in forms that have not been influenced by human activity ((USEPA 1989)[130], (USEPA 2002)[131]). This definition has been adopted by some states and programs.

Some states, such as Maine, Vermont, Wisconsin, Washington, Florida, and Texas, have defined natural soil background in their guidance documents, and in some cases have placed the definitions into rule. Other states, such as Idaho, have definitions for natural (and anthropogenic) background that acknowledge the need for representative measurement to determine background.

2.2 Anthropogenic Ambient Soil Background

For purposes of this document, anthropogenic ambient soil background is defined as the concentration of a substance, or family of closely related substances (for example, similar element species or similar chemical compounds), present in soil due to anthropogenic nonpoint sources, especially when chemicals have the ability to be transported long distances. "Anthropogenic" describes the activity that caused the release while "ambient" describes the widespread distribution of the chemical. This is potentially most relevant to some recalcitrant organic compounds, such as polycyclic aromatic hydrocarbons (PAHs) and polychlorinated dibenzo-p-dioxins and dibenzofurans (PCDD/Fs), which can be present in soil at low concentrations because of their persistence, their ubiquity, or their ability to be transported long distances. Additionally, this could be relevant to metals with a widespread urban source, such as the long-distance transport of mercury. Additional information about these chemicals can be found in Section 7.

What constitutes anthropogenic input and how that is allowed within the context of a project or site has been subject to divergent interpretation. Anthropogenic soil background, anthropogenic ambient soil background, and ambient soil background as defined in various state and federal guidance documents may not be consistent with the definition used in this document. For instance, USEPA defines anthropogenic soil background as natural and human-made substances present in the environment as a result of human activities, not specifically related to the site release in question ((USEPA 1989)[130], (USEPA 2002)[131]), therefore, this definition would include nonpoint ambient sources (or sources that come from many diffuse sources). USEPA's definition of anthropogenic soil background includes both natural and anthropogenic soil background that is not associated with a release from the site that is being evaluated.

When establishing anthropogenic ambient soil background as defined in this document, it is important to carefully consider the intended purpose and clearly define anthropogenic ambient soil background to identify which sources should and should not be included. The definition for anthropogenic ambient soil background varies more widely between regulatory agencies and other entities than that for natural soil background. In most cases, it is defined as including both natural background and diffuse sources of chemicals that can be transported long distances and are present in similar concentrations across a large area (for example, dioxins or PAHs). Local direct or indirect release sources such as those from a specific facility or stormwater runoff are excluded.

For example, when investigating lead, the areas near roadways may be excluded since the impacts of lead-based gasoline may not be uniform throughout the area. However, when investigating lead impacts from an air emission source, it may be necessary to understand anthropogenic ambient soil background near roadways to discern the contribution from the air emission source compared to lead related to gasoline engine emissions on the roads.

Some studies and projects may use alternative definitions of anthropogenic ambient soil background to fit the goals of the specific project. For example, some historical diffuse releases, such as vehicle emissions, might be included in an urban background study if the entire area being included has been subjected to this source and it aligns with the goals of the study and regulatory authority requirements. Some nonpoint or point sources, such as urban runoff, might be included for a site-specific project if the entire area being included has been subjected to this source and it aligns with the goals of the project and regulatory authority requirements. It is generally not appropriate to include releases from industrial activities or reuse of contaminated fill. However, it is noted that some states, such as New Jersey and Massachusetts, recognize historic fill as an anthropogenic ambient soil background condition; historic fill as a background condition is not addressed in this document. In cases where anthropogenic ambient soil background is being characterized for a large area it may not be appropriate to include local diffuse releases or nonpoint or point sources that do not impact the entirety of that area.

2.3 Additional Background Definitions

In addition to natural and anthropogenic ambient soil background, two other types of background definitions are used in this guidance document. Default and site-specific soil background are defined below and are discussed in more detail in subsequent sections (Section 3).

2.3.1 Default soil background

Default soil background is generally established by regulatory agencies for a larger area (for example, a state, a region, or a unique geological area) that generally shares similar physical, chemical, geological, and biological characteristics. Because default soil background is intended to be used to evaluate a large number of sites, it is generally established to be conservative. Default soil background can be established for both natural and anthropogenic ambient soil background concentrations.

Regulatory agencies use default soil background as a screening tool to determine whether contaminant concentrations at an individual site are within the background concentrations of the larger area. A single value (also known as a background threshold value (BTV)) is most often used to represent soil background since this is easier to use in screening. Many states have default soil background values relevant to the entire state or different regions of the state that are used to compare to cleanup site concentrations to determine whether site contaminant concentrations are considered background (site contaminant concentrations are less than the BTV used as the default soil background). For the purpose of this document, default soil background will be referred to as a single value, which is consistent with findings from the state survey and investigation of regulatory guidance conducted to produce this document (Section 12 and Section 13).

Default soil background may be established using a dataset from either:

- an existing soil background study that was conducted for another purpose if it has been evaluated to ensure that it is appropriate to use or
- a soil background study conducted specifically to establish the default soil background.

2.3.2 Site-specific soil background

Site-specific soil background is generally established by a responsible party for an area of limited geographic scope that represents one specific site (for example, an incinerator cleanup site, a railroad yard cleanup site). This is generally a more accurate way to evaluate whether site chemical concentrations are representative of background since it uses information relevant to a specific site in a limited geographical area. Site-specific soil background can be established for both natural and anthropogenic ambient soil background.

If the soil chemical concentrations at a site exceed default soil background values or there is reason to believe that default soil background is not applicable to the site, most regulatory agencies allow responsible parties to complete a more refined evaluation by establishing site-specific background. An area that has similar physical, chemical, geological, and biological characteristics as the cleanup site being evaluated, but has not been subjected to the same chemical releases as the cleanup site, is used to represent site-specific background. The physical, chemical, geological, and biological characteristics of the site being evaluated and of the background reference area used to establish site-specific background are generally more comprehensively characterized when establishing site-specific background.

Site-specific soil background may be established using a dataset from either:

- an existing soil background study that was conducted for another purpose if it has been evaluated to ensure that it is appropriate to use
- a site-specific soil background study conducted specifically to establish soil background for the site being evaluated

The site-specific soil background dataset can be used to:

- establish a site-specific soil background value (for example, a BTV)
- compare the site-specific soil background dataset to the site dataset

Establishing a site-specific soil BTV and comparing it to site contaminant concentrations provides an upper-end estimate of the site concentrations compared to soil background concentrations. Comparing a site-specific background dataset to a site dataset provides an estimate of whether the site concentrations are similar to soil background concentrations or if they exhibit a positive bias that can be attributed to contamination of the site soil.

2.3.3 Soil background reference area

Soil background reference areas are the areas identified as appropriate for collection of samples used to ultimately determine a soil background concentration or range and are also used in ecological risk assessment as described below. Thus, a soil background reference area may serve multiple purposes.

Background reference areas used in ecological risk assessment should be chosen to closely represent natural and geochemical conditions for the site that is being evaluated for potential ecological risks. USEPA states that a soil background reference area for ecological risk assessments is intended to "mirror the physical, climatic, chemical, and biological aspects of the Superfund site" (USEPA 1994)[132]. A well-developed conceptual site model helps to inform the appropriate selection of ecological background reference areas. When conducting an ecological risk assessment, ecological background reference areas may be used in two ways:

- To evaluate impacts on community composition: Biological surveys of flora and fauna present at the background reference area are conducted and compared to ecological receptors present at the site being evaluated. This approach may be used to determine whether the presence of the contaminant has impacted a specific species, or to compare the composition of the community on the sites. For this type of comparison, similarity of the habitat and community composition is generally the most important selection characteristic.

- For selection of COPC. Soil samples from the ecological background reference area can be analyzed to determine background soil concentrations, which may be used to identify COPC needing further evaluation. It is assumed that wildlife, which has evolved in the presence of naturally occurring background, is not impacted by these background concentrations. For this type of comparison, comparable soil characteristics and geologic formation are generally the most important selection characteristics.

Background reference areas used in developing soil background concentrations are chosen to closely represent background conditions for the site being evaluated for potential human health and ecological risks. In some cases, the background reference area is located off site as close to the site as possible, but at larger investigative sites it could also be an area of the site that has not been subjected to site releases. Regulatory agencies have different requirements for identifying soil background reference areas.

3 ESTABLISHING SOIL BACKGROUND

3.1 Introduction

This section includes information on establishing both default soil background and site-specific soil background. Default background studies are typically conducted by regulatory agencies or owners of multiple project sites to address the need for understanding background across a broad area. In contrast, site-specific background is collected specific to a particular cleanup site. When background concentrations are greater than risk-based concentrations, a comparison of site and background concentrations may be used to make decisions concerning appropriate remedial actions, including evaluating potential risks that can be reduced or controlled by remedial actions and what risks will likely remain due to soil background concentrations.

Section 9, Section 10 and Section 11 provide descriptions of sampling, laboratory analytical methods and statistical methods that are relevant to establishing soil background. Please reference Framework 1 and Framework 2, which are provided to depict the process generally used to establish default soil background and site-specific soil background, respectively. Other items important when establishing soil background and using it in risk assessment are a conceptual site model (CSM) and data quality objectives (DQOs) (Section 8). This section is intended to highlight key considerations specific to establishing default soil background and site-specific background, and references details in these other sections when appropriate.

3.1.1 Default soil background

Default soil background is generally established by regulatory agencies for a larger area (for example, state, urban region, or unique geologic zone) that shares similar physical, chemical, geological, and biological characteristics. Regulatory agencies use default soil background as a screening tool to determine whether contaminant concentrations at an individual site are generally within the background concentrations of the larger area. A single value (BTV) is often used to represent soil background since this is a simple way to screen sites, although soil background is more properly described by a range of values.

Many states have default soil background values relevant to the entire state or different regions of the state that can be compared to cleanup site concentrations to determine whether site contaminant concentrations are consistent with background conditions. Most regulatory agencies do not require remedial action for contaminants consistent with appropriate background concentrations (that is, site concentrations are at or below background concentrations). For this document, default soil background will be described as a single value, which is consistent with findings from our state survey and investigation of regulatory guidance (Section 12 and Section 13). Since default soil background values will be used to evaluate a wide range of sites, it is typically established using conservative assumptions or statistical parameters. Default soil background can be established for both natural and anthropogenic ambient soil background concentrations.

Conducting a default soil background study to derive default background values tailored to the information needs is optimal, but not always feasible since this requires significant time and resources. It may be appropriate to use an existing soil background study to establish default soil background if the existing study design and data objectives meet the informational needs of the background study. As reflected in the States Survey (Section 12 and Section 13), not every jurisdiction allows use of anthropogenic background to evaluate site conditions. Please reference Framework 1, which depicts the process generally used to establish default soil background.

3.1.2 Site-specific soil background

Site-specific soil background is generally established for an area of limited geographic scope that represents one specific project site (for example, an incinerator cleanup site, a railroad yard cleanup site). This is generally a more accurate way to evaluate whether site chemical concentrations are representative of background since it uses information relevant to a specific site in a limited geographical area. In many cases, site-specific soil background can be established for both natural and anthropogenic ambient soil background concentrations. As reflected in the States Survey (Section 12 and Section 13), not every jurisdiction allows use of anthropogenic background to evaluate site conditions. Please reference Framework 2, which depicts the process generally used to establish default soil background.

If the soil chemical concentrations at a site exceed default soil background values, most regulatory agencies allow responsible parties to complete a more refined evaluation to establish site-specific background. An area that has similar physical, chemical, geological, and biological characteristics as the cleanup site being evaluated, but has not been subjected to the same chemical releases as the cleanup site is used to represent site-specific background. The physical, chemical, geological, and biological characteristics of the site being evaluated and the soil background reference area used to establish site-specific background are generally more comprehensively characterized when establishing site-specific background.

Site-specific soil background may be established using a dataset from either:

- an existing soil background study that was conducted for another purpose if it has been evaluated to ensure it is appropriate to use
- a site-specific soil background study conducted specifically to establish the soil background for the site being evaluated

Once identified, a site-specific soil background dataset can be used in several ways, including to:

- establish a site-specific soil background threshold value (BTV)
- compare a site-specific soil background dataset to a site investigation dataset

The appropriate study design will depend on project goals and regulatory agency requirements. When conducting a site-specific background evaluation, it is common to both establish a BTV and compare the central tendencies of the background and site datasets. Establishing a site-specific soil BTV and comparing it to site contaminant concentrations can determine if the maximum site concentrations are within the range of soil background concentrations and can help identify potential localized contamination (hot spots) for further investigation. In contrast, comparing the central tendencies of a site-specific dataset to a site background dataset can determine if there may be slight but pervasive contamination. The two procedures are therefore complementary, as they test for the presence of different types of contamination, and they can be performed together. If a given chemical in a site dataset fails either test, then it can be examined further using geochemical evaluation or environmental forensics to confirm or rule out the actual presence of site-related contamination.

Conducting a site-specific soil background study to derive background values designed to achieve project goals is optimal, but not always feasible since this requires substantial time and resources and can have significant administrative hurdles (for example, site access). An alternative is to use an existing soil background study to establish default soil background if the existing study design and data objectives also meet project needs (Section 3.6).

3.2 Conducting a Soil Background Study

Conducting a background study for the purpose of establishing default or site-specific soil background is preferred to the use of an existing study since it allows the study to be designed to achieve project goals. It is important to perform adequate planning to ensure the collected data will address project goals and regulatory requirements. The following items, which are discussed more fully in this section, should be considered when designing a soil background study intended to determine default or site-specific soil background:

- What type of soil background is being obtained, natural and/or anthropogenic ambient?
- What are the definitions of natural and/or anthropogenic ambient soil background? These will impact what types of areas are included and excluded from sampling.
- Does the soil background reference area have sufficiently similar physical, chemical, geological, and biological characteristics to the cleanup site(s)?
- How is the obtained data intended to be used? Will it be compared to a large number of cleanup sites throughout a state or a more limited area such as a region, city, or county?
- Are sampling design and collection methods comparable? When possible, use the same methods to obtain the data for the cleanup sites that were used for the default soil background samples.
- Are laboratory sample preparation and analytical methods comparable? When possible, the same laboratory sample preparation and analytical methods should be used so the concentrations may be compared to one another. This may not always be possible due to logistic or contract laboratory constraints. If different methods are used, differences in results from those methods need to be considered before deciding whether it is appropriate to use the data.

3.2.1 Natural background

When establishing *natural soil background*, it is important to carefully consider (1) the intended purpose, (2) the applicable definition of natural background, and (3) which sources will and won’t be included. Natural background soil concentrations can differ depending on soil type and geologic location and origin ((Chen, Hoogeweg, and Harris 2001)[173], (De Oliveira et al.

2014)[233]). In most cases, samples to establish natural background will be obtained only from areas that have not been influenced by discrete/point source releases (for example, hazardous waste or petroleum releases) or diffuse/nonpoint sources (for example, smelter or lead gasoline emissions). Although there may be some cases where a regulatory agency's definition of natural background may differ slightly from this, for purposes of this document we will use this definition, which is consistent with the definition in Section 2.

There may also be cases where a regulatory agency will allow samples from areas that are not natural background to be included with natural background since they are not from the site under evaluation. In this specific case the definition of background is changed to encompass a release not associated with the site being evaluated plus background. This is also the definition used by the USEPA Superfund program. For the purposes of this document, that would not be considered natural background or anthropogenic ambient soil background, rather it would be considered natural background plus anthropogenic ambient soil background, including point and nonpoint sources not released by the site being evaluated. Samples in proximity to these sources, such as another cleanup site release; stormwater runoff; lead from lead-based gasoline, smelters, or lead-based paint; or other direct or indirect local releases may be included as background samples if allowed by the regulatory agency with authority over the site but clearly should not be used for sites that are not impacted by them. A geochemical evaluation (Section 5) can assist in distinguishing between natural variability and low-level anthropogenic sources in a background dataset.

To ensure that a study appropriately represents the natural background of a selected area (regardless of size), ensure that selected sample locations are unlikely to have been impacted by human activities (Section 9.1). Soil background reference areas typically avoided include roadways, developed areas, industrial areas, and identified local anthropogenic releases. In some cases, it may be difficult to exclude all anthropogenic sources within the soil background reference area of interest. These sources may not be obvious when identifying sampling locations but become obvious when the data are analyzed. The normal heterogeneous nature of soil creates natural variability that may mask anthropogenic sources. The more specific and thorough the sampling criteria are developed to exclude anthropogenic inputs, the stronger the background dataset will be.

It is useful to establish minimum distances between sample sites and anthropogenic sources when developing the sampling plan. For example, in the U.S. Geological Survey (USGS) continental U.S. soil background study, the following distances from anthropogenic sources were used (Smith et al. 2013)[40]:

- more than 200 m (656 ft) from a major highway
- more than 50 m (164 ft) from a rural road
- more than 100 m (328 ft) from a building or structure
- more than 5 km (3.1 mi) downwind of active major industrial activities (for example, power plants or smelters)

Another USGS study to determine natural default soil background trace element concentrations in Wisconsin used the following criteria (Stensvold 2012)[482]:

- must be in a forested lot, permanent pasture, or otherwise undisturbed area at least 6 m (20 ft) away from a fence line.
- must not be within 1.6 km (1 mi) of any other study sample site.
- must not be within 8 km (5 mi) of any other sample from the same soil group.
- must not be within 30.5 m (100 ft) of existing known historical construction site or disturbed area (such as roads, dumps, pits, pipelines, or homesites).
- must not be within 91.4 m (300 ft) of a potential source of contamination (for example, past or present orchard or vegetable-growing area; cattle-dipping site; wood preservation activities; grasshopper bait; land that has had poultry or swine manure, sewage waste, or paper mill sludge applied to it; areas with known releases listed by the Wisconsin Department of Natural Resources Bureau for Remediation and Redevelopment Tracking System).

3.2.2 Anthropogenic ambient soil background

When establishing *anthropogenic ambient soil background,* it is important to carefully consider the intended purpose and clearly define anthropogenic ambient soil background to identify which sources should and should not be included. For purposes of this document, anthropogenic ambient soil background is defined in Section 2.2; however, the definition for anthropogenic ambient soil background varies more widely among regulatory agencies and other entities than that for natural soil background. In most cases, it is defined as including both natural background and diffuse sources of chemicals that can be transported long distances and are present in similar concentrations across a large area (for example, dioxins or PAHs). Local direct or indirect release sources such as those from a specific facility or a stormwater outfall are excluded, which is consistent with anthropogenic ambient soil background as defined in Section 2.

For example, when investigating lead, the areas near roadways may be excluded since the impacts of lead-based gasoline may not be uniform throughout the area. However, when investigating lead impacts from an air emission source, it may be necessary to understand anthropogenic background near roadways to discern the contribution from the air emission source compared to lead related to emissions on the roads.

USEPA Region 4 and Southeastern States Urban Background Study Example

USEPA Region 4 and southeastern states conducted a collaborative urban background soil study to document background concentrations of surface soils in an urban setting. As seen in the table below, the average lead concentrations in surface soils in the cities sampled varies from as low as 14 mg/kg to 213 mg/kg. Each city's mean lead concentration is below USEPA's current residential screening level (400 mg/kg). The variability of lead concentrations in the cities sampled represents the varying concentrations of lead that can be present in an urban setting. These data can aid in understanding when there may be contaminant releases versus anthropogenic ambient background. More information on the urban background study can be found at https://www.epa.gov/risk/regional-urban-background-study.

City	# Samples	Minimum (mg/kg)	Maximum (mg/kg)	Mean (mg/kg)	SD (mg/kg)
Chattanooga, TN	50	14	580	94.8	119.4
Columbia, SC	50	1.7	200	39.9	37.9
Gainesville, FL	50	2	110	14.5	18.2
Lexington, LY	50	18	420	84.3	89.3
Louisville, KY	50	25	1100	163.7	190.5
Memphis, TN	50	13	1000	122.5	199.7
Raleigh, NC	50	7.2	180	32.9	36.4
Winston-Salem, NC	50	20	1400	213.8	241.2

Anthropogenic ambient soil background may also not include emissions from a current local source, such as a smelter or a refinery or areas near a stormwater conveyance. Historic sources may also need to be considered. In urban areas where industrial activity has taken place throughout the region with facilities coming and going over decades, exclusion of all local sources may be more difficult. Scenarios where other sources might be included in anthropogenic ambient soil background are discussed further herein.

It may not be the intent of the default background study to avoid all anthropogenic sources, but rather to obtain samples that reasonably represent conditions in an area or region. The objective might be to include all sources that have been released to the same area or region even if they would not be considered diffuse sources from long range transport. Areas where fill has been placed are generally not sampled to determine soil background. Historic records can guide site selection to avoid fill material and soil borings can be collected to confirm the presence of soil horizons that match those mapped in the United States Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS) soil surveys. If an entire area has been filled or the landscape reshaped prior to development, it may be necessary to analyze soil boring samples collected to below the fill/disturbed layer to establish whether the fill material was contaminated.

PAH anthropogenic ambient background

Several state agencies have recognized the ubiquitous presence of PAHs from atmospheric deposition in various guidance documents and provided methodologies for sampling, analysis, and evaluation ((Cal DTSC 2009)[201], (MADEP 2002)[337]). Recently, a PAH study was performed in Milwaukee, Wisconsin, in collaboration with the Wisconsin Department of Natural Resources (Siemering and Thiboldeaux 2021)[212], where samples were collected and analyzed from park areas demonstrated to be undisturbed and undeveloped for more than 80 years that had no evidence of fill material (confirmed by soil borings) and met the following criteria:

- 6 m (20 ft) from roadways and any parking lots,
- 6 m from other asphalt surfaces (bike/walk paths potentially coal tar sealant–coated),
- >1.5 m (5 ft) from any concrete sidewalks or any other structures,
- exclude athletic fields (if necessary, only along edges thereof), areas that suggest overland soil runoff/deposition from rainwater, and areas where street-cleared snow is piled.

At the remaining areas sampled, atmospheric deposition was considered the only potential source of PAHs, supported by the finding that maximum and median concentrations were vastly higher at the surface (0 to 7 cm) than at depth (15 to 30 cm), dropping to below detection limits for most PAHs. PAH concentrations showed no spatial gradients, and statistical analysis indicated the 3-6 ring PAHs were from the same diffuse local sources with the 2-ring compounds being transported in from outside the region.

The contaminants present in anthropogenic soil background samples can be the result of aerial deposition from diffuse sources (for example, PAHs, dioxins, furans), including smelters, power plants, past home heating with coal, home heating oil, backyard burn barrels, etc., as well as natural sources such as grass or forest fires. In cases such as this, comparisons of specific chemical/congener ratios and principal component analysis can be used as signatures to identify contaminant sources (Siemering and Thiboldeaux 2021)[212] and allow for comparisons to the same ratios in site-specific data. There are also source-specific differences between urban and rural anthropogenic activities. Soils found in densely populated urban areas with long histories of industrial activity will have very different PAH and dioxin mixtures (regardless of concentration) than those found in less populated rural areas.

3.3 Choosing an Area for a Soil Background Study

The primary objective of selecting an area for soil background study is to find a background location that is free of chemical impacts from the site under investigation and has similar characteristics to the study area.

- A background reference area should be in the vicinity of the cleanup sites being evaluated but should exclude any cleanup sites and local releases.

- If there are regional anthropogenic ambient soil background sources, then they should affect the site and background reference areas similarly.

- Contaminant fate and transport pathways should be similar (for example, potential for runoff).

- Vegetation type should be similar (for example, forested vs. scrub-shrub).

Regardless of proximity, the background reference area and cleanup site should share as many physical, chemical, geological, and biological characteristics as possible. Site similarities are crucial to ensuring that the soil background established is relevant and conservative enough to be used to screen the cleanup sites that will be evaluated against it. Details regarding geologic and hydrologic conditions are presented herein. Section 9.1 provides a more detailed discussion of factors to consider when selecting a background reference area.

3.3.1 Geologic and geochemical considerations

When conducting background studies, it is important to ascertain that soil samples of similar characteristics and origin are being compared.

Geologic variability in soil parent materials plays a crucial role in the elemental composition of the soils in a soil background reference area (further discussed in Section 5). The larger a soil background reference area becomes, the more likely that soil types of varying chemical concentration will be encountered. Parameters that are indicative of soil geochemical composition include:

- lithology
- mineralogy
- soil type
- soil salinity
- cation exchange capacity (CEC)
- percent organic carbon
- soil density
- soil porosity
- soil pH and redox potential

Published soil and geologic surveys can typically provide sufficient information on site attributes. Test soil borings can be used to confirm sampling site uniformity of soil on site with the soil background reference area soils. Boring materials may also be tested to confirm soil geochemical attributes. Collection of this information during soil background sampling can aid in further analysis, such as identifying causes for minor differences between a site-specific background site and the study site, or further differentiating background from sites sampled to determine regional default background. Soil background samples should be collected from a designated soil background reference area that reflects the scope for the study. For example, if establishing default soil background for the state of Florida, obtain samples throughout Florida or a representative portion of the state to establish default soil background value. Site concentrations from samples in Florida can be compared to this default soil background value for Florida, particularly if they share the same geology.

If establishing soil background for areas with significantly different geological regions, consider whether it is necessary to establish different default soil background values for each distinct geological region or if regional data will be applied to the entire state. In the 2012 Wisconsin USGS (Stensvold 2012)[482] study, the highest measured default background value for each of the 16 elements investigated was applied to the entire state to simplify regulatory application. Southwest Wisconsin is part of a historic lead and zinc mining district with higher Pb and Zn background values than the rest of the state (52 mg/kg Pb versus 20–30 mg/kg Pb) but creating a separate standard for this region and clearly delineating soils impacted by the ore body was determined to be unfeasible.

Geochemical evaluations employ field observations and analytical data, such as total metals concentrations in discrete soil samples. For example, high metal concentrations may be associated with a specific geologic area (for example, mineralized area) and much lower concentrations are associated with another geologic area. So, in addition to the geological areas being different, the measured concentrations in each area represent unique background concentrations. It is also possible that distinct geologic areas are sufficiently similar chemically. In this scenario, they may not present a unique chemical profile for most metals and may be considered in aggregate for developing background datasets. Geochemistry is discussed further in Section 5.

3.3.2 Hydrogeologic conditions

Saturated soils can affect the concentration of chemicals in soil when chemicals are soluble as well as create greater distinctions between dry weight and wet weight measurements. It is important to understand if saturated soils are present in the soil background reference area (for example, wetlands, creeks, very shallow groundwater) and to avoid these features if the goal of sampling is soil background concentrations. Sampling areas with these features to understand soil background specific to these conditions would be more appropriately a site-specific dataset (Section 3.8). Conditions such as a wetland or a creek should be mapped and called out in the sampling plan. Other conditions to monitor include precipitation and evapotranspiration.

3.4 **Sampling**

Before planning to sample, it is important to develop the context in which a dataset will be used. Default soil background data will generally be used to screen a large number of sites and will be broadly applied. It is designed to be more conservative than site-specific background data to ensure sites that are not truly background are not screened out. Also consider whether default soil background may be applicable to a rural, suburban, and/or urban area. What caveats should be established to guide future use of the default data, such as applicable geologies?

For example, Ohio's Environmental Protection Agency is developing background soil concentrations by county with the intention of using the background values in site assessments in those counties. Individual reports are published for each county that describe the process of collecting, analyzing, and publishing background values.

When choosing the most appropriate site-specific sampling design be sure to consider the project DQOs and how the soil background dataset will be used and what comparisons may be made. It is also important to ensure a sampling method is used that is comparable to the sampling method used for area default soil background determination. In a perfect scenario the same sampling method would be used for the default soil background study and the sites that are being screened using it. This is not always possible since sampling design and methods may differ from site to site for different reasons. A detailed discussion regarding sampling activities is included in Section 9. Key factors in sampling include sampling design (for example, randomized, stratified randomized), numbers of samples, sample type (for example, discrete, composite, or incremental sampling methodology), and sample depth (for example, surface or deeper horizons).

The selection of sampling design is dependent on the goals of the study. If different geologies will be evaluated independently then a stratified random sampling approach is more appropriate. If all samples will be used collectively to develop a single background sample location, then a simple random sampling approach may be more appropriate. Pros and cons for different sample types are described in Section 9.4.

Sampling depth between site-specific and default soil background datasets should also be comparable. In some cases, there may be areas where this is not possible due to underlying rock, but the goal should be to be as consistent as possible. As an example, the USGS dataset (Smith et al. 2013)[40] included surface samples (0 to 5 cm), and two soil horizons: uppermost mineral zone and partially weathered parent material.

3.5 Laboratory Analysis

It is important to ensure usage of analytical methods that give substantively similar results across default background, site-specific background, and site investigation studies. Cleanup sites typically use USEPA's analytical methods, which should also be used for soil background studies. Note that some entities use analytical methods that differ from those validated by the USEPA (see Sections 10.4 and Section 14.1). Analytical methods are discussed further in Section 10.

3.6 Using an Existing Soil Background Study

If it is not feasible to conduct a site-specific soil background study, it may be appropriate to use data from a previously conducted soil background study to establish default soil background for a larger area (for example, a state, region, or unique geological area). Section 13 contains a list of some previously conducted studies that might be appropriate to use. This list is not intended to be exhaustive and there are other studies that can be used. Use of an existing study will reduce investigation costs and time. However, using an existing study that is not appropriate may lead to inappropriate site decisions as well as additional delays or cost if it is decided later that a soil background study specific to the needs is necessary. Thus, an existing study must be critically evaluated to ensure it is appropriate to use.

Site attributes and sampling methods may not completely agree between the default background dataset and the cleanup sites being evaluated. Even so, these studies are not necessarily inappropriate for determining background values, as long as the sampling design and methods are adequately explained and comparable. The implications of differing sample design or methods should be well documented, along with any uncertainties in the comparison with background and investigative site values.

It is important to ensure the study uses adequate documentation, including:

- sampling design and methods (for example, anthropogenic versus natural background, sample depths)
- site topography and soil sample locations
- soil boring logs, including composition, stratigraphy, and depth to water (vadose zone thickness)
- quality assurance and quality control (QA/QC) sampling ((IDEQ 2018)[475])

The specific considerations for using a default background dataset are similar to the considerations for developing a default background dataset discussed in Section 11.1.1. These considerations are identified in this section, but precautions addressed in Section 11.1.1 should be reviewed as well.

- Purpose of existing study
    - What was the purpose of the existing study and how were the data intended to be used?
    - Does the intended purpose match the purpose of establishing default soil background?
    - Does the study include areas and types of samples that would be included if a site-specific background study were to be conducted?
    - Is the area included in the study representative of the cleanup sites that will be using the data for screening?
    - How old is the study? Are the data and results still relevant and representative?
- Type of soil background
    - Was the goal to sample natural background or anthropogenic ambient soil background?
        - Natural background—Were samples collected outside areas influenced by diffuse anthropogenic sources and point or direct sources?
        - Anthropogenic ambient soil background—Were samples collected in areas affected by diffuse anthropogenic sources but at sufficient distance

from direct and indirect releases if these are not included per the regulatory agency (for example, outfalls, roadways, industrial activities)?

- Geologic and geochemical considerations
    - What geologic areas are represented by the soil background study?
    - Are the existing study and the cleanup sites sufficiently similar in physical, chemical, geological, and biological characteristics?
- Hydrogeologic conditions
    - Does the existing study contain samples in sediments or wetlands, or other areas influenced by water?
    - Is this different from the cleanup sites that will be screened?
- Sampling
    - Are the sampling methods used in the existing study sufficiently similar to those used at the cleanup sites that will be screened?
    - Are the sample depths used in the existing study similar to those that will be used at the cleanup sites that will be screened?
- Analytical methods
    - Are the extraction/digestion methods and analytical methods used for the existing study similar to the cleanup sites that will be screened?
    - Are the particle sizes analyzed in the existing study the same as those that will be analyzed in the cleanup sites to be screened?
    - Are the measurements (wet weight or dry weight) used in the existing study the same as those that will be used at the cleanup sites to be screened?

3.7 Background Dataset Analysis

Once a suitable background dataset has been collected or identified careful data analysis must also take place. Of primary importance are data distribution, how outliers are handled, and statistical software selection. These topics and other data analysis topics are covered in depth in Section 11.

Data distribution is a descriptive statistic, often represented by a graphed curve, which describes all the values within a dataset and the frequency at which those values occur. Not all data are distributed in the same manner, and categories have been developed to describe common data distributions. The most recognized distribution is the normal distribution (Section 11.1, Table 11-1). Statistical tests often have underlying assumptions regarding sampling distribution.

An outlier or an outlying observation refers to an extreme observation in either direction that appears to deviate markedly in value from other measurements of the dataset in which it appears. In practice, only outliers that are demonstrably erroneous or belonging to populations not representative of background conditions should be excluded from the background dataset. In background investigations, typical sources of error that can result in outliers include: (a) transcription error, (b) sampling error, (c) laboratory error, and (d) sampling of media not representative of background conditions as determined by forensic and geochemical analyses. Outliers are discussed in depth in Section 11.5.

Selecting the statistical software package that will be used to analyze the background value dataset will significantly impact the background value determination process. There are many readily available software packages that can be useful for background data analysis (see Section 11.9, Table 11-6). While most of the statistical analysis programs listed will have the capability to conduct a majority of the analytical methods required for background statistical analysis, not all programs will be able to easily conduct all methods.

3.8 Establishing Default or Site-Specific Soil Background

Default or site-specific BTV can be an upper bound comparison value generated from the soil background dataset. To calculate a BTV for soil, it is important to review the dataset and understand the distribution of the data (Section 11.2), determine how nondetect values will be handled (Section 11.3) , present the data graphically (Section 11.4), and identify and remove extreme, isolated outlier value(s) (Section 11.5).

Once the background dataset is established, several statistical values are available for use as the BTV for site data comparison. These statistical values are upper bound estimates of the background dataset (definitions from (USEPA 2015)[197]):

- upper percentiles—value below which a specific percentage of the population occurs (for example, 95th percentile).
- upper prediction limits (UPL)—the predicted upper bound value for a single comparison value.
- upper tolerance limit (UTL)—an upper confidence limit on a percentile of the population. For example, a 95-95 UTL is the value below which 95% of the population will fall with 95% confidence.
- upper simultaneous limit (USL)—the upper boundary of the largest value in a background dataset.
- maximum detected value—may result in false positives (for example, the sample set may not be large enough to have fully measured the higher end values), particularly in a small sample set.

Less frequently, the single statistical value provided for a background dataset is an upper confidence limit (UCL). A UCL represents an upper bound estimate of the mean and, if used, should be compared to a mean value for the site dataset; this is useful information to determine whether there is pervasive, low-level contamination of the site soil. It is not appropriate to use the UCL of the mean for a background dataset in point-to-point comparisons with site data, since the UCL of the mean does not represent an upper bound of the soil background concentration (BTV).

It is important to understand the basis for any published background values when making a comparison to site data. Section 11 provides a more detailed discussion of the pros and cons of selecting one of these values as the background value. As also discussed in Section 11, there are several statistical software packages used to evaluate reference datasets and calculate BTVs, including USEPA's ProUCL software. USEPA's ProUCL software is most often used by the regulatory community because it is well documented, relatively easy to use, and specific to the types of statistics that are relevant to the environmental field.

A single sample or a few samples above these upper bound values may not indicate a potential impact to soil above background. As one increases the number of comparisons between site data and the background dataset, the possibility of a true background value in the site dataset exceeding the single background statistic (false positive error) increases.

3.8.1 Using a site-specific background dataset

Comparisons of a site-specific background dataset and a contaminated site dataset also can be made using both datasets rather than (or in addition to) comparison of upper end concentrations to a BTV. These comparisons can be made using statistical methods described in Section 11, such as Student's *t*-test for normal datasets or Wilcoxon rank sum (WRS) test (also referred to as the "Wilcoxon-Mann-Whitney test" or "Mann Whitney U test") for datasets that are not normally distributed. These methods are based on comparisons of the central tendencies of these datasets rather than just the upper end of the distribution. The central tendency comparisons are more reflective of the potential for exposure, which is based on an upper estimate of the mean of a dataset. If the statistical tests indicate that the datasets are different, that suggests that exposure to typical concentrations in one dataset is different from exposure to typical concentrations in the other dataset. Use of both BTV comparison and a dataset comparison provides a more complete analysis of the site and background datasets. As noted previously (Section 3.1.2), the two procedures are complementary, because they test for the presence of different types of contamination. If a given chemical in a site dataset fails either test, then it can be examined further using geochemical evaluation or environmental forensics to confirm or rule out the actual presence of site-related contamination.

3.8.2 Advanced methods

If it is found that the methods discussed above cannot be used to establish site-specific background, there are several advanced statistical methods that may be used to extract a soil background set from a site dataset. A more significant degree of professional judgment is necessary when using these methods, which may lead to significant uncertainty. If it is decided to use one of these methods, it is essential that an expert statistician be included on the project team and understand the underlying uncertainties. These methods are not intended to be used by a risk assessor or risk manager without the assistance of a statistician. Use of one of these methods also requires a larger dataset than the statistical methods described previously in this document. Some advanced statistical methods (discussed in Section 11) that can be useful in some situations include:

- iterative graphical approach (Section 3.9)
- multivariate methods
    - principal component analysis
    - discriminant analyses
    - polytopic vector analysis
    - soft independent modeling of class analogy

3.9 Extracting Site-Specific Background Dataset from an On-site Dataset

The site-specific background dataset extraction approach described in this section briefly addresses technical issues for environmental scientists and managers faced with how to determine site-specific background level analyte concentrations. The site-specific background dataset extraction approach represents an iterative graphical approach to build consensus when site background cannot be determined following standard methodologies and policies described in guidance documents developed by the USEPA and state agencies. The background extraction approach tends to yield a defensible background dataset of reasonable size (often much larger than the one that is collected by traditional sampling of off-site background reference areas) with geological and anthropogenic (when present) influences comparable to those of the site under study. Several new terms and phrases have been used in this section, which are described in Appendix E. This section describes the reality that "background" is often a negotiated estimate rather than a strict statistical or scientific one. Just as the conceptual site model is "evergreen," the understanding of background may evolve during the entire course of an environmental project and should be revisited accordingly.

The USEPA's Office of Solid Waste and Emergency Response (OSWER) has developed several guidance documents (for example, (USEPA 1989)[130], (USEPA 1992)[249], (USEPA 1992)[250], (USEPA 2002)[131], (USEPA 2002)[154] and (USEPA 1995)[220]) covering how a traditional background dataset is to be sampled/collected, how the data are analyzed, and when background data are necessary to perform site and background evaluations. This section is not intended to address federal and/or state agencies' policy-related decisions on when to collect background samples or how to use background data to achieve cleanup levels/achieve applicable or relevant and appropriate requirements (ARARs). It is emphasized that the background extraction approach should be used only when methods following USEPA policies have failed and/or it is not possible to collect a sufficient amount of traditional background data from unimpacted off-site locations. Additional useful information about background evaluations in soils and extraction of site-specific background can also be found in (USDON 2002)[36] and the ASTM E3242-20 (ASTM 2020)[146] document.

3.9.1 Need for background data

Site managers and risk assessors need to determine whether an analyte at the site is present due to 1) site-related chemical releases; 2) non-site-related anthropogenic sources and influences; and/or 3) inherent natural background variability. Determining a site-specific background with natural and anthropogenic influences comparable to the site is an important aspect of performing exposure and risk assessments and of establishing the scope of site-related releases and determining COPC. Some examples of non-site-related anthropogenic activities may include ubiquitous polycyclic aromatic hydrocarbon (PAHs) compounds formed during the incomplete burning of organic materials; vehicular exhaust, and emissions from wearing of tires; domestic heating; or pesticide runoff from agricultural practices at other site areas.

3.9.2 When to use the background extraction approach

In the presence of anthropogenic influences and variable site geology, it becomes challenging to identify an off-site background reference area not impacted by site-related activities because of the confounding factors of non-site-related chemical releases and inherent natural variability. In such complex situations, an iterative quantile-quantile (Q-Q) plots-based background extraction approach following a population partitioning method (for example, (Singh, Singh, and Flatman 1994)[91]) can possibly be used to extract analyte concentrations from a broader on-site mixture dataset representing a site-specific background dataset with geological formations and anthropogenic influences comparable to those of the site. It should be noted that there are other population partitioning approaches available in the statistical literature that can also be used to extract a background dataset from an on-site mixture dataset. Specifically, in the multivariate setting (evaluating several analytes simultaneously), multivariate methods including principal component and discriminant analyses ((Anderson 2003)[230], (Johnson and Wichren 2015)[239], (McLachlan 2004)[468], (Wolfe 2010)[477]) can be used to tease out multiple populations potentially present in an on-site mixture dataset and determine the background subpopulation.

In this section, the univariate Q-Q plots-based iterative approach has been used only to extract a site-specific background dataset from the on-site dataset. It should be pointed out that no attempt has been made to determine and specify potential intermediate subpopulations present in the on-site dataset as described in (Singh, Singh, and Flatman 1994)[91]. It should also be noted that normal Q-Q plots (as used in this section) are routinely used as an exploratory tool (for example, (Tukey 1977)[437], (Hoaglin, Mosteller, and Tukey 1983)[434]) to identify outliers and multiple populations potentially present in a dataset. In this iterative process, no Gaussian model is used to draw any statistical inference, including estimation and hypothesis testing. Therefore, the reader should not assume that the background extraction process described in this section can be used only when the mixture on-site dataset follows a normal/Gaussian distribution. Once a background dataset of a sufficient amount of observations has been established (extracted or traditional), background data distribution is determined using goodness of fit (GOF) tests; many GOF tests are available in the ProUCL 5.1 software. Depending upon the data distribution of the extracted background dataset, hypothesis testing approaches and BTV estimations are used to perform background versus site comparisons. BTVs can be used as screening values to identify COPC and determine site locations exceeding background level concentrations. A BTV represents a parameter in the upper tail of the background population distribution; some statistics used to estimate BTVs include upper prediction limits (UPLs), upper tolerance limits (UTLs), and upper simultaneous limits (USLs). There is no consensus about the use of an upper limit to estimate a BTV. A brief description of these upper limits is presented in Section 11.7 with additional information described in Appendix A. Additional theoretical details can be found in the ProUCL 5.1 Technical Guide (USEPA 2015)[197]. An on-site concentration less than the BTV estimate may be intrepreted as representative of an unimpacted background location, and an on-site concentration exceeding a BTV estimate is viewed as coming from a potentially contaminated site area.

3.9.2.1 *Sites amenable to the use of the iterative background extraction process*

The iterative background extraction process may be used at a contaminated site with the following characteristics:

Sites (for example, federal facilities, industrial complexes, mining sites)

- consisting of heterogeneous areas with high natural variability,
- known to have had many on-site releases with many areas of concern (AOCs) and operating units, and
- with known urban development and other anthropogenic activities (for example, farming, use of petroleum products, training and testing performed by the U.S. military).

Attempts to collect an off-site traditional background dataset following methods described in USEPA guidance were found to be deficient for reasons such as inability to identify relevant unimpacted areas and to collect enough off-site background data appropriate to perform statistical background versus site evaluations.

A database already exists consisting of a large number of analytical results (data points) collected over a defined time providing sufficient coverage to the site AOCs, and all stakeholders agree that it is reasonable to assume that the database consists of concentrations that can be used to represent unimpacted background locations. The size of the existing on-site dataset depends upon the sampling efforts performed at the site and size of the site. For smaller sites, the availability of an on-site dataset of size 200–250 data points may be sufficient; however, for larger sites (for example, federal facilities), it is desirable to have the availability of larger (for example, 300–400 data points or greater) on-site datasets providing sufficient coverage to all AOCs present at the site.

3.9.2.2 *Assumptions and involvement of the project team*

The available on-site data should be large enough to provide sufficient coverage to the site AOCs. In addition to containing concentrations representing locations impacted by site-related releases, the dataset also contains concentrations representing unimpacted locations. Based on this prerequisite, that within all environmental site datasets exist background level concentrations, non-site-related anthropogenic concentrations (may or may not be present), and concentrations indicative of site releases, a normal Q-Q plots-based iterative method represents a viable approach to extract an anthropogenic site-specific background dataset from a broader mixture on-site dataset. As noted above, depending upon the site size, an existing on-site dataset containing at least 200–400 data points may be sufficient to use the background extraction method provided stakeholders agree that it is reasonable to assume that the dataset also contains concentrations representing locations unimpacted by site-related releases. Since most sampling at a site is performed on suspected contaminated areas, the stakeholders need to take into consideration this fact when making an assumption as to the reasonableness of having unimpacted data points in their on-site dataset. Hence this methodology's recommendation for large datasets where multiple attempts at finding representative background samples have already occurred and failed. The more site samples and attempts at determining background through established guidelines, the more likely the dataset contains unimpacted data points within the numerous site samples while still not being easily discernible as representative background.

The involvement of the project team and site experts is essential for successful application of the iterative normal Q-Q plots-based method to determine an appropriate background breakpoint (BP) and extract a site-specific background dataset from a broader on-site dataset. In this section, a background BP represents a value that distinguishes between background level concentrations and concentrations representing impacted site locations. The background BP is determined using Q-Q plots generated iteratively on the on-site dataset. Because of the inherent subjective/expert decision in determining outliers and multiple populations, this method must be performed with sufficient input and agreement from all stakeholders. Based upon the information provided by iteratively generated Q-Q plots, the project team makes the final determination about an appropriate background BP distinguishing between the concentrations representing a background population and contaminated population representing impacted site locations. From the statistical point of view, the approach can be used on any on-site dataset collected from any environmental medium. However, the applicability of the approach may also depend upon the analyte of interest (for example, PAHs) and the site medium (for example, soil) under investigation. It is recommended that the project team consult experts (for example, soil scientists, geochemists) before using the approach on some datasets, such as PAHs in soil.

3.9.2.3 *Treatment of nondetect observations*

Nondetects (NDs) do not represent impacted locations if their detection limits are sufficient to identify concentrations of interest. Sometimes, detection limits (DLs)/reporting limits (RLs) associated with ND observations are significantly higher values (for example, PAHs, metals in soil) than the detected observations. The use of NDs with elevated DLs tends to mask detected observations representing contaminated locations. Elevated NDs exceeding the detected observations interfere with the proper determination of a background BP, therefore causing difficulties in the proper extraction of a site-specific background dataset. In most cases, NDs with elevated DLs should be excluded from the extraction process. It is emphasized that only NDs with elevated DLs need to be excluded from the extraction process; all other NDs may stay in the pooled dataset used to extract a background dataset. Once a background BP has been determined, all detect and nondetect observations less than or equal to the background BP are included in the extracted site-specific background dataset.

3.9.3 Using the existing off-site background data—highly recommended when available

The background extraction approach is used when a representative traditional background dataset of adequate size is not available; guidance about the size of the background dataset is provided in Section 9 and Section 11 of this document. If the team is not confident enough to use the existing background dataset to perform background evaluations and wants to use the extraction process on the existing on-site dataset, the extraction process should be used on the combined on-site and the available off-site background data. In this scenario there is no need to separately evaluate (for example, identify outliers) the background dataset. The iterative process on the combined dataset takes care of outliers (if any) present in the existing background dataset.

3.9.4 Statistical approach

Environmental scientists have borrowed the normal probability plots/normal Q-Q plots-based approach from geochemical and mining applications (for example, (Sinclair 1974)[65], (Sinclair 1976)[66] ,(Sinclair 1983)[67], (Sinclair 1991)[90], (Fleischhauer and Korte 1990)[53], (Halil and Sarac 1988)[54], (Papastergios et al. 2011)[58]). The probability plot/Q-Q plot-based background extraction approach has been used on on-site datasets collected from the various environmental media, including groundwater (for example, (Kim et al. 2015)[56], (Panno et al. 2006)[59], (Panno et al. 2007)[60]), sediments (for example, (Halil and Sarac 1988)[54]), and soils (for example, (Cook 1998)[47], (Matschullat, Ottenstein, and Reimann 2000)[57], (Reimann, Filzmoser, and Garrett 2005)[62], (Reimann and Garrett 2005)[354], (Renez et al. 2011)[64], (Cal DTSC 2009)[52], (HI DOH 2012)[235], (BC Environment 2001)[232]). In related documents available in the literature (and some cited above), the normal Q-Q plots/normal probability plots-based approach has been used as an exploratory tool only to identify outliers and multiple populations present in a mixture dataset.

The exploratory probability plots ((Sinclair 1974)[65], (Sinclair 1976)[66], (Fleischhauer and Korte 1990)[53]) or equivalent Q-Q plots ((Singh, Singh, and Flatman 1994)[91], (Reimann and Garrett 2005)[354]) based method is used to extract a site-specific background dataset from a broader on-site dataset with anthropogenic and geological conditions comparable to those of the site under study. The approach is used on raw untransformed datasets and does not require that the dataset should be normally distributed. In the context of deriving a background dataset from a mixture on-site dataset, a probability plot/Q-Q plot is used as an exploratory tool ((Tukey 1977)[437], (Hoaglin, Mosteller, and Tukey 1983)[434]) to identify multiple populations (and outliers) rather than using it to assess the data distribution. Whether the data are normally or lognormally distributed or follow some other distribution, a normal probability plot in the original raw scale represents a useful tool for exploring the presence of multiple populations and outliers in a dataset.

Normal Q-Q plots are used iteratively to identify locations that can be used to represent site background. Depending on data variability and on-site dataset size, several iterations may be required to determine a subset of lower concentrations that can be used to represent a site background dataset. The discontinuities and inflection points in a Q-Q plot are considered to represent transition between different populations, possibly representing different site areas with varying degree of contamination. When using an on-site dataset consisting of observations from multiple populations, a background BP is selected at a relatively low concentration level (for example, (Sinclair 1974)[65], (Sinclair 1976)[66]), which is determined by the project team using the information provided by the iteratively generated normal Q-Q plots. The inflection points are not always self-evident. In those cases, their identification may rely on expert judgment and that should be recognized and acknowledged by the project team prior to undertaking the process.

Starting from the top of the initial Q-Q plot generated using all data values, discontinuities and inflection points are identified, and new Q-Q plots are generated without using values greater than the inflection point and/or the point of discontinuity. A continuous (without discontinuities and/or inflection points) Q-Q plot (not necessarily exhibiting a straight line) suggests that the dataset comes from a single population. If a Q-Q plot does not represent a continuous graph, the process should be repeated iteratively, removing higher concentrations at each iteration. The iterative process stops when a Q-Q plot displays a continuous pattern without inflection points and/or discontinuities of considerable magnitude as determined by the project team. Based upon continuity, inflection points, and breaks of considerable magnitude present in iteratively computed Q-Q plots, the project team determines a background BP, distinguishing between concentrations representing a background dataset and site data representing impacted site locations. Fleischhauer and Korte (1990)[53] demonstrated that small variations in the estimation of the position of the background BP or the inflection point on a probability plot are unlikely to significantly influence the resulting background concentration breakpoint. By using the iterative Q-Q plots-based approach on a pooled dataset consisting of on-site and off-site concentrations, many on-site locations exhibiting lower concentrations (for example, less than the background BP) will be considered as representing background locations, and background locations exhibiting elevated concentrations (for example, outliers) will not be included in the extracted background dataset.

Once a background BP has been agreed upon by all parties and members of the project team, all observations (detects and nondetects) in the pooled on-site dataset less than or equal to the background BP may be considered to represent an extracted site-specific background dataset. The final Q-Q plot of the extracted background data should be fairly continuous and without inflection points representing a single population. Decision-making statistics such as UTLs are computed based upon a dataset representing a single statistical background population (fundamental assumption). Statistical goodness of fit (GOF) tests are performed to determine the distribution of the extracted background dataset. Depending upon the probability distribution of the resultant background dataset, a parametric or a nonparametric upper limit (for example, UTL, USL) is computed to estimate the BTV. Also, depending upon the project status, project objectives, and data needs, background versus site comparisons may also be performed using graphical displays and hypothesis testing approaches described in USEPA guidance documents ((USEPA 2002)[131], (USEPA 2006)[134]) and available in the ProUCL 5.1 software.

3.9.4.1 *Step-by-step summary of the iterative background extraction method*

A step-by-step summary of the iterative process used to determine a background BP and to extract and establish a site-specific background dataset is described as follows.

1) Use exploratory graphical displays (for example, box plots, index plots, and Q-Q plots) and/or hypotheses testing approaches to determine whether there are significant differences in constituent concentrations in the various strata (surface versus subsurface)

of an environmental medium (for example, soils, sediments). For constituents with statistically significant differences in surface and subsurface soil concentrations, separate background datasets may be extracted for each stratum; otherwise, one background dataset for all strata combined would be extracted. However, it is up to the project team to decide whether separate background datasets would be extracted even when the concentrations of the two or more (for example, soil types) strata are comparable. Statistical methods and graphical displays needed to perform tests listed in this step are available in ProUCL 5.1.

2) Use exploratory iterative normal Q-Q plots on the pooled mixture on-site data to determine a background BP, separating background concentrations/locations representing unimpacted locations and concentrations potentially representing locations contaminated by on-site chemical releases. When elevated DLs are associated with NDs, only detected observations, or all detects and NDs (except those with elevated DLs), may be used in this step. However, it is possible that the true background threshold concentration is below *all* the DLs, and that the detectable concentrations may contain *only* contaminated data. This determination must be made by the project team.

3) NDs may be present in a background dataset; after a background BP has been identified, use all detects and NDs in the pooled dataset less than or equal to the background BP to establish a site-specific extracted anthropogenic ambient soil background dataset.

4) Perform GOF tests on the extracted background dataset. Depending upon the data distribution, compute parametric or nonparametric upper limits (UTLs, USLs) to compute BTV estimates. A brief description of UPLs, UTLs, and USLs is provided in Appendix A.

5) Optional: Use color-coded index plots to compare impacted on-site data and extracted background data. A color-coded index plot representing a snapshot of the entire on-site area with many AOCs and extracted background data provides added insight to the site managers and the responsible party and helps them make informed cleanup decisions.

The approach described here has been illustrated using an arsenic dataset collected from surface and subsurface soils of a real polluted site. A brief description of the computation of upper limits is provided in Appendix A, a description of index plots is provided in Appendix B, and the terminology used is summarized in Appendix E.

3.9.5 Extracting background-level arsenic concentrations from the on-site soil dataset of a Superfund site

This real dataset example illustrates the site-specific background extraction method described in Section 3.9. The dataset used in this example comes from surface (SS) and subsurface (SB) soils of a large Superfund site (Site) containing many AOCs. The Site is very heterogeneous with varying geology and soil types. The on-site SS and SB soils data were collected from many AOCs: a1, a6, a7, a10, a11, a12, a14, a20, a22, a23, s1, s2, s2, s4, s5, and s6 present at the Site. The Site AOCs are contaminated due to site-related releases as well as non-site-related anthropogenic activities (for example, farming). It is also likely that concentrations of the COPC in different

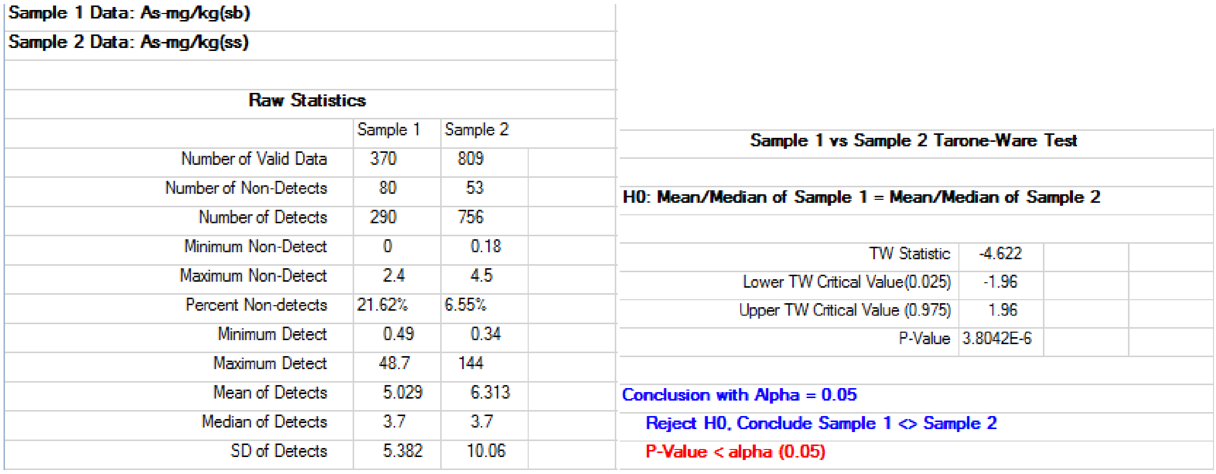
AOCs vary due to natural/inherent variability in Site geology and soil types. A limited amount of off-site background data (denoted by *bk* in graphs) was also available. However, due to natural geological variability and the presence of anthropogenic activities, the project team and the state personnel were not confident that the available background data could be used to perform defensible background evaluations. The project team was concerned that additional traditional background data with inherently comparable and anthropogenic site conditions could not be collected following standard USEPA practices. Therefore, the project team decided to use the iterative Q-Q plot-based method to extract site-specific background datasets to establish sitewide background datasets for the COPC. The existing on-site arsenic dataset collected from soils of the AOCs, and off-site background locations has been used to extract a sitewide arsenic background dataset. This example walks through the background extraction approach used to extract and establish sitewide background datasets and compute BTV estimates based upon the extracted background dataset.

The first step is to determine whether arsenic concentrations in surface soil and subsurface soil are comparable; the data may be combined only if they are not statistically significantly different. Figure 3-1 displays multiple Q-Q plots comparing arsenic in surface and subsurface soils and Figure 3-2 displays an index plot comparing arsenic in surface and subsurface soils (also see Appendix B). In this section, normal Q-Q plots are used to identify multiple populations present in a pooled on-site dataset and determine if the subset consisting of the lowest set of concentrations can be used to represent a site-specific background dataset. Discontinuities (breaks, jumps) and inflection points on a Q-Q plot suggest the presence of data from multiple populations.

The Tarone-Ware test results comparing arsenic in surface soil and subsurface soil are summarized in Table 3-1. The graphical displays shown in Figure 3-1 and Figure 3-2**,** and the Tarone-Ware test results of Table 3-1 suggest that arsenic concentrations in surface and subsurface soils differ significantly (p-value << 0.05). Therefore, the project team decided to extract separate background datasets for arsenic in surface and subsurface soils. The process used to extract and establish a sitewide arsenic background dataset for surface soils is described as follows.

Table 3-1. Tarone-Ware test results comparing arsenic in surface and subsurface soils

Source: Anita Singh ADI-NV Inc.



Sample 1 Data: As-mg/kg(sb)
Sample 2 Data: As-mg/kg(ss)

Raw Statistics

	Sample 1	Sample 2
Number of Valid Data	370	809
Number of Non-Detects	80	53
Number of Detects	290	756
Minimum Non-Detect	0	0.18
Maximum Non-Detect	2.4	4.5
Percent Non-detects	21.62%	6.55%
Minimum Detect	0.49	0.34
Maximum Detect	48.7	144
Mean of Detects	5.029	6.313
Median of Detects	3.7	3.7
SD of Detects	5.382	10.06

Sample 1 vs Sample 2 Tarone-Ware Test

H0: Mean/Median of Sample 1 = Mean/Median of Sample 2

TW Statistic	-4.622
Lower TW Critical Value(0.025)	-1.96
Upper TW Critical Value (0.975)	1.96
P-Value	3.8042E-6

Conclusion with Alpha = 0.05

Reject H0, Conclude Sample 1 <> Sample 2

P-Value < alpha (0.05)

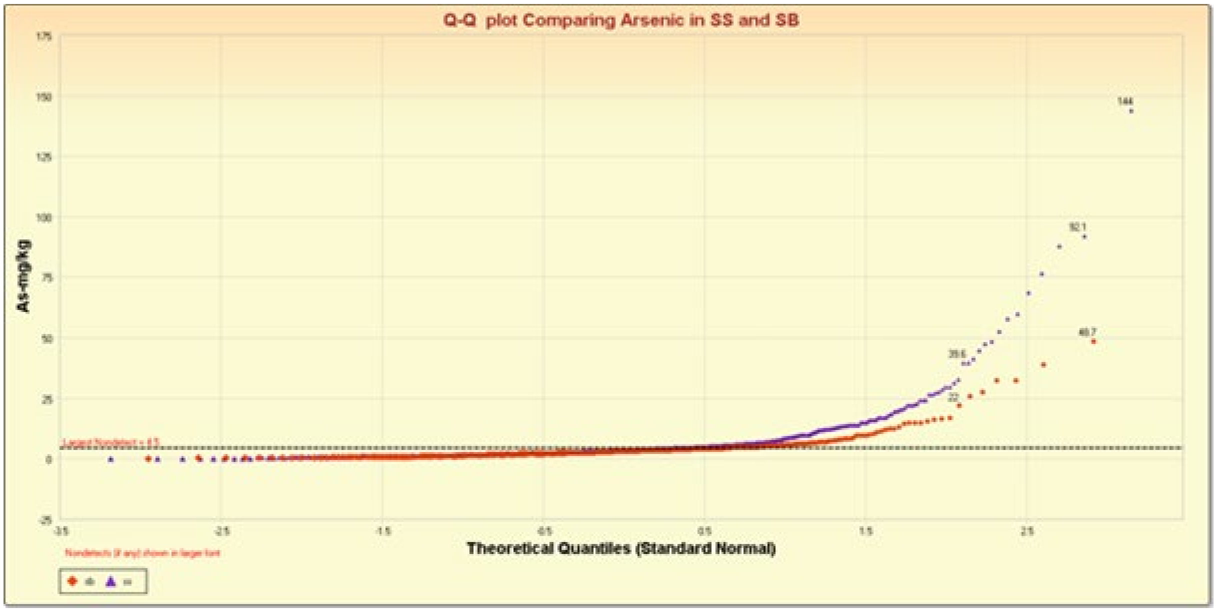


Figure 3-1. Q-Q plots comparing arsenic in surface and subsurface soils. Using the pooled redataset consisting of arsenic concentrations of the existing background and various AOCs; a horizontal line is displayed at the largest detection limit.

Source: Anita Singh ADI-NV Inc.

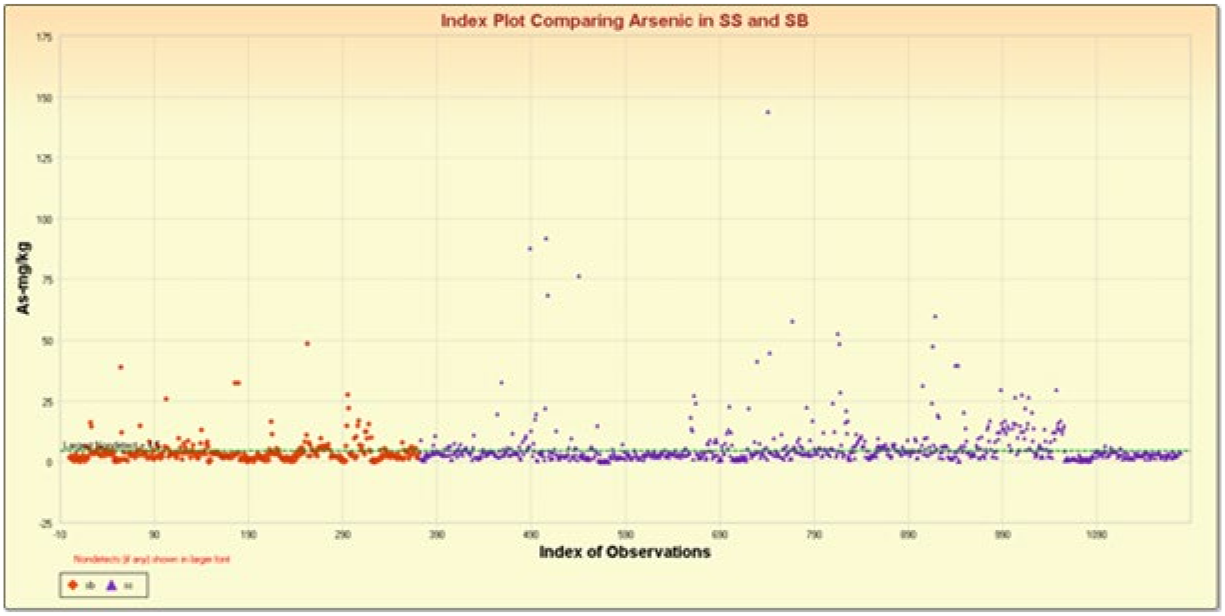


Figure 3-2. Index plot comparing arsenic in surface and subsurface soils. Using the pooled dataset consisting of arsenic concentrations of the existing background and various AOCs; a horizontal line is displayed at the largest detection limit.

Source: Anita Singh ADI-NV Inc.

An examination of the graphical displays shown in Figure 3-1 and Figure 3-2 (and results of Table 3-1) suggests that surface soil of the Site exhibits greater arsenic concentrations than subsurface soil.

Next, iterative Q-Q plots are generated. Figure 3-3 has the initial normal Q-Q plot generated using all detected arsenic concentrations collected from surface soil of the AOCs and the existing background (bk) areas. Note that NDs are excluded from the extraction process but will be included in the extracted background dataset. From Figure 3-3, a large break in the normal Q-Q plot was noted at 92.1 mg/kg and another break was noted around the arsenic concentration of 32.8 mg/kg. To determine (magnify) the magnitude of these discontinuities, another Q-Q plot was generated using arsenic values less than 33 mg/kg as shown in Figure 3-4.

Figure 3-3. Normal Q-Q plot of detected arsenic concentrations in the pooled dataset consisting of the existing background (bk) and AOCs surface soil data.

Source: Anita Singh ADI-NV Inc.

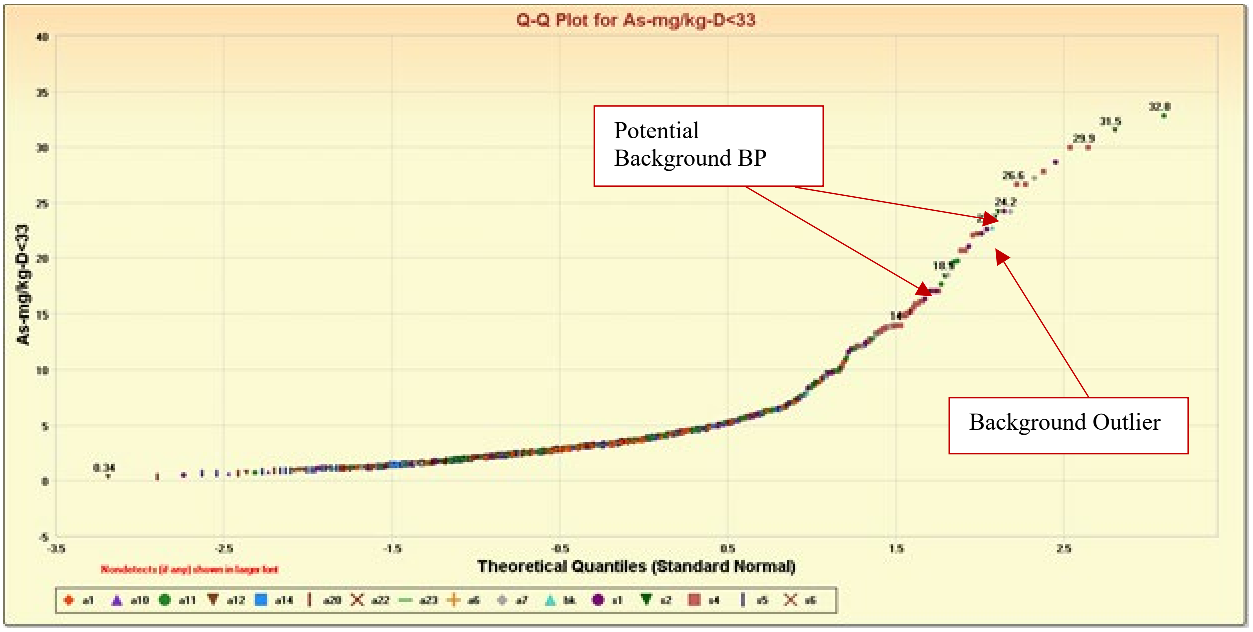
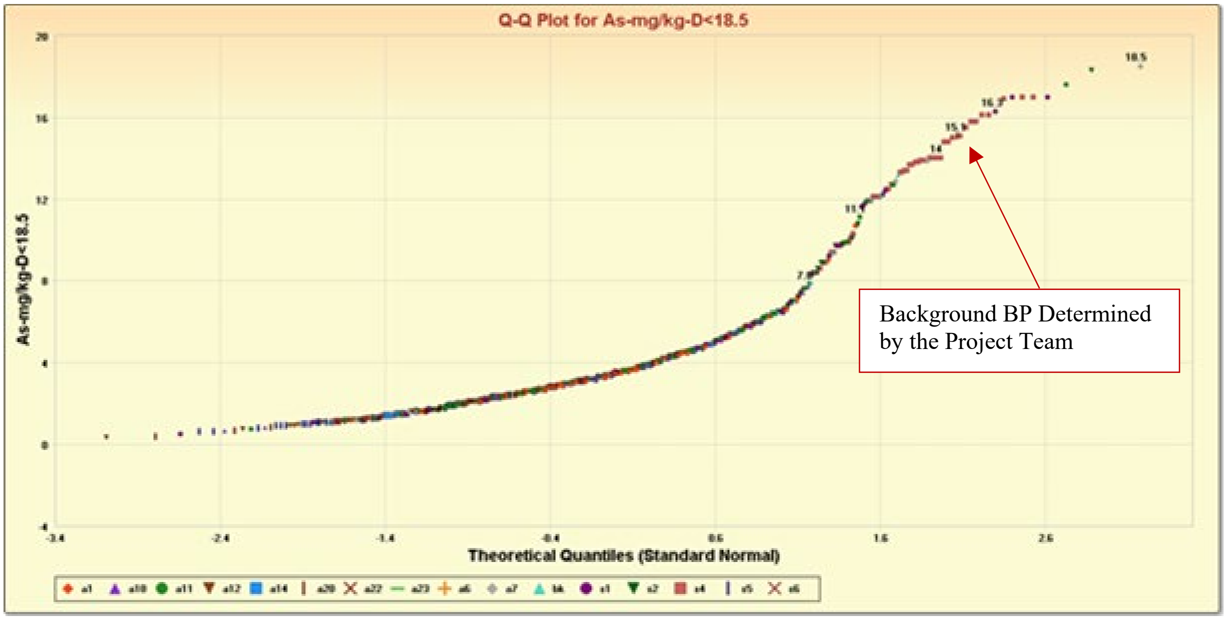


Figure 3-4. Normal Q-Q plot of detected arsenic concentrations <33 mg/kg in the pooled dataset consisting of the existing background (bk) and AOCs surface soil dataset.

Source: Anita Singh ADI-NV Inc.

Observation 22.7 mg/kg shown in Figure 3-4 comes from the existing background dataset (bk) and it represents an outlier in that existing background dataset. After examining the Q-Q plot shown in Figure 3-4, the project team (using available expert site knowledge) determined that 18.5 mg/kg represents a potential background BP. To determine the continuity of the Q-Q plot (with input from the project team) based upon arsenic concentrations ≤ 18.5 mg/kg, another Q-Q plot shown in Figure 3-5 was generated using arsenic values ≤ 18.5 mg/kg. In this figure, a few breaks were noted in the upper part of the Q-Q plot with arsenic values > 15.1 mg/kg; and the lower part of the Q-Q plot with arsenic values ≤15.1 mg/kg appeared to represent a reasonably continuous graph. To confirm these observations, another Q-Q plot shown in Figure 3-6 was generated using detected arsenic values ≤ 15.1 mg/kg.



Background BP Determined by the Project Team

Figure 3-5. Normal Q-Q plot of detected arsenic concentrations ≤18.5 mg/kg in the pooled dataset consisting of the existing background (bk) and AOCs surface soil dataset.

.

Source: Anita Singh ADI-NV Inc.

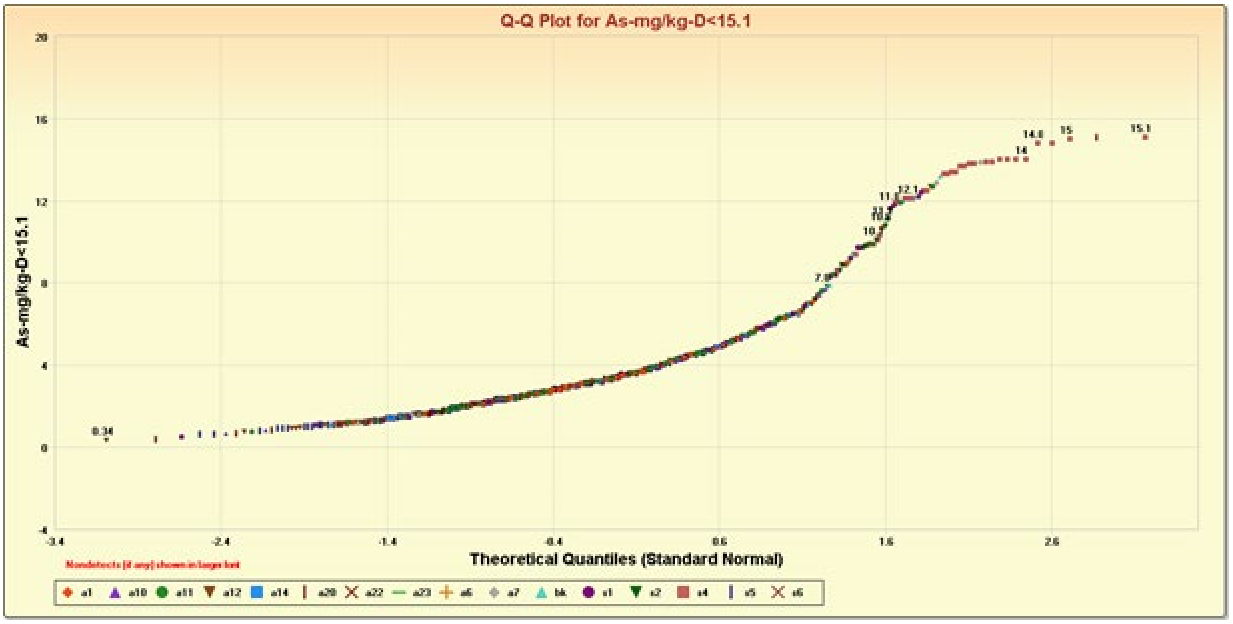


Figure 3-6. Normal Q-Q plot of detected arsenic concentrations ≤ 15.1 mg/kg in the pooled dataset consisting of the existing background (bk) and AOCs surface soil data.

Source: Anita Singh ADI-NV Inc.

The graph shown in Figure 3-6 represents a fairly continuous graph. Input from the project team played an important role at this step. Based upon the information provided by the iterative Q-Q plots shown in Figure 3-3 through Figure 3-6, and taking the conceptual site model (CSM) into consideration, the project team decided to use 15.1 mg/kg as a background BP distinguishing between site-specific background and contaminated on-site concentrations.

<u>Site-specific Background BP and Extracted Background Data (*bk-extrct*):</u> All surface soil arsenic concentrations (detects and NDs) less than or equal to the background BP, 15.1 mg/kg, were used to establish a site-specific background dataset. Figure 3-7 exhibits an exploratory normal Q-Q plot (including detects and NDs) based upon the extracted arsenic background dataset, which is labeled as *bk-extrct* in the graphs.

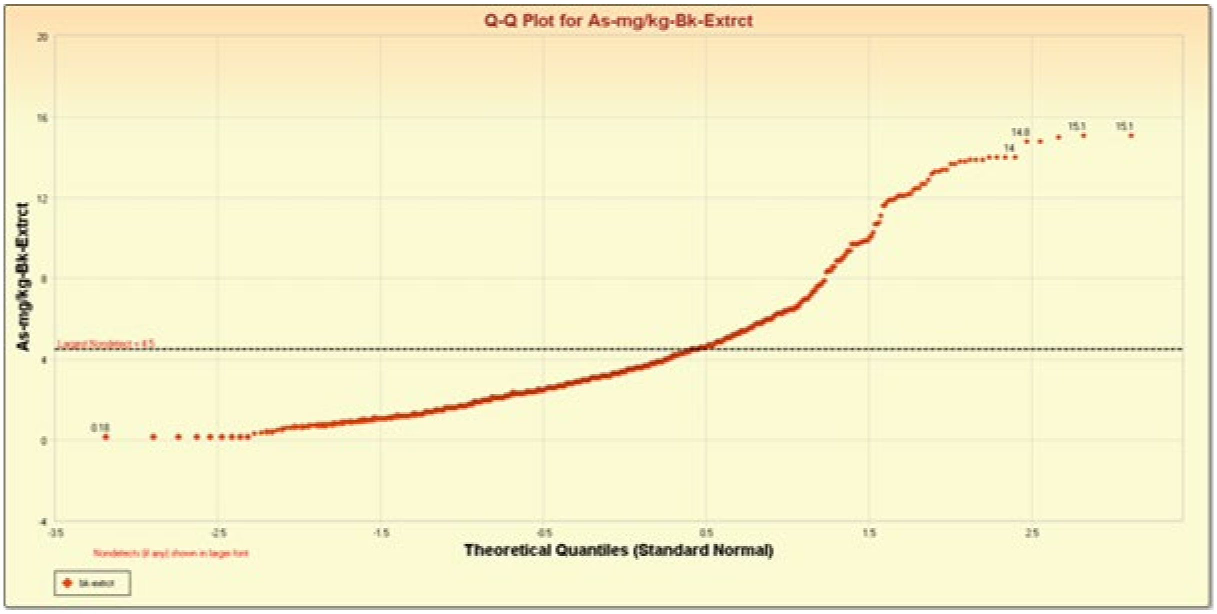


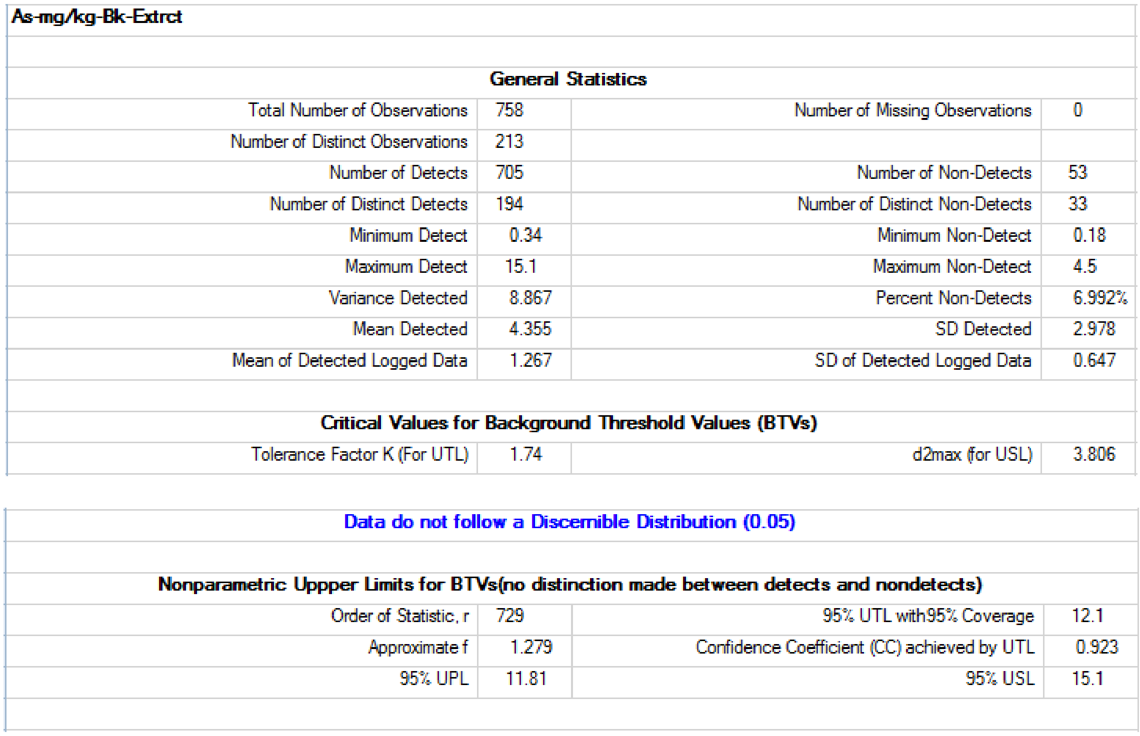
Figure 3-7. Normal Q-Q plot based upon the extracted background data (*bk-extrct*) for arsenic in surface soil with concentrations ≤ 15.1 mg/kg (detects and nondetects).

Source: Anita Singh ADI-NV Inc.

<u>Computing BTV Estimates for Arsenic in Surface Soil:</u> Summary statistics and BTV estimates based upon the extracted background dataset *(bk-extrct)* are summarized in Table 3-2. The detected background data shown in Figure 3-7 do not follow a discernible distribution; nonparametric statistics were used to estimate BTV. In this case, the project team decided to use a 95% USL (=15.1 mg/kg) as an estimate of the BTV.

Table 3-2. Calculation of BTV estimates for arsenic in surface soils

Source: Anita Singh ADI-NV Inc.



As-mg/kg-Bk-Extrct

	General Statistics		
Total Number of Observations	758	Number of Missing Observations	0
Number of Distinct Observations	213		
Number of Detects	705	Number of Non-Detects	53
Number of Distinct Detects	194	Number of Distinct Non-Detects	33
Minimum Detect	0.34	Minimum Non-Detect	0.18
Maximum Detect	15.1	Maximum Non-Detect	4.5
Variance Detected	8.867	Percent Non-Detects	6.992%
Mean Detected	4.355	SD Detected	2.978
Mean of Detected Logged Data	1.267	SD of Detected Logged Data	0.647
	Critical Values for Background Threshold Values (BTVs)		
Tolerance Factor K (For UTL)	1.74	d2max (for USL)	3.806

Data do not follow a Discernible Distribution (0.05)

	Nonparametric Uppper Limits for BTVs(no distinction made between detects and nondetects)		
Order of Statistic, r	729	95% UTL with95% Coverage	12.1
Approximate f	1.279	Confidence Coefficient (CC) achieved by UTL	0.923
95% UPL	11.81	95% USL	15.1

An index plot comparing extracted background arsenic data (in blue) with concentrations of the Site AOCs (not part of the extracted background) is shown in Figure 3-8.

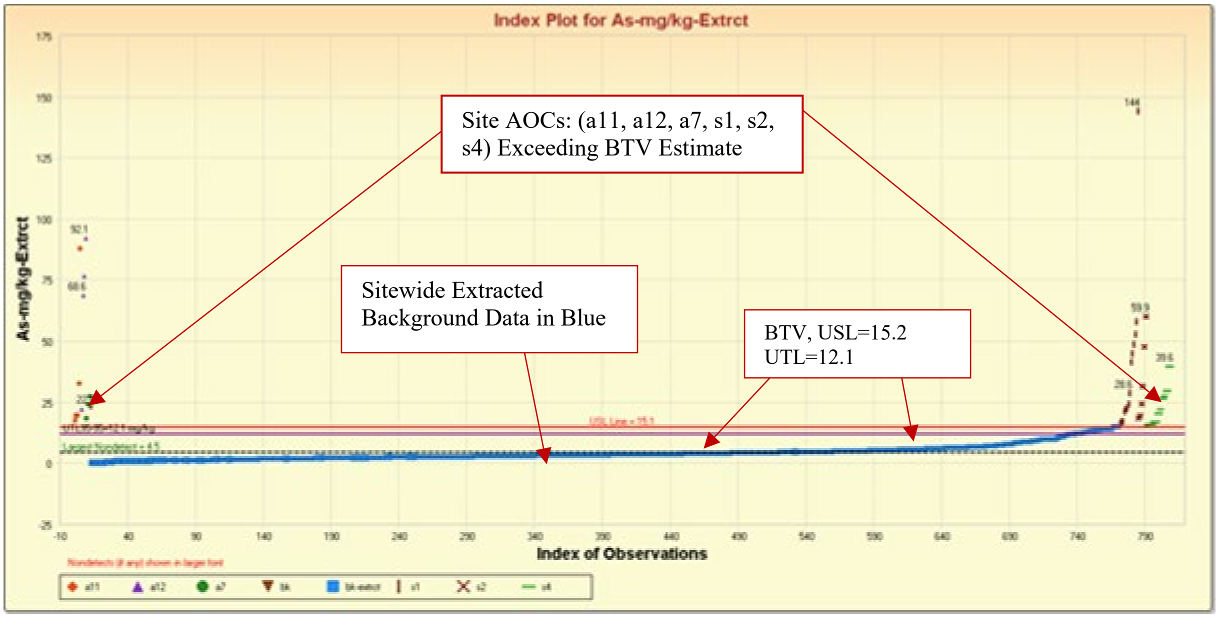


Figure 3-8. Index plot of arsenic in surface soil comparing AOCs data with extracted arsenic background data (*bk-extrct*) and BTV estimates: 95-95 UTL = 12.1 and 95 USL = 15.1 mg/kg.

Source: Anita Singh ADI-NV Inc.

A single index plot of a COPC represents a comprehensive snapshot of the entire on-site dataset by identifying on-site locations with concentrations exceeding the BTV estimates, AOCs exhibiting elevated constituent concentrations, and AOCs having higher concentrations in comparison with the various other AOCs. From Figure 3-8, it is noted that AOCs a1, a10, a20, a23, a6, s5, and s6 exhibit lower arsenic concentrations that are considered to represent site-specific background, and the remaining data from AOCs a11, a12, a7, s1, s2, and s4 exhibit concentrations much higher than those of the extracted site-specific background data (*bk-extrct* shown in blue) and BTV estimates. These kinds of graphical displays help the site managers and the responsible party in making informed decisions to move ahead to make cleanup decisions.

Optional Exercise: For the present Site, GPS coordinates were also available; therefore, a post plot displaying concentrations of the extracted background data and various AOCs was generated. The optional post plot shown in Figure 3-9 separates unimpacted and impacted locations, which was supported by the Site CSM. The generation of post plots is optional because it requires the availability of GPS coordinates of the sampled locations.

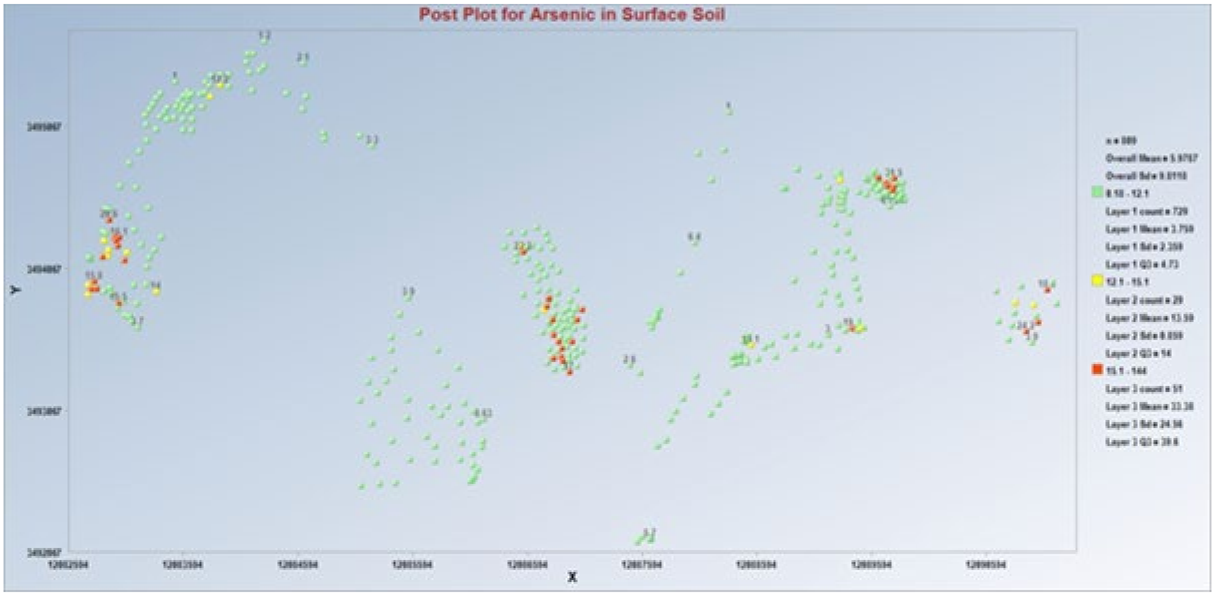


Figure 3-9. Post plot of arsenic in surface soil showing locations exhibiting background level concentrations (green), intermediate arsenic concentrations (yellow lying between 12.1 mg/kg and 15.1 mg/kg) and concentrations potentially representing impacted site locations (red).

Source: Anita Singh ADI-NV Inc.

4 USING SOIL BACKGROUND IN RISK ASSESSMENT

This section describes how default and site-specific background for natural or anthropogenic ambient soil background may be used in human health and ecological risk assessment. Most regulatory agencies allow default and site-specific soil background to be used in human health risk assessment to identify whether chemical concentrations posing excess risk at a site may be attributable to site activities or to soil background conditions.

Different regulatory frameworks use different human health risk assessment procedures. For human health, risk assessments can be conducted by comparing site concentrations to a human health risk-based soil value or by using site concentrations to calculate cancer and noncancer health risk estimates (excess lifetime cancer risks and hazard quotients). Similarly, ecological risk assessments can be conducted by comparing site concentrations to an ecological risk-based soil value or by using site concentrations in food chain models to calculate risk estimates (hazard quotients).

State or federal regulatory frameworks recommend the use of either or both of these risk assessment methods, along with the evaluation soil background. When calculating risk estimates, some agencies may require that the portions of the risks attributable to the site and background be presented separately.

Regardless of the receptor considered (human or ecological), many regulatory agencies allow for an initial screening, which typically involves comparison of a site concentration to a risk-based screening value. When the site concentration is greater than a risk-based screening level, it is not unusual to compare the site concentrations to default soil background values, when available and applicable. Chemicals that are present at concentrations greater than the corresponding human health or ecological risk-based screening values and the default background values are typically carried forward in the risk assessment.

For the purposes of this document, a screening risk assessment is defined as the comparison of site concentrations to human health and ecological risk-based screening values and soil background values. When dealing with a single chemical, a conclusion that concentrations are less than those values indicates that there is no need for further evaluation. When dealing with multiple chemicals at a site, different agencies may use different strategies for retaining or excluding chemicals as COPC. In addition, for the purposes of this document, site-specific risk assessments may proceed into more complex methods to estimate human health and ecological risk with chemical toxicity, site-specific exposures, and in the case of ecological assessments, ecological species comparisons between a reference site and the site being evaluated, and may include site-specific background evaluations.

Therefore, the general applications of soil background in the human health and ecological risk assessment process are as follows:

- Screening risk assessment: To potentially eliminate chemicals that are present in soil at concentrations greater than risk-based screening values but less than default soil background from further evaluation.

- Site-specific risk assessment: For human health, risk characterization is used as well as site-specific soil background information to determine whether site concentrations are representative of background and in some cases differentiate risks associated with site activities from those associated with background conditions.

In addition, a site-specific ecological risk assessment may also characterize site risks by comparing receptors and site conditions with conditions at a soil background reference area.

Moreover, soil background is often used to:

- refine the conceptual site model that is the basis for the human health and ecological risk assessments by using site-specific background information to determine whether the presence of a chemical is due to a release and should be carried forward in the risk assessment or if it is consistent with background conditions and the chemical should not be included among the COPC
- establish remedial goals that are consistent with soil background

These applications described above are intended to show a basic overview of how soil background can be used in the risk assessment process. They are not intended to cover any specific steps included in any one regulatory process.

Comparison of site concentrations to background can be completed using one or more of the following approaches:

- Compare site concentrations to a default background threshold value (BTV) during screening
- Compare site concentrations to a site-specific BTV during a site-specific risk assessment
- Statistically compare (using a two-sample hypothesis test) site and site-specific background concentration datasets during a site-specific risk assessment
- Compare ecological receptors and conditions present at a soil background reference area to ecological receptors and conditions present at the site being evaluated
- Incorporate geochemical evaluations to provide an additional line of evidence whether an inorganic chemical at a site does or does not reflect background concentrations (Sections 5 and Section 6)
- Incorporate environmental forensics to provide an additional line of evidence whether an organic chemical at a site does or does not reflect background concentrations (Section 7)

In some cases (for example, no appropriate default background to use, a highly complex site is being evaluated), a site-specific BTV (or background dataset) might be used during screening. If site-specific background is used in the screening process, it is necessary to ensure that the same criteria outlined in this guidance for site-specific background are followed, including data quality requirements. This type of information is often not readily available during the screening step.

This section describes statistical and non-statistical approaches that can be used to compare a representative site concentration to a default or site-specific BTV, statistically compare a site concentration dataset to a site-specific soil background dataset, calculate potential risks from background concentrations, compare ecological receptors and conditions from a soil background reference area to a site being evaluated, and determine how soil background can be used to establish remedial goals. Please reference Framework 3, which depicts the process for using soil background in risk assessment.

4.1 Representative Site Concentration to Compare to a BTV

When using a default or site-specific BTV in human health or ecological risk assessments, a decision regarding the appropriate site concentration to compare to the BTV must be made. When comparing a site concentration to a risk-based soil value derived based on exposure, toxicity values, and chemical-specific parameters, it is generally accepted to use the 95% upper confidence limit (95 UCL) of the mean concentration (an upper limit of the average site concentrations). This is not the case when comparing site concentrations to soil background to determine whether an area is contaminated since the soil BTV is established using an upper limit value. An upper limit site concentration should be compared to a BTV.

To ensure the chosen site concentration is appropriate, the following items should be considered:

- Type of statistic used to establish the BTV ((USEPA 2015)[197] and Section 11)
    - As discussed in Section 11, each type of statistical value (upper percentile, upper prediction limit, upper tolerance limit, upper simultaneous limit) has its advantages and disadvantages in different situations.
- How the BTV will be used
    - If it is a default soil BTV, it will be used to compare to a large number of sites.
    - If it is a site-specific BTV, it will be used to compare to one site or a small number of similar sites.
- Policies of the regulatory agency
    - Many regulatory agencies have guidance on the representative site concentration they will allow to be compared to a default or site-specific BTV.

For comparisons with BTV, estimates used to represent site concentrations should be consistent with the statistic used to establish the BTV. Possible site concentrations that are appropriate to use to compare to a default or site-specific BTV are maximum and 95th percentile concentrations, which are discussed in Section 4.1.1.1 and Section 4.1.1.2.

4.1.1.1 *Maximum*

The maximum site concentration is generally used for an initial comparison to a BTV, regardless of which type of statistic is used to establish the BTV (for example, 95-95 UTL, 95-95 UPL). The maximum detected site concentration of a chemical in soil represents a conservative, high-end (or upper limit) concentration. Using a maximum to compare to a BTV provides a high degree of certainty that the site concentration is attributable to background if the maximum concentration of the site is below the BTV. Alternatively, in some cases, using the maximum value could result in characterizing site concentrations as not likely to be representative of background when they are actually background. Other important factors for determining the use of the maximum site concentration should be considered, including the sample size and its relationship with the 95th percentile. For some regulatory programs, the maximum may be appropriate when there is large variability of the site data or there is an inadequate sample size.

An additional line of evidence can be gained by comparing each individual site concentration to the BTV by performing a point-by-point comparison. This allows presentation of how many of the individual site concentrations exceed the BTV. For example, results showing that one out of 20 site concentrations exceed a BTV by less than an order of magnitude or results showing that 10 out of 20 site concentrations exceed a BTV by two orders of magnitude may lead to two very different decisions by a regulatory agency as to whether site concentrations represent background. Completing a point-by-point comparison provides additional information that can be used by risk assessors and risk managers to make more informed decisions.

If the BTV is calculated based on the USL, the maximum concentration should be used. As discussed in Section 11.7.5 and Section 11.8, the USL represents a limit that no background concentration should exceed and addresses the false positive error. Therefore, if site concentrations and background are similar, then it should be assumed that the maximum site concentration should similarly not exceed the USL BTV.

4.1.1.2 *95th percentile*

A 95th percentile is an upper limit of the dataset that represents the value below which 95% of the individual data points will fall. In other words, only 5% of the time will a value from that dataset be above the 95th percentile. Section 11.7.1 notes that upper percentiles are reliable for a dataset that is large and representative of a single population; the confidence in this estimate is sample size–dependent. This statistic may be appropriate to use in some cases rather than the maximum but only when the BTV is calculated based on similar statistics, such as a 95th percentile or 95-95 UTL. It is not recommended to be used with a USL for reasons described in Section 4.1.1.1

4.2 **Using Default Background**

For state programs that use default BTVs, most human health and ecological risk assessments will incorporate them during screening. Due to the short- and long-term cost and potential complexity associated with developing a site-specific BTV and a site-specific background dataset, decisions to collect site-specific background data and/or perform geochemical or environmental forensic evaluations are often not made until it has been determined that a site concentration exceeds a default BTV and/or is associated with risks above regulatory limits.

While default BTVs are used in many states and under many circumstances, there are some states that developed different ways to address the issue of soil background in site cleanup. These approaches do not necessarily fall under either default or site-specific background evaluation. For example, Pennsylvania has a background cleanup standard that, although not part of the risk assessment, addresses soil background through the analysis of soil samples of regulated substances present, but not related to, a release at the site (25 Pa. Code § 250.202(b)). Alabama considered soil background during site characterization but uses a process different from those included in this guidance (ADEM 2017)[472]. In addition, some sites are rather complex and may require site-specific background in screening. A thorough understanding of the conceptual site model (Section 8.1) may facilitate identification of the need for site-specific background early in the site investigation process. In extreme cases, because of the complexity of the geology, there could be the need for more than one BTV for a single site.

4.2.1 Comparing a representative site concentration to a default BTV

The first step in comparing a representative site concentration to a BTV is choosing the most appropriate site concentration to use for that purpose (Section 4.1), which includes the maximum and point-by-point concentration comparisons, and upper percentiles. The choice will depend on the statistic that was used to establish the BTV, CSM, DQOs, and the regulatory agency program providing oversight (Section 8 and Section 11.1). Comparing a representative site concentration to a default BTV shows whether the maximum site concentration or an upper bound estimate (upper percentile) is within the range of soil background concentrations and can help identify localized contamination (for example, upper percentile or maximum; see Section 4.1.1.1 and Section 4.1.1.2).

4.2.2 Test for proportions

When the representative site concentration exceeds the default BTV and the regulatory agency requires further assessment, a next step could be comparing the entire site dataset (each individual site concentration not identified as an outlier) to the BTV using the test for proportions available in USEPA's ProUCL software (USEPA 2015)[199] or in other appropriate software packages. The test for proportions evaluates whether the rate of exceedances above the default BTV is significantly different from zero (not attributed to chance). Some regulatory agencies recommend making such comparison before establishing site-specific background since it is easy to do and does not require obtaining any additional data.

As mentioned above, when performing the test for proportions, it is important to eliminate outliers from the site dataset. The decision to eliminate an outlier should not be made based on the outlier test alone. If results from a statistical test show a potential outlier, an investigation should be performed to determine whether it is a true outlier or simply a characteristic of the heterogeneity of soil. Some examples described in Section 11.5 include an error in recording the numerical value or an error by the sampling crew or laboratory in following appropriate quality assurance/quality control procedures.

The test for proportions determines whether an allowable proportion (for example, 1%, 5%, 10%) of the site dataset would likely exceed the BTV (Table 11-3). USEPA's ProUCL User Guide (USEPA 2015)[199]) and Technical Guide (USEPA 2015)[197]) include details regarding how to perform the test and other items to consider, as well as the potential for errors.

4.3 Using Site-Specific Background

When conducting a site-specific human health or ecological risk assessment, a site-specific background dataset can be obtained and used to:

- establish a site-specific BTV as discussed in Section 3.7 for comparison to a representative site concentration (Section 4.3.1)
- compare the site-specific background concentration dataset to the site concentration dataset

In some cases (for example, no appropriate default background to use, or a highly complex site is being evaluated), a site-specific BTV (or background dataset) might be used during screening. Site-specific BTVs provide a more accurate representation of background concentrations in the vicinity of the site than default background concentrations. Thus, when such a value is available (or can feasibly be derived) and it is acceptable within the site's regulatory context, it should be preferred over a default BTV.

When using site-specific background datasets, it is common to:

- establish a site-specific BTV to compare to a representative site concentration to understand if the site concentrations are within the distribution of soil background and to identify localized contamination (for example, a BTV is often compared to the site maximum as part of a screening-level step)
- compare the central tendencies of the site and background datasets using two-sample hypothesis statistical tests (for example, a *t*-test). By comparing their central tendencies and variances, it is possible to identify if there may be slight but pervasive contamination.

The two procedures are therefore complementary, as they test for the presence of different types of contamination, and they can be performed together. If a given chemical in a site dataset fails either test, it may be examined further using geochemical evaluation (for inorganics) or environmental forensics (for organics) to confirm or rule out the actual presence of site-related contamination, if necessary.

4.3.1 Comparing a representative site concentration to a site-specific BTV

A site-specific BTV can be established and used to take a more site-specific look at whether site concentrations are reflective of background conditions. Site-specific BTVs are established using a dataset collected at a background area that has been determined to reasonably match the characteristics of the site being evaluated. Default BTVs are generally applicable to a larger number of sites over a larger area and include the site being evaluated plus many others. Comparing site concentrations to a site-specific BTV determines whether the site concentrations are within the range of soil background concentrations and can help identify localized contamination. The appropriate measures of a site concentration for comparison to a BTV include the maximum and upper percentiles. The choice will depend on the statistic that was used to establish the BTV and the CSM, DQOs, and the regulatory program requirements (Section 8 and Section 11.1). Professional judgment informed by site history can also be used to support decision-making at this point.

If the site concentration is above the site-specific soil BTV, the site dataset may be compared to the BTV using a hypothesis test (for example, test for proportions). The test for proportions is available in USEPA's ProUCL software ((USEPA 2015)[199]) discussed in Section 4.2.2. Other software (for example, R) can also be used to conduct the proportions test or other relevant tests to compare site data to background data (as discussed in Section 4.3.2).

4.3.2 Comparison of site concentration dataset to site-specific background concentration dataset

A statistical comparison of a site dataset's and site-specific background datasets' central tendencies can be conducted using the two-sample hypothesis tests listed in Table 11-5, which also describes the advantages and disadvantages of using each test. These tests can be used to determine whether site concentrations are generally higher than background; in this way, slight but pervasive contamination at the site can be detected statistically. Results from a statistical test alone do not necessarily indicate whether site concentrations are within background. Multiple lines of evidence (for example, Q-Q plots, box plots, geochemical evaluation, environmental forensics) and professional judgment (based on knowledge of site history) should be used to determine whether site concentrations represent background (Section 11.1.5 and Section 11.4). Section 14.6 provides an example of using forensic methods combined with site history and spatial tools to identify site concentrations that are inconsistent with background concentrations and chemical fingerprints and that were potentially influenced by known or suspected sources.

Statistical test results are only as good as the quality of the data used and the validity of the underlying assumptions. Statistical methods are very useful to test the hypothesis that two groups of samples belong to the same population; however, other lines of evidence can be used to support the conclusions drawn from test results. Weight of evidence approaches are useful to reduce uncertainty in risk assessment and decision-making.

4.3.3 Community composition comparison

While this guidance is focused on the chemical aspect of background conditions, ecological risk assessments also often consider one or more biological or ecological background conditions (or "reference conditions"). Ecological risk assessors often use a weight of evidence approach to characterize risk, including various chemical, biological, and toxicological lines of evidence (for example, bioaccumulation modeling or soil toxicity tests). Together these lines of evidence will either point together toward an impact or they will point in different directions, indicating uncertainty. The biological communities at one or more suitable background sites can provide a reference against which to judge the condition of the biological community at the site. Concepts, methods, and limitations of biological surveys and the use of related data in risk assessments are outside the scope of this guidance; relevant guidance should be sought elsewhere. The use of biological background or a "reference condition" is only one of several lines of evidence that can be considered in addition to chemical background (described herein).

4.3.4 Geochemical evaluations and environmental forensics

Geochemical evaluation for inorganics and environmental forensics evaluation for organics are additional tools that can help determine whether site concentrations are from background only or whether there is contamination. In some situations, it may make sense to perform these evaluations early on, in the initial screening step or at the start of the site-specific step, especially for highly complex sites, but in many cases they are not accomplished until the end of the site-specific risk assessment (Section 5, Section 6, and Section 7).

4.3.4.1 *Geochemical evaluations*

Geochemical evaluations can be used to provide an additional line of evidence to support whether or not site inorganic concentrations represent background during human health and ecological risk assessments. Geochemical evaluations are based in part on selected elemental ratios, and they are used to identify the processes controlling element concentrations in soil and to confirm or rule out the presence of contamination in individual samples (Section 5). Accordingly, geochemical evaluations can be used during background studies, such as to verify whether statistical outliers should be retained in a candidate background dataset, and they can be used during comparisons of site versus background datasets. This topic is covered in Section 5 and Section 6. Although these methods may be used at any step in the risk assessment process, in some cases they are not accomplished early in the process due to their complexity, which requires an expert as well as additional resources and funding.

4.3.4.2 *Environmental forensics*

Environmental forensics evaluations can be used to provide an additional line of evidence to support whether or not site organic concentrations represent background during human health and ecological risk assessments. Certain classes of organic chemicals such as polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), and dioxins are complex mixtures of structurally related compounds that, depending on source, occur in unique relative amounts forming distinctive compositional patterns (Section 7). Environmental forensics can be used to determine whether such organic chemicals in candidate background samples share characteristic compositional features consistent with background conditions, and the techniques can also be used to compare site and background datasets. This topic is covered in Section 7. Although these methods may be used at any step in the risk assessment process, in some cases they are not accomplished early in the process due to their complexity, which requires an expert as well as additional resources and funding.

4.3.5 Characterizing risks from background

The primary goal of many environmental regulatory agencies, and of the statutes under which such agencies work, is to protect public health and the environment from current and potential threats posed by uncontrolled or illegal releases of hazardous substances, pollutants, and contaminants. Some of these chemicals may be present in the site soil not only because of releases but also from other natural or anthropogenic ambient sources that are not attributable to site activities, or to other anthropogenic point sources (for example, a release from an adjacent facility). When the total cancer risks and noncancer hazards assessed for a specific receptor at the site exceed either risk-based values (for example, risk-based screening values, ARARs) or their regulatory target risk level (incremental excess lifetime cancer risk (ECR) or hazard quotient (HQ)), it might be useful to evaluate what portion of the risk is associated with exposure to background concentrations and what portion is associated with exposure to site concentrations. The USEPA OSWER 9285.6-07P states that "COPC that have both release-related and background-related sources should be included in the risk assessment." USEPA further states that "if data are available, the contribution of background to site concentrations should be distinguished." It also states "when concentrations of naturally occurring elements at a site exceed risk-based screening values, that information should be discussed qualitatively in the risk characterization" (USEPA 2002)[131]. This can also be accomplished quantitatively, which is the topic of this section. Understanding the potential risk associated with exposure to naturally occurring and anthropogenic ambient background concentrations may provide risk assessors and managers with an additional line of evidence to facilitate informed decisions.

Under CERCLA, USEPA assesses and uses total risk at Superfund sites, including risks from all CERCLA hazardous substances present on site, to establish the baseline risk levels, determine a basis of action, and select an appropriate remedy. The Hazard Ranking System Final Rule defines the site as any "[a]rea(s) where a hazardous substance has been deposited, stored, disposed, or placed, or has otherwise come to be located" (40 CFR 300). Hazardous substances present on the site from non-CERCLA sources are commingled with a CERCLA release, contribute to the total risk at the site, and therefore are included in the total risk for the site. Reporting total site risk levels by subtracting background risk levels from the site risk levels is inconsistent with the CERCLA definition of site risks and the basis for action and would generally not be appropriate for use at a Superfund site. Distinguishing risks contributed by background from risks contributed by a site, however, is consistent with CERCLA guidance ((USEPA 2002)[131], (USEPA 2018)[252]). As noted in USEPA guidance "[if] background risk might be a concern, it should be calculated separately from site-related risk" ((USEPA 1989)[130], (USEPA 2018)[252]). Background levels are relevant to the CERCLA risk assessment and risk management where the risk-based cleanup levels are lower than the background concentrations. In such situations, USEPA generally selects background as the cleanup level. Other federal cleanup programs, such as the Resource Conservation and Recovery Act (RCRA) Corrective Action Program, use this same approach to background. In addition, quantitative risk assessment of background concentrations may provide useful information for communicating site risks to community members and other stakeholders. State and local regulatory agencies may follow different approaches in using background in risk assessment and risk management that permit the approaches discussed below. Also, USEPA regions may vary regarding how they handle calculating risks from soil background. It is best to discuss the approach to background with the lead regulatory agency early in the risk assessment process.

4.3.5.1 *Incremental risk analysis of soil background*

Risk characterization used in human health risk assessment summarizes and integrates the outputs of toxicity and exposure assessments to characterize baseline risk, using both quantitative expressions and qualitative statements (USEPA, 1989). Quantitative estimates for carcinogenic COPC are presented as the ECR, which is an increase in probability of developing cancer over a lifetime exposure to COPC. For noncarcinogenic COPC, quantitative estimates are presented as an HQ, which is an estimate of the potential for an individual to develop adverse health effects. Cumulative ECRs are calculated by summing individual ECRs for all COPC. Hazard indexes (HI) are calculated by summing all HQs for all the COPC that act on the same target organ, or by the same mechanism, considering all potentially complete exposure pathways and media for each receptor considered in the CSM.

When COPC that are related to site-specific background condition are included in the risk characterization, the calculated site ECRs and HIs may overestimate risks resulting solely from site-related releases or operations. Incremental risk analysis provides insight on the magnitude of the relative contribution that exposure to background COPC has on the calculated total cancer risks and noncancer hazards.

In risk characterization equations, the receptor-specific exposure point concentration (EPC) for a COPC directly affects the ECR or HQ for that receptor. If all other parameters stay the same, a change in the EPC will typically result in a proportional change in the cancer risk and HQs for the COPC.

When comparing the risk associated with exposure to background COPC to that from exposure to COPC present because of site-related activities, the EPCs for both the site-related releases and background should be based on estimates of central tendencies (for example, 95% UCL of the mean).

The incremental risk analysis method is based on a statistical relationship between the risk estimates (ECR and HQ) from exposure to site-related COPCs and exposure to site-specific background COPC. Site-specific EPCs for both the site-related releases and background are used in incremental risk analysis to determine the relative contribution of background to the estimated total site risk. It is important that the EPCs for background used in the incremental risk analysis are based on site-specific background data collected using a similar sampling design as the site data and collected from similar soil material and depth for each complete receptor exposure pathway. In addition, when comparing background risk to site-related risk, the reasonable maximum exposures should be calculated using the same type of EPC. The EPC for background should be developed consistent with the methodology used to calculate the EPCs for the risk characterization using site data (ITRC 2015)[237]. Refer to Section 9 for specifics on sampling design.

The following sections discuss specific approaches for using site-specific background in incremental risk analysis.

COPC identification

Incremental risk analysis begins with identification of constituents that are a background condition. Statistical tests discussed in Section 11 and Section 4.3.2 can be used to identify site-related COPC that are consistent with a background condition. In addition, geochemical evaluations and environmental forensics discussed in Section 5, Section 6, and Section 7 may be used to identify COPC to provide a line of evidence on site-specific background condition. The geochemical evaluation determines whether site-related metals data are consistent with site-specific background conditions, and if the metals data needs to be evaluated further in the incremental risk analysis. Geochemical evaluation can assist in distinguishing low-level site-related COPC concentrations from naturally occurring background COPC concentrations, which allows the risk assessor or risk management team to decide whether a COPC should be retained as a COPC for risk characterization. Environmental forensics in the context of this guidance is focused on determining whether the observed chemical concentrations in soils may be representative of natural or anthropogenic ambient soil background.

Based on the outcome of these evaluations, if selected COPC are identified as statistically correlated to the site background condition, then a site-specific background EPC can be calculated for the selected COPC. When assessing the risk, the reasonable maximal exposure should be developed using an exposure point concentration that is an estimate of the central tendency (for example, 95% UCL of the mean) and not an upper bound threshold value (for example, 95-95 UTL, 95th percentile) for both the site-related releases and background, as pointed out in ((ITRC 2015)[237], Section 6.2.5.1).

Cumulative Risk Approaches

The cumulative incremental risk analysis method evaluates the potential contribution of site-specific background to the total site ECRs and HI. Total site risks are calculated using an EPC for the site, and background-related risks are calculated using EPCs for background. The same exposure model, exposure equations and assumptions, and toxicity values are used in calculation of the total site and background ECR and HIs. Refer to specific regulatory agency guidance on inputs to exposure equations and toxicity values.

Once total site risks and background-related risks are calculated, the risks are compared using one of or any combination of the following methods:

- Subtraction method—the total ECR and HI calculated for background are subtracted from the total site ECR and HI to differentiate the portion of the cancer risks and noncancer hazards that are from site-related releases and background ((ITRC 2015)[237], Section 6.2.5.2).
- Percent contribution method—calculates the potential percent contribution of background ECR and HQ to the total site ECR and HQ to differentiate the portion of the cancer risks and noncancer hazards that are from site-related releases and background.
- Comparison method—compares the total site ECR and HI to the ECR and HI calculated for background. In this method, the difference between total site risk and background risk is not quantified.

Note that regulatory agencies typically require reporting of the total site risks, so the emphasis of these methods is on comparison of total site risks and risks related to background condition. The acceptability and specific approach of applying these methods may be USEPA region-, state-, and/or project-specific. It is best to discuss the approach for evaluating distinction of background and site-related risks with the lead regulatory agency. The subtraction method and the percent contribution method would generally be inappropriate for use under CERCLA or RCRA and would not be accepted by USEPA. Information gained from the comparison method may be useful for risk communication, but generally would not be appropriate for CERCLA or RCRA decision-making (USEPA 2018)[252].

Cumulative Risk Example (USEPA generally does not accept this method)

An incremental risk analysis for a site identified that excess cancer risk (ECR) for an industrial worker and hazard index (HI) for a construction worker exceeded the regulatory agency ECR and HI targets of 1×10^{-6} and 1 (target risk levels may vary; please consult the lead regulatory agency regarding the appropriate target risk level to use), respectively. Potentially complete exposure pathways for the industrial worker and construction worker included exposure to COPCs in soil via direct contact (for example, incidental ingestion and dermal contact) and inhalation of volatiles and particulates.

Using the site-specific background exposure point concentration (EPC) developed for the site, the site-specific background cumulative cancer risks for the industrial worker ($ECR_{background}$) and noncancer hazards for the construction worker ($HQ_{background}$) were calculated. The same exposure assumptions, exposure equations, and toxicity values used to calculate total risk were used to develop the background incremental excess lifetime cancer risk ($IELCR_{background}$) and $HI_{background}$ for each receptor. A site-specific background EPC was developed using the same statistical method (95UCL of the mean) as the EPC in the calculation of the site cumulative ECR (ECR_{Site}) and HI (HI_{Site}).

The cumulative ECR_{Site} calculated for the industrial worker was 8×10^{-6} and the HI_{Site} was 3 for the construction worker. The $ECR_{background}$ for the industrial worker was calculated to be 1×10^{-6} and the $HI_{background}$ for the construction worker was 2.

The following presents an analysis of the incremental risk of site-specific background on the cumulative ECR_{Site} and HI_{Site} for the construction worker.

Industrial Worker

Subtraction Method

$$Incremental\ Risk\ Without\ Background = ECR_{Site} - ECR_{background}$$

$$Incremental\ Risk\ Without\ Background = 8 \times 10^{-6} - 1 \times 10^{-6} = 7 \times 10^{-6}$$

Percent Contribution Method

$$\frac{ELR_{background}}{ECR_{Site}} = \frac{1 \times 10^{-6}}{8 \times 10^{-6}} \times 100 = 0.125\%$$

Comparison Method

Total Site ECR	Background ECR
$8\,x\,10^{-6}$	$1\,x\,10^{-6}$

Construction Worker

Subtraction Method

$$IIncremental\ Risk\ Without\ Background = \llbracket HI \rrbracket\ \_Site - \llbracket HI \rrbracket\ \_background$$

$$Incremental\ Risk\ Without\ Background = 3 - 2 = 1$$

Percent Contribution Method

$$\frac{HI_{background}}{HI_{Site}} = \frac{2}{3} \times 100 = 66.7\%$$

Comparison Method

Total Site HI	Background HI
3	2

Based on the incremental risk analysis, the site-specific background condition is not significantly contributing to the cumulative ECR for the industrial worker scenario. The contribution of the site-specific background ECR ($ECR_{background}$) to the ECR for the site (ECR_{site}) is 0.125%. When the ECR calculated for the site-specific background condition $ECR_{background}$ is subtracted from the ECR calculated based on the EPC (ECR_{Site}), the ECR still exceeds the regulatory agency target level of 1×10^{-6}.

However, for the construction worker, the contribution of the site-specific background HI ($HI_{background}$) is 66.7% and when the $HI_{background}$ is subtracted from the HI calculated based on the EPC (HI_{Site}), the cumulative HI is reduced to 1 and does not exceed the regulatory agency target level of 1.

Therefore, ECR for the industrial worker is site-related and may warrant remedial action. For the construction worker, the majority of the elevated HI is attributable to the site-specific background condition and may not require remedial action.

Risk Drivers

Another incremental risk analysis method is focused on potential contribution of site-specific background to the risk drivers identified in the risk characterization step. Risk drivers are defined as those chemicals that pose an unacceptable risk to human or environmental receptors and potentially trigger a need for response action ((USEPA 2001)[441], (USEPA 2001)[442]). The same calculations discussed under Cumulative Risk Approaches are used to compare the ECRs and HQs for the COPC-specific risk drivers using the EPCs for the site and background.

This method is a streamlined approach and is focused solely on risk drivers for potentially complete exposure pathways and does not evaluate the cumulative risk from all site-specific background data. A benefit from using this approach includes decreased level of effort (time and costs) to identify contribution of site-specific background to COPC that may require remediation because of a site-specific risk assessment. Risk managers may then incorporate this information in risk management decisions.

Risk Driver Example (USEPA generally does not accept this method)

From the previous example, arsenic and benzo(a)pyrene in surface soil were identified as risk drivers for the industrial worker exposure scenario. Site-specific background EPCs developed for these COPCs were used to evaluate potential contribution of background conditions to risk drivers using the percent contribution method. Both background and site EPCs were based on the 95% UCL of the mean.

Value	Arsenic	Benzo(a)pyrene
COPC ECR	1 x 10^{-6}	4 x 10^{-6}
Site-related EPC	4.22 mg/kg	10.4 mg/kg
Background EPC	3.55 mg/kg	0.208 mg/kg

The arsenic background EPC is divided by the arsenic total site EPC to determine a potential contribution to the EPC from the site-specific background condition.

$$\frac{Background\ EPC_{arsenic}}{Total\ EPC_{arsenic}} = \frac{3.55\ {}^{mg}/_{kg}}{4.22\ {}^{mg}/_{kg}} \times 100 = 84\%$$

The site-specific background condition for arsenic may be contributing up to 84 percent of the total EPC for arsenic. To calculate a potential incremental risk from the site-specific background condition, the COPC ECR is multiplied by the percent contribution of risk from background. It should be noted that $\text{Background}_{\text{incremental}}$ risk is rounded to 1 significant figure ((USEPA 1989)[130]).

$$Background_{incremental\ risk} = COPC\ IELCR\ \times 84\%$$

$$Backgound_{incremental\ risk} = 1\ x\ 10^{-6} \times 84\% = 1\ x\ 10^{-6}$$

Similar to the arsenic evaluation, the potential contribution to the ECR from the site-specific background condition is calculated using the following equations for benzo(a)pryene.

$$\frac{Background_{benzo(a)pryene}}{Site - related\ EPC_{benzo(a)pryene}} = \frac{0.208\ {}^{mg}/_{kg}}{10.4\ {}^{mg}/_{kg}} \times 100 = 2\%$$

$$Background_{incremental\ risk} = COPC\ IELCR\ \times 2\%$$

$$Background_{incremental\ risk} = 4\ x\ 10^{-6} \times 2\% = 8\ x\ 10^{-8}$$

This evaluation provides support that the majority of the risk associated with arsenic is a background condition and that arsenic is not a site-related COPC. Conversely, this evaluation suggests that the majority of risk associated with benzo(a)pyrene is not a background condition and that benzo(a)pyrene is a site-related COPC rather than attributable to a site-specific background condition. Therefore, this risk driver evaluation helps inform the risk manager of potential response action that may be warranted based on the concentrations of benzo(a)pyrene detected in soils at the site.

4.4 Use of Background for Remedial Goals

Typically, when site conditions are associated with unacceptable risks, the lead regulatory agency requires some form of response action. Some response action may include numeric remedial goals, which are frequently based on site-specific risk-based values. When a risk-based remedial goal is within the range of the soil background concentrations of the same chemical, the soil background concentration may be used as the achievable remedial goal instead of the risk-based concentration.

There are two ways to use soil background to set a remedial goal:

- use a default or site-specific BTV
- compare the soil background dataset to the site dataset to determine which areas may require a response action

Remedial goals can be based on two types of BTVs: those that are based on a central tendency BTV, and those based on an upper limit. The type of BTV to use is a risk management decision, not a risk assessment decision. On one hand, using a central tendency BTV as a remedial goal is a more conservative approach than using an upper limit BTV. On the other hand, a remedial goal based on a central tendency BTV is more likely to give false positive errors (concluding that a soil sample is contaminated when it is well within the bounds of background) than one based on an upper limit BTV. Conversely, using an upper limit BTV may be more susceptible to false negative errors (concluding that a soil sample is not contaminated when it is contaminated).

Regardless of which type of BTV is used, it is important to ensure that both the BTV and the site concentration used to compare to the BTV are of the same type of statistics. For example, if an upper limit BTV is used as the remedial goal, then the site concentration to compare to the upper limit BTV should also be an upper limit. If a central tendency BTV is used as the remedial goal, then the site concentration to compare to the central tendency BTV should also be a central tendency.

One possible approach, often adopted as a first-tier comparison, is comparing site concentrations to a remedial goal set to a BTVs on a point-by-point basis. This can focus remediation planning on those areas where concentrations exceed the BTV (for example, hot spots of contamination). For example, all areas of the site with concentrations exceeding the BTV might be remediated to background levels. There are alternative approaches for comparing site concentration to a BTV, including comparing a representative site concentration to a BTV (Section 4.2.1 and 4.3.1) or a test for proportions (Section 4.2.2).

A site dataset and soil background dataset may also be compared to determine which areas require cleanup using statistics. Hypothesis testing may be used to indicate whether hot spot removal can reduce site concentrations to a level similar to background without remediating all areas of the site. Hypothesis testing may also be used to support the decision to remediate soil across the entire site when low-level, pervasive contamination is present. In cases where low-level, pervasive contamination is present, hot spot removal would not reduce site concentrations to background levels.

Ultimately, the choice of which BTV to use for setting a remedial goal will depend on many different factors, including which regulatory agency is providing oversight, the applicable regulatory framework, and negotiation among stakeholders. The rationale for selecting a specific type of BTV for a given site should be clearly explained and documented to ensure transparency in the decision-making process.

As noted in previous sections, it may be prudent to incorporate geochemical evaluation and/or chemical forensic analyses at this stage or earlier in the risk assessment process to confirm whether site and background datasets are chemically representative of background (not just statistically comparable). When possible, it is recommended that these analytical tools be incorporated earlier in the risk assessment process to ensure that background characterization and risk characterization are well supported.

4.5 Additional Considerations

When considering soil background in risk assessment, additional considerations such as bioavailability of chemicals and uncertainties should be accounted for and addressed, when possible and if necessary.

4.5.1 Bioavailability of chemicals in site and background soil

Bioavailability is important to understand when using soil background because it makes the chemical more (or less) available for uptake. The relative bioavailability of a chemical in the background soil may be different from the relative bioavailability of the same chemical in the exposure area (where there is contamination). How to account for background bioavailability is described in detail in Section 9.2.3 of the ITRC's Bioavailability of Contaminants in Soil Guidance document (ITRC 2017)[206]. For human or ecological exposures, contaminants in soil may not be as bioavailable as the forms used in the toxicity tests used to develop risk-based criteria. When calculating the relative bioavailability of a COPC in the background reference area, it is important to make sure that the conditions (for example, sample depth, soil type) are comparable to those on the site. Although bioavailability is better understood for lead, arsenic, and PAHs ((USEPA 2007)[223], (USEPA 2012)[224], (ITRC 2017)[206]), particularly for the incidental ingestion of soil, states such as Hawaii recommend bioavailability testing when site concentrations exceed default BTVs (HI DOH 2011)[439].

If the bioavailability of a chemical was considered or modified in generating a BTV, the conditions of the site must closely match the background study. This is of the utmost importance because bioavailability is greatly influenced by soil characteristics, including mineralogy, grain size, pH, and soil organic matter.

Although it is unlikely that bioavailability of chemicals will be considered in developing a BTV, methods for some chemicals and receptor types have been developed that simultaneously consider background chemical concentrations and bioavailability in site soils. One approach developed for ecological receptors (plants and invertebrates) and metals, for example, is to develop soil cleanup values based on background concentrations and bioavailability modeling of "added metal" (the metal concentration greater than background) (Checkai et al. 2014)[202]. An example of a freely available and flexible tool for developing bioavailability-based soil guidelines that incorporates background soil concentrations of metals is available from Arche Consulting (2020)[200].

When site-specific BTVs are generated, due to site concentrations exceeding a risk-based screening level or default BTVs, an assessment of a contaminant's bioavailability at a site may be considered.

4.5.2 Uncertainty and assumptions

A common thread among all the assumptions in using a default or developing the site-specific soil BTV is to lean toward being acceptably conservative to be protective of the most sensitive receptors. The consideration of soil background in a risk assessment will allow a focus on the risks actually associated with a release or site. Therefore, it is important to clearly identify and explain the assumptions and uncertainties made when establishing default or site-specific BTVs. In contrast with default BTVs, site-specific BTVs are developed for each individual site and are often not readily applicable elsewhere. The use of a BTV, especially a site-specific BTV, should be approved by the proper regulatory agency prior to the completion of the risk assessment.

4.5.2.1 *Physical setting*

When using soil BTVs for a specific area, there is uncertainty about the physical setting because geochemical and physical processes are important to consider along with geological characteristics.

To use default soil background values, both physical and chemical parameters need to be similar at both the site and the locations used during the background study. Based on the amount of information needed, the level of detail required to characterize a highly variable site does not typically allow the comparison to a default BTV. For sites with highly variable soil properties, equally detailed site-specific BTV are most appropriate for comparison. Background concentrations will be inherently variable depending on spatial distribution of the samples. This variability is rooted largely in the heterogeneous nature of soil. Before using a site-specific BTV in an area of interest that has high natural variability, confirm that the variability pairs well with the data used to generate the site-specific BTV (Section 3.7).

Assuming the geochemical and physical processes that determine background, the chemical concentration for a constituent of interest can guide how the background concentration is determined. An assumption of aerial deposition may lead to the collection of surface soil samples in the study used to develop default BTV. On the other hand, knowing that chemicals such as metals, are naturally occurring, one could assume that metals are widely spread throughout the soil column. In that case, it would, therefore, be appropriate to compare samples from various sample depths to a default BTV (while also considering soil depths that are appropriate to the exposure of the human or ecological receptors being evaluated in the risk assessment). Comparing site data to a BTV should follow the same assumptions and sampling schemes. This may limit the applicability of a default BTV.

It is important to consider if the geology of the location where a default BTV was developed is similar to the location that is being evaluated to ensure that the value was established from an area with similar geological characteristics. In many cases, this will be less of an issue with site-specific soil BTV development due to a more narrowed focus on choosing an area that has similar geology. If the area of interest falls in an area with a high degree of geological variation, the development or use of site-specific BTV may not be approved by the lead oversight agency. The remainder of geological considerations should have previously been addressed during both the sampling (Section 9) and establishment of default soil BTV (Section 3.7).

There is also uncertainty in potentially changing conditions of background reference areas over time due to unexpected events (for example, flooding, wildfires, runoff). Areas prone to natural disasters or in a dynamic environment will experience temporal changes. Data collected at one of these locations may represent only a snapshot in time and the applicability of a BTV may be limited. Additionally, unknown off-site activities that increase or decrease background concentrations may underestimate or overestimate naturally occurring constituents. Site-specific BTV may be needed; however, site-specific values generated from previous studies carry additional uncertainties that increase with the age of the data and whenever different parties conduct the studies. Additional care in the form of geochemical evaluations (Section 5) or environmental forensics (Section 7) may be needed to determine whether the site data are representative of site-specific background.

4.5.2.2 *Methods*

The statistical methods used will influence the calculated BTV (refer to Section 11). The study's handling of nondetect values (Section 11.3), outliers (Section 11.5), and sample location each carry uncertainty into the risk assessment. Methods and assumptions used in determining a BTV balance between two outcomes: false positive or false negative error rates. Methods to estimate BTV typically attempt to reduce the number of false positives or to reduce the number of false negatives. Understanding how the BTV is established and the assumptions made are essential for their scientifically sound use. It is important to ensure that the analytical chemistry method used to quantitate COPC in a soil sample is sensitive enough to detect reliably amounts of COPC smaller than the correspondent risk-based concentrations (for example, screening levels, remedial goals). Much uncertainty can come from the exclusion of certain chemicals due to missing or unknown risk-based screening values or lack of sampling for those chemicals. If chemicals were previously eliminated due to data limitations, the consequences for excluding those chemicals should be discussed (USEPA 1989)[130].

As technology and laboratories' ability to detect lower concentrations continually improve, previously established default BTV can become out of date. For more information on detection limits and their incorporation in soil background risk assessments, refer to Section 10.

Soil collection method is a consideration that is made early in the DQO process and will impact how site data can be used in the risk assessment. According to ITRC (2020)[431], although discrete and ISM samples could be combined with some caveats, in practice, there are no established methods for combining discrete and ISM data. The underlying theory between the two types of sampling methods is different enough that the results of the average would conceptually not make sense. Section 9.4 lists advantages and disadvantages of different sample types.

4.5.3 Professional judgment

A large part of professional judgment will be addressing the assumptions noted in Section 4.5.2. Professional judgment should be used to make decisions regarding sampling or whether a default background study is appropriate for comparison. If site conditions fail to meet the parameters/assumptions of the default BTV study and are, therefore, unable to be used, developing site-specific values is an option.

Section 4.1 describes options available for comparison of site concentration and a default BTV. The data and considerations used in the development of the default BTV and the site's dataset for a compound will identify some options as inappropriate for use. Professional judgment and communication with the regulatory agency are essential when considering background concentrations of chemicals in a risk assessment.

5 GEOCHEMICAL EVALUATIONS

Geochemical evaluation is a technique based on selected elemental ratios that is used to identify processes controlling element concentrations in soil and confirm or rule out the presence of contamination in individual samples. These ratios reflect trace elements' affinities to adsorb on the surfaces of specific minerals in soil. An anomalous ratio can indicate the presence of excess trace element from a contaminant source. While commonly used during site investigations to delineate contaminated areas and refine lists of COPC ((Myers and Thorbjornsen 2004)[469], (Thorbjornsen and Myers 2007)[267]), geochemical evaluation is also important during background studies, where the goal is to characterize representative background concentrations and conditions ((Thorbjornsen 2008)[269], (Geiselbrecht et al. 2019)[15]).

A notable feature of geochemical evaluation is that it employs the field observations and existing analytical data, such as environmentally available metals concentrations in discrete soil samples, that are acquired during standard environmental investigations. This minimizes overall project cost, because specialized analytical data are not required (Thorbjornsen and Myers 2007)[267]. Elemental ratios have long been used as a geochemical prospecting technique ((Hawkes and Webb 1962)[258], (Boyle 1974)[255], (Levinson 1974)[259], (Whitney 1975)[270]) and are well suited for the background applications discussed below.

The geochemical evaluation described in this guidance does not include stable isotope analyses, sequential extraction procedures, or other specialized analytical techniques. Not only are such techniques outside the scope of this guidance, but they are also not necessary for background determination at most sites. Geochemical evaluation is also not a statistical technique (Section 5.1), but rather it complements statistics by evaluating analytical results from a different perspective, thereby reducing decision errors that are inherent to any single methodology. Geochemical evaluations should be performed by experienced geochemists, because many potential geochemical processes need to be considered when evaluating soil datasets.

5.1 Geochemistry Is Not Statistics

There are many statistical procedures that are applied to environmental data, but they are not used during geochemical evaluation, which has as its aim the identification of geochemical mechanisms controlling element concentrations in soil. It is not a hypothesis test with associated probabilities; it does not assume linear relationships between elements, nor do linear models apply, for the reasons explained in Thorbjornsen and Myers (2007)[268] and ASTM E3242-20 (ASTM 2020)[146]. A properly performed geochemical evaluation identifies geochemical mechanisms controlling element concentrations in individual soil samples, which procedures such as multivariate statistics cannot do. Geochemical considerations are an important complement to statistical evaluations and serve as an independent line of evidence when characterizing background concentrations in soil and during comparisons of site versus background data.

5.2 Uses of Geochemical Evaluations

Geochemical evaluations are used for the following purposes:

- to identify the processes controlling element concentrations
- to identify contaminated samples
- to determine whether statistical outliers should be retained in the background datasets
- to confirm or rule out the grouping of candidate background datasets

These purposes are described in the following sections.

5.2.1 Identify the processes controlling element concentrations

Many geochemical mechanisms can control element concentrations in soil, and they are highly localized phenomena due to effects such as pH and oxidation-reduction (redox) in soil pore fluid. It is important to understand these processes so that elevated concentrations, which may be perceived as unrepresentative of background, can be explained. Section 5.5 provides further discussion of some of these geochemical mechanisms.

5.2.2 Identify contaminated samples

Geochemical evaluation identifies samples with element concentrations that are anomalous relative to uncontaminated samples. These anomalous concentrations are not always obviously elevated in terms of their absolute concentrations. Moderately contaminated and even slightly contaminated samples can be identified using geochemical evaluation. This includes concentrations that would pass statistical outlier tests and/or lie below regulatory screening levels. Exclusion of all such contaminated samples results in a dataset that more appropriately represents background geochemical conditions. This becomes especially important later in the project life cycle, when site-to-background comparisons are performed that incorporate geochemical evaluation. Retaining a contaminated sample with anomalously high elemental ratio(s) in the background dataset means that contaminated site samples are more likely to be erroneously declared as uncontaminated.

5.2.3 Evaluate statistical outliers

Background data screening processes may include statistical outlier tests (Section 11.5) to identify unusually high or low concentrations that do not fit a mathematical model, typically the normal distribution. Such "unexpected" concentrations may look suspicious but should not be excluded without further inspection to verify that they are not representative of background. Removal of naturally elevated concentrations would bias descriptive and inferential statistics toward lower, unrepresentative values, and it would censor important geochemical information about site soils. As noted by USEPA (2006)[192]: "These [outlier] tests should only be used to identify data points that require further investigation. The tests alone cannot determine whether a statistical outlier should be discarded or corrected within a dataset. This decision should be based on judgmental or scientific grounds."

Geochemical evaluation is a critical tool for determining whether statistical outliers represent contamination, in which case they should be removed from a background dataset, or have a natural source or are due to inherent variability, in which case they should be retained

(Thorbjornsen 2008)[269]. Section 14.3, Section 14.4, and Section 14.5 provide case studies of geochemical evaluations performed following statistical outlier tests.

5.2.4 Confirm or rule out the grouping of candidate background datasets

A common assumption is that background soil samples must be separated on the basis of soil type (such as the U.S. Department of Agriculture's soil survey map units), location, or depth. Assumptions such as these should be tested using geochemical evaluation of selected elemental ratios (Thorbjornsen 2008)[269]. Similar elemental ratios provide a line of evidence for pooling soil background datasets that might otherwise be segregated based on a priori assumptions, at least during the geochemical evaluation phase of site-to-background comparisons if not also during the background study itself. Trace element concentrations in soil at a given site are controlled primarily by adsorption on fine-grained minerals such as clays and iron oxides, so grain-size effects (Section 5.5) have the largest influence on their concentrations. This means the various a priori classification schemes are often irrelevant for background purposes and separating samples confers no advantage.

An example of using geochemical evaluation to test the validity of grouping samples follows. At a Georgia facility the candidate background dataset included surface soil samples collected from 0 to 2 feet below ground surface and subsurface soil samples collected at 1-foot and 2-foot intervals with starting depths ranging from 2 feet below ground surface (the shallowest subsurface sample) to 42 feet below ground surface (the deepest subsurface sample). Of interest was whether the surface and subsurface samples could be combined (pooled), because during subsequent site investigations the project risk assessors would be evaluating hypothetical receptors' exposure to “total” (surface and subsurface) soil. Geochemical evaluation was performed that included comparisons of elemental ratios. The two sets of background samples exhibited consistent elemental ratios for the evaluated elements, including cobalt (Figure 5-1 and Figure 5-2). Natural cobalt concentrations in soil at this facility are controlled primarily by adsorption on manganese oxide minerals, which is reflected as consistent Co/Mn ratios regardless of absolute cobalt concentration. The surface soil Co/Mn ratios span the same range as the subsurface soil ratios (Figure 5-2), even though the depth intervals vary greatly. It was concluded in this situation that a combined (“total”) background soil dataset was appropriate for the facility, in addition to separate background surface soil and background subsurface soil datasets. Descriptive statistics and inferential statistics were subsequently provided for the three background datasets.

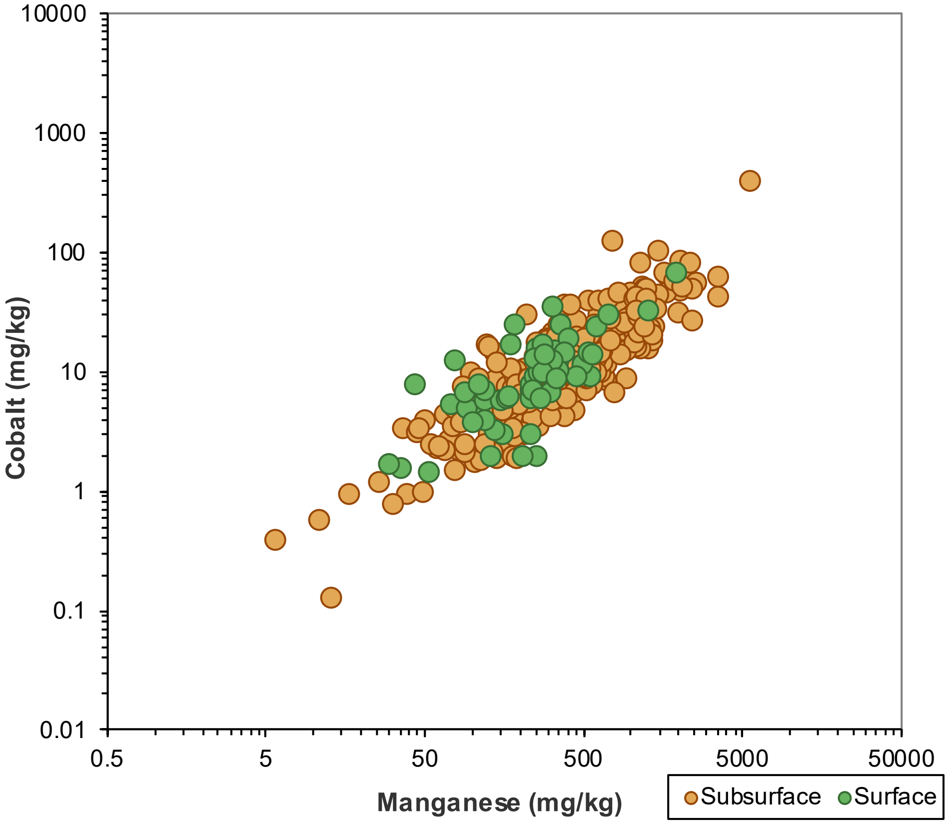


Figure 5-1. Cobalt vs. manganese in background surface and subsurface soil samples.

Source: Karen Thorbjornsen, APTIM.

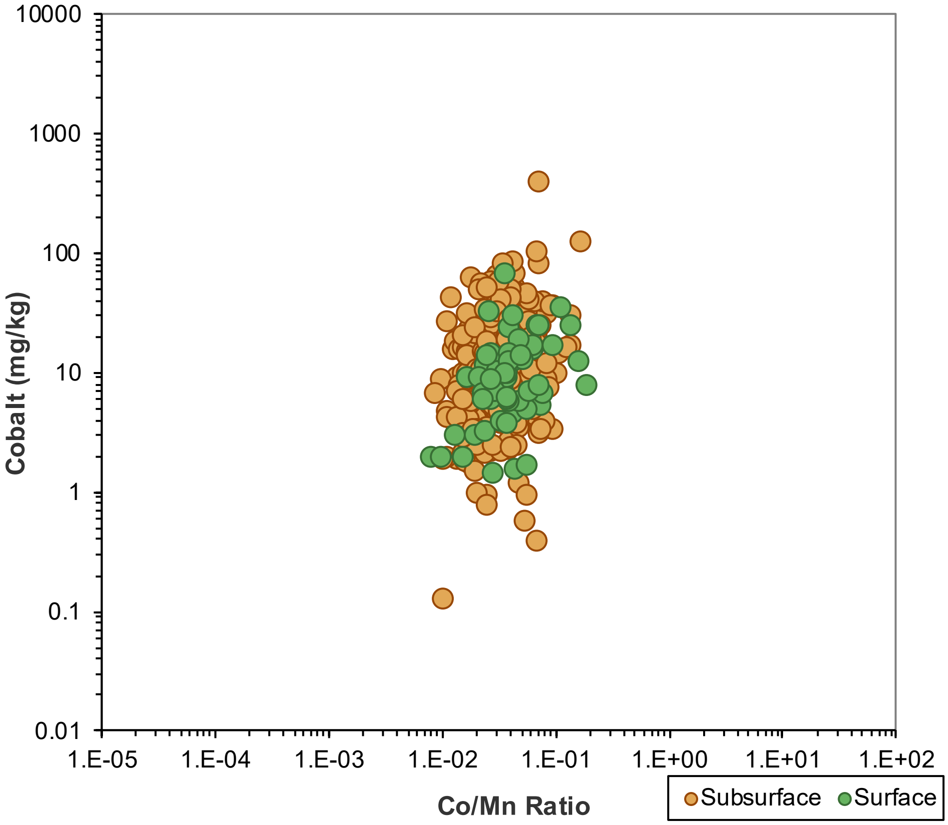


Figure 5-2. Cobalt vs. Co/Mn ratios in background surface and subsurface soil samples.

Source: Karen Thorbjornsen, APTIM.

Geochemical evaluation can complement statistical tests that compare the absolute concentrations of groups of data (for example, Wilcoxon rank sum test [also referred to as the "Wilcoxon-Mann-Whitney test" or "Mann Whitney U test"] or Kruskal-Wallis test). When the statistical tests indicate no statistically significant difference among the data groups *and* geochemical evaluation indicates similar elemental ratios among the data groups, then that provides multiple lines of evidence that the data groups can be combined into one background dataset. This is preferable to relying only on a statistical approach.

An advantage of combining datasets (where demonstrated to be appropriate) is that the summary background statistics will be based on a larger number of samples, thus providing greater confidence in the estimates of the population properties.

5.3 General Methodology

Geochemical evaluation is based on natural associations between elements and minerals in soil and employs all available field observations, analytical data, and information about the facility or area. Major element analyses (including aluminum, calcium, iron, magnesium, and manganese) are required to properly evaluate the trace elements of interest (such as arsenic, chromium, and lead). In soil, "major" elements are defined as having concentrations greater than 100 mg/kg and "trace" elements have concentrations less than 100 mg/kg (Sposito 2016)[263]. Where project budgets allow, the USEPA Target Analyte List of 23 elements should be analyzed because it includes most of these elements, thereby permitting a more thorough evaluation. If the analyte list is too restrictive (for example, only the trace element of interest and one major element, such as arsenic and iron), then important information may be lost that would otherwise explain the concentrations of the element of interest.

In addition to the examination of trace versus major element ratios (described below), information that should be examined includes field geologists' soil descriptions (found on drilling logs, sample collection forms, or field logs), site photographs, soil pH measurements, site history and physical setting, and interviews with personnel familiar with the site or facility operations. All lines of evidence should be brought to bear when drawing conclusions about the presence or absence of inorganic contamination in soil samples.

In support of the geochemical evaluation, scatter plots and ratio plots are constructed to explore the elemental associations and identify potentially contaminated samples ((Myers and Thorbjornsen 2004)[469], (Thorbjornsen and Myers 2007)[267]). A useful starting point is to examine the relationship between aluminum and iron concentrations in the soil samples. Covariance of these major elements does not reflect a geochemical association (such as adsorption effects) but rather a grain-size effect. In oxic soils, samples with high aluminum and iron concentrations have a high proportion of clay minerals and iron oxide minerals, which are very fine-grained. As noted in Section 5.5, these minerals have affinities to adsorb specific trace elements, so samples with more of these minerals will contain higher background concentrations of the associated trace elements, including higher natural background concentrations. Coarser-grained soils will generally contain lower background concentrations of trace elements.

To evaluate trace elements, scatter plots are constructed to depict the detected concentrations of the trace element of interest against the detected concentrations of the major element that represents the mineral to which the trace element may be adsorbed. In the absence of contamination, the samples will tend to exhibit a common trend and similar elemental ratios. Uncontaminated samples with higher trace element concentrations will have proportionally higher major element concentrations and lie on the background trend. Contaminated samples will have anomalously high elemental ratios; on the scatter plot they will lie above the trend formed by the other samples. Such samples contain more trace element than can be explained by the soil mineral content, and they may contain a component of contamination. Ratio plots depicting trace element concentrations versus the corresponding trace versus major element ratios are prepared to accompany the scatter plots. The major elements against which trace elements are evaluated reflect the affinities that the trace elements have for specific minerals, which is a function of the specific trace element and geochemical environment. This is discussed in more detail in Section 5.5, which also includes example scatter and ratio plots.

5.4 Nondetects

The intent of geochemical evaluation is to determine whether detected element concentrations have a natural source or are impacted by anthropogenic or site releases and/or activities, by considering the geochemical mechanisms controlling the concentrations in soil. Nondetects, while typically indicating very low (below the detection limit) concentrations, are a function of a laboratory process; they have no meaning in the geochemical context and provide no geochemical information for the site. Nondetects are therefore excluded from the evaluation (Myers and Thorbjornsen 2004)[469]. Estimated concentrations, such as “J” flagged (or “qualified”) concentrations, are included in geochemical evaluation datasets. Although they have more analytical uncertainty than unqualified concentrations, estimated concentrations still contain useful information. The higher analytical uncertainty associated with their values may result in additional variability in the data presented in scatter plots and greater spread in ratio plots, particularly at the low end of the trace element concentration range, where most J-qualified concentrations tend to reside.

5.5 Key Geochemical Processes

Trace element concentrations in soil are controlled by multiple processes, including solubility (which is a function of pH, redox, temperature, etc.) and adsorption/desorption (“sorption”) reactions. Typically, the most important control is sorption reactions on specific mineral surfaces, which are driven by solute and surface charges. A detailed discussion of geochemical mechanisms is outside the scope of this guidance, but they are more fully explained in other sources, including Stumm and Morgan. (1996)[264], Cornell and Schwertmann (2003)[256], and Sposito (2016)[263].

Trace elements have affinities to adsorb on specific soil-forming minerals (such as clays), represented by major elements (such as aluminum). These adsorption affinities are manifested as consistent trace versus major element ratios in a set of uncontaminated soil samples. Samples with excess trace element from a contaminant source are identified by their anomalously high elemental ratios. Section 5.5.1 through Section 5.5.3 provide a few common element associations that are explored through the use of scatter and ratio plots. Other associations should be explored beyond those listed below, to ensure a thorough evaluation of concentration data for soil at a given site.

5.5.1 Clay minerals

Clay minerals contain aluminum as a primary component. They tend to maintain net negative surface charges under circumneutral pH conditions and attract cationic species such as barium (Ba^{2+}), cadmium (Cd^{2+}), and zinc (Zn^{2+}), among others (Kabata-Pendias 2010)[151]. Concentrations of barium versus aluminum, cadmium versus aluminum, and zinc versus aluminum are typically examined. In the example scatter plot of Figure 5-3, barium concentrations are plotted along the *y*-axis and the corresponding aluminum concentrations are plotted along the *x*-axis.

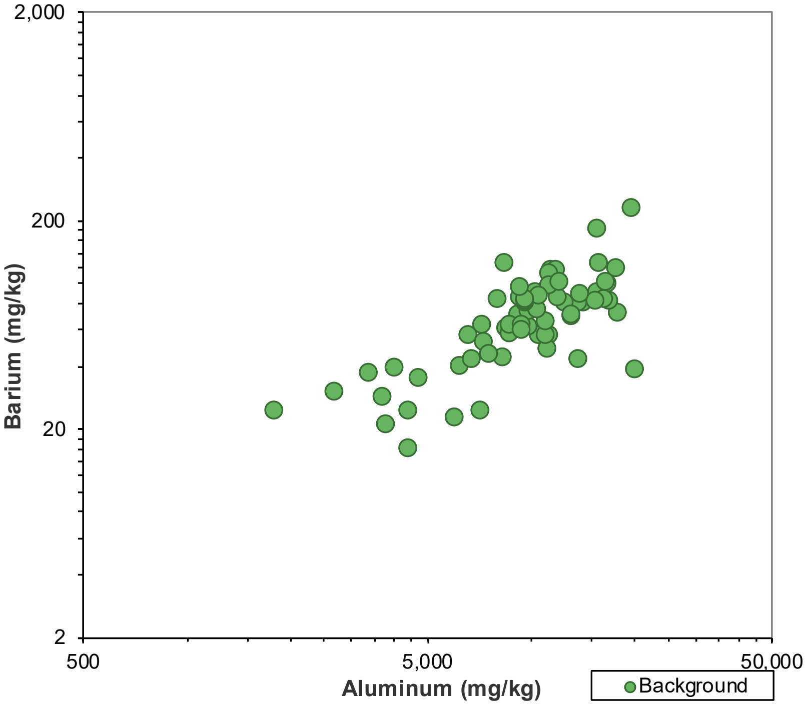


Figure 5-3. Barium vs. aluminum in background soil samples.

Source: Karen Thorbjornsen, APTIM.

The covariance of barium and aluminum concentrations exhibited by these samples (barium concentrations tend to increase as aluminum concentrations increase) suggests that the barium concentrations are controlled by adsorption on clay minerals. Another way of looking at the same data is provided by a ratio plot (Figure 5-4), in which the barium concentrations are plotted along the *y*-axis and the corresponding Ba/Al ratio for each sample is plotted along the *x*-axis.

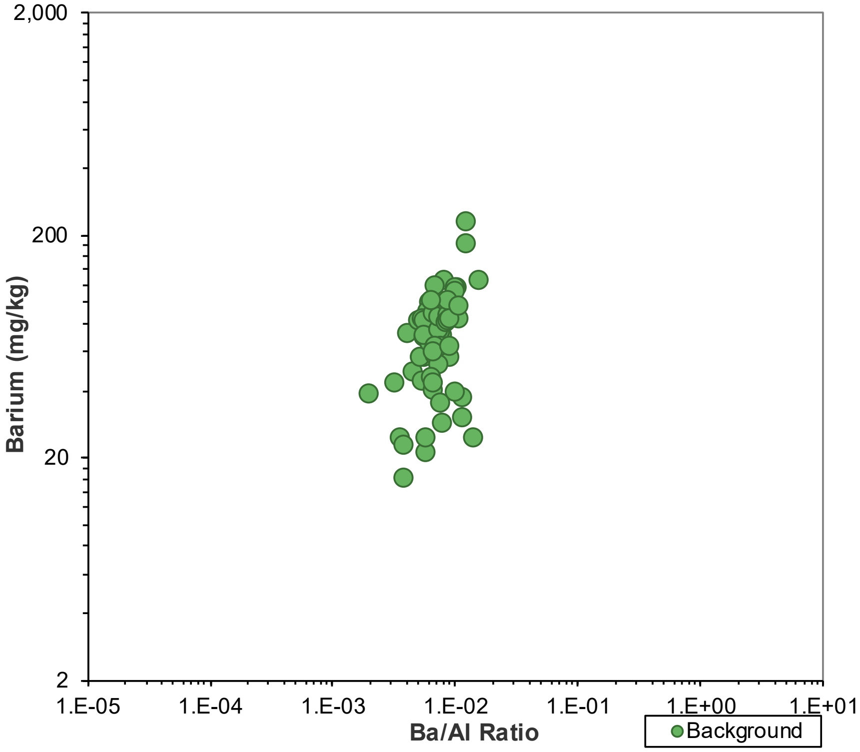


Figure 5-4. Barium vs. Ba/Al ratios in background soil samples.

Source: Karen Thorbjornsen, APTIM.

The background samples exhibit a relatively narrow range of Ba/Al ratios (Figure 5-4), which is expected in the absence of barium contamination. Anomalously high ratios (which are not present in this example dataset) are sometimes more readily seen on ratio plots than on scatter plots; samples with anomalously high ratios lie to the right of uncontaminated samples in a ratio plot. Because the ratios are calculated, they are reproducible, objective, and can be individually inspected in spreadsheets. Some scatter is expected at the low end of the trace element concentration range, due in part to the analytical uncertainty surrounding estimated concentrations.

5.5.2 Iron oxides

Iron oxides are a large group of minerals that contain iron as a primary component (Cornell and Schwertmann 2003)[256]. They tend to maintain net positive surface charges under circumneutral pH and oxidizing redox conditions and attract oxyanion species such as arsenic ($HAsO_4^{2-}$, $H_2AsO_4^{-}$) and vanadium ($H_2VO_4^{-}$, HVO_4^{2-}), among others. Concentrations of arsenic versus iron and vanadium versus iron would be examined, for example.

A scatter plot of vanadium concentrations versus iron concentrations (Figure 5-5) in a set of site-specific background soil samples reveals a common trend with a positive slope. The samples' V/Fe ratios span a narrow range (0.0013 to 0.0043; Figure 5-6), which is consistent with a natural source for the vanadium concentrations. A sample with excess vanadium from a contaminant source would lie above the trend in Figure 5-5 and would be shifted to the right of the other samples in Figure 5-6, but that is not observed for this dataset.

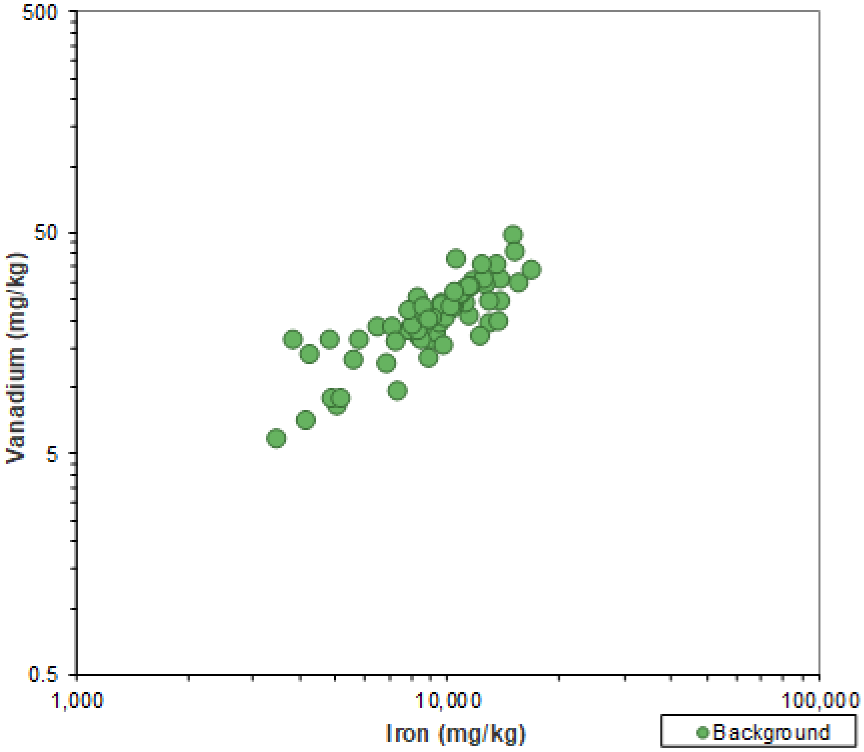


Figure 5-5. Vanadium vs. iron in background soil samples.

Source: Karen Thorbjornsen, APTIM.

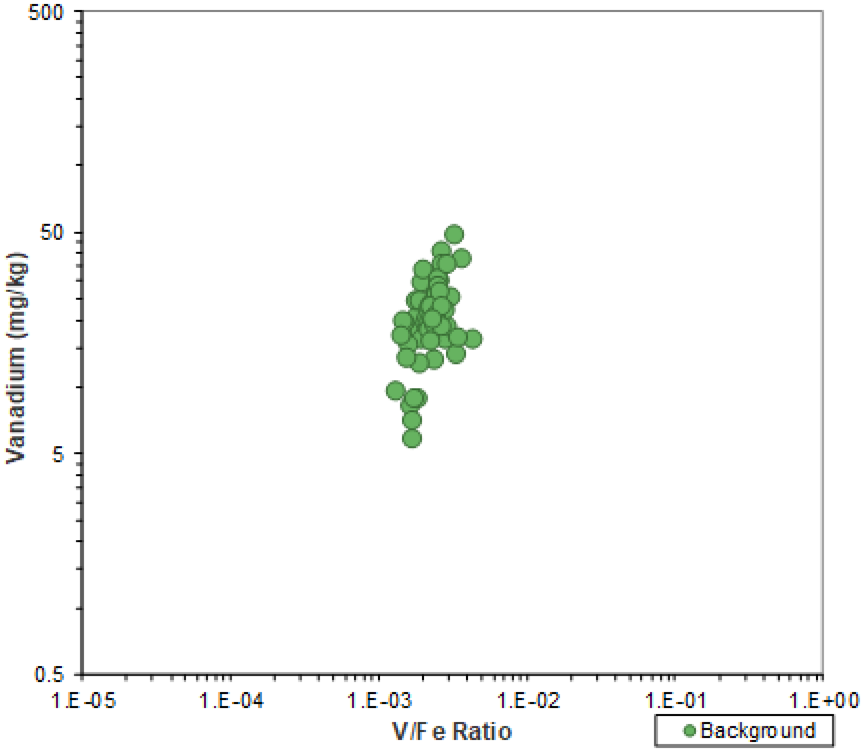


Figure 5-6. Vanadium vs. V/Fe ratios in background soil samples.

Source: Karen Thorbjornsen, APTIM.

5.5.3 Manganese oxides

Manganese oxides are a large group of minerals that contain manganese as a primary component; where present in oxic soils, they tend to maintain strong negative surface charges and attract positively charged species such as cobalt (Co^{2+}) and lead (Pb^{2+}), among others (Kabata-Pendias 2010)[151]. Concentrations of cobalt versus manganese and lead versus manganese would be examined, for example.

In the example below, a site-specific background soil dataset contains two samples with relatively high cobalt concentrations (Figure 5-7); however, the two samples also contain proportionally higher manganese. The two samples were determined to be preferentially enriched in manganese oxide minerals and therefore contain naturally higher concentrations of associated trace elements, such as cobalt.

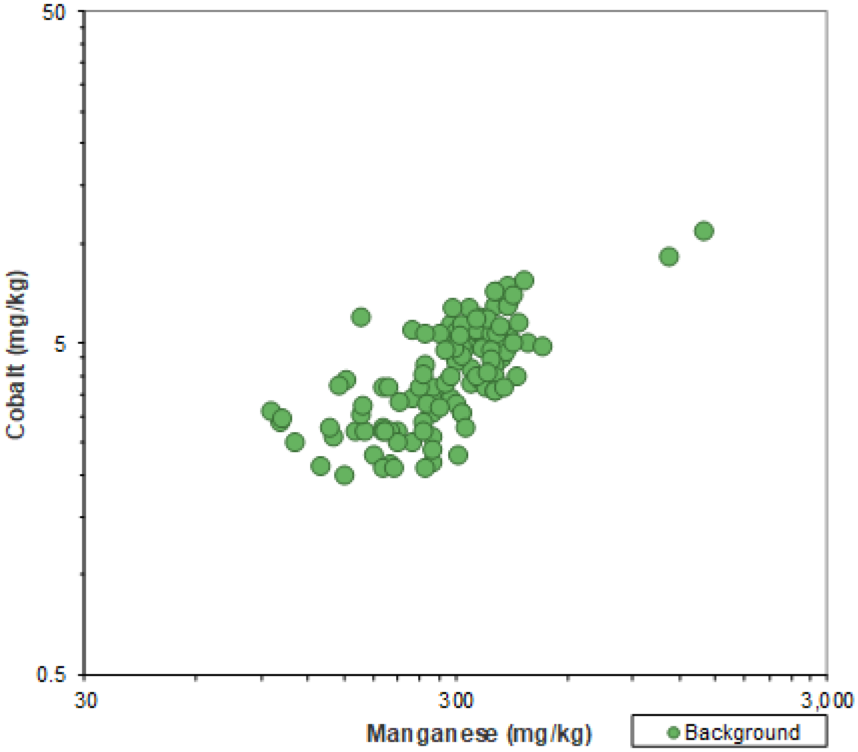


Figure 5-7. Cobalt vs. manganese in background soil samples.

Source: Karen Thorbjornsen, APTIM.

The Co/Mn ratios of the two samples are consistent with those of the other background samples (Figure 5-8), which is indicative of a natural source for the cobalt detections. If the samples contained excess cobalt from a contaminant source, then their Co/Mn ratios would be elevated relative to those of the other background samples, but that is not observed. The two samples reflect the natural variability of the site soils.

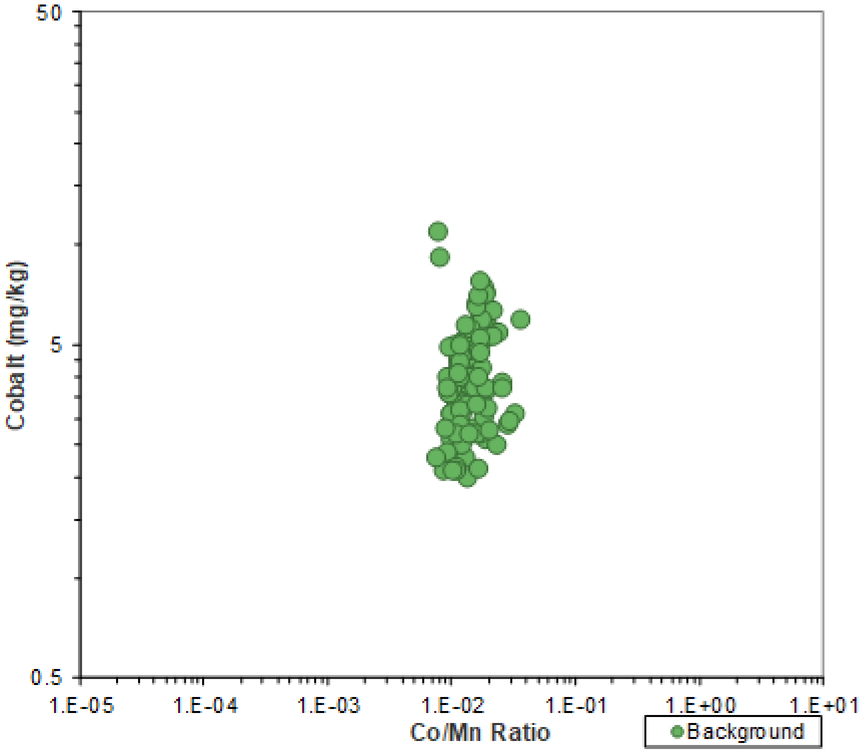


Figure 5-8. Cobalt vs. Co/Mn ratios in background soil samples.

Source: Karen Thorbjornsen, APTIM.

5.5.4 Additional geochemical processes

Additional geochemical processes need to be considered, along with relevant non-geochemical processes. They include:

- other trace versus major element associations not described above, such as adsorption of cadmium and zinc on iron oxides (Kabata-Pendias 2010)[151]
- adsorption of specific trace elements such as mercury and copper on organic particles, which would be evaluated via comparison of trace element concentrations versus total organic carbon concentrations (for example, see Xue et al. (2019)[271])
- the presence of evaporite minerals in arid soils
- the effects of weathering of certain rock types, such as mafic and ultramafic units (for example, see Morrison et al. (2009)[260])
- the presence of mineralized zones that may have localized areas with high concentrations of specific elements
- strong soil profile development that may cause vertical redistribution of elements

- the effects of bioconcentration, such as microbial or “biofilm” coatings on soil particles (for example, see Pal and Paul (2008)[262])
- physical weathering of rock (for soils that are dominated by mechanical fragments rather than chemical weathering products)
- the effects of low redox in hydric soils

Geochemical evaluation is not simply a graphical approach. All available data and information must be examined when evaluating the concentration data, as noted in Section 5.3, and the evaluation should be documented in writing. It is recommended that an experienced geochemist perform the geochemical evaluation because of the complexity that may be encountered at any given site. Conclusions must be explained and supported by applicable citations from the scientific literature, and they must often be defended during regulatory review.

5.6 Extracting Background Data from Existing Data

Geochemical evaluation should be used to examine candidate background samples, in addition to the use of statistical procedures. Extracting background data from existing datasets can minimize the number of new background samples that may need to be collected, thereby saving money. Multiple procedures are typically performed when background data are extracted, with the goal of identifying uncontaminated samples that represent background conditions. Statistical procedures are most commonly performed, but those procedures can only identify elevated concentrations or attempt to distinguish “subpopulations” among a set of concentrations without consideration of their geochemical and scientific context. Some statistical procedures erroneously assume that data must fit a statistical distribution (such as the normal or gamma distribution); violations of the distribution or failures of statistical tests are often assumed to indicate the presence of contamination or other undesirable attributes. This is an unreasonable assumption for naturally heterogeneous geogenic materials such as soil. Statistics, being based in mathematics, cannot explain the reasons for detected concentrations of elements in soil. That is the province of geochemistry and related sciences. Geochemical evaluation can be used to screen out contaminated samples, if present, in the candidate background dataset ((Thorbjornsen 2008)[269], (ASTM 2020)[146]). Note that soil samples will need to be analyzed for major elements (for example, aluminum, calcium, iron, magnesium, manganese) in addition to trace elements of interest to support geochemical evaluation of metals.

6 USING GEOCHEMICAL EVALUATIONS IN RISK ASSESSMENT

The results of geochemical evaluations provide insight into metals concentrations in soil in two main ways:

- using a background dataset to identify which detected concentrations are likely attributable to background (see Section 5, Section 14.4, and Section 14.5)
- evaluating a site dataset to determine which of the detected concentrations, if any, are consistent with background conditions

These results can be used at two key points where background is considered in the risk assessment process:

- COPC selection
- risk characterization

The above depend on when the applicable risk assessment guidance allows consideration of background.

6.1 Using Geochemical Evaluations During COPC Selection

Typically, a geochemical evaluation is performed after an initial screening for COPC has occurred because concentrations below screening levels would not typically warrant a geochemical evaluation. Depending on the applicable guidance, soil chemical concentrations are compared to risk-based screening values alone or to the higher of two screening values: the risk-based screening value or background (default or site-specific). Any chemicals with site concentrations that exceed their associated screening values are then included in the site-specific geochemical evaluation. The geochemical evaluation is typically used to determine whether the site concentrations are consistent with background conditions (“background-related”), and thus that the element need not be evaluated further in the risk assessment even if concentrations exceed risk-based screening values.

As discussed in Section 5.2, moderate or low-level contamination may be present in the site dataset that is not representative of background. Such concentrations may be lower than some of the higher naturally occurring background concentrations. Geochemical evaluation can assist in distinguishing low-level contamination from naturally occurring background, allowing the risk assessor or risk management team to decide whether a chemical should be retained as a COPC for risk evaluation.

Some *advantages* of geochemical evaluations at the screening (COPC selection) stage in the process are:

- Apparent statistical outliers (“false outliers”; Section 11.5) that are determined to be representative of background conditions (based on geochemical evaluation) can be recognized and retained in the background dataset to refine the estimation of a background concentration (for example, BTV) (Section 5.2) and provide a more representative background dataset.

- Multiple populations of background data, if present, can be identified geochemically (Section 5.2), and the knowledge of such populations can be used to guide additional sampling (if appropriate) and support the interpretation of results of the site assessment/risk assessment.

- Consideration of the relevant geochemical processes controlling element concentrations in soil can enhance the project team's understanding of uncertainties inherent in quantifying soil background concentrations and natural mechanisms that affect soil background concentrations, as they apply to the site.

- COPC may be screened out of the risk assessment process at an earlier stage as attributable to background, thus allowing for focus on COPC not attributable to background and a more streamlined risk assessment.

Some *challenges* when considering geochemical evaluations at the screening (COPC selection) stage are:

- Although a geochemical evaluation at this early stage of risk assessment process is a best practice, the importance of background and the complexity of the site may not be sufficiently understood by all stakeholders at that point such that geochemical evaluations are not considered or used effectively.

- Geochemistry expertise is needed, along with additional cost, to perform the geochemical evaluation, and, in some cases, there can be a lack of expertise at regulatory agencies and thus lack of understanding and acceptance of these methods.

6.1.1 Example #1: Evaluate anomalous concentrations further

This simplified, hypothetical example demonstrates that if anomalous samples are identified in a site dataset, then those samples should be scrutinized further to determine whether they represent potential contamination. As shown in Figure 6-1 and Figure 6-2, most samples exhibit covariance between Trace Element A and Major Element B, including the samples with the highest trace element concentrations (up to approximately 1,000 mg/kg). The trace element concentrations are most likely naturally occurring in most samples and are deemed representative of background. Three samples have anomalously high A/B ratios relative to the others (Figure 6-2). If these samples were a site dataset, then it might be concluded that most site samples are consistent with background conditions, with the exception of the three samples that lie above the general trend in the scatter plot and to the right of the other samples in the ratio plot (see arrows). However, additional lines of evidence regarding potential geochemical and physical processes should be evaluated to determine whether there is another explanation for the anomalously high ratios (Section 5.5). For example, are the trace element concentrations controlled primarily by adsorption on another soil-forming mineral that is not represented by the major element depicted here? The additional lines of evidence may support a more defensible conclusion as to whether these samples represent a component of site-related contamination, or their concentrations are consistent with background.

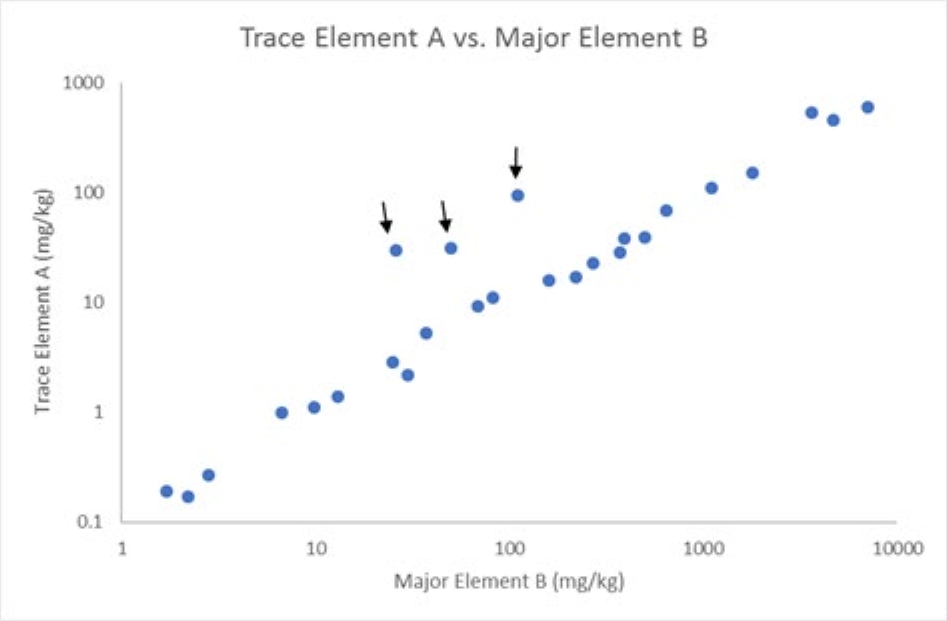


Figure 6-1. Scatter plot of Trace Element A vs. Major Element B shows anomalously high concentrations of Trace Element A in specific samples.

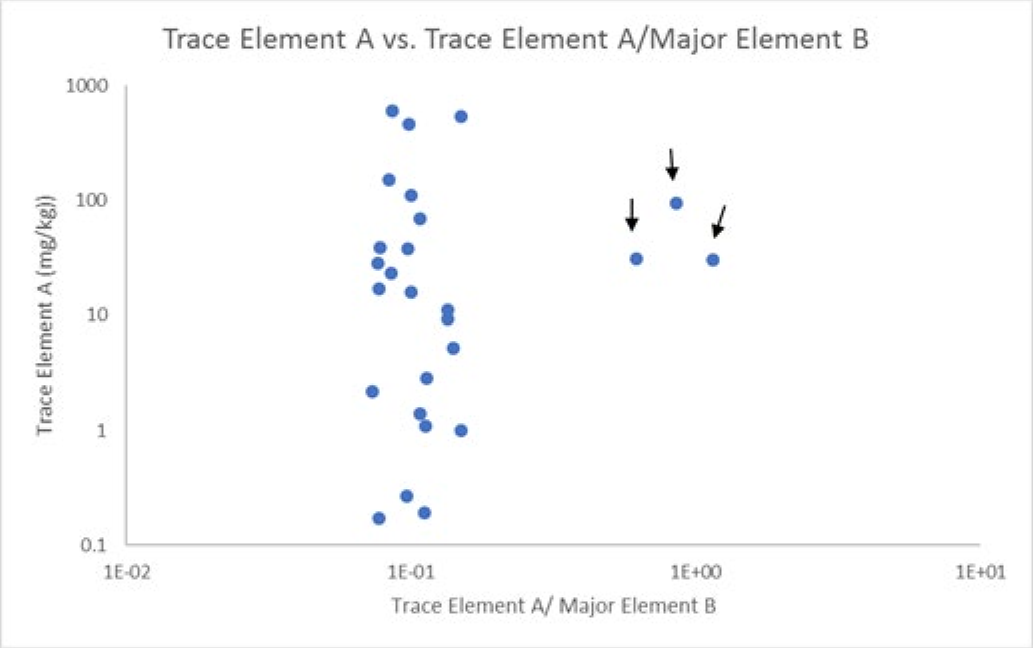


Figure 6-2. Ratio plot of Trace Element A vs. Trace Element A/Major Element B.

6.1.2 Example #2: Samples confirmed to be unrepresentative of background

In this simplified example shown in Figure 6-3 and Figure 6-4, the majority of samples were identified as being consistent with background based on the covariance of trace element A concentrations with aluminum concentrations; however, a subset of the samples (circled here on Figure 6-3 and Figure 6-4 for illustration purposes) lies above the trend formed by most samples in the scatter plot and is shifted to the right of the other samples in the ratio plot. They have anomalously high elemental ratios, indicating potential input of trace element from site-related contamination. In this example scenario, further evaluation of the suspect samples had confirmed that there was no natural source to explain their anomalous ratios (Sections 5.3 and Section 5.5. Therefore, their trace element concentrations could be compared to risk-based screening criteria in the COPC selection step, and the other data (along the main trend) could be excluded from the COPC selection process as being consistent with background. This example also illustrates how low-level site-related chemical concentrations may be distinguished from background concentrations that have a much wider concentration range.

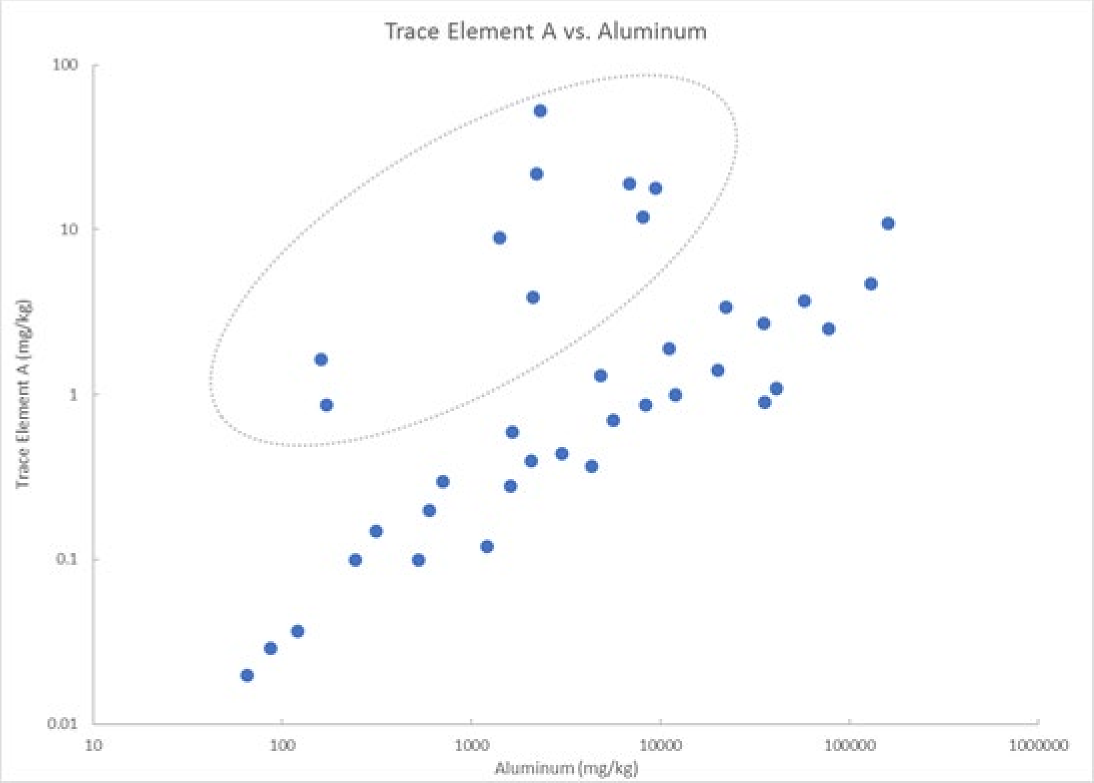

Trace Element A vs. Aluminum
Trace Element A (mg/kg)
100
10
1
0.1
0.01
10
100
1000
10000
100000
1000000
Aluminum (mg/kg)

Figure 6-3. Scatter plot of Trace Element A vs. aluminum.

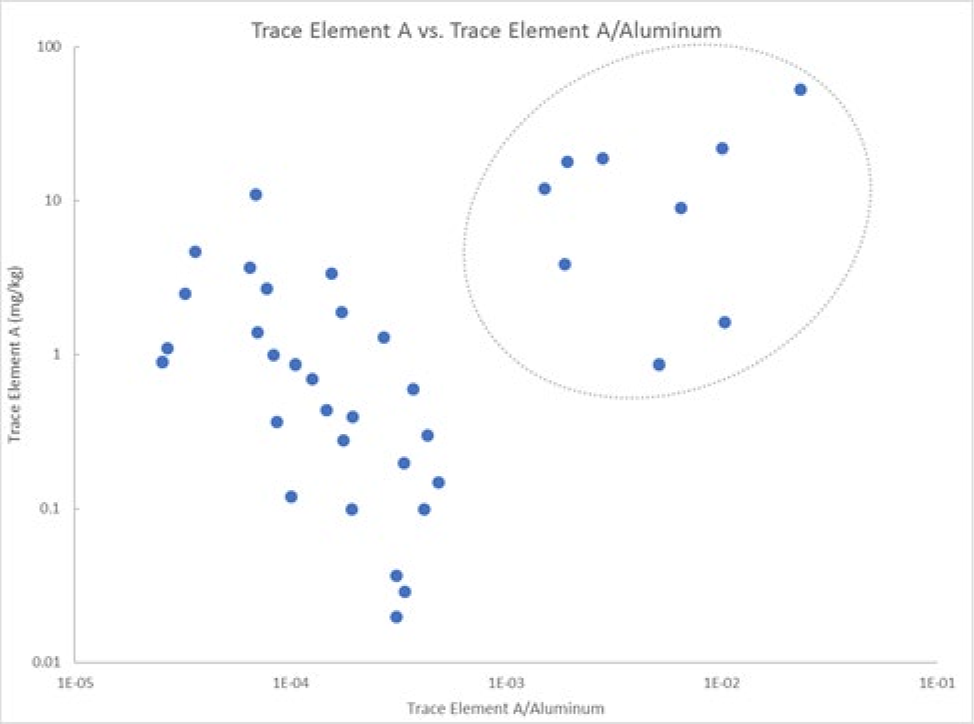


Figure 6-4. Ratio plot of Trace Element A vs. Trace Element A/aluminum.

6.2 Using Geochemical Evaluations During Risk Characterization

If geochemical evaluations are not used during the screening stage to select COPC, then geochemical evaluation results can be used during the risk characterization stage. The overall goal is to determine whether element-related risks can be attributed to site impacts or represent naturally occurring or anthropogenic ambient conditions.

Using the typical human health risk assessment outline terminology of ITRC-Risk-3 (ITRC 2015)[237], this evaluation would likely occur as a subsequent refinement presentation step following initial presentation of risk estimates. In the ecological risk assessment outline terminology ((USEPA 1997)[272], (USEPA 1998)[273]), this synthesis of risk estimation occurs under risk description. In either case, the geochemical evaluation is typically provided as its own stand-alone chapter or appendix in the risk assessment, or in the referenced remedial investigation report or background study report.

Some *advantages* of considering geochemical evaluation results at the risk characterization stage are:

- At this stage of the risk assessment process, the need for, and implications of, geochemical evaluation may be well understood, resulting in a focused evaluation and efficient use of resources.

- These results aid in understanding the results of the risk assessment, which is a primary goal of risk characterization (ITRC 2015)[237], including understanding what risks are with and without the element(s) included.
- Results assist risk managers in scoping appropriate remedial action for a site.

Some potential *challenges* when applying the results of geochemical evaluations in risk characterization are:

- These results add complexity to risk results presentation, which might hinder risk communication with stakeholders and risk managers.
- In some cases, geochemical evaluation conclusions may result in modification of the CSM (for example, through revision of background concentrations or redefinition of impacted areas). It is a best practice to conduct geochemical evaluations early, in order to focus the site assessment/risk assessment process and avoid revising the CSM at a late stage of the risk assessment process, which can result in inefficiency and additional cost.
- Geochemistry expertise is needed, along with additional cost, to perform the geochemical evaluation; in some cases, there can be a lack of expertise at regulatory agencies and thus a lack of understanding and acceptance of these methods.

These potential challenges can be addressed by involving an experienced geochemist in planning, evaluation, and presentation stages of the risk assessment process. The exact approach to incorporating the results of geochemical evaluations into risk characterization will differ depending on how the risk assessment plan is developed, as well as site-specific considerations. Some example scenarios are presented in the following sections. Depending on the applicable guidance and specific circumstances, these approaches could be incorporated within the primary presentation of risk results or presented as subsequent evaluations for comparative purposes. The documentation of the geochemical evaluation could then be included in an appendix to the risk assessment report or in a stand-alone chapter.

6.2.1 Example #3: Point-by-point spatial evaluation for delineation

On a 100-acre site, sampling is conducted on a half-acre grid to evaluate a hypothetical residential exposure risk scenario. The end-product of risk characterization is a spatial grid of cancer risks for arsenic. A geochemical evaluation is performed, and it is determined that iron is the applicable major element with which to evaluate arsenic (Section 5.5.2). A map is prepared depicting sample locations, arsenic and iron concentrations, and the corresponding As/Fe ratios. The map shows that the As/Fe ratios inconsistent with background (anomalously high As/Fe ratios) are spatially clustered, whereas those that are consistent with background are distributed throughout the site. The spatial evaluation complements the geochemical evaluation and helps delineate areas of site-related arsenic impacts.

6.2.2 Example #4: Incorporation into site-specific risk-based cleanup levels or preliminary remediation goals

For some sites, the risk characterization portion of the risk assessment may conclude by establishing site-specific risk-based cleanup levels or preliminary remediation goals. To continue with the hypothetical site in Example 3 (Section 6.2.1), a separate examination of the site data finds that approximately 10% of the site samples show elevated arsenic concentrations greater than a site-specific BTV. The geochemical evaluation indicates that there are some locations where elevated arsenic concentrations appear to result from site impacts, but that there are other locations where arsenic concentrations are consistent with background. Because the geochemical evaluation identified additional locations as being consistent with background, both the data from these locations and the background dataset can be used to establish arsenic's site-specific cleanup levels (or preliminary remediation goals) for the site. The cleanup level based on this combined background dataset would then be compared to a risk-based cleanup level, and the higher of the two values would be used as the final cleanup level.

6.3 **Considerations**

Risk assessment guidance (for example, (ITRC 2015)[237] incorporates the concept that level of effort in developing components of a risk assessment should be "fit for purpose." When contemplating incorporation of the results of a geochemical evaluation in a risk assessment, the following considerations are important (see also Section 5):

- Are the necessary analytical data (Section 5.3) available to perform geochemical evaluation?
- Is the sampling design robust and does it capture site and background geochemical variability?
- How does the geochemical evaluation contribute to the risk assessment, and to what extent does it inform the results and conclusions of the risk assessment? Does the geochemical evaluation have any implications relative to the usefulness of the risk assessment results?
- How do the results of the geochemical evaluation fit with the existing CSM? Are there inconsistencies that could lead to uncertainty in the risk assessment or to a reconsideration of what chemicals are site-related?
- Does assignment of data to "impacted" or "background," based on the results of a geochemical evaluation, change the overall risk assessment conclusions or interpretation?
- Based on professional judgment of the project team, are metals results and risk assessment results consistent with expected impacts (or lack thereof) based on site history or experience at similar sites?

- How does the uncertainty resulting from inclusion/exclusion of data being considered for reassignment as background based on geochemical evaluation compare to other sources of uncertainty in the risk assessment?

Discussion of some of these same considerations/questions in the risk characterization or uncertainty analysis sections of a risk assessment may be useful as lines of evidence supporting the results of a geochemical evaluation.

7 ENVIRONMENTAL FORENSICS RELATED TO SOIL BACKGROUND

7.1 Introduction

Environmental forensics is a well-established discipline that considers scientific, operational, and historic information to interpret the potential sources and/or ages of contamination detected at a site, typically at anomalous concentrations. Although a common objective of an environmental forensic evaluation is to identify causation and timing of contamination (Morrison and Murphy 2006)[344], in the context of this guidance its value is focused on determining whether the observed chemical concentrations in soils may be representative of natural or anthropogenic ambient background or are due to site-related contamination. Using environmental forensics to determine whether a contaminant is representative of background requires an expert who is knowledgeable about these methods and the chemistry of the contaminant being evaluated.

Contamination requiring risk assessment can take many forms. Some forms are compositionally simple and their presence in the environment can be measured by only one or a few individual chemicals or elements (for example, arsenic or metals). Assessments of these simpler forms of contamination relative to background can often be achieved using numerous statistical tests (Section 11) and geochemical evaluations (Section 5). Other forms of contamination are compositionally much more complex, being comprised of scores of different chemicals, some with related chemical structures and similar but not identical chemical properties. Risk assessment of these more complex forms of contamination relative to background can often benefit from environmental forensics. Specifically, the complexity of these contaminants not only provides a basis to distinguish among different sources of the same contaminant but also, and of relevance herein, to distinguish contamination from background, particularly at low contaminant concentrations.

In this section, five different groups of chemicals that occur in the environment as both contamination and background are considered, namely:

- polycyclic aromatic hydrocarbons (PAHs; Section 7.2)
- total petroleum hydrocarbons (TPH; Section 7.3)
- polychlorinated biphenyls (PCBs; Section 7.4)
- polychlorinated dibenzo-p-dioxins and dibenzofurans (PCDD/F; Section 7.5)
- per- and polyfluoroalkyl substances (PFAS; Section 7.6)

The varying compositions of these five groups of chemicals in different contaminant sources can be exploited to distinguish contamination from background. As such, the application of environmental forensics is clearly relevant, albeit perhaps underused, in the human health and ecological risk assessment of soils, and therefore, warrants consideration by risk assessors.

In addition, this section also includes information on remote sensing (Section 7.7), whose application relies not on the chemical complexity of different source contaminants, but rather on the spatial and temporal differences that can be recognized in a study area using various remote (noninvasive) techniques.

7.2 Polycyclic Aromatic Hydrocarbons

PAHs are ubiquitous in soils worldwide (Nam, Sweetman, and Jones 2009)[347]. Some PAHs are acutely toxic and have carcinogenic properties (Menzie, Potocki, and Santodonato 1992)[340], and as a result their occurrence in soils is regulated in most jurisdictions. Their ubiquity arises because (1) PAHs are formed by many natural and anthropogenic processes, and (2) once formed, PAHs can be spread throughout the environment, often by atmospheric transport. Some PAHs form naturally from the degradation of plant debris in soil (biogenic PAHs) or during the formation of crude oil (petrogenic PAHs) over geologic time. The latter can enter the environment both naturally (crude oil seeps) or due to anthropogenic spillage/leakage of crude oils and refined petroleum. Most PAHs, however, are formed by the incomplete combustion of organic matter (pyrogenic PAHs), such as wood, petroleum, coal, or even garbage. Again, combustion can be either natural (for example, wildfires) or anthropogenic (for example, fossil fuel burning or residential wood burning), and the resulting emissions and residues can be spread through the environment via runoff, atmosphere transport, or as redistributed contaminated fill.

The variety and widespread occurrence of PAHs' sources and their persistence during transport result in a certain level of PAHs in all soil that is attributable to background. In some soils, these PAHs can be derived predominantly to exclusively from natural processes, which impart a natural background. In other soils, PAHs can be derived from the diffuse occurrence of PAHs from anthropogenic ambient sources, such as atmospheric deposition, or from other sources that are not a result of a site release, such as runoff or fill.

Distinguishing PAHs in soils attributable to natural and/or anthropogenic ambient soil background from those PAHs due to contamination (for example, spilled oil or coal tar) is important with respect to soil risk-based values used in risk assessment. When PAH concentrations are high, background threshold values used in lieu of the soil risk-based value can sometimes be established through statistical assessments of PAH concentrations in large datasets (Section 11). This distinction, however, is much more difficult in soils containing relatively low concentrations of contaminants, in which case-specific, sometimes subtle, chemical differences among different PAH sources can be useful. Environmental forensic assessments can aid in assessing the character of statistical outliers to determine whether they represent background or contamination.

In the following section, the basic methods used in the forensic analysis of PAHs in soils are presented. For additional information the reader is directed to the multitude of studies conducted over the past 50 years that assessed PAH sources (Appendix C).

7.2.1 Forensic analysis of PAHs in soils

Assessments of PAH in soils are conducted with the intention of determining impacts of local anthropogenic sources (contamination), both historic and current. This can sometimes be achieved through the use of the descriptive and inferential statistical techniques described in

Section 11, but forensic analysis, which can include multivariate statistical analyses, may be warranted to support or refute any statistical determinations regarding background. Forensic analysis requires not only an understanding of the PAHs present in specific forms of contamination, but also the ability to distinguish PAHs attributable to contamination from those attributable to representative background based upon chemical patterns.

There is no single method for achieving this objective as no two datasets or study areas are the same. Instead, distinguishing PAHs in soils due to contamination from those due to representative background is best achieved through careful analysis, which often relies upon some combination of the following four forms of forensic analyses:

- Pattern recognition includes visual, qualitative comparisons of the available PAH concentration histograms (and corresponding TPH chromatograms), including comparison to known standards in a sample library and/or site-specific samples of known origin(s)
- Diagnostic ratios based upon parent and/or alkylated PAH concentration data, again sometimes including comparisons to known standards and/or site-specific samples of known origin(s)
- Spatial and/or temporal analysis of PAH patterns, diagnostic ratios, and/or concentrations, including a comparison to historical information compiled for the area (as part of the CSM)
- Quantitative source apportionment analysis that serves to allocate contributions between multiple PAH sources present, such as principal component analysis (PCA), chemical mass balance (CMB), and positive matrix factorization (PMF)

There is extensive literature on all of these methods, so each is only briefly discussed in the following paragraphs.

7.2.1.1 *Pattern recognition*

The human eye remains a powerful tool and, although experience in viewing PAH histograms can reveal subtle differences, even an inexperienced interpreter can visually compare the PAH histograms of different samples and assess any likely relationships (Stout et al. 2002)[371]. Because PAHs are usually derived from sources that also contain other (non-PAH) chemicals, the value of TPH chromatograms also cannot be underestimated. Good gas chromatography with flame ionization detection (GC/FID) (for example, USEPA Method 8015) of all the extractable compounds or hydrocarbons in a soil can often aid in understanding the source(s) of the PAHs alone (Section 7.3).

Qualitative fingerprinting is perhaps most useful in instances wherein marked differences exist between contamination and background PAH fingerprints. For example, combustion-derived particles in soils that are attributable to anthropogenic ambient soil background are dominated by 4- to 6-ring PAHs, which can be readily distinguished graphically from an impact of petroleum, such as used diesel fuel.

The long-recognized differences in the skewed versus bell-shaped PAH homolog profiles (Blumer 1976)[279] can be easily recognized through qualitative inspection (if alkylated PAH data are available (see (Douglas et al. 2015)[114] for details on alkylated PAH analysis). For example, Figure 7-1 shows the PAH distributions and concentrations for PAH sources often found in soils proximal to roadways, namely, abraded asphalt and soot (urban dust). The former exhibits bell-shaped PAH homologue patterns dominated by alkylated PAHs, whereas the latter exhibits skewed patterns dominated by the parent (non-alkylated) PAH. These serve to demonstrate the differences between petrogenic and pyrogenic PAH patterns, which are made more obvious with the availability of the alkylated PAH data. Of course, the influences of weathering on the PAH distributions must always be considered, which requires greater understanding of the effects of evaporation, water-washing, and biodegradation on PAH fingerprints (for example, (Elmendorf et al. 1994)[303]).

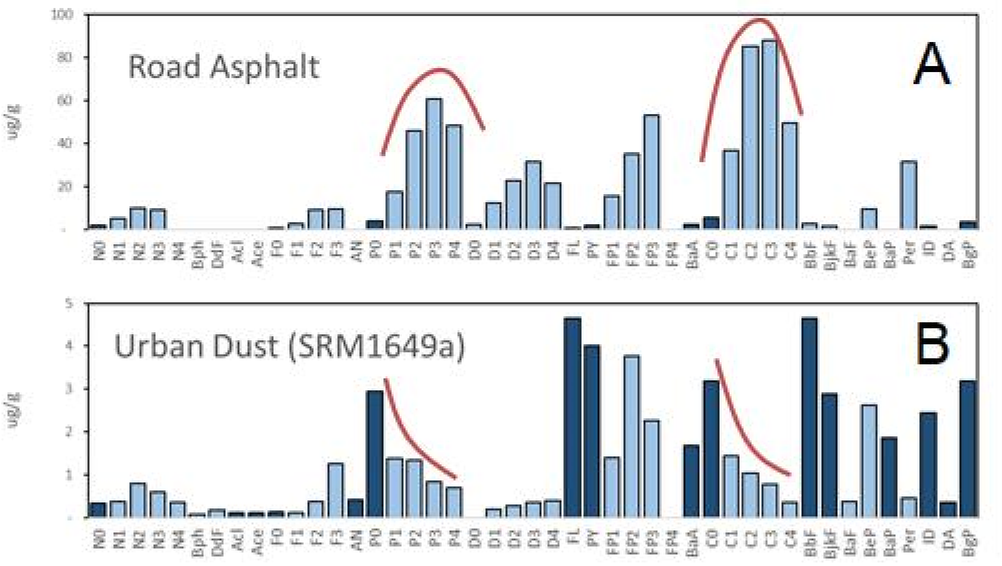


Figure 7-1. PAH histograms for materials that can contribute to anthropogenic ambient soil background. (A) road asphalt and (B) urban dust (NIST SRM1649a). Dark blue bars represent USEPA priority pollutant PAHs; light blue bars represent alkylated PAHs often used in forensic assessments; red lines depict (A) bell-shaped and (B) skewed homologue profiles (see (Stout et al. 2015)[369] for additional details).

Source: (Stout, Uhler, and Emsbo-Mattingly 2004)[370].

7.2.1.2 *Diagnostic ratios*

Qualitative differences between PAH patterns can become less clear when a greater number of potential PAH sources are involved. In such situations, or in studies involving large numbers of samples, comparisons among the PAH fingerprints can be achieved through the use of diagnostic ratios. (See Appendix C for additional sources of information). Most diagnostic ratios involve PAH isomers (same molecular mass, but different molecular structure) and thereby similar physico-chemical properties, which better ensure that the ratios remain stable during weathering (as the isomers undergo similar environmental fates). As such, differences in diagnostic ratios among samples are commonly attributed to different PAH sources. Table 7-1 provides a partial inventory of some common PAH diagnostic ratios reported in two widely referenced studies.

Table 7-1. Diagnostic ratios used to assess PAH sources

Source: Scott A. Stout, NewFields Environmental Forensics Practice, LLC.

Ratio	Value/Range	Reported Source
AN/(P0+AN)	<0.1	Petrogenic
	>0.1	Pyrogenic
FL/(PY+FL)	<0.4	Petrogenic
	0.4–0.5	Petroleum combustion
	>0.5	Coal & biomass combustion
BaA/(C0+BaA)	<0.2	Petrogenic
	0.2–0.35	Coal combustion
	>0.35	Petroleum combustion
IND/(GHI+IND)	<0.2	Petrogenic
	0.2–0.5	Petroleum combustion
	>0.5	Coal & biomass combustion
See text for cautions regarding use of diagnostic ratios. PAH abbreviations: AN-anthracene; P0-phenanthrene; FL-fluoranthrene; PY-pyrene; BaA-benz[*a*]anthracene; C0-chrysene; IND-indeno[*1,2,3-cd*]pyrene; GHI-benzo[*g,h,i*]perylene.		

The simplicity of calculating and plotting diagnostic ratios among PAHs unfortunately can lead to the unquestioning application of this method in PAH source identification—or the ratios from one set of samples or site being inappropriately applied elsewhere (Boehm et al. 2018)[99]. There is a generally large degree of scatter in the ratios among samples from large PAH datasets with no clear cutoff in values, although cutoffs are often identified to exist (Table 7-1). The diversity of fuels (gasoline, diesel, biomass) or different combustion conditions can result in large variations in the ratios in combustion emissions reported among pyrogenic and petrogenic sources (Lima, Farrington, and Reddy 2005)[334]. As such, diagnostic ratios from one study may not reflect results obtained when different source materials are involved. An example of this was demonstrated by Lima, Farrington, and Reddy (2005)[334], who showed that a cross-plot of the ratios of fluoranthene/pyrene and phenanthrene/anthracene ratios for a large number of known sources did not accurately reflect those predicted by published ratios (Table 7-1). Furthermore, some diagnostic ratios (for example, AN/(P0+AN) appear to be unstable during some forms of weathering (Uhler and Emsbo-Mattingly 2006)[375].

As such, diagnostic ratios' validity in source differentiation is often questioned ((Lima, Farrington, and Reddy 2005)[334], (Galarneau 2008)[119], (Katsoyiannis and Breivik 2014)[325], (Tobiszewski and Namieśnik 2012)[471], (Zou, Wang, and Christensen 2015)[402]) and they should not be blindly applied—and used only with caution and only within the context of the area under study ((Lima, Farrington, and Reddy 2005)[334], (Boehm et al. 2018)[99]).

7.2.1.3 *Spatial and temporal analysis*

Spatial analysis of PAH concentrations or diagnostic ratios can further reveal lateral or vertical trends that can help in distinguishing contamination from background. More importantly, these techniques can help identify PAH concentration gradients or hot spots, either of which could suggest the location of a possible point source of PAHs relative to overall diffuse background conditions. The identification of possible point source locations can be bolstered if the locations happen to correspond to existing or former facilities at which PAH-containing materials were handled or produced. Conversely, if spatial displays of PAH data do not reveal hot spots or concentration gradients associated with a specific point source under investigation, then the contribution of the suspected source should be re-evaluated relative to representative background conditions in the study area.

7.2.1.4 *Quantitative source apportionment*

Finally, forensic evaluations of PAH datasets using PCA, CMB, or PMF provide bases to evaluate large datasets in an unbiased manner, in ways that are not confounded by the number of candidate sources. PCA is likely the most widely used multivariate analysis to identify the different profiles of the principal components determined, that can then be attributed to specific sources. As an example, the factor score plot produced from a PCA analysis of the 16 priority pollutant PAHs (EPAPAH16) concentrations in 350 sediment samples from an urban river is shown in Figure 7-2 (Stout and Graan 2010)[364]. The results revealed variably weathered creosote in some samples, whereas most contained mixtures of weathered creosote and PAHs derived from background.

PCA's factor scores for each sample (Figure 7-2) subsequently can be used to calculate the percent contributions of the major sources that are identified using multiple linear regression ((Larsen and Baker 2003)[330], (Zuo et al. 2007)[403], (Wang et al. 2010)[384]). CMB fits the soil data with pre-defined (known or presumed) PAH source fingerprints (Li, Jang, and Scheff 2003)[333]. In contrast, PMF generates candidate PAH source fingerprints, which can then be identified through comparison to known source fingerprints and used to allocate the contributions from each ((Wang et al. 2009)[385], (Stout and Graan 2010)[364]).

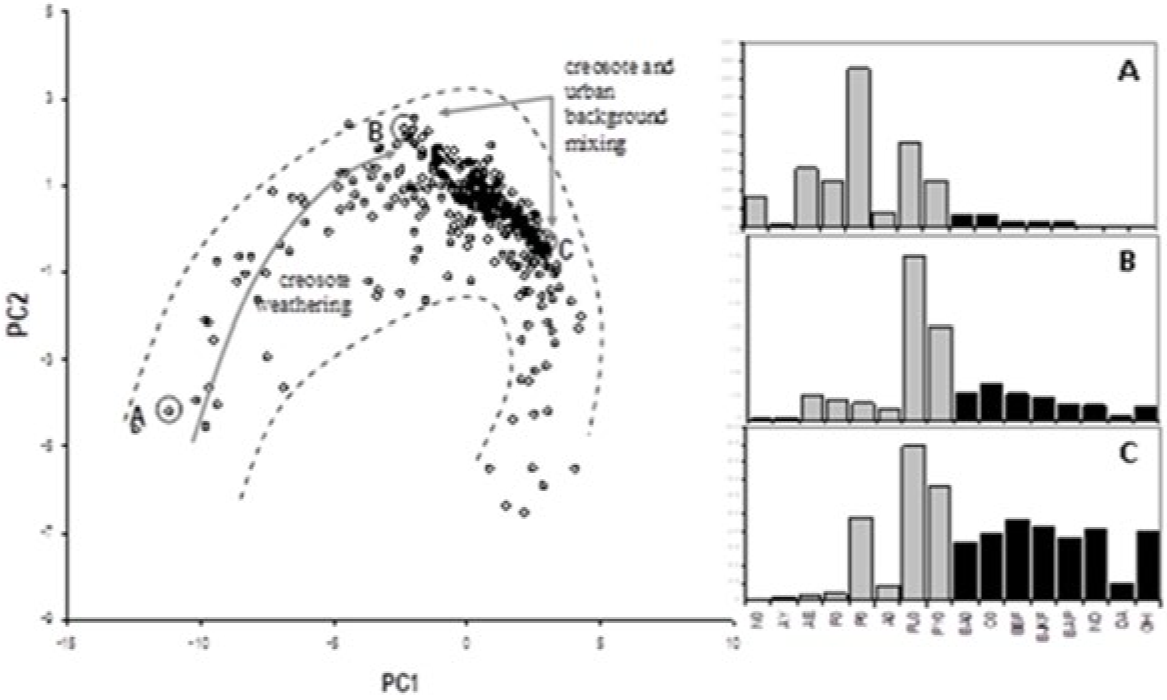


Figure 7-2. Principal component factor score plot. PC1 vs. PC2 (left) showing overall trends among PAHs in 350 sediment samples. Most samples appear as mixtures of weathered creosote and urban (anthropogenic ambient) soil background. PAH profiles for representative samples (A-C) are shown (right).

Source: Reprinted with permission from Stout and Graan (2010)[364]. Copyright 2010, American Chemical Society.

7.3 Total Petroleum Hydrocarbons (TPH)

ITRC has developed guidance on TPH and its risk at petroleum-contaminated sites (ITRC 2018)[316]. This earlier document provides a thorough assessment of TPH as it relates to petroleum. However, not all hydrocarbons in soils are attributable to petroleum (crude oil, fuels, lubricants, or oil-related wastes). Hydrocarbons in soil can also be derived from natural and anthropogenic ambient soil background sources, as well as from forms of contamination *other* than petroleum, for example, coal tar, creosote, or pitch. Thus, "TPH" is a misnomer since its detection in soils does not necessarily indicate the presence of petroleum. This section specifically focuses on natural and/or anthropogenic ambient background hydrocarbon in soils due to biogenic materials and particulate coal.

7.3.1.1 *Applications and limitations of different types of TPH data*

The TPH concentration in soil is a method-defined analytical measurement. The applications and limitations of different TPH analyses are described in ITRC (2018)[316] (for example, Table 5-4 and Section 5.9 of ITRC (2018)[316]). TPH concentrations in soils are most frequently determined by solvent extraction followed by gas chromatography with flame ionization detection (GC/FID) per USEPA Method 8015, or comparable chromatographic method (for example, in Canada, (CCME 2007)[444]). This method allows for both the quantification of how much TPH is present in soil and, because of the GC/FID chromatogram produced, an assessment of the boiling range and character of that TPH. Because USEPA Method 8015 measurements are based on the response from the FID, all chromatographable chemicals within a soil sample's extract will be measured. These will include resolved compounds, which appear as peaks on the chromatogram, and unresolved compounds, which appear as "humps" on the chromatogram's baseline. The latter is often referred to as the unresolved complex mixture.

Before measurement of TPH concentrations are made by USEPA Method 8015, the extractable materials in a soil sample are removed via extraction with methylene chloride. In addition to *hydrocarbons*, chemicals containing only hydrogen and carbon atoms, this solvent (and most others) will also remove *non-hydrocarbons*, including polar compounds (containing chemicals with atoms of sulfur, nitrogen, or oxygen) present within the soil.

Naturally occurring organic matter in soils can contain both hydrocarbons and non-hydrocarbons (polars). Their presence in any soil extract will be measured as TPH, and unwittingly considered as contamination rather than background if left unrecognized ((Stout and Uhler 2003)[365], (Wang et al. 2012)[386]). The application of cleanup steps (for example, silica gel cleanup) that separate polar and nonpolar components prior to analysis of TPH is described in Section 5.10 of ITRC (2018)[316].

7.3.1.2 *Characterization of background TPH*

Biogenic Organic Compounds

The most common naturally occurring organic background material encountered in soils is due to the presence of microbial and vascular (land) plant debris, collectively referred to as *biogenic organic compounds* (BOC) or *naturally occurring materials* (NOM). BOC is used herein. BOC consists of mixtures of biochemicals synthesized by living organisms that undergo modest diagenetic changes during their preservation and accumulation in soils and sediments ((Stout and Uhler 2003)[365], (Wang et al. 2009)[385]). Examples of BOCs include n-alkanes, terpenoids, steroids, fatty acids, alcohols, and esters.

The BOC component within soils containing microbial and plant debris can be significant, particularly in moist, highly vegetated environments where peat or other organic-rich soil accumulates (or had in the past). The following studies address contributions of BOC to TPH measurements: (Dworian (1996)[115], Wang et al. (2012)[386], Stout and Wang (2008)[366]).

Particulate Coal

Although rarer than BOC, some soils may contain naturally occurring TPH associated with the organic compounds derived from coal (or organic-rich shales) that has been eroded from sedimentary rock outcrops. Thus, natural background TPH in soils in certain geologic settings where these precursor rock outcroppings exist must be considered ((Stout and Emsbo-Mattingly 2008)[363], (Achten and Hofmann 2009)[93]). Separately, in regions where coal is mined, used, stored, or transported, decades of operations can distribute particulate coal and coal dust in nearby soils (Hindersmann and Achten 2018)[125]. In such areas, coal would be considered as an anthropogenic ambient background material, but perhaps difficult to distinguish from natural background TPH due to coal.

The examples described above underscore the importance of having an experienced analytical chemist review the chromatograms from TPH analyses prior to use of TPH data in risk assessment (ITRC 2018)[316].

7.4 Polychlorinated Biphenyls (PCBs)

Please refer to the USEPA's Polychlorinated Biphenyls (PCBs) website as a resource for technical information pertaining to PCBs.

PCBs are a mixture of chlorinated analogs of the biphenyl molecule. During chemical synthesis, mixtures of PCB molecules are produced that can contain between one and ten chlorine atoms (Figure 7-3), resulting in 209 possible PCB congeners. Each congener differs in terms of the number and position of the chlorines on the biphenyl molecule. PCBs are synthetic organic compounds; there are no natural sources of PCB in the environment.

In the United States, PCBs were manufactured almost exclusively by Monsanto Corporation and marketed under the trade name Aroclor. The production of Aroclors ran from 1929 until 1977, when Monsanto voluntarily terminated production because of environmental concerns about PCBs. The USEPA banned the manufacture of PCBs and began phasing out most uses of PCBs in 1979 (USEPA 1979)[404].

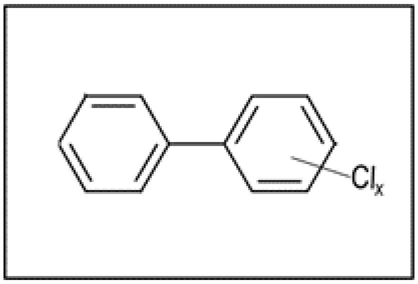


Figure 7-3. General structure for polychlorinated biphenyls. The molecule can contain one to ten chlorine atoms (x), resulting in 209 possible polychlorinated congeners.

Source: A. Uhler, NewFields Companies LLC (unpublished).

The most common Aroclors manufactured by Monsanto were:

- Aroclor 1016
- Aroclor 1242
- Aroclor 1260
- Aroclor 1221
- Aroclor 1248
- Aroclor 1262
- Aroclor 1232
- Aroclor 1254
- Aroclor 1268

The nomenclature used to identify specific Aroclors was the format "Aroclor 12XX," where the prefix 12 represented the biphenyl molecule (12 carbons), and the second set of digits (XX) represented the approximate percentage of chlorine that composed the Aroclor. For example, Aroclor 1248 is a chlorinated biphenyl mixture containing approximately 48% chlorine. By contrast, Aroclor 1016 was prepared by the fractional distillation of Aroclor 1242, which excluded the higher boiling (more highly chlorinated) congeners (Mayes et al. 1998)[454].

The congener composition of Aroclors has been well studied, and this compositional data can be very useful in forensic evaluation of PCBs in soils. For example, it can be used to determine the compositional feature of PCBs in both background and site-impacted soil samples. Each of the commercial Aroclors contains a distinct distribution of congeners, which varies as a function of total chlorine content (Figure 7-4). The degree of chlorination for the congeners depicted in this figure increases from left to right. These distinctive patterns of congeners (and potential mixtures thereof) form the basis for forensic identification of PCBs in soil samples.

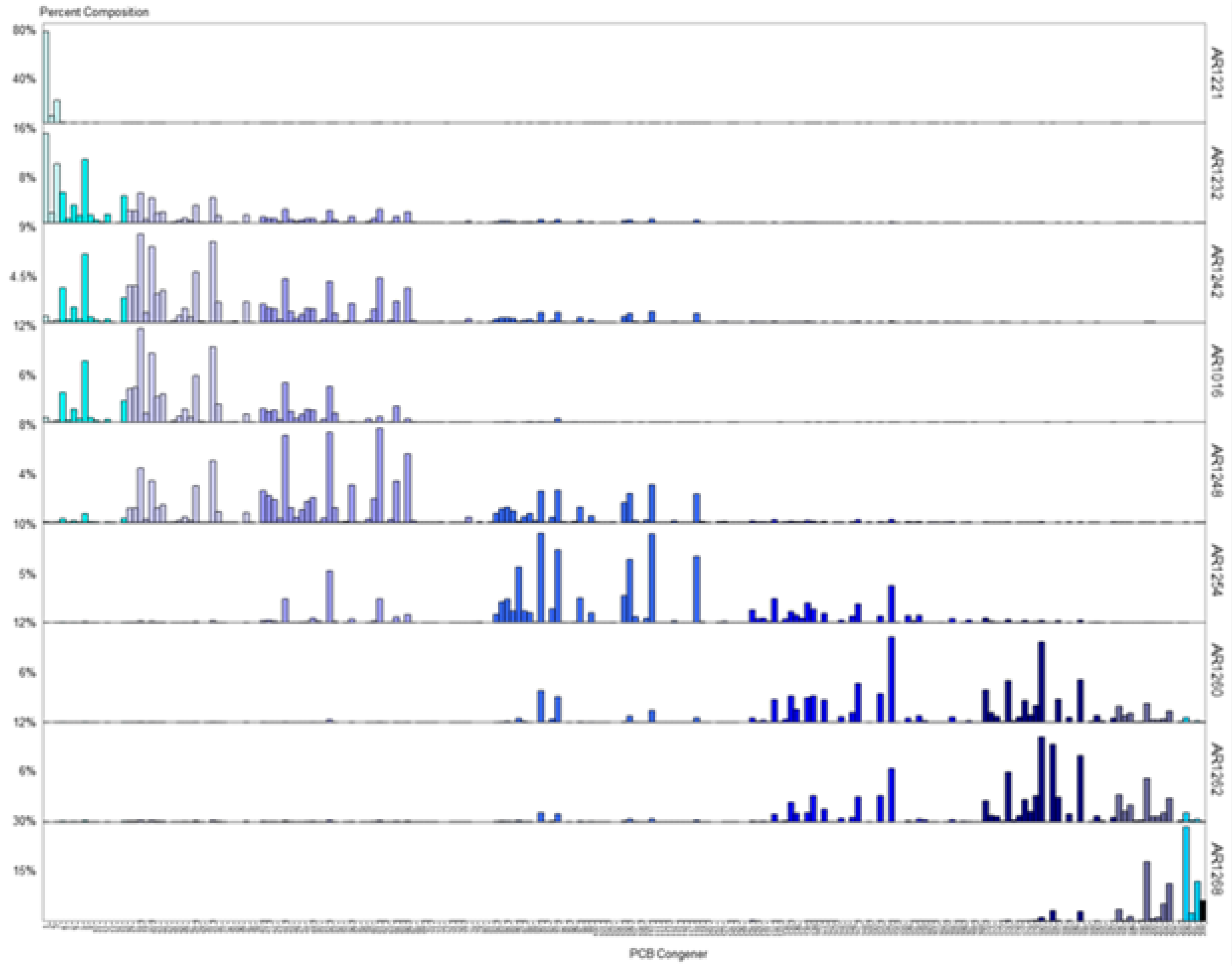


Figure 7-4. The PCB congener distribution for commercial Aroclors. Congener number (and number of chlorines per molecule) and LOC increase from left to right.

Source: A. Uhler, NewFields Companies LLC (unpublished).

Groups of congeners containing the same number of chlorine atoms are referred to as congeners of the same levels of chlorination (LOC). For example, congeners containing one chlorine atom make up LOC1, congeners containing two chlorine atoms make up LOC2, etc., ending with congeners containing 10 chlorine atoms (LOC10). The congeners that make up the various Aroclors are mixtures of differing relative proportions of these LOC (Table 7-2). LOC data are very useful for classifying PCBs measured in soil samples and provide a convenient means to compare the overall PCBs compositional character of background and site-impacted soil samples. This section discussed specific environmental forensics for assessing background in soils. Please refer to USEPA (1979)[404], ATSDR (2000)[275], Douglas et al. (2015)[114], Johnson et al. (2006)[448] for details regarding background information about PCBs that may be helpful for some readers who are not familiar with this information.

Table 7-2. Composition of Aroclors by percent level of chlorination

Source: Scott A. Stout, NewFields Environmental Forensics Practice LLC.

Aroclor	LOC1	LOC2	LOC3	LOC4	LOC5	LOC6	LOC7	LOC8	LOC9	LOC10
1221	97.5	0.69	0.41	0.86	0.21	0.10	0.12	0.05	0.01	0.00
1232	24.0	26.7	26.4	18.5	3.76	0.42	0.16	0.04	0.01	0.00
1242	0.70	13.7	43.4	34.2	7.27	0.66	0.04	0.00	0.00	0.00
1248	0.06	1.40	20.7	55.6	20.2	1.78	0.14	0.03	0.01	0.00
1254	0.02	0.21	0.83	16.1	54.6	25.5	2.71	0.09	0.00	0.00
1260	0.03	0.15	0.40	0.69	9.91	43.6	35.7	8.66	0.82	0.09
1262	0.02	0.34	1.21	1.09	3.44	27.3	45.5	19.3	1.75	0.03
1268	0.00	0.00	0.03	0.09	0.09	0.11	4.89	42.5	45.8	6.46

7.4.1 Forensic analysis of PCBs in soil

Forensic analysis of PCBs in soils, including both those identified as candidate background samples and suspected impacted samples, involves data exploration techniques that combine evaluation of PCB concentrations and PCB pattern analysis. Often, the results of pattern or numerical analyses are paired with geospatial analysis to depict relative geographic locations of samples with similar (or different) PCB compositional patterns.

Evaluation of the distribution of PCB soil concentration data are used to identify subsets of data representative of background and of impacted conditions. The statistical methods to carry out that analysis are described in Section 3 and Section 11 of this guidance. Examples of investigations of background levels of PCBs in soils can be found in ATSDR (2000)[275], Meijer et al. (2003)[339], Creaser et al. (1989)[296], USEPA (2007)[406].

Once candidate background and suspected impacted soil samples are identified, an evaluation of the pattern(s) of PCBs for those classes of samples can be conducted. The steps involved in PCB pattern analysis involve data pretreatment, followed by graphical analysis of diagnostic PCB congener patterns that are descriptive of the composition of the samples, thereby allowing the investigator to describe and contrast the PCB composition of both background and soils suspected of being impacted by PCBs.

Selection of data evaluation methods and/or selection of diagnostic metrics used in forensic analysis is accomplished through critical exploration of the dataset under investigation. The goal of such initial data exploration is to identify compositional characteristics descriptive of the samples and, where the data allow, distinguishing PCB compositional characteristics of background soil samples from those of impacted soils.

7.4.2 Data screening and treatment of nondetects

Effective forensic evaluation of PCB data relies on both the method of chemical analysis used to produce the data and the inherent quality of the dataset. In theory, the PCB congener-specific data produced from USEPA Method 1668 provides more PCB compositional information than the less data-rich USEPA Method 8082 or simple Aroclor-only data produced by USEPA Method 8080. However, the underlying quality of data often dictates how reliable a dataset will be for forensic analysis. The data user is cautioned to carefully screen data quality prior to use in forensic evaluations. The most common systematic problems with PCB datasets are elevated detection limits (for congeners, LOC results, or Aroclor concentrations) and significant co-elution of congener peaks, which can lead to difficulty in interpreting the patterns of PCB congeners, and/or introduces significant biases in numerical analysis of PCB LOC or congener data.

Elevated detection limits that are the result of laboratory artifacts or simply very low PCB concentrations in certain samples can lead to a significant number of nondetects for reported congener or LOC data. Replacing the nondetected values with either the detection limit or one-half the detection limit is a common data preprocessing approach and will standardize the influence of nondetects in multivariate data analysis. However, if certain samples contain a large number of nondetected results, the congener results will be dominated by the substituted detection limit values. Data users should conduct sensitivity analysis to identify and exclude samples where the nondetects bias or skew the results of subsequent forensic analysis of PCB congener patterns. In multivariate analysis or diagnostic ratio analysis, samples biased by detection limit substitutions are usually clearly evident and graphically cluster together, remote from samples based upon reliable congener or LOC data.

Combining PCB data from different sources or from different analytical methods can lead to significant biases in the combined dataset. For example, gas chromatographic traces of PCBs developed by USEPA SW-846 Method 8080 or Method 8082 can vary significantly in resolution, and thus quality, from laboratory to laboratory. Such differences in chromatographic quality can make compositional comparisons problematic. Total PCB concentrations can vary significantly for the same samples analyzed using different analytical methods because of “double counting” of overlapping chromatographic peaks that can occur in traditional Aroclor analysis—a phenomenon that does not affect high-resolution congener analysis data

(Wischkaemper, Beliveau, and R.W. Henderson 2013)[483]. Similarly, congener data produced using the lower resolution USEPA SW-846 Method 8082 can be significantly affected by co-elution interferences compared to congener data produced using the high-resolution USEPA Method 1668. This can lead to apparent congener compositional differences among samples that are actually nothing more than artifacts caused by congener co-elution.

In summary, it is incumbent upon the environmental professional evaluating PCB data to carefully review data for uniformity in reporting limits and resolution quality of PCB congener and/or LOC data prior to conducting forensic evaluation of PCB data. Based on that review, the geochemist can eliminate low quality data or take other remedial pre-evaluation steps (for example, combining potential co-eluting congener data into congener data pairs) prior to conducting forensic evaluation.

7.4.3 Forensic PCB compositional pattern analysis techniques

There are numerous techniques for evaluating the PCB composition of background and impacted soils that range from simple (for example, Aroclor identification) to complex (for example, multivariate data analysis). As the degree of complexity of the analysis increases, more subtle distinctions are recognized regarding the PCB composition among background samples or between background and impacted samples. The goals of such forensic evaluations are to (a) describe the background soils; (b) identify potential compositional outliers in background soil datasets; and (c) catalog the compositional features of impacted site soils, and contrast such features with those of the background dataset.

- Aroclor identification. The simplest—and lowest resolution—PCB data evaluation method is based on laboratory-supplied Aroclor identification, which is the most basic output from PCB analytical results such as USEPA Method 8080 or USEPA Method 8082. In this approach, the PCB characteristics of soil samples are simply organized by Aroclor type. Figure 7-5 is an example of how to evaluate such data. In this figure, Aroclor identification data are binned by concentration, and the Aroclor character of the candidate background and potentially impacted soil samples can be evaluated.

- Level of chlorination histograms and cross-plots. Small datasets are amenable to straightforward comparisons of histograms of the relative LOC among samples from data developed from USEPA Method 8082 or USEPA Method 1668. LOC concentration data for samples are normalized and plotted in histogram form from LOC1 through LOC10. Figure 7-6 illustrates that qualitative similarities or differences in LOC composition among samples can be readily observed using this method. In this figure, the average LOC was calculated and used as a descriptive statistic to compare background soil sample data to potentially impacted soil samples.

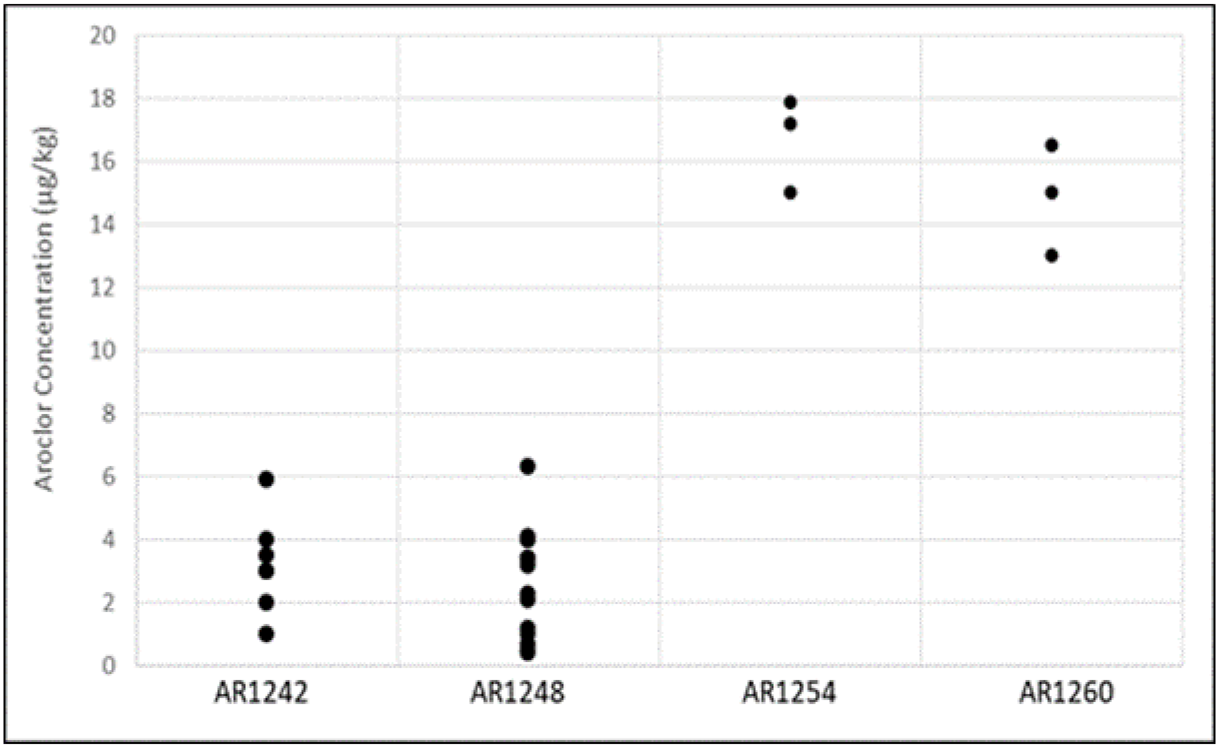


Figure 7-5. Candidate background soil samples organized by Aroclor type and concentration. Sixty-five percent of the samples have Aroclor concentrations less than 7 µg/kg and contained either Aroclor 1242 or Aroclor 1248. Higher concentration (>12 µg/kg) soil samples contained higher molecular weight Aroclor 1254 and Aroclor 1260, suggesting potential point source impacts.

Source: A. Uhler, NewFields Companies LLC (unpublished).

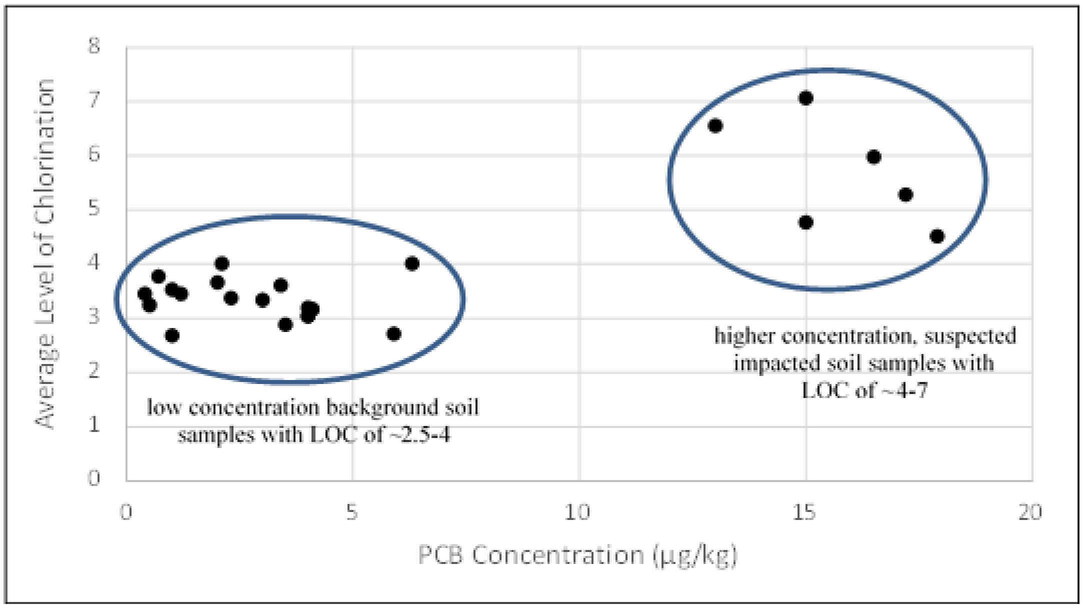


Figure 7-6. Low PCB concentration background soil samples have a lower LOC composition than suspected impacted site soil samples.

Source: A. Uhler, NewFields Companies LLC (unpublished).

- PCB congener histograms. Similar to LOC histograms, PCB congener composition histograms provide a graphical depiction of the PCB composition of samples. Histograms analogous to those shown in Figure 7-4 can be prepared for all samples, showing the relative PCB congener concentrations in congener number order by increasing LOC (Douglas et al. 2007)[485]. Qualitative PCB compositional characteristics of background and potentially impacted soil samples can be observed and categorized using this technique.

- Multivariate numerical analysis of LOC and PCB congener datasets. PCB datasets containing LOC or congener data are amenable to a variety of numerical data analysis techniques that provide a means to categorize PCB compositional data using multivariate techniques. An excellent review of multivariate data analysis methods for forensic evaluation is provided by Johnson et al. (2007)[323]. Examples of multivariate data analysis techniques available to the investigator include:

    - Principal component analysis (PCA). A mathematical method used to reduce the dimensionality of large datasets by transforming a large set of variables into a smaller one that still contains the majority of the information of the original data (Bro and Smilde 2014)[285]. PCA transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called principal components. The principal components are rank-ordered by the amount of variance in the data they explain. Principal components represent the directions of the data that explain a maximal amount of variance among the data. Factor scores plots are visual projects of the principal component analysis. Factor scores that plot close to one another share similar chemical variability; sample scores that plot distant have different chemical variability (composition). Compositional characteristics of samples can be inferred from the factor scores and corresponding component loading charts.

- Polytopic vector analysis (PVA). A statistical pattern recognition technique that allows determination of compositional end-members and proportions of end-members found among samples (Noémi, Goovaerts, and Adriaens 2004)[350]. The PVA method generates three basic parameters: 1) the number of end-members (sources), 2) the composition of each end-member (signature of each source), and 3) the relative proportions of each end-member (source) in every sample.
- Soft independent modeling of class analogy (SIMCA). PCA is performed on a reference dataset, and reference class characteristics are identified. Unknown samples are then analyzed and compared to the reference data and classified according to best fits to the reference data, or mixtures thereof (Dunn, Stalling, and Wold 1984)[301].

The advantages of these multivariate numerical methods are that they provide a convenient means to analyze large datasets and reduce the output into relatively straightforward graphical output that can then be interpreted, such as for identifying the compositional features of background soil samples and comparing those chemical signatures to potentially impacted soil samples.

7.5 Polychlorinated Dibenzo-p-Dioxins and Dibenzofurans (PCDD/F)

Please refer to the USEPA Dioxin website as a resource for technical information pertaining to polychlorinated dibenzo-p-dioxins and dibenzofurans (PCDD/Fs).

PCDD/Fs are complex mixtures of chlorinated congeners of two benzene rings connected by either one or two oxygen molecules (PCDFs and PCDDs, respectively). These molecules can contain one to eight chlorine atoms (Figure 7-7). Arrangement of the one to eight chlorine atoms around each molecule yields 75 possible PCDDs congeners and 135 possible congeners of PCDF (Shields et al. 2006)[360].

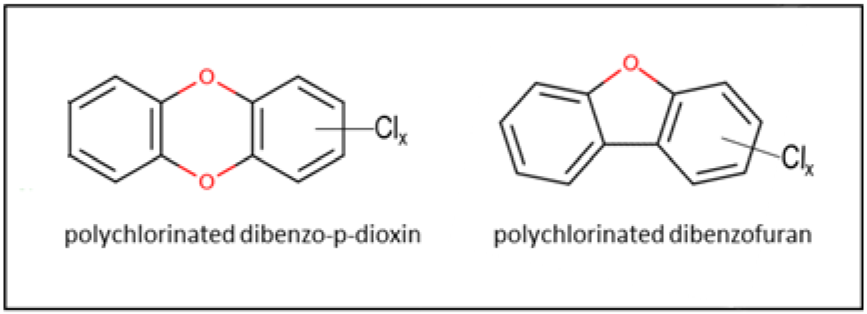


Figure 7-7. Generalized structures for polychlorinated dibenzo-p-dioxins and dibenzofurans. The number of chlorine atoms (x) can range from one to eight.

Source: A. Uhler, NewFields Companies LLC (unpublished).

Of the 210 possible congeners of PCDD and PCDF, 17 are commonly measured in environmental investigations, along with the summed total of the congeners of the same level of chlorination for the tetra- through hepta-chlorinated compounds (Table 7-3). The USEPA recommends the use of the consensus toxicity equivalency factor (TEF) values for 2,3,7,8-tetrachlorodibenzo-p-dioxin and dioxin like compounds published in 2005 by the World Health Organization (USEPA 2021)[478]. The concentration of the 17 PCDD/Fs congeners is commonly converted to a 2,3,7,8-TCDD ("dioxin") toxicity equivalent (2,3,7,8-TCDD TEQ) by multiplying each of the 17 congener concentrations by a congener-specific TEF, which is a fraction of the measured toxicity of 2,3,7,8-TCDD, the most toxic of the 17 PCDD/F congeners. The individual results of this transformation are then summed to form the 2,3,7,8-TCDD TEQ and used to facilitate risk assessment of exposure to the mixture of multiple PCDD/Fs. The TEFs listed in Table 7-3 are those published by the World Health Organization (2005). Among the benefits of converting raw dioxin concentration data to TEQ is the ability to compare measured TEQ concentrations with published benchmarks for media, including soil. This section discusses specific environmental forensics for assessing background in soils. Please refer to USEPA (2003)[409] and Shields et al. (2006)[360] for details regarding background knowledge about PCDD/F that may be helpful for some readers who are not familiar with this information.

Table 7-3. Most commonly reported PCDD and PCDF congeners

Source: (Van den Berg et al. 2006)[396].

Analyte	Toxicity Equivalency Factor (TEF)
Dioxin congeners	
2,3,7,8-TCDD	1
1,2,3,7,8-PeCDD	1
1,2,3,4,7,8-HxCDD	0.1
1,2,3,6,7,8-HxCDD	0.1
1,2,3,7,8,9-HxCDD	0.1
1,2,3,4,6,7,8-HpCDD	0.01
OCDD	0.0003
Furan congeners	
2,3,7,8-TCDF	0.1
1,2,3,7,8-PeCDF	0.03
2,3,4,7,8-PeCDF	0.3
1,2,3,4,7,8-HxCDF	0.1
1,2,3,6,7,8-HxCDF	0.1
1,2,3,7,8,9-HxCDF	0.1
2,3,4,6,7,8-HxCDF	0.1
1,2,3,4,6,7,8-HpCDF	0.01
1,2,3,4,7,8,9-HpCDF	0.01
OCDF	0.0003
Level of chlorination-sums	
Total tetrachlorodibenzo-p-dioxin (TCDD)	N/A
Total pentachlorodibenzo-p-dioxin (PeCDD)	N/A
Total hexachlorodibenzo-p-dioxin (HxCDD)	N/A
Total heptachlorodibenzo-p-dioxin (HpCDD)	N/A
Total tetrachlorodibenzofuran (TCDF)	N/A
Total pentachlorodibenzofuran (PeCDF)	N/A
Total hexachlorodibenzofuran (HxCDF)	N/A
Total heptachlorodibenzofuran (HpCDF)	N/A

7.5.1 Data screening and treatment of nondetects

Prior to forensic analysis, raw PCDD/F analytical data should be carefully screened to identify and censor data that are inaccurate or biased. This screening includes both candidate background and site investigation data. Data users should evaluate and exclude data that suffer from elevated detection limits; contain laboratory flags indicating that data may be affected by field or laboratory blank interferences, matrix interferences, or instrument detection limitations (detector saturation); or lack acceptable mass spectrometer second ion confirmation for reported PCDD/F congeners. Useful data quality assessment guidelines for PCDD/F analytical data are provided in the USEPA *National Functional Guidelines for High Resolution Superfund Methods Data Review* (USEPA 2016)[167].

Many PCDD/F datasets contain nondetect results. Nondetects can be handled in a number of ways. Replacing the nondetect with either the detection limit or one-half the detection limit is a common approach and will standardize the influence of nondetects in subsequent data treatment (Shields et al. 2006)[360]. However, as the total PCDD/F concentration in a sample approaches the detection limit, a greater proportion of a sample's PCDD/F composition will be accounted for by nondetects, and the transformed PCDD/F congener results will be dominated by the substituted detection limit values. Data users should conduct sensitivity analysis to identify and exclude samples where the nondetects bias or skew the results of subsequent forensic analysis of PCDD/F congener patterns. In multivariate analysis or diagnostic ratio analysis, samples biased by detection limit substitutions are usually clearly evident and graphically cluster together, remote from samples based upon reliable PCDD/F data.

7.5.2 Forensic data analysis of PCDD/F in soil

Forensic analysis of PCDD/F in soils, including both those identified as candidate background samples and suspected impacted samples, involves data exploration techniques that combine evaluation of PCDD/F concentration and PCDD/F pattern analysis. Often, the results of diagnostic numerical analyses are paired with geospatial analysis to depict relative geographic locations of samples with similar (or different) PCDD/F compositional patterns.

Evaluation of the distribution of PCDD/F soil concentration data is used to identify subsets of data representative of background and of impacted conditions. The statistical methods to carry out that analysis are described elsewhere in this guidance. An excellent compilation of background investigations of PCDD/F in soil can be found in USEPA (2003)[409].

Once candidate background and suspected impacted soil samples are identified, an evaluation of the pattern(s) of PCDD/F isomers for those classes of samples can be conducted. The steps involved in PCDD/F pattern analysis involve data pretreatment, followed by graphical analysis of diagnostic PCDD/F congener patterns that are descriptive of the composition of the samples, thereby allowing the investigator to describe and contrast the PCDD/F composition of both background and soils suspected of being impacted by PCDD/F. Selection of data evaluation methods and/or selection of diagnostic metrics used in forensic analysis is accomplished through critical exploration of the dataset under investigation. The goal of such initial data exploration is to identify compositional characteristics descriptive of the samples and, where the data allow, to distinguish PCDD/F compositional characteristics of background soil samples from those of impacted soils. Refer to USEPA (2003)[409] and USEPA (2006)[410] for information pertaining to PCDD/F patterns for background soils and for documented PCDD/F sources, respectively.

PCDD/F patterns based on congener concentration data alone can be used for initial data exploration of candidate background soil samples and suspected impacted soil samples to evaluate obvious chemical patterns that describe and/or distinguish certain samples or groups of samples. Depiction of PCDD/F congener concentration data in basic graphical format such as bar charts or radar plots allows the investigator to quickly evaluate key PCDD/F compositional features of samples. Examples of such bar charts include those shown in Figure 7-8.

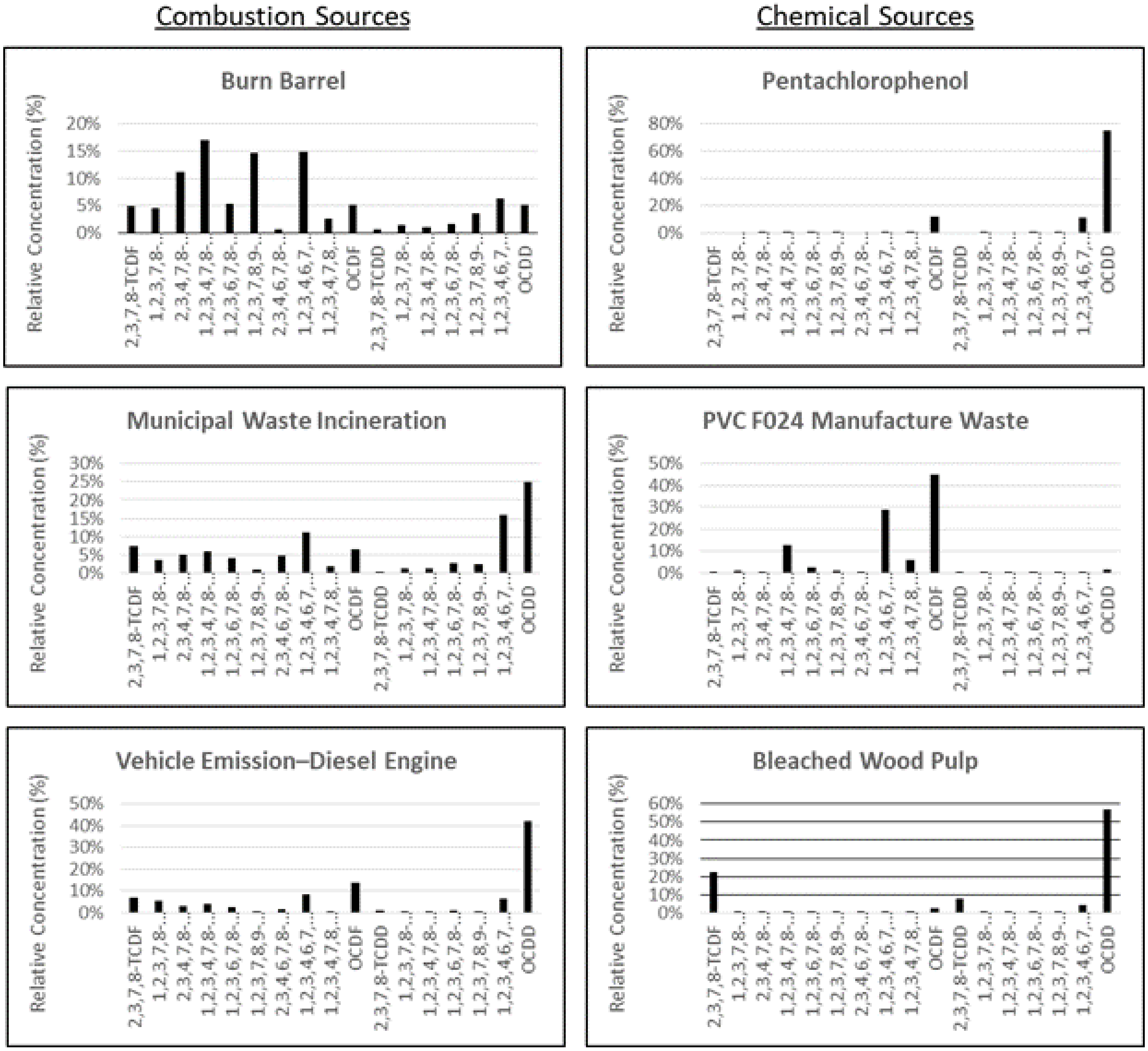


Figure 7-8. Examples of combustion and chemically produced PCDD/F congener distributions.

Source: (USEPA 2006)[410].

Soil data from PCDD/F investigations often contain samples of highly varied total concentration. In addition, the relative concentration of PCDD/F congeners within a sample can vary by several orders of magnitude. Both total concentration and disparity in relative PCDD/F congener concentration can lead to significant biases during data analysis. Hence, forensic evaluation of PCDD/F compositional features benefits from pretreatment of the data prior to exploratory analysis.

Data normalization helps attenuate the effect of these potential concentration-driven biases. Shields et al. (2006)[360] describes a number of data pretreatment steps that are useful for forensic evaluation of PCDD/F patterns and are applicable to evaluation of data related to soil samples. These include:

- Sum standardization. The concentration for each of the commonly reported 17 PCDD/F congeners is divided by the sum total concentration of the 17 congeners and multiplied by 100. This normalizes the relative PCDD/F concentrations on a scale of 0–100 for all samples, eliminating any concentration bias that may be introduced into compositional analysis. This approach does not adjust for significant differences in relative concentration among congeners, and the user must be aware of the limitations of this approach for certain types of data analysis—for example, correlation analysis.

- Relative homologue standardization. The concentrations of the commonly reported 17 individual PCDD/F congeners are divided by the respective total homologue concentration and multiplied by 100. This approach attenuates some differences in relative concentration among PCDD/F congeners. It is a useful data processing step for initial analysis of the data using certain multivariate data analysis techniques such as PCA or hierarchical cluster analysis, the latter of which identifies structure (similarity) among groups of samples within a dataset (McKillup 2012)[456].

- Relative TEQ standardization. The concentrations of the commonly reported 17 individual PCDD/F isomers are multiplied by the congener's respective TEF. This approach has the benefit of attenuating the dominance of relatively high PCDD/F congeners such as OCDD, OCDF, and other higher chlorinated congeners, yielding datasets that are reasonably well scaled and amenable to a variety of graphical and numerical analyses.

- Total homologue standardization. The concentration of each homologue class (for example, total tetrachlorodioxin isomers, TCDD) is divided by the sum of all homologue classes and multiplied by 100. Shields et al. (2006)[360] reported that this approach reveals gross differences in PCDD/F composition. The technique is applicable for basic graphical analysis (for example, bar plots) and useful for comparing PCDD/F composition among samples.

Evaluating the compositional features of PCDD/F in soil can be accomplished by a number of approaches, ranging from basic to more complex:

- Graphical (PCDD/F isomer distribution) analysis. The most straightforward approach for evaluating PCDD/F compositional features for soil samples is through basic PCDD/F pattern analysis of PCDD/F composition using bar plots or star plots (Shields et al. 2006)[360]. Transformed data, using one of the techniques described above, yields PCDD/F compositional depictions that can be compared among samples. This technique is often useful for small datasets where the investigator is not encumbered by the need to inter-compare a large number of compositional plots.

- Diagnostic ratios. The composition of PCDD/F isomers arising from many types of both combustion and chemical sources often has distinct proportions of PCDD/F isomers, which lend themselves to diagnostic ratio analysis. This analysis involves cross-plotting diagnostic ratio pairs, or a single diagnostic ratio versus total PCDD/F concentration. In such cross-plots, samples of similar source character share similar Cartesian space, while samples containing PCDD/F of different chemical character plot remote from the otherwise similar samples ((Horstmann, McLachlan, and Reissinger 1993)[312], (Townsend 1983)[374]).

- Multivariate analyses. Mathematical, multivariate methods can be used to identify samples that contain PCDD/F with similar (or different) PCDD/F compositional patterns. A variety of multivariate techniques are available to the investigator and include principal components analysis (PCA), hierarchical cluster analysis, and other receptor-based models (Johnson et al. 2007)[323]. These methods are particularly well suited to large sample datasets, where otherwise simpler methods of analysis—for example, comparison of histograms—would be too unwieldy and time-consuming.

The goal of PCDD/F forensic analysis—regardless of the data analysis method used—is to identify the pattern(s) of PCDD/F isomers that are embodied by the statistically identified group of samples that is emblematic of background. The characteristic pattern or range of patterns assigned to background conditions can then be compared to samples suspected of being impacted by site-related PCDD/F. Differences in the chemical patterns between background and impacted samples lend insight into the nature and geographic extent of suspected PCDD/F impacts. An example comparing the PCDD/F compositional pattern for off-site background soil samples to PCDD/F impacted site soil samples is shown in Figure 7-9.

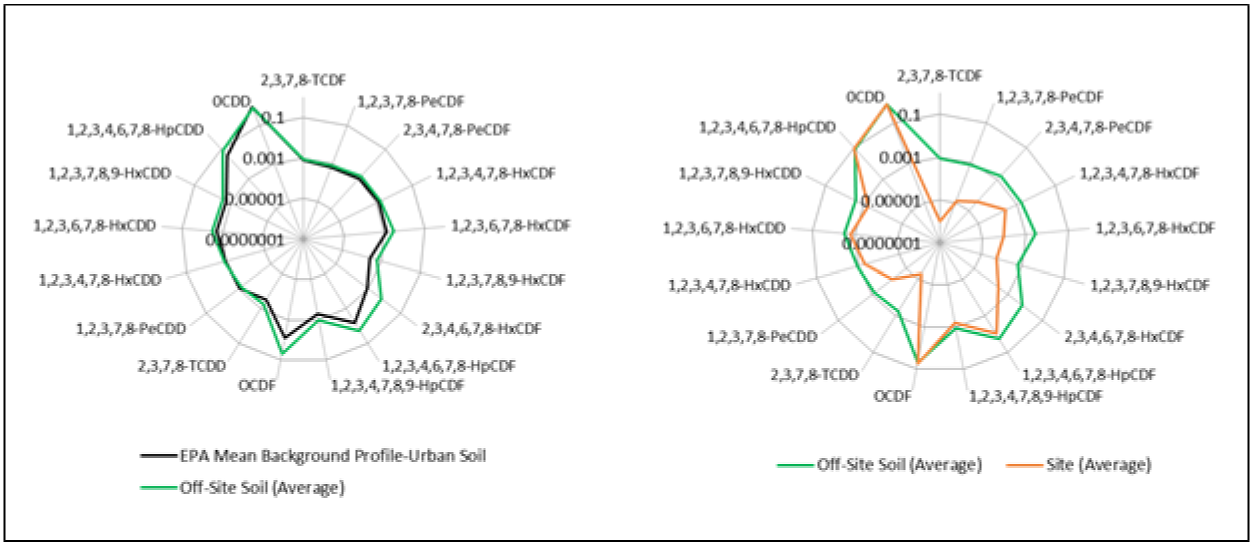


Figure 7-9. Sum-standardized star plots showing comparison of PCDD/F pattern for off-site background soil samples with the average PCDD/F pattern for impacted soils. The plot on the left shows the PCDD/F pattern for off-site background soil samples that compare favorably with the pattern for USEPA mean urban background soil, while the plot on the right shows that the average PCDD/F pattern for impacted soils from the related industrial site exhibits a clearly different PCDD/F pattern compared to background.

Source: A. Uhler (unpublished).

7.6 Perfluoroalkyl Substances (PFAS)

Per- and polyfluoroalkyl substances (PFAS) are synthetic fluorinated aliphatic compounds. They are more water-soluble than other classes of contaminants, such as PAHs and PCBs, and are therefore most commonly recognized as groundwater contaminants. However, some of the PFAS released to the environment will adsorb to solids, and PFAS are found in soils, which can act as a reservoir for groundwater contamination. A detailed discussion of the PFAS sources, physiochemical properties, environmental fate and transport, regulations, site characterization methods, sampling and analysis considerations, and treatment technologies is presented in the ITRC *PFAS Technical and Regulatory Guidance Document* (ITRC 2021)[447]. This section focuses on the measurement, data handling, and data analysis techniques that are relevant to forensic analysis of PFAS substances in background and site-impacted soils.

7.6.1 PFAS uses

The unique properties of PFAS have allowed them to find use in many different industrial and consumer applications. For example, PFAS are:

- key components of Class B fluorine-containing firefighting foams, including aqueous film forming foam (AFFF), fluoroprotein (FP), and film-forming fluoroprotein foam (FFFP)
- used as stain- and water-resistant coatings and treatments for textiles and leather

- used to treat paper products for packaging or serving food
- used in metal plating and etching, as they can maintain their efficacy under the high temperatures and acidic conditions
- used in wire manufacturing to coat wires and cables
- used to manufacture semiconductors in the photolithography process
- used in manufacturing, as industrial surfactants, to manufacture plastics and fluoropolymers, as mold release coatings, or in other applications
- used in consumer products, including as nonstick surfaces in cookware; to treat carpets or automotive upholstery; in outdoor textiles and sporting equipment; in cleaning agents and fabric softeners; in polishes and waxes; in medical products; in varnishes, dyes, and inks; and in personal care products such as shampoo, hair conditioner, sunscreen, cosmetics, toothpaste, and dental floss

In some cases, PFAS materials may be directly discharged to the environment (for example, as with firefighting with AFFF or atmospheric releases from manufacturing facilities). Given the range of industrial and consumer uses of PFAS materials, PFAS is also common in various waste streams; wastewater treatment plants and landfills are common sources of PFAS to the environment. For a more detailed description of the history and use of PFAS, see the ITRC's fact sheet entitled *History and Use of Per- and Polyfluoroalkyl Substances (PFAS)* (ITRC 2020)[318] and Section 2.5 of the ITRC's online *PFAS Technical and Regulatory Guidance Document (ITRC 2021)[447]*. This section discusses specific environmental forensics for assessing background in soils. Please refer to the ITRC fact sheet entitled *Naming Conventions and Physical and Chemical Properties of Per- and Polyfluoroalkyl Substances (PFAS)* (ITRC 2020)[319] as well as Buck et al. (2011)[288] or additional information regarding background knowledge about PFAS that may be helpful for some readers who are not familiar with this information.

7.6.2 Forensic PFAS compositional pattern analysis techniques

The field of PFAS forensics is in its nascent stages. The understanding of the concentrations of target PFAS or forensic signatures associated with source materials is evolving. Research is being conducted to understand fate and transport in the environment and to evaluate tools to investigate changes in PFAS composition for the purposes of source attribution. The reader is referred to Brusseau, Anderson, and Guo (2020)[286] and Rauert et al. (2018)[353] as examples of investigations pertinent to findings of PFAS in soils.

Benotti et al. (2020)[277] have suggested a tiered approach for characterizing the PFAS signatures associated with source materials and have suggested that some or all aspects of this approach may be useful when evaluating PFAS contamination in soil or water samples. However, until a library of source materials is available, and until researchers understand the full scope of forensic information that may help elucidate source, the best approach for associating or dissociating environmental contamination from one or more sources is to compare PFAS information in a soil or water sample to similar information from the site-specific source area/s. Information on PFAS forensics is provided in Section 10.5 of the ITRC's online *PFAS Technical and Regulatory Guidance Document (ITRC 2021)[447]*.

7.7 Remote Sensing

7.7.1 Overview

Remote sensing isn't a single tool, but a collection of tools that can detect properties of materials without direct contact. The ones of primary interest to background evaluations are satellite imagery, aerial photography, optical or electromagnetic spectrum techniques, such as optical sensing with specific bands (for example, SPOT, ASTER, Landsat), optical sensing based on hyperspectral imaging (based on 100 or more contiguous spectral bands), and optical sensing using the thermal infrared band. For determining suitable background reference areas, the most useful remote sensing tools are historical imagery and delineation of contaminated soils by correlating soil attributes with spectral features across a range of spectral bands.

Remote sensing can include radar, satellite imagery, geographical information systems, sonar, imaging through spectral bands, and light detection and ranging (lidar). For a given material, the amount of solar radiation that is reflected, absorbed, transmitted, and emitted varies with wavelength. When this radiation or electromagnetic energy is plotted over a range of wavelengths, the resulting curve is called the material's spectral signature. Multispectral imaging, based on several spectral signatures, and image classification can be used to generate thematic maps. More information regarding remote sensing and sensors can be found in these references: ((ITRC 2019)[317], (Gholizadeh et al. 2018)[479] ,(Slonecker and Fisher 2014)[362], (Singh 2016)[361]).

7.7.2 Items to consider

The reflected or emitted radiation values observed by remote sensing tools are products of geophysical properties and conditions at or near the earth's surface (or shallow subsurface). Certain spectral signatures can be interpreted with multiple explanations and multiple metals and contaminants at a site can result in overlapping signatures that can be complex to sort out. It is important to have ground truthing of remote sensing data for proper interpretation.

Scale needs to be considered. For example, spatial resolution can vary from the millimeter scale (drone-based cameras) to meters (SPOT and Landsat) or greater, so the appropriate data source must be chosen to meet the needs of the project at hand. Remote sensing data are often in a raster format and the size of the pixels must be compared to the needs of the project. Modern advances, such as deploying remote sensing equipment with unmanned aerial vehicles (drones), may be helpful for smaller scale projects.

Satellite remote sensing data with lower spectral resolution may be inappropriate. Hyper- and super-spectral satellite sensors have recently been developed and may provide better tools for environmental data.

Remote sensing based on spectral reflectance techniques requires exposure to soils, so areas with extensive pavement or tarmac may not profit from these tools, unless the goal is to delineate extent of pavement or roads. The presence of vegetation may also impede the use of spectral reflectance techniques.

7.7.3 Application

Although there is no published research on using remote sensing techniques for establishing background concentrations or risk assessments, remote sensing is a tool that could be very helpful in assisting with such assessments, such as evaluating whether an area should be considered as background. An important use of remote sensing in establishing background concentrations is for examining images of the selected background reference area in the past. Historical satellite imagery and aerial photographs can be used to identify areas that currently appear barren or untouched by human activity but that historically were subject to activities that could have affected background concentrations in soil (for example, areas surrounding a closed or abandoned mine). This can occur both in a range of areas where vegetation growth is quick and broadly dispersed to areas that have remained undeveloped once historical structures were removed.

Remote sensing has been used for many years for evaluating spatial distribution of soil properties, such as soil texture and soil moisture, and for mineral exploration, but its use for environmental investigations has been sparse. Its strength is to provide spatial information over large areas.

The ability of remote sensing to delineate the spatial distribution of metals or other contaminants would help determine whether the concentrations at a site are locally restrictive, hence site-related, or extensive and more than likely background. The spatial relationship of potential contaminants, such as metals and PAHs in soil, can be useful in improving risk management.

Arsenic contamination in soils can be mapped by relating the reflectance spectrum of plants with arsenic content in soils and mapping those spectra (Asmaryan et al. 2014)[480]. Arsenic in soil can be predicted from hyperspectral imaging (Wei et al. 2020)[389]. Remote sensing is a fast-evolving field and may have greater use in the future for determination of chemical concentrations.

8 CONCEPTUAL SITE MODEL AND DATA QUALITY OBJECTIVES

Two general items that are important when establishing soil background and using it in risk assessment are a conceptual site model (CSM) and data quality objectives (DQOs). The descriptions below are not intended to be comprehensive or detailed, rather they are intended to introduce these concepts and provide key references for additional information.

8.1 Conceptual Site Model

A CSM is the integrated representation of the physical and environmental context, the potential fate and transport of COPC, and the complete and potentially complete exposure pathways associated with each receptor at a cleanup site that is being evaluated (ASTM E2616-09(2014), (ASTM 2014)[135], ASTM E1689-20 (ASTM 2020)[136], ASTM D6169/D6169M (ASTM 2013)[141]. The goal of a CSM is to provide an understanding of relevant site features and conditions to understand the extent of COPC and the critical exposure pathways for evaluation in risk assessments. A CSM should provide a thorough understanding of the physical characteristics of the site, as well as the sources of site contamination, potentially contaminated media, contaminant transport pathways, and exposure pathways applicable to receptors. A detailed understanding of a cleanup site that is being evaluated informs the development of an accurate CSM, which in turn provides important information that may be used to determine whether site chemical concentrations represent soil background and to identify potential soil background reference areas. Important concepts that should be included in a CSM are contribution of natural and diffuse anthropogenic inputs versus site-associated releases because all might be present at the site. CSMs may evolve throughout the duration of a project as more information becomes available (USEPA 2011)[462]. As noted, a well-developed site CSM aids in selection of background reference areas by ensuring that the site context—including land use, geochemical, and chemical considerations—is understood and background reference areas can be selected that reflect similar inputs, minus site release impacts.

As described in the ITRC TPH document ((ITRC 2018)[316], information on the development of a CSM is readily available in several guidance documents including the following:

- ITRC Triad Implementation Guide, ITRC SCM-3 (ITRC 2007)[488]
- Guidance for Conducting Remedial Investigations and Feasibility Studies Under CERCLA (USEPA 1992)[489]
- Data Quality Objectives Process for Hazardous Waste Site Investigations: Final Guidance (USEPA 2000)[490]
- Standard Guide for Developing Conceptual Site Models for Contaminated Sites, ASTM E1689-20 (ASTM 2020)[136]
- Environmental Cleanup Best Management Practices: Effective Use of the Project Life Cycle Conceptual Site Model (USEPA 2011)[462]

8.2 Data Quality Objectives

A systematic planning process, such as USEPA's DQO process (USEPA 2006)[134], is a key step in developing a successful sampling and analysis program to ensure the appropriate sampling, analyses, and data evaluations are conducted to meet program objectives. The DQO process is important to use whenever sampling and analysis are conducted as part of establishing soil background and using soil background in risk assessment because this leads to consensus on the type, quality, and quantity of data needed to meet project goals (USEPA 2006)[134].

The DQO process may be used when designing a soil background study intended either to characterize soil background to establish default background for a larger area (for example, a state, a region, or a unique geological area) or to obtain site-specific soil background for a site being evaluated. It also aids in the planning process by establishing the scope and scale of the study required, which can aid in resource planning.

USEPA's DQO process should be applied to soil background studies to help ensure that background data are comparable to data from the site being evaluated and the background dataset is suitable for its intended use in risk assessment.

Details on the USEPA's DQO process can be found in *Guidance on Systematic Planning Using the Data Quality Objective Process* (USEPA 2006)[134]. For a complete list of USEPA guidance documents on quality programs for environmental sampling, refer to https://www.epa.gov/quality/agency-wide-quality-program-documents. The initial five steps of the DQO process are focused on defining qualitative criteria, such as the nature of the problem, the decisions or estimates that need to be made, the types of data needed, and a "decision rule" process that defines the logic for how the data will be used to draw conclusions from the study findings. This aids in defining the characteristics and spatial extent of the sampling program necessary to determine soil background.

The sixth step establishes acceptable quantitative criteria on the quality and quantity of the data to be collected. These criteria are known as performance or acceptance criteria, or DQOs. This step is used to define whether data collected or otherwise used are acceptable or unacceptable for use. Performance criteria form the basis of a project's quality assurance project plan (QAPP), which relates back to the DQOs and ensures that the data are suitable for the intended use (for example, verifying that the analytical reporting limits are lower than background values). The final step of the DQO process involves data collection design that will meet both the qualitative and quantitative criteria to ensure that sampling design and the analytical program generate data acceptable for use. This final step, developing the plan for obtaining the data, is the basis for the sampling and analysis plan (SAP).

SAPs and QAPPs are important for conducting a soil background study, as they document the project's DQOs so that data collection and data analyses generate data suitable for use. The QAPP and SAP are discussed further below.

8.2.1 Quality assurance project plan

The QAPP should define the procedures to collect, preserve, and analyze samples, as well as store and manage analytical data, to ensure that the data collected are of sufficient quality to meet project needs. The QAPP should include at a minimum:

- requirements for field quality control (QC) samples (for example, field duplicates, field blanks)
- methods to prevent cross contamination (for example, field decontamination procedures)
- field equipment, including calibration and maintenance of that equipment
- field documentation methods
- number of samples to be collected for each evaluation, and justification for sample number
- measurement performance criteria (such as bias, precision, and completeness) for each test method proposed

Additional details on QAPPs can be found in guidance documents published by USEPA, including QA/G-5 (USEPA 2002)[133], and in USEPA's Quality Assurance Project Plan Development Tool.

8.2.2 Sampling and analysis plan

A well-designed sampling program is critical when conducting a study to determine soil background concentrations. SAPs are sometimes called field sampling plans (FSPs) and are generally required for projects performed under regulatory orders or as part of a regulatory process. Please refer to Section 9 for additional details.

As outlined in USEPA (2006)[134], a SAP should include the following components:

- rationale for each sample or group of samples based on the project DQOs
- number of samples, along with justification for the number of samples to be collected
- sample type (composite vs. discrete samples)
- sample locations and design, along with justification for how sample locations were selected and the area or quantity that each individual sample represents
- sample collection method (for example, surficial soil sampling versus drilling)
- protocols for sample collection, preservation, handling, and shipping
- analytical methods

- statistical sampling plan

9 SAMPLING

A well-designed sampling program is critical when conducting a study to determine default or site-specific soil background or obtaining soil samples from an investigative site that is being evaluated to compare to default or site-specific soil background. This section provides an overview of important considerations when designing a soil background sampling program, including:

- areas for collecting representative soil samples
- sample depth
- sample size
- sampling methods (discrete, composite, incremental sampling methodology (ISM))
- sampling design
- sample collection methods
- sample handling

Whenever possible, existing guidance will be referenced because this section is not intended to be a detailed description of soil sampling procedures but is intended to provide an overview of how the procedures relate to soil background. Field sampling practices are a source of variability in reported results. This variability can be minimized, but not eliminated, by using a high-quality sampling plan and experienced personnel to perform the field work.

9.1 Background Reference Areas

When selecting areas to collect soil background samples, it is very important that locations should be as similar as possible to the site that is being evaluated in physical, chemical, geological, biological, and ecological characteristics; the rationale for this is discussed further in Section 9.1.3. Land use of the background reference area, compared to the site, should also be considered. The background reference areas selected should not be affected by site releases or site activities. Additionally, where transport processes are likely to deposit new soils on the site (such as floodplains, alluvial fans, or areas with high dust creation and deposition), the background reference location should be located upgradient or upwind of the site to avoid contamination from the site and to provide information on contaminants that may be transported to the site. Multiple background reference areas could be required for the collection of sufficient background samples to determine representative background concentrations if the site exhibits a wide range of physical, chemical, geological, and biological characteristics.

Typically, background reference areas are located off site. In some cases, it may be difficult to find a suitable background reference area in an industrialized area. In these cases, an on-site area may be used as a background reference area if the area has not been affected by site releases or site activities (USEPA 2002)[154]. Thus, background reference areas are not limited to natural areas or undeveloped lands.

Review of the history of the site and surrounding areas, as well as current and historical aerial photographs, geological maps, soil maps, and vegetation maps, can assist in selecting candidate background reference areas. Soils maps, available geotechnical and environmental drilling borehole data, and satellite imagery can also provide useful insights. Areas that generally should be avoided when collecting background samples include industrial areas, roadways, stormwater ditches, areas of local anthropogenic releases, and areas with fill. Areas with fill are generally unsuitable to use as a soil background reference area; however, some jurisdictions may allow for the exclusion of contaminants in fill materials if these contaminants are not site-related. Before considering the use of fill in a background study, the regulatory agency should be consulted.

Alternatively, existing background studies from sites located adjacent to the cleanup site being evaluated, with similar soil properties, and collected using similar sampling methods, can often be used for representative background values with high confidence (USEPA 1995)[251]; however, use of the adjacent site as a background reference area still needs to be evaluated and technically justified.

Once preliminary background reference area sampling locations have been chosen, it is important to compare these locations to available remediation site databases (for example, USEPA's Cleanups in My Community or state websites such as California's CalEnviroScreen) to ensure that sample locations are not close to known potential contamination sources. The data resulting from samples taken from the background reference area could be evaluated using statistical (Section 11), geochemical evaluation (Section 5), or forensic (Section 7) methodologies to ensure that the background reference area has not been impacted by localized contamination sources (sometimes referred to as "hot spots"). Some complex situations, such as the presence of anthropogenic influences and variable site geology, can make it challenging to identify background reference areas not impacted by on- or off-site activities. In such situations, it might be appropriate to establish site-specific background data by extracting it from an on-site dataset (Section 3.8).

9.1.1 Natural soil background

When selecting samples to determine natural soil background, sample locations should be chosen from areas that generally share similar physical, chemical, geological, and biological characteristics and are unlikely to have been impacted by human activities. Public lands, such as state or national forests, are often good candidate areas for obtaining background samples, assuming such areas have not been previously impacted by anthropogenic activities (including agricultural activities).

Practitioners should document the technical rationale for choosing each of the sample locations.

9.1.2 Anthropogenic ambient soil background

When selecting samples to determine anthropogenic ambient soil background as defined by this document, sample locations should be chosen from areas that generally share similar physical, chemical, geological, and biological characteristics and are unlikely to be impacted by local releases. These soil background reference areas could be affected by anthropogenic nonpoint sources of chemicals that are present in soil (often, but not always at lower concentrations)—not because of local anthropogenic sources, but because of their persistence, ubiquity, and/or ability to be transported long distances. Examples include PAHs, polychlorinated dibenzo-p-dioxins (PCDDs), persistent pesticides (such as DDT), and mercury.

What is included in anthropogenic ambient soil background depends on what the lead regulatory agency allows and the goals of the project or the site-specific situation. In some cases, the lead regulatory agency allows the inclusion of local releases that are not a result of releases from the site that is being evaluated to be included in anthropogenic ambient soil background (for example, the release of metals related to local mining or smelting operations). In some jurisdictions, metals might be included as a widespread urban source related to the historical association of lead with vehicle emissions (the contribution of this source has decreased over time, as leaded automotive gasoline use has been phased out). In these cases, the level of urbanization and intensity of anthropogenic land use are very relevant. Sample locations should be selected from an area with a similar level of urbanization to the site and as proximate to the site as possible.

Practitioners should document the technical rationale for choosing each of the sample locations.

9.1.3 Geological and other considerations

Naturally occurring concentrations of elements are an important consideration in soil background studies. Naturally occurring elements detected in soil are derived from parent material (bedrock) that was chemically and physically weathered. Climate, physical and chemical erosion processes, and the composition of the parent material determine the minerals that occur. Additional detail on the elements commonly found, their parent material, and effects of weathering and climate on chemical composition can be found in the reference *Chemistry of Soils* (Sposito 2016)[263]. Additional discussion of geochemical considerations is provided in Section 5.

Geotechnical sampling and testing procedures (for example, grain-size distribution) should be included as part of the SAP ((Winegardner 2019)[158], (ITRC 2020)[431]), along with the COPC. Mineralogy and biotic and abiotic weathering are important factors in determining grain-size distribution and are underlying factors that can contribute to environmental availability of naturally occurring elements. In addition, soil texture can have a substantial effect on the distribution of COPC in soil. For example, fine-grained soils can have a greater sorption capacity for some COPC (for example volatile organic compounds (VOCs), semivolatile organic compounds, and metals) due to greater surface area. Laboratory analysis to ensure comparable grain-size distributions between site datasets and background datasets should be conducted to ensure any COPC concentration differences between the datasets are not just a result of differing grain-size distributions.

In some soils, vegetation can affect the distribution of elements. Some metals (for example, arsenic, cadmium, copper, mercury, lead) can be taken up by plants or adsorbed to organic matter. When the plant dies, these elements can be concentrated in the surface soil. Information regarding plant communities and differences in vegetation density should be considered as part of background study design (Kabata-Pendias 2010)[151].

9.2 Sample Depth

Soil sample depth should be considered when conducting background studies to ensure that datasets from the site and the background reference area are comparable. The soil sample depths chosen for both the site and background reference areas should be consistent with the receptors (as described in the CSM) that are expected to have contact with the soil. As noted above, background and site sample locations should be similar in physical, chemical, geological, biological, and ecological characteristics, which can be affected by depth. For example, surface soil can be affected by atmospheric influences, so the use of variable sample depths to define surface soil (for example, 0–15 cm versus 0–60 cm) can yield quite different results for contaminants that are deposited from the atmosphere. It should be noted that the definition of surface soil can vary by jurisdiction, with depths ranging from surface to 2, 15, 30, or 60 cm. For example, the USEPA's Soil Screening Guidance (USEPA 1996)[427] defines surface soil as the top 2 cm.

Sample depth should also take into account other site-specific factors that may vary with depth. These factors include the depth to the water table, perched water zones, or soil stratigraphy (for example, sand lenses and fracturing).

For risk assessment, the selected sample depth needs to reflect the type of exposures and receptors expected. Surface soil samples generally target the data needed to evaluate human health exposures via direct ingestion, dermal contact or inhalation (via dust) pathways, as well as some ecological exposures. Subsurface soil sampling generally targets the data needed to evaluate human health exposure to soil during construction or utility work and migration to groundwater or animals burrowing at depth.

Site-specific considerations must be taken into account when evaluating receptors and exposure pathways. For example, ecological receptors could be exposed to soils at depth while burrowing, depending on the species present at the site. Human health exposure to subsurface soils (as well as shallow soil) may need to be considered if there is soil disturbance at a site. Soil disturbance can be caused by construction (for example, basements, utility lines, or residential pool installation), landscaping, children digging while playing, gardening, soil erosion, or recreational activities (for example, ATVs).

As noted in Section 8, the CSM should document all relevant complete and potentially complete and complete exposure pathways for the receptors at the site, and that information should be used to help inform the relevant sampling depths for soil background.

9.3 Sample Size

Adequate background dataset sample size will need to be determined on a project-specific basis in accordance with the project's DQOs. For example, in the context of defining DQOs to support use of background in risk assessment, the DQOs for developing a site-specific background dataset would differ from DQOs for developing a default background dataset. The number of samples necessary to determine representative background, along with the rationale for the sampling locations and depths chosen, must be outlined in the SAP. The intended application, assumptions about the underlying distribution of the concentrations of the COPC, tolerable error rates (maximum acceptable error rate set by the decision maker), and sample design need to be considered when determining sample size. Inadequate background sample size can lead to unreliable or erroneous conclusions (ASTM E3242-20; (ASTM 2020)[146]).

Size of the background reference area and expected variability should also be accounted for when determining adequate sample size to ensure the data are representative. A larger number of samples may be required to adequately represent a larger and more variable background reference area. This is discussed in more detail in Section 11.1.3.

Additional information on determining adequate sample size can be found in USEPA (2002)[153] and USEPA (2002)[154]. Many software packages and tools can be used to determine an adequate sample size; two are discussed in detail here.

9.3.1 Sample size tools

Information on statistical software is also provided in Section 11.9.

9.3.1.1 *Visual Sample Plan software*

The free Visual Sample Plan software (VSP Development Team 2020)[430] is a recommended tool for use in calculations to determine sample size. The VSP software is specifically designed to provide an output to support the project DQOs. The software selects “the number and location of samples so that the results of statistical tests performed on the data…have the required confidence for decision making” (Matzke et al. 2014)[425]. The software can also recommend a minimum number of samples taking into account budget and sample design. Using the DQO process and this software ensures that the background study goals are well defined up-front, while identifying and minimizing uncertainty. VSP calculates DQO-based sample sizes for a wide range of applications, including the estimation of the mean or median, calculation of confidence intervals on the mean, one-sample and two-sample hypothesis tests, and location of hot spot areas, among many other statistics. Site maps and aerial/satellite photographs can be imported, with the resulting sample locations displayed on the maps or photographs, accompanied by geographic information system (GIS) coordinates for precise sample placement in the field. VSP has the support of a wide range of government agencies, including the Department of Energy, Department of Defense, USEPA, and Department of Homeland Security, among others.

9.3.1.2 *ProUCL*

ProUCL is another free statistical tool that can be used to determine sample size (USEPA 2015)[199]. Its sample-size module allows the user to calculate DQO-based sample sizes to support estimations of the mean, as well as one-sample and two-sample hypothesis tests (for example, *t*-test or Wilcoxon rank sum test). The software can also be used for comparing background and site concentrations. ProUCL is easy to use and accepted by regulatory agencies.

Its disadvantages include a limited number of scenarios for sample-size calculations, inability to import maps, and inability to depict sample placement.

9.4 Sample Methods

Sampling methods can consist of discrete, composite, or incremental sampling methodology (ISM). The sampling method chosen will depend on project goals.

Soil background samples should ideally be collected using the same, or similar, sampling methods as the samples from the site that are being evaluated to ensure comparability of the datasets. At a minimum, composited samples, which are a physical representation of "average" conditions, should not (except under select conditions) be directly compared to discrete sample data. If the same sampling methods cannot be used, a comparison of the methods should be conducted to ensure that substantively different results will not be obtained. For example, the comparison of data collected using ISM to data collected using a discrete sampling methodology should be done with an understanding of the potential error in the mean based on the ISM result (ITRC 2020)[431]. Note that comparing ISM and discrete datasets may be possible, but this should be done by a qualified statistician.

9.4.1 Discrete samples

Discrete soil sampling is "the process of collecting a single soil sample from a specific location and depth interval" (EPA 2013)[152].

9.4.1.1 *Advantages*

- Discrete soil sampling allows for less labor-intensive collection of representative data for volatile organic compounds than other sampling methods that involve some form of sample compositing.

- This method can be used to understand the distribution of background concentrations and to aid in identifying localized areas of elevated concentrations if microscale heterogeneity is sufficiently managed (information may be masked with composite samples if they represent a larger area and ISM samples if the decision units selected are too large).

- Data can be used to calculate representative point BTVs (a measure of the upper threshold of "point" background concentrations). In this context, "point" implies a discrete sample, as opposed to an area-average composite or ISM sample.

- Discrete soil samples allow the examination of data via geochemical evaluation (Section 5).

- Discrete samples are relatively easy to collect and do not require any specialized equipment or sampling skills.

9.4.1.2 *Disadvantages*

- Discrete soil sampling can bring potential increased effort and cost, as a larger number (relative to composite or ISM) of discrete samples may be needed to adequately characterize an area. This especially true in the case where soils are very heterogeneous.
- Discrete samples may represent a small mass compared to area or volume of the target population being analyzed, so the sample may miss contamination or may not be representative (ITRC 2012)[148]. The DQO process should be followed to ensure that discrete samples are representative.

9.4.2 Composite samples

A composite soil sample is a sample composed of several smaller subsamples that are physically mixed to create a single homogenous sample. The subsamples must be the same volume. The composite sample should be representative of the entire composite area or volume.

9.4.2.1 *Advantages*

- Composite soil sampling can produce an estimate of the mean, with fewer analyses and lower cost compared to discrete sampling (USEPA 2002)[153].
- This method can reduce errors due to soil heterogeneity, because a single composite sample can be more representative of a defined area than a single discrete sample—when the goal is to characterize the mean (ITRC 2020)[431]. Note that not all composite samples are equally effective in their ability to reduce variability.

9.4.2.2 *Disadvantages*

- Composite samples can be prepared for analyzing volatile compounds. To minimize contaminant losses from volatilization when preparing the composite sample, methodologies must be used that are more time-consuming than those used for discrete samples.
- Composite sampling can introduce additional error related to weighing than with discrete samples (more weights are being taken, so there is more error). Soils that are mostly clay are difficult to homogenize and tend to be poor candidates for composite sampling.
- Composite sampling yields a reduced amount of information on variability, and information on spatial trends can be masked and diluted.
- Summary statistics from discrete sample results (for example, individual site measurements from historical sampling programs) are not directly comparable to summary statistics from composite sample results; comparison is only possible using statistical methods. While discrete samples are analogous to point values, composite and ISM measurements are representative of area averages.

- Information on the upper tail of the distribution may be lost; elevated concentrations might not be identified, because the data represent a physical average. Since composite samples usually represent larger areal extents, their data are typically not used to calculate a BTV or in geochemical evaluations.

9.4.3 Incremental sampling methodology (ISM)

ISM is a structured sampling and processing protocol that reduces data variability and is a superior methodology to provide an estimate of mean contaminant concentrations in a defined volume of soil. ISM provides representative contaminant concentrations in samples from specific soil volumes, defined as decision units or sampling units), by collecting numerous increments of soil that are combined, processed, and subsampled for laboratory analysis (ITRC 2020)[431]. ISM was developed by mining corporations because standard sampling methods often missed ore bodies that were indicative of profitable levels of metallic ore present in rock formations. Since then, these sampling concepts have been applied to agricultural, food, drug, and environmental sampling. In environmental investigations, incremental sampling was originally intended for clearing large tracts of land at former bombing ranges to determine whether more focused sampling was required, but its use has since been expanded (ITRC 2020)[431]. Use of ISM during background studies, as with any sampling design, should be carefully considered and justified.

9.4.3.1 *Advantages*

- ISM yields more consistent and reproducible results when characterizing the mean than that which is obtained by discrete or composite sampling approaches (ITRC 2012)[148].

- ISM reduces data variability and increases sample representativeness when calculating mean values for a specified volume of soil by designing and accounting for soil heterogeneity, so fewer samples are typically needed to obtain the same statistical power (ITRC 2012)[148].

- This method provides less biased and more precise estimates of the mean than discrete sampling plans, which typically have much lower sample density (ITRC 2012)[148].

- ISM is more cost-effective (for shallow soil sampling programs) than moderate- to high-density discrete sampling plans that provide a comparable level of decision quality when the goal is to characterize the mean (ITRC 2012)[148].

- Samples can be collected for analyses of VOC compounds, but refinements to the procedures are required. These may include the use of special bottle ware, not drying or milling samples before laboratory analysis, and direct preservation of soil in increments in methanol or collections of increments by specialized samplers for processing in the lab (ITRC 2020)[149].

9.4.3.2 *Disadvantages*

- Although regulatory acceptance of ISM is growing, some regulatory agencies may still not accept ISM.

- Although the methanol preservation approach for VOC analyses is effective at minimizing the loss of volatile contaminants from a soil sample, methanol preservation can result in lower analytical sensitivity. "The methanol dilution step causes elevated analytical detection limits compared to the direct soil purge-and-trap and low concentration method techniques. Analytical detection limits could be elevated above relevant screening levels for certain targeted contaminants" (ITRC 2020)[431].

- Summary statistics from discrete sample results (for example, individual site measurements from historical sampling programs) are not directly comparable to summary statistics from ISM sample results. A more reliable comparison is obtained when both the site and background datasets are collected and analyzed using the same methodologies (ITRC 2012)[148].

- Information on the upper tail of the distribution is lost because the goal of ISM is to characterize the mean (elevated concentrations might not be identified, depending on the size of the decision unit chosen).

- Because the data represent a physical average, a BTV calculated from the data will likely not be representative of the upper tail of the background data distribution.

- ISM prevents the use of geochemical evaluation, because the compositing of incremental sampling obscures the natural variability in element concentrations and will diminish the ability to distinguish anomalously high elemental ratios relative to background elemental ratios. Note that geochemical evaluation data could be collected by concurrently taking discrete samples during an ISM sampling program.

9.5 **Sampling Design**

There are two main types of sampling designs: judgmental (also known as targeted biased sampling) and probability-based (or statistical). Probability-based sampling design includes simple random sampling, systematic sampling/grid sampling, and stratified systematic sampling.

The selection of (and rationale for) the sampling design should be determined in accordance with the project's DQOs, as outlined in the SAP. A robust sampling design is required to develop representative and defensible soil background concentrations. For example, DQOs could include determining the presence/absence of a given chemical; determining soil background levels; and evaluating the human health risks associated with site-related chemicals (for example, is the incremental risk greater than background).

An overview of various sampling designs, as well as advantages and disadvantages of each, is available in the USEPA Guidance for Choosing a Sampling Design for Environmental Data Collection (USEPA 2002)[154].

9.5.1 Judgmental (targeted or biased) sampling design

With judgmental sampling design, the number of samples and locations are selected at the discretion of a qualified person, based on their judgment/expertise. Judgmental sampling plans do not allow for a full characterization of uncertainty. Statistical analysis of data collected using judgmental sampling design cannot be used to make any type of scientifically defensible probabilistic statements about the target population. Conclusions drawn about the data are made solely based on judgment and depend entirely on the validity and accuracy of this judgment (USEPA 2006)[134]. Because judgmental sampling design is based on nonrandom sampling, this results in datasets that can be biased, clustered, and correlated, so conclusions based on the dataset cannot be extrapolated to the whole site (USEPA 2002)[153]. Since judgmental sampling is not amenable to statistical analysis, it is not recommended for establishing a soil background dataset.

9.5.2 Probability-based (statistical) sampling design

With probability-based (or statistical) sampling design, each possible sampling location has a known probability of being selected, and only those sampling locations selected are sampled and used in the development of background concentrations. Advantages of probability-based design include the ability to account for uncertainty in the data, to draw conclusions about the target population, and to properly express uncertainty in these conclusions (USEPA 2006)[134]. Probability-based sampling designs are recommended for background studies because statistical analyses can be applied and bias is reduced.

9.6 Sample Collection Methods

Soil background samples should be collected using the same sampling methodologies as the site data, to the extent reasonable and appropriate, to ensure comparability of the datasets.

Considerations when determining the method for soil sample collection include:

- target depth for the soil samples (surface/shallow versus deep)
- whether discrete, composite, or ISM samples are required
- COPC (plus any additional analytes necessary for geochemical evaluation or forensic analysis)
- soil conditions
- site conditions for access of equipment (for example, space limitations, surface covering)

Various methods for soil sample collection are available and can be broken into the three main categories as outlined in Table 9-1. Additional information on soil sample collection methods can be found in the following references: ASTM D1452 (ASTM 2009)[139], ASTM D4700-15 (ASTM 2015)[144], ASTM D6151/D6151M-15 (ASTM 2015)[140], ASTM D6169/D6169M-13 (ASTM 2013)[141], ASTM D6282/D6282M-14 (ASTM 2014)[142], ASTM D6286/D6286M-20 (ASTM 2020)[143], ASTM D6914/D6914M-16 (ASTM 2016)[145], (BC Environment 2020)[147], (USEPA 2012)[428], (USEPA 2014)[429] and (ITRC 2020)[431], as well as local, state, and USEPA regional guidance.

9.7 Sample Handling

Once soil samples have been obtained, they should be placed in the clean sample containers provided by the analytical laboratory and appropriate to the parameters being analyzed. During sample collection in the field, precautions must be taken to avoid cross contamination between samples (for example, the use of appropriate field decontamination procedures). Recommended sample containers and sample preservation for soil samples being submitted to the laboratory for different analytes or analyte groups are outlined in the appropriate USEPA reference method (see Section 10); these should be confirmed with the selected laboratory prior to sample collection.

Proper sampling handling and quality controls, as outlined in the QAPP (Section 8.2.1), should be followed to maintain sample integrity.

Table 9-1. Soil sample collection methods

Source: (USEPA 2002)[153].

Description	Equipment Examples	Advantages/When Appropriate to Use	Disadvantages/Precautions
Surface Soil Sampling Collect shallow/surface soil samples.	• Trier • Trowel • Shovel • Soil probe • Hand auger (manual or powered) • Core sampler	• Cost-effective • Rapid data collection • Relatively easy to use • Less intrusive and disruptive	• Limited depth range • Not ideal for rocky, dense, or hard soil conditions • Depending on method, soil type, and depth required, discrete soil samples can be difficult to collect due to sloughing
Test Pitting Collect shallow to intermediate soil samples. Samples are collected from the bucket of the equipment.	• Rubber-tire backhoe • Tracked excavator	• Cost-effective • Depth range up to 3–4 m with backhoe and 5–6 m with excavator • Provides good visualization of soil samples • Large sample volumes collected	• Disruptive, so not ideal for paved or developed sites • Difficult to collect discrete soil samples in unstable soil conditions (for example, sand) • More uncertainty in depth interval sample is being collected from than when drilling methods are used

Description	Equipment Examples	Advantages/When Appropriate to Use	Disadvantages/Precautions
Drilling Collect shallow to deep soil samples. Samples can be collected in all types of geologic material.	• Auger drilling (for example, solid stem augers and hollow stem augers) • Core sampling devices (for example, split spoons or Shelby tubes) • Direct push • Sonic drilling • Percussion drilling (for example, cable tool)	• Truck or tracked-mount rigs allow access to variety of sites • Collect samples at depths >100 m • Collect samples in any geologic conditions • Some methods allow for collection of discrete, undisturbed samples • Continuous, undisturbed samples can be collected	• Higher cost • Height and space requirements for rig can limit use • Waste (soil cuttings) is usually derived • Solid stem augers not preferred, as confidence in soil sample depth is lower • Use of drilling muds or fluids can introduce cross contamination • Sample volume potentially limited depending on method and soil conditions • Sample area per location is limited based on the diameter of auger/sampling device used (most common for hollow stem augers is 6–31 cm, inside diameter)

10 ANALYTICAL METHODS

10.1 Introduction

A critical component in establishing soil background, whether it be default or site-specific, is to ensure that the soil samples are analyzed by laboratory methodologies that generate high-quality analytical data that meet the data quality objectives (DQOs) of the soil background study and are comparable to the site data being evaluated. Soil sample concentrations reported by the laboratory can be influenced by the soil sample collection and preservation methods, laboratory sample preparation methods (this includes soil sample preprocessing, digestion, or extraction), and laboratory analytical methods used.

When using data from an existing study to establish soil background, laboratory sample preparation and analytical test methods that were used in the existing study should be evaluated to ensure that they provide substantively equivalent results to the laboratory method used at the investigative site(s) being evaluated. Different jurisdictions may have various definitions of what “substantively equivalent” means. However, it generally means that the two test methods being compared give results for the contaminants being analyzed in certified or standard reference materials that have a small allowable bias between the reported results for the two test methods. The magnitude of the allowable bias can vary by test method and jurisdiction.

Data generated using different laboratory methods may (or may not) be comparable. In cases where there is a need to use data analyzed using different laboratory methods, it is important to evaluate the potential difference between the results generated by the two methods, clearly understand the uncertainties involved, and consider this in risk assessment results and risk management decisions.

When conducting a study to establish soil background, to ensure comparability, it is important that soil samples collected from the area considered background and the area being evaluated are collected and preserved using the same techniques and analyzed using the same sample preparation and analytical test methods (or sample preparation and test methods that provide substantively equivalent results) for each analyte or analyte group. If practical, when conducting a site-specific background evaluation, site and background samples should be analyzed by the same laboratory (and if at all possible, in the same analytical batches) to reduce the potential for test method bias between the site and background datasets.

In some cases, the background dataset is compiled from multiple existing soil background studies. In these cases, the different source datasets should be examined to determine whether they were generated using sample collection, preservation, preparation, and analytical methods that provide substantively equivalent results. Datasets generated from sample collection, preservation, preparation, or analytical test methods (or a combination of any of these factors) that provide substantively different results should not be included in a compiled background dataset. Even if the sample collection, preservation, preparation, or analytical test methods used for different datasets provide substantively similar results, a compiled background dataset from multiple sources should not be created until such a grouping is demonstrated to be technically acceptable using statistical methods discussed in Section 11 and geochemical evaluation methods discussed in Section 5.

10.2 **Obtaining Reliable Analytical Data**

10.2.1 Data quality

Choosing the laboratory methods to be used in a soil background study is part of the USEPA DQO process, which is discussed in Section 8.2.

A soil background study, whether it is an existing study or one that will be conducted, should have a quality assurance plan. USEPA recommends having a quality assurance project plan (QAPP). This plan will specify the laboratory sample preparation method(s) and analytical method to be used for every analyte or analyte group. The quality assurance plan will also specify DQOs, such as measurement performance criteria (for example, various acceptable bias, precision, and analytical limit criteria) for every test method. Typically, the completeness (the number of analyses meeting all measurement performance criteria) for each analytical parameter and the entire analytical program are DQOs specified in the quality assurance plan. A complete list of laboratory DQO elements is provided by USEPA (USEPA 2002)[133].

10.2.2 Test method bias and precision

For laboratory test methods, ASTM International provides definitions for bias and precision (ASTM E177 (ASTM 2019)[127], ASTM E456-13A(2017)e4 (ASTM 2017)[128]). Test method bias is "the difference between the expectation of the test results and an accepted reference value" ((ASTM 2017)[128]), while test method precision is defined as a measure of "the closeness of agreement between independent test results obtained under stipulated conditions."

The quality assurance plan will contain performance criteria that evaluate test method bias and precision for data generated by every test method. A number of QC samples will be evaluated for data quality indicators such as bias (for example, method blanks, matrix spikes, and laboratory control samples) and precision (for example, laboratory duplicates, matrix spike duplicates, and laboratory control sample duplicates). Surrogates will be used to indicate the potential bias in the analysis of individual samples for organic analytes. Results for these data quality indicators will be evaluated during the data validation stage (Section 10.2.4).

10.2.3 Laboratory quality system and analyte group accreditation

To ensure that high-quality analytical data with low test method bias and high test method precision are being generated, samples should be analyzed by laboratories whose quality systems have been accredited. Examples of items that are included in laboratory quality systems are their processes for calibration, calibration verification, and laboratory quality control.

Environmental laboratories are generally accredited to ensure that their quality system meets NELAC (National Environmental Laboratory Accreditation Conference) requirements (NELAC 2016)[8]. NELAC requirements meet the quality system requirements for ISO Standard 17025 (ISO 2017)[6], in addition to country-specific requirements for the United States. Laboratories are accredited by accreditation bodies (either state regulatory agencies or select nongovernmental organizations) that perform on-site quality system assessments. Note that a state or stakeholder agency (for example, the Department of Defense) may have their own laboratory accreditation programs and/or requirements.

An accredited quality system means that the system has met certain minimum standards; it does not ensure that for individual test methods, the data reported by the laboratory have minimal test method bias. However, an accredited quality system makes this outcome more likely. It should be noted that as part of many accredited quality systems, the laboratory provides an estimate of the measurement uncertainty for every analytical test method in the laboratory.

Proficiency testing programs are offered to the laboratory by the accreditation agency and involve the regular analysis of samples with an unknown concentration. The laboratory must participate in the proficiency testing program for any analyte groups for which it wishes to be accredited. If results for a proficiency testing sample are outside the acceptance limits (that is, the test method bias is unacceptably high or low), the laboratory can lose its accreditation for that analyte group. For soil background studies, the laboratory should be currently accredited (if accreditation is offered) for any analyte groups that are anticipated to be sent to the laboratory for analysis. Analysis should not be performed at a laboratory if the accreditation has been suspended for a specific analyte group, to avoid the generation of potentially suspect data.

10.2.4 Data validation

Data reported by an analytical laboratory need to be independently reviewed to assess whether the data are fit for the intended purpose (such as the comparison of background and site soil analytical data). Often, this review is performed using the data validation process. This process is an assessment of data quality, tracing the history of the sample from collection through sample storage, sample preparation, instrumental analysis, and data reduction; this process ensures that the resulting data from the sample's analysis are accurate, traceable, and appropriately qualified if any data quality issues were discovered in the validation process.

Validation is a structured, documented review of the data. A qualified analytical chemistry data validation expert who is independent of the laboratory generating the data should perform this review, typically using USEPA data validation methodologies ((USEPA 2016)[167], (USEPA 2017)[164], (USEPA 2017)[165]). Note that a state or stakeholder agency (for example, the Department of Defense) may have their own data validation requirements.

10.3 **Analytical Limits**

Environmental datasets often contain "nondetect" and/or "estimated" results, based on the limited sensitivity of laboratory methods used to measure contaminant concentrations (note that estimated results can often be based on other factors besides the limited sensitivity of laboratory methods). This sensitivity can be described using two general types of thresholds (depicted in Figure 10-1)—the detection limit (DL) and the quantitation limit (QL). The reporting limit (RL), which is a surrogate for the QL, is also depicted in this figure. A review discussing detection, quantitation, and reporting limits in straightforward language has been published online by the American Industrial Hygiene Association (Brisson and Popp 2017)[4]. USEPA has also published an overview (USEPA 2006)[161]. A much more technically detailed review of analytical limits was the subject of an Advisory Committee on Water Information webinar (van Buuren 2017)[9].

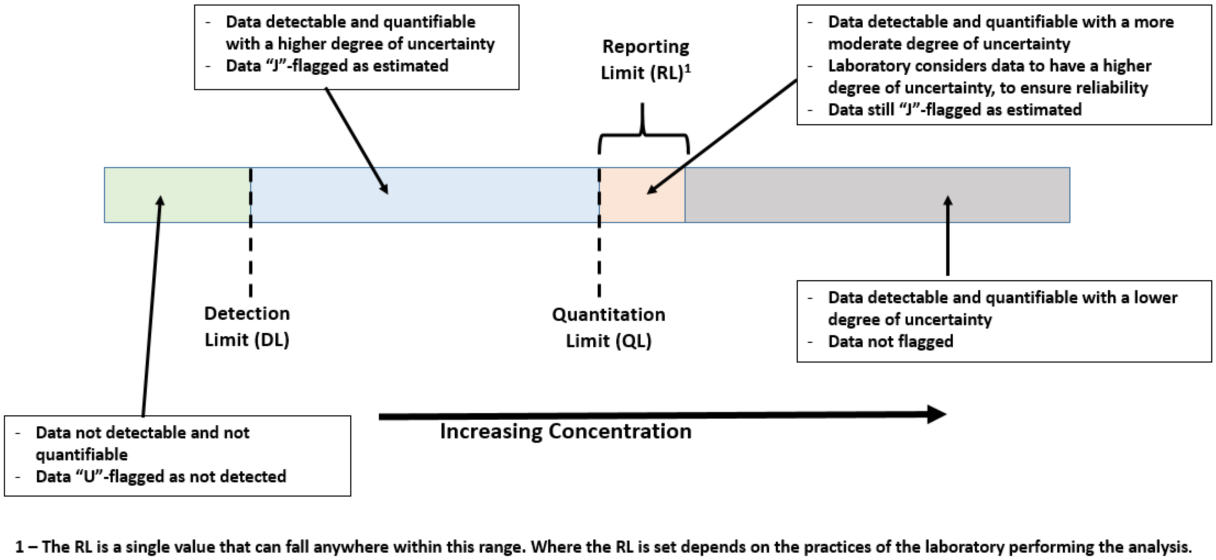


Figure 10-1. Relationship between various analytical limits.

Source: Doug Blue (ExxonMobil) and Shahrokh Rouhani (NewFields).

Briefly, these analytical limits are described in more detail:

- Results that fall below the DL (termed nondetects) are indistinguishable from blank results.
    - Nondetects are a form of censored data, referred to as left-censored data, because they are always reported as being less than the DL.
    - Data that fall below the relevant DL are flagged (or "qualified") by the analytical laboratory with a "U" code, which allows data users to identify such measurements.
    - Most laboratory Certificates of Analysis will report the method detection limit (MDL) for a sample. Alternatively, the limit of detection (LOD) may be reported.
    - USEPA does specify a methodology to set the MDL (this is detailed below).
    - MDLs are laboratory-specific and instrument-specific; they can vary between different analytical instruments (even for the same manufacturer and model number) performing the same method in the same laboratory. However, many commercial laboratories use data pooled from all similar instruments, so MDLs are consistent within that instrument group.

  - The treatment of results less than the MDL in statistical analysis of the data is discussed in detail in Section 11.3.

- Results that fall between the DL and QL are detected but are quantified with a higher degree of uncertainty.
  - Values between the DL and QL are considered to be estimated.
  - Data that fall between the relevant DL and QL are flagged (or "qualified") by the analytical laboratory with a "J" code, which allows data users to identify such measurements.
  - Most laboratory Certificates of Analysis do not report the QL for a sample; instead, an RL is reported.
  - USEPA does not specify a methodology to set the QL.
  - QLs are laboratory-specific and instrument-specific; the QL can vary between different analytical instruments (even for the same manufacturer and model number) performing exactly the same the method in the same laboratory. However, many commercial laboratories use data pooled from all similar instruments, so QLs are consistent within that instrument group.
  - RLs are sometimes used by laboratories as a surrogate for the QL (see discussion below).

The definition of the MDL has evolved over time (USEPA 2007)[162] and the definition and calculation methodology have recently changed. The MDL is now defined as "the minimum measured concentration that can be reported with 99% confidence that the measured concentration is distinguishable from method blank results" (USEPA 2016)[168]. The MDL accounts for aspects of measurement such as the instrumentation, sample preparation, matrix effects, and laboratory reagents. The limit of detection (LOD) is the measure of an analytical method to detect the presence of an analyte with a 99% level of confidence; it does not provide information on the quantification of an analyte. The LOD is a concept similar to the MDL, in that it is a measure of the DL, although it is calculated differently than the MDL.

Instead of a QL, a laboratory will sometimes report an RL on their Certificates of Analysis. Like a QL, the RL is the lowest concentration of analyte in a sample that can be reported with a defined degree of analytical test method bias and precision. The USEPA does not specify a methodology to set the RL and there are multiple definitions of the RL (van Buuren 2017)[9]. Often, the laboratory sets the RL value at the QL, plus an added margin of safety to account for variations in the test method that may occur over time, or variation in instrument performance. However, in some cases, the laboratory will set the RL equal to the QL, which is why the RL is shown to have a range of possible values in Figure 10-1. For example, consider a laboratory that has three instruments (same model and manufacturer) performing metals analysis for soil samples. For copper, the QLs for the three instruments are determined to be 3.7 mg/kg, 3.8

mg/kg, and 4.1 mg/kg. In this case, the laboratory may set the RL for copper to 5 mg/kg to avoid having to report different QLs for different instruments on the Certificate of Analysis.

For soil background studies, it is important that the analytical methods used meet the project DQOs. The RLs of the analytical methods used must meet the project DQOs, so they will be low enough to detect and quantify the analytes of interest, as well as minimizing (to the extent possible) left-censoring of the data. For many contaminants (especially organic contaminants), background concentrations may be very low. To minimize the number of data points less than the RL in the background dataset for these contaminants, it is often required to use "low-level" test methods (the test method that provides the lowest reporting limits for the analytes).

Typically, reported results greater than the DL are used "as is" (even if results are estimated) when calculating soil background values. There are several recommended procedures for treatment of data less than the DL when calculating soil background values (see Section 11.3 for further discussion).

Accredited laboratories should ensure sample preparation and analytical methods can generate the appropriate MDL and QL or RL, as per the project-specific DQOs. Early coordination with the selected laboratory can ensure that project objectives are met; the laboratory should review the project quality assurance plan to ensure accuracy and achievability before samples are submitted for analysis. Note that a state or stakeholder agency (for example, the Department of Defense) may have their own detection limit requirements.

10.4 Sample Preparation

Sample preparation broadly covers the procedures performed on the soil sample from the time of its receipt by the laboratory up until instrumental analysis is performed. Organic analytes are typically extracted from the soil using an organic solvent and it is the extract that is instrumentally analyzed. For metals, the soil sample is digested using acids and it is this digestate that is instrumentally analyzed. However, soil samples typically require preprocessing before extraction or digestion; this preprocessing is also considered part of sample preparation. This preprocessing can include steps such as sample drying (air, oven, or chemical), disaggregation, sieving and milling, or pulverizing, as well as subsampling of the preprocessed soil to obtain an aliquot for digestion or extraction.

Typically, the largest variability in the reported results is due to the sample preparation methods used for the soil sample, not the analytical method used to obtain the reported result. Different sample preparation methods can produce very different results for the same sample, so results may not be comparable if different sample preparation methods are used (this may also be true if different analytical methods are used). For that reason, common sample preparation methods for metals and various organic analyte groups are discussed in Table 10-1; the table gives a brief synopsis of the sample preparation method (including details on sample preprocessing and extraction or digestion details) and also discusses whether the sample preparation method is suitable for use in soil background studies. For a more detailed description of the sample preparation methods, the reference methods identified in Table 10-1 should be consulted.

Note that many USEPA SW-846 methods are referenced in Table 10-1 and Table 10-2. The USEPA periodically updates methods and changes the letter after the numerical method designation to reflect this revision. In general, use of the newest revision of each method is recommended, though there might be project-specific reasons to use an older version of a method. To facilitate the application of this guidance over time, the revision letters have been omitted from the method numbers in these tables. However, project-specific documents should always clearly identify the applicable revision number being used.

Soil sample preparation considerations when compiling soil background datasets include:

- For metals, soil sample preparation differs, depending on whether the goal is to determine the total metal concentrations in the sample, or just the environmentally available concentration of these metals. Sample digestion for total metals typically involves the use of hydrofluoric acid to more fully dissolve the aluminosilicate soil matrix and liberate more metals. It is harsher than the sample digestion procedures for environmentally available metals. Thus, greater soil metals concentrations will be obtained if the same sample undergoes total metals digestion versus digestion for environmentally available metals. For risk assessment purposes (see further discussion in Section 4), it is the environmentally available concentration of metals that should be quantified, not the total concentration.
    - For example, the USGS performed a low sample density study (one sample per 1,600 km^2) to determine elemental soil background values across the conterminous United States (Smith et al. 2014)[213]. USGS used a total metals digestion of the <0.15 mm fraction using four acids (including hydrofluoric acid) as the digestion method, which yields higher concentrations for metals than the less aggressive USEPA digestion methods used to determine environmentally available concentrations. Metals data from USGS studies (where total metals are quantitated) cannot be directly compared to data generated for environmentally available metals using USEPA methods, so they should not be used in soil background studies (Brooks 2020)[160].
    - For metals analysis, sample digestion targeting just environmentally available metals (USEPA Method 3050 or USEPA Method 3051) can give reported concentrations up to an order of magnitude less than when a more aggressive total metals digestion method (USEPA Method 3052) is used for the same sample (Ames and Prych 1995)[1]. Soil samples digested using the USGS methodology will give even higher metal concentrations results than USEPA Method 3052, since it uses a more aggressive acid digestion and analyzes either the <2 mm or <0.15 mm soil fraction (see discussion in next bullet).
    - Critically, different sample preprocessing and/or digestion methods (but using the same analytical instrumentation) can give a much larger difference in the reported metals results for a soil sample versus if the same soil sample underwent the same sample preparation procedures but was analyzed using inductively coupled plasma (ICP)/mass spectrometry (MS) rather than ICP/atomic emission spectroscopy (AES) instrumentation.

- Sample preprocessing can affect the reported concentrations of environmentally available metals. USEPA digestion methods do not specify what preprocessing is to be performed, so different laboratories use different options. Some of these options (and their influence on the reported soil metals concentrations) are discussed in detail below (note that this list does not include all possible preprocessing options):

    - Option 1—a wet soil sample typically just has the largest stones manually picked out of the sample and the sample is digested. The metals concentrations for these samples are calculated on a dry weight basis using results of the moisture analysis of that sample. With no sample preprocessing that affects the particle size distribution of the soil sample being digested for analysis, this option will provide the lowest environmentally available metals concentrations for the soil sample.

    - Option 2—a wet soil typically has the largest stones manually picked out of the sample. The soil sample is air dried (or dried in an oven at 30°C) and disaggregated (broken apart), and dry sample is digested. With no sample preprocessing that affects the particle size distribution of the same sample being digested for analysis, this option will provide environmentally available metals concentrations for the soil sample similar to Option 1.

    - Option 3—a wet soil has the largest stones manually picked out of the sample. The soil sample is air dried (or dried in an oven at 30°C) and disaggregated. The disaggregated sample is passed through a 10-mesh (<2 mm) sieve and the fraction passing through the sieve is digested. This sample preprocessing results in soil with smaller particle sizes and higher surface area being digested for analysis. This option will provide environmentally available metals concentrations for the soil sample directionally higher than Options 1 and 2.

    - Option 4—a wet soil has the largest stones manually picked out of the sample. The soil sample is air dried (or dried in an oven at 30°C) and disaggregated. The sample is passed through a 10-mesh (<2 mm) sieve, with the fraction passing through this sieve then further milled/pulverized. After milling, the sample is passed through a 100-mesh (<0.15 mm) sieve and the fraction passing this sieve is digested. This sample preprocessing results in soil with even smaller particle sizes and higher surface area being digested for analysis than for Option 3. This option will provide environmentally available metals concentrations for the soil sample directionally higher than Option 3.

    - Soil sample preprocessing methodologies used prior to digestion are a key factor to consider when determining whether the metals results from two datasets are substantively the same or not.

    - Sample preprocessing methods should be tailored to fit the intended use of the analytical data. For example, pulverizing of soil is generally not appropriate when the dermal exposure pathway is being evaluated.

- For organic contaminants, sample preparation involves the extraction of the target analytes from the soil sample. The concentration of analyte obtained from analysis can vary widely, depending on the solvent chosen to extract the analytes from the soil matrix. Generally, organic analyte data from datasets generated using two different extraction solvents should not be combined in the same background dataset.
    - In addition to extraction solvent, if different methods are used to clean up the solvent extract before analysis, this can influence the reported results. The magnitude of this influence is typically less than the size of the effect observed from using different extraction solvents.

- Incremental sampling methodology (ISM) can be applicable for background studies. The ITRC ISM-2 guidance ((ITRC 2012)[148], (ITRC 2020)[431]) describes the sample preprocessing options of air drying, disaggregation, sieving, milling, and two-dimensional slabcake subsampling. See Section 5 of the ISM-2 guidance for details on using project objectives to select among the sample preprocessing options if ISM is being used and how to implement these options at the laboratory.

- To completely understand the differences in sample preparation (especially sample preprocessing before digestion or extraction) between the laboratories used in two studies, just comparing the reference methods used will not suffice, since there is often some flexibility provided in the reference method. It is recommended that the standard operating procedures used by both laboratories be examined in detail to see if sample preparation methods could result in substantive differences in the reported results between the two studies. Standard operating procedures are typically appended to the QAPP (if the QAPP is available for review) or can be provided (upon request) by the laboratory that performed the sample analysis.

Table 10-1. Sample preparation

Sources: (USEPA 2020)[166] and (Taggart 2002)[463].

Chemical	Reference Method	Summary	Comments
Metals	USEPA Method 3050 (Heating Block Digestion)	Soil is preprocessed using a number of options (see the text of Section 10.4 for a full discussion). The preprocessed soil is digested at 90–95°C on a hot plate or heating block. Digestion uses nitric acid, hydrogen peroxide, and typically hydrochloric acid (HCl always used for ICP/AES and can be used with some ICP/MS systems).	**Suitable for soil background studies**. Will dissolve all environmentally available metals, but not aluminosilicate-bound metals that are not environmentally available.
	USEPA Method 3051 (Microwave Digestion)	Mimics USEPA 3050B, except it uses microwave heating of sample and hydrogen peroxide is not used.	**Suitable for soil background studies**. Same comments as for 3050; provides similar results as 3050. Has shorter digestion times than 3050 and higher precision (better temperature control, versus heating block).
	USEPA Method 3052 (Total Digestion)	Similar to USEPA 3051, except the sample is microwave-digested at 180±5°C, using a mixture of nitric and hydrofluoric acids. Goal is the total decomposition of the sample, including all aluminosilicate and organic matrices. Analysis of the digestate yields a total metals value.	**Not suitable for soil background studies**. Dissolves all metals, including silicate-bound metals that are not environmentally available. Provides soil metals results biased high compared to USEPA Method 3050 and USEPA Method 3051.

Chemical	Reference Method	Summary	Comments
	USGS Q030, T01 & T20 (Total Digestion)	Soil is air dried; fraction passing through a 10-mesh sieve (the <2 mm fraction) is retained for analysis. Optionally, the <2 mm fraction can be further pulverized and only the fraction passing through a 100-mesh sieve (<0.15 mm fraction) is retained for analysis. The sample is digested with four acids (hydrochloric, nitric, perchloric, and hydrofluoric) on a hot plate or heating block for several hours at temperatures up to 160°C. The goal is the total decomposition of the sample, including all aluminosilicate and organic matrices, to yield a total metals value.	**Not suitable for soil background studies**. Dissolves all metals, including silicate-bound metals that are not environmentally available. Provides soil metals results biased high compared to USEPA digestion methods for environmentally available metals (USEPA Method 3050 and USEPA Method 3051).
Mercury	USEPA Method 7471 (Mercury Digestion)	Sample digested in 3:1 $HCl:HNO_3$ (aqua regia); oxidized with potassium permanganate.	**Suitable for soil background studies**. Will dissolve all environmentally available mercury.
OCP, PAH, PCB, PCDD/F TPH	USEPA Method 3540 (Soxhlet Extraction)	A soil sample is chemically dried with anhydrous sodium sulfate, placed in a thimble, and extracted using the appropriate solvent in a Soxhlet extractor. If necessary, the extract is further processed (for example, dried, concentrated, and cleaned up) before analysis.	**Suitable for soil background studies**. Rigorous and rugged reference method to which all other methods are compared. Uses large solvent volume and long extraction time (16–24 hours). Preparation methodologies with shorter extraction times are more typically used and provide substantively equivalent results that are suitable for background studies.
	USEPA Method 3541 (Automated Soxhlet Extraction)	Soil sample dried with anhydrous sodium sulfate is immersed in boiling solvent, then Soxhlet extracted (similar to USEPA 3540) and finally concentrated. If necessary, the extract is further processed (for example, dried, concentrated, and/or cleaned up) before analysis.	**Suitable for soil background studies**. Uses shorter extraction times (~2 hours) and smaller solvent volume than USEPA Method 3540, while still giving analyte recoveries similar to that method.

Chemical	Reference Method	Summary	Comments
	USEPA Method 3545 (Pressurized Fluid Extraction)	Chemically dried soil sample is placed in extraction cell and heated (temperature and time depend on analyte). Sample is pressurized (1,500–2,000 psi) with appropriate solvent. Multiple extraction cycles used for some analytes. If necessary, the extract is further processed (for example, dried, concentrated, and/or cleaned up) before analysis.	**Suitable for soil background studies**. Typical extraction cycle is 5–10 minutes. Small solvent volumes used in extraction.
	USEPA Method 3546 (Microwave Extraction)	Chemically dried soil sample is placed in extraction cell with the appropriate solvent and heated via microwave (temperature and pressure depend on the analyte). One extraction cycle used. The extract is filtered to remove solids. If necessary, the extract is further processed (for example, dried, concentrated, and/or cleaned up) before analysis.	**Suitable for soil background studies**. Uses shorter extraction times (10–20 minutes) and smaller solvent volume than USEPA Method 3540, while still giving analyte recoveries similar to that method.
OCP, PAH, PCB, PCDD/F TPH	USEPA Method 3550 (Ultrasonic Extraction)	Chemically dried soil sample is placed in an ultrasonic cell with the appropriate solvent and the mixture is extracted with solvent three times. The extract is separated from the soils by filtration or centrifugation. If necessary, the extract is further processed (for example, dried, concentrated, and/or cleaned up) before analysis.	**Not suitable for soil background studies**. Uses shorter extraction times than USEPA 3540, but still uses relatively large solvent volumes. Method states it “might not be as rigorous as other extraction methods for soils” and that recoveries for some analytes are low. Not recommended for environmental soil background studies, due to potential for low recoveries for some analytes.
VOCs	USEPA Method 5035 (Purge and Trap)	The methanol preservation option, with a 1 g soil to 1 mL methanol ratio in a preweighed vial, is recommended for best analyte recovery. An aliquot of the extract is diluted in reagent water and purged onto the trap of the analytical instrument. The low concentration options using aqueous preservation solutions can produce low biased results when the VOCs are strongly bound to the soil particles.	**Suitable for soil background studies**. Reporting limits are typically higher for methanol-preserved samples than the low-level options. The methanol option is preferable for soil background studies, if the analysis is performed with enhanced mass spectrometer sensitivity (for example, using selected ion monitoring) to compensate for the required dilution of the methanol extract.

Notes:
OCP—organochlorine pesticides
PAH—polycyclic aromatic hydrocarbons
PCB—polychlorinated biphenyls
PCDD/DF—polychlorinated dibenzo-p-dioxins/dibenzofurans
TPH—total petroleum hydrocarbons
VOC—volatile organic compounds

10.5 Analytical Test Methods

For common metals and various organic analyte groups for contaminants of concern at sites, the analytical test methods used are discussed in Table 10-2; the table gives a brief synopsis of the test method and discusses whether the analytical method is suitable for use in soil background studies. For a more detailed description of the analytical method, the reference method identified in Table 10-2 should be consulted.

In risk assessments (including contaminant fate and transport modeling), analysis of soil samples for various other physical and chemical properties is useful, so collecting this data in a background study may be warranted. These parameters include (but are not limited to) grain-size distribution, pH, total organic carbon, and cation exchange capacity (Section 9.1.3). Discussion of the sample preparation and analytical test methods for these parameters is beyond the scope of this document.

Table 10-2. Analytical methods

Sources: (USEPA 2020)[166], (Taggart 2002)[463].

Chemical	Analytical Method(s)	Summary	Comments
Metals	USEPA Method 6010 USGS T01 (ICP/AES)	A digested sample is nebulized into an ICP, where the metal atoms are ionized. The metal ions are quantitated using AES.	**Suitable for soil background studies if RLs are low enough**. ICP/AES analysis is marginally less expensive than ICP/MS but has elevated RLs for some metals (for example, silver, thallium, and mercury).
	USEPA Method 6020 USGS T20 (ICP/MS)	A digested sample is nebulized into an inductively coupled plasma (ICP), where the metal atoms are ionized. The metal ions are quantitated using mass spectrometry (MS).	**Suitable for soil background studies.** ICP/MS typically has lower RLs than ICP/AES, so use of ICP/MS is preferred for soil background studies (to lower the nondetect frequency for some trace metals).
Mercury	USEPA Method 7471 (CVAA)	A digested sample is chemically reduced, converting divalent mercury to elemental mercury, which is aerated to vaporize the mercury. The cold vapor (CV) passes through an atomic absorption (AA) spectrometer, where the mercury is quantitated.	**Suitable for soil background studies.** Instrumentation is typically more readily available than ICP/MS.

Chemical	Analytical Method(s)	Summary	Comments
OCP	USEPA Method 8081 (GC/ECD)	Sample is extracted using any of USEPA Methods 3540, 3541, 3545, or 3546. Extraction solvents typically used are 1:1 hexane/acetone or 1:1 methylene chloride/acetone. Extracts are cleaned up [for example, alumina (USEPA Method 3610), Florisil (USEPA Method 3620), or silica gel (USEPA (3630)]. After cleanup, the extract is analyzed by injecting into a capillary gas chromatograph, equipped with an electron capture detector (GC/ECD).	**Suitable for soil background studies.** Currently, GC/ECD use is preferred over GC/MS for soil background studies because of lower RLs. With GC/ECD, careful evaluation of low-level detections is recommended because of the potential for false positives.
PAH	USEPA Method 8270 (GC/MS) (Full Scan or Selected Ion Monitoring (SIM) mode)	Sample is extracted using any of USEPA Method 3540, 3541, or 3545. Methylene chloride/acetone is the extraction solvent typically used, with extract cleanup typically not performed. The extract is analyzed by injection into a capillary gas chromatograph, equipped with mass spectrometer detector (GC/MS) operated in either full scan or SIM mode.	**Suitable for soil background studies.** Price difference between full scan and SIM analysis is small. In soil background studies it is preferred to use methods with lower reporting limits (for example, SIM).
PCB	USEPA Method 8082 (GC/ECD)	Sample is extracted using any of USEPA Methods 3540, 3541, 3545, or 3546, using 1:1 hexane/acetone or 1:1 methylene chloride/acetone. Extracts are cleaned up using sulfuric acid/potassium permanganate (USEPA Method 3665). After cleanup, the extract is analyzed by injecting into a capillary gas chromatograph, equipped with an electron capture detector (GC/ECD).	**May be suitable for soil background studies.** Much less expensive than USEPA Method 1668. Only quantitates Aroclors and select congeners. Higher RLs than congener analysis. Suitable for soil background studies, if data for all congeners not needed and RLs meet DQOs.
	USEPA Method 1668 (HRGC/HRMS)	Sample is typically extracted using USEPA Method 3540, using hexane as the extraction solvent. Extracts are cleaned up [for example, Florisil (USEPA Method 3620)]. After cleanup, the extract is analyzed by injection into a high-resolution gas chromatograph, equipped with a high-resolution mass spectrometer detector (HRGC/HRMS).	**Suitable for soil background studies.** Much more expensive than USEPA Method 8082. Can individually detect most congeners at lower levels than Aroclor analysis.

Chemical	Analytical Method(s)	Summary	Comments
PCDD/DF	USEPA Method 8290 (HRGC/HRMS)	Sample is extracted using USEPA Method 3540 or USEPA Method 3545 using toluene as the solvent. Extracts then acid/base washed, dried, and cleaned up using a column containing alumina, silica gel, and activated carbon. After cleanup, the extract is analyzed using a high-resolution gas chromatograph, equipped with a high-resolution mass spectrometer detector (HRGC/HRMS).	**Suitable for soil background studies.** Able to detect PCDD/DF congeners at very low levels (ng/kg levels).
TPH (GRO and DRO)	USEPA Method 8015 (GC/FID)	For DRO, sample is extracted using any of USEPA Methods 3540, 3541, 3545, or 3546, typically using a methylene chloride/acetone solvent that is cleaned up with silica gel (USEPA Method 3630) when removal of polar sample components is appropriate. For GRO, purge & trap (USEPA Method 5035) or static headspace (USEPA Method 5021) are often used. Quantitation is via capillary gas chromatograph, equipped with a flame ionization detector (GC/FID). Gasoline range organics (GRO) quantitates C_6-C_{10}, while diesel range organics (DRO) typically quantitates C_{10}-C_{28}.	**Suitable for soil background studies.** RLs less than USEPA Method 1664 (gravimetric determination of hexane extractable materials) are possible.
VOC	USEPA Method 8260 (GC/MS) (Full Scan or SIM mode)	Purge & trap (USEPA Method 5035) or static headspace (USEPA Method 5021) are often used. The vapor generated from the sample is analyzed using a capillary gas chromatograph, equipped with mass spectrometer (GC/MS) operated in either Full Scan or SIM mode.	**Suitable for soil background studies.** Price difference between full scan and SIM analysis is negligible. Methods with lower RLs, such as SIM, are preferable.

<u>Notes</u>:
DRO—diesel range organics
GRO—gasoline range organics
OCP—organochlorine pesticides
PAH—polycyclic aromatic hydrocarbons
PCB—polychlorinated biphenyls

PCDD/DF—polychlorinated dibenzodioxins/dibenzofurans
TPH—total petroleum hydrocarbons
VOC—volatile organic compounds

Analytical method considerations when compiling soil background datasets include:

- Soil results for organic and inorganic methods are typically reported on a dry weight basis. However, in some studies (especially older studies) results might be reported on a wet weight basis. If results were reported on a wet weight basis, convert these values to a dry weight basis before including them in a background dataset. If it is not possible to perform this conversion (there is no soil moisture analysis available for the sample for which data were reported on a wet weight basis), the wet weight data should not be used in a background dataset containing dry weight values.

- Metals field screening data generated using portable x-ray fluorescence (XRF; USEPA Method 6200) should not be included in the background dataset generated by the laboratory using ICP/AES, ICP/MS, or cold vapor atomic absorption (CVAA) spectroscopy methods. These field XRF results do not have the strict QA/QC used to generate the environmentally available metals data in the laboratory; quantitation is based on a method with entirely different sample preparation. In addition, XRF quantitates total metals and not the environmentally available metals concentrations considered in soil background studies. However, portable XRF data can be useful to field-screen soil samples, to select samples to be submitted for laboratory analysis for metals.

    - Note that laboratory-based XRF instruments do generate data under strict QA/QC protocols. However, because XRF is a measure of the total metals concentration in the soil sample, laboratory data generated from this analytical method should not be included in background datasets for environmentally available metals.

- If a geochemical evaluation to establish soil background is planned, refer to Section 5 for guidance on the necessary metals to include in the analysis. USEPA's target analyte list (TAL) of 23 metals contains all the reference elements (aluminum, iron, manganese, calcium, magnesium, sodium, and potassium) and most of the trace elements (antimony, arsenic, barium, beryllium, cadmium, chromium, cobalt, copper, lead, mercury, nickel, selenium, silver, thallium, vanadium, and zinc) needed to perform a geochemical evaluation. In some cases, additional trace metals (for example, molybdenum) may need to be added to this list, if these elements are COPC in the soil background study.

- Some jurisdictions have total PCB criteria for soils. Total PCB values calculated from the Aroclor data (USEPA Method 8082) will give a different result (typically lower) than the total PCB concentrations calculated using congener data (USEPA Method 1668). When determining background values for total PCB, the mixing of data generated using these two test methods in the background dataset should be avoided.

11 STATISTICS

This section is intended to provide a general overview of the statistics used in establishing soil background values, comparing the site data to a background threshold value, or comparing site data to background data. The appropriate statistical approach will be influenced by the project goals and site-specific needs. Project stakeholders should identify and agree to an appropriate statistical approach when designing the soil background study. Many of the methods and procedures outlined in this section are discussed in greater detail elsewhere (for example, USEPA ProUCL Technical Guide (USEPA 2015)[197]).

11.1 **Data Requirements**

11.1.1 Precautions

Data are essential in any scientific analysis, particularly statistical analysis. However, not all data are created equal and prior to any analysis, statistical or otherwise, the source and quality of data should be examined. Any conclusions are only as strong as the data used in the analysis. Therefore, it is essential to ensure data are not only of suitable quality, but also reflective of site-specific or regional conditions so that a representative background value is determined. Section 9 describes appropriate sampling methods and Section 10 describes appropriate laboratory methods, including data qualification. Further, many data quality issues are mitigated by developing a working conceptual site model (CSM) prior to any field sampling (ITRC 2013)[137]. As part of developing a CSM, DQOs can be developed that inform necessary quantity and quality of data (ITRC 2013)[137]. Developing a CSM and DQOs is discussed further in Section 8.

11.1.2 Appropriate use of statistics

Statistical analysis is a powerful technique that provides quantitative results to inform background and site decisions (ITRC 2013)[137]. Selection of the appropriate statistical analyses depends on the CSM, the study design, the sample design, and the data characteristics. Further, while statistical tests produce quantitative results, these results also have associated uncertainty. It is inherently infeasible to sample an entire volume of soil to obtain the exact background value; therefore, any statistically derived soil background value will also have inherent uncertainty. However, many statistical tests are calculated to a confidence level or interval, typically 95%, which provides some level of certainty. Statistical tests are sensitive to the confidence level or interval, which can produce false positive or false negative errors. Statistical background values, and these types of errors, are discussed in detail in Section 11.6. All statistical tests also have underlying assumptions. Violating these assumptions, or misapplying a statistical test, can produce erroneous and misleading results and degrade the defensibility of any decision at the site.

When applied appropriately, statistical methods provide quantitative results to define soil background concentrations and address project objectives (ITRC 2013)[137]. However, statistics are only one component of any scientific analysis. Statistics should not be used to compensate for low data quantity or poor data quality. Any scientifically defensible decision generally requires multiple lines of evidence in addition to statistical analyses (ITRC 2013)[137]. Additional examples of investigative methods include geochemical evaluation (Section 5) and environmental forensics (Section 7).

11.1.3 Minimum sample size

The appropriate application of statistics requires understanding the underlying assumptions and uncertainties of any statistical test, including the number of required samples. The appropriate sample size depends on the question being asked and the desired level of certainty. The minimum sample size is the number of individual samples required to conduct a specific statistical test or to calculate a specific statistical parameter with an acceptable level of uncertainty. Generally, a minimum sample size of 8–10 samples is required to do this ((USEPA 2009)[198], (USEPA 2015)[195]). This recommendation is applied in many state agency guidance documents ((IDEM 2012)[176], (IDEQ 2014)[177], (KDHE 2010)[179]). However, depending on site characteristics and conditions or desired statistical power, more samples may be required. Some state agencies require a minimum sample size of 20 samples ((KEEC 2004)[180], (MTDEQ 2005)[473]), due to the typical heterogeneity of soils (MTDEQ 2005)[473]).

In general, large datasets will provide better estimates of background concentrations and result in more powerful statistical tests (Cal DTSC 2008)[464]. However, the results of any statistical analysis are only as strong as the data. A strong dataset is a representative sample of the target population. A target population is the entire presence or distribution of a constituent in the soil (for example, arsenic concentrations in a soil type) (ITRC 2013)[137]. Generally, target populations cannot be fully characterized (ITRC 2013)[137], as it is technically or financially infeasible to sample and analyze every segment of soil at a site. Therefore, each soil sample is only a small portion of the entire target population. A representative sample then is a dataset of soil samples that have statistical characteristics congruent with the statistical characteristics of the target population (ITRC 2013)[137].

Therefore, a dataset must come from a representative sample of the target population, collected and analyzed using the appropriate methods discussed in Section 9 and Section 10. While a large dataset is desirable, nonrepresentative data (for example, collected using different sampling techniques or from a different target population) should not be included to arbitrarily increase the dataset size. Namely, a large dataset is not the same as a large dataset suitable for statistical analysis (ITRC 2013)[137].

Ultimately, site conditions or characteristics and the chosen statistical test should inform the necessary number of samples. The statistical tests described in this section identify the minimum number of samples required, if appropriate. It is recommended that the appropriate regulatory agency also be consulted when determining the minimum number of samples ((IDEM 2012)[176], (IDEQ 2014)[177]).

11.1.4 Data selection

In addition to a minimum sample size, many statistical tests also have requirements regarding data selection or sampling. An underlying assumption of all statistics is that the collected measurements, or soil samples, are a random sample of the targeted population that is also free of any bias (IDEQ 2014)[177]. Random sampling is achieved through appropriate sampling techniques that are informed by the CSM. Random sampling does not necessarily imply that the entire soil at the site is sampled randomly, but that each target population is sampled randomly. Any investigative site will possibly contain several soil subgroups (for example, shallow or deep soils). Stratified sampling is the identification and sampling of each subgroup randomly. Section 9 of this document provides more discussion on appropriate sampling methods.

11.1.5 Target population

One key aspect of developing a CSM is determining the target population and whether the site includes one or more populations (ITRC 2013)[137]. A target population is the presence or distribution of a constituent in a soil (for example, arsenic concentrations in a soil type) (ITRC 2013)[137]. Target populations may be specific to a site, different soil groups, or background conditions (natural or anthropogenic ambient). As discussed previously in this section, it is imperative that all data for any statistical analysis be representative of the target population. Analysis of nonrepresentative data could lead to erroneous results. For example, a background value calculated from a dataset containing several unrelated populations will result in an erroneous result. However, the presence of multiple populations does not necessarily indicate a nonrepresentative dataset, as some background scenarios may inherently contain more than one subpopulation (for example, a site with several anthropogenic sources of the same constituent). Regardless, prior to any statistical analysis, the dataset must be evaluated to ensure it is representative of only one target population.

Site characteristics such as geology, soil depth, or geochemistry may indicate differing soil populations. For instance, deep soils may be representative of natural background conditions, while shallow soils would be more representative of anthropogenic ambient background conditions. In addition, graphical displays such as a quantile-quantile plot (Q-Q plot) or histogram can be used to determine the presence of data from several populations (Section 11.4). While these methods can reveal multiple populations, the selection of target population data from multiple populations can be subjective, requiring multiple lines of evidence and professional judgment. For example, in some cases for practical reasons and based on professional judgment, multiple background subpopulations may be combined in a statistical analysis.

Graphical methods (for example, box plots) and statistical methods (for example, Student's *t*-test) can be used to evaluate and confirm the presence of multiple populations. These methods are discussed further in Section 11.4. Geochemical evaluations and forensic methods, discussed in Section 5, Section 6, and Section 7, can also be used to discern multiple populations within the dataset.

11.1.6 Underlying assumptions

Dataset size and representativeness are not the only requirements when applying statistics. Fundamentally, all statistics rely on three assumptions:

- The sampled data represent a random sample of the target population that is free from bias.
- The sampled data are representative of the target population.

- Each sample is random and not influenced by other measurements (each sample is independent of the other samples (IDEQ 2014)[177]).

Further, specific statistical methods may have individual assumptions. For instance, some statistics have underlying assumptions regarding the distribution of the dataset (discussed in detail in Section 11.2). Prior to application of any statistical test, the assumptions of that method need to be understood.

11.2 **Data Distribution**

Data distribution is a descriptive statistic, often represented by a graphed curve, that describes all the values within a dataset and the frequency at which those values occur. Not all data are distributed in the same manner, and categories have been developed to describe common data distributions. Generally, statistical tests have underlying assumptions regarding sampling distribution. A prominent theorem in probability is the Central Limit Theorem, which states that the sampling distribution of the sample mean of a large number of independent random samples from a population is normal regardless of the distribution of the sampled population. Many common statistical methods rely on this theorem, and statistical tests that rely on distributional assumptions are referred to as parametric tests. However, in practice, environmental datasets may not have data distributions suitable for analysis using parametric methods. Therefore, prior to applying any statistical test, the data distribution must be understood. Applying a statistical test to a dataset that does not meet its distributional assumption will possibly produce erroneous and/or indefensible results (ITRC 2013)[137].

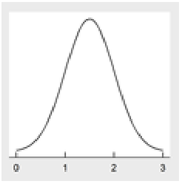
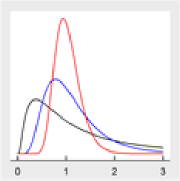
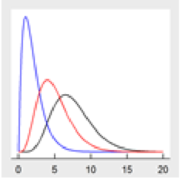
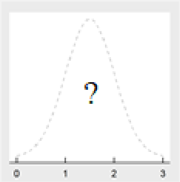
The most recognized distribution is the normal distribution (Table 11-1). The mathematical model of a normal distribution produces a smooth and symmetrical bell-shaped curve (ITRC 2013)[137]. The shape of the curve (or bell-shape) is determined by the mean and standard deviation of the dataset (ITRC 2013)[137]. The shape of the curve is directly related to the standard deviation; a small standard deviation produces a narrow distribution, while a larger standard deviation produces a wider distribution (ITRC 2013)[137].

While a normal distribution is desirable for applying parametric statistics, many datasets have distributions that are not normal. Some common factors producing non-normal distributions include outliers and spatial or temporal changes (ITRC 2013)[137]. For example, natural background soil concentrations are often distributed asymmetrically (Chen, Hoogeweg, and Harris 2001)[173] and are positively skewed (USEPA 2015)[199].

While there could be additional considerations and guidance on the treatment of different data distributions, USEPA’s Statistical Program ProUCL considers two distributions that describe data with skewed distributions: lognormal and gamma (Table 11-1; refer to USEPA’s ProUCL Technical Guide for in-depth discussion regarding theory and application of these distributions). Typically, skewed data are logarithmically transformed and tested for normality. If the transformed data distribution is normal, then the data was considered lognormal (IDEQ 2014)[177]. Traditionally, the log-transformed data were then analyzed using parametric statistical methods. However, the use of lognormal statistics can cause bias and unjustifiably elevated values for the mean and test statistics, especially for smaller datasets (<20–30 samples ((ITRC 2013)[137], (USEPA 2015)[195], (USEPA 2015)[199], (Singh, Singh, and Iaci

2002)[189]). Skewed data can also be described by the gamma distribution (Table 11-1). The theory and application of the gamma distribution to skewed data were developed by Singh, Singh, and Iaci (2002)[189] and incorporated into earlier versions of ProUCL. The application of the gamma distribution to skewed data and associated parametric methods is recommended in USEPA's Unified Guidance, subsequent state agency guidance, and the current version of ProUCL ((USEPA 2009)[198], (USEPA 2015)[195], (USEPA 2015)[199], (IDEQ 2014)[177]).

Table 11-1. Common distributions of environmental data

	Distribution Curve	Basic Properties	Statistical Analysis Methods
	Normal	Data distribution is not skewed and centered around the mean	Analyze data set using parametric statistical methods
	Lognormal	Data distribution is skewed and log transforming the data produces a normal distribution	Analyze data set using lognormal statistical methods only if data cannot be modeled by the normal or gamma distributions and the data set is not small (<15-20 samples) and highly skewed
	Gamma	Data distribution is skewed and modeled by the gamma distribution	Analyze data set using gamma statistical methods
	Nonparametric	Data distribution cannot be modeled as a normal, lognormal, or gamma distribution	Analyze data set using nonparametric statistical methods

11.2.1 Parametric and nonparametric statistical methods

Parametric statistics can achieve high levels of confidence with a relatively small number of data points ((USEPA 2015)[195], (USEPA 2015)[199], (ITRC 2013)[137]). However, environmental datasets often do not match the distributional assumptions for parametric methods, or the distribution cannot be determined because of insufficient sample size or a large number of nondetects (IDEQ 2014)[177]. In these instances, nonparametric methods may be more appropriate to use since they do not make distributional assumptions. Regardless, nonparametric methods still require a sufficient dataset and generally require larger datasets to achieve a similar level of confidence compared to parametric methods (ITRC 2013)[137].

The first step in any statistical analysis is to determine the data distribution. Numerous methods are available that determine data distribution, typically by testing for normality. Normality can be examined graphically using normal probability plots or mathematical goodness of fit tests. Some common goodness of fit tests are listed below, including the distribution they can be used to test for ((USEPA 2015)[195], (ITRC 2013)[137]):

- coefficient of skewness and variation (normal and skewed distribution)
- Shapiro-Wilk test (normal distribution)
- Shapiro-Francia normality test (simplified version of Shapiro-Wilk for large, independent samples)
- Kolmogorov-Smirnov test (gamma distribution in USEPA's ProUCL)
- Anderson-Darling test (gamma distribution in USEPA's ProUCL)

Datasets determined to be normally distributed are generally analyzed using parametric methods. Parametric methods are the most appropriate for normally distributed data. Datasets with skewed distributions were traditionally analyzed using lognormal methods; logarithmically transformed data that resulted in a normal distribution were then analyzed using parametric methods. However, as discussed previously, Singh, Singh, and Iaci (2002)[189] developed alternative methods using the gamma distribution that produce more accurate statistical results. Provided the dataset follows a gamma distribution, gamma statistical methods should be used to describe skewed datasets (USEPA 2015)[199]. USEPA's ProUCL Technical Guide recommends that lognormal methods should be applied only if the dataset cannot be modeled by normal or gamma distributions and the dataset is not highly skewed and of small size (<15–20 data points). In general, if the data cannot be modeled by any distribution suitable for a parametric analysis (for example, normal, lognormal, gamma), or if the dataset is insufficient to determine the distribution, then the dataset is considered to follow a nonparametric distribution (Table 11-1). If the dataset does not follow a distribution with available parametric methods, then nonparametric statistical methods should be applied (IDEQ 2014)[177].

11.2.2 Uncertainties

As with any analysis, any determination of data distribution has uncertainties. Tests for normality depend on the values at the ends or tails of the ordered data, or the maximum and minimum values in the dataset (ITRC 2013)[137]. Therefore, numerous nondetects (more that 10% of the dataset) or outliers can affect normality tests. While nondetects and high concentrations may affect normality tests, these values may be typical and expected for a skewed environmental dataset that is truly not distributed normally. In general, several methods should be applied to determine data distribution, not just tests for normality. Goodness of fit tests can help identify the most reasonable distribution type, with statistical software providing the ability to easily test for several distributions and visually inspect the test results via Q-Q plots (see Section 11.4 for further discussion of graphic methods).

Additionally, large datasets collected over space or time can result in non-stationarity (data with a trend or changing variance). Specifically, soil samples collected over a large area for determining default soil background may include data that are not in the same distribution, causing the hypothesis of normality to be rejected by the distribution test even if the sample subsets are normally distributed (ITRC 2013)[137]. Trend tests or analysis of variance tests should be used if non-stationarity is suspected. Trend tests also rely on assumptions regarding data distribution; An analysis of variance test is a parametric test, and an equivalent nonparametric test is the Kruskal-Wallis test. In instances where non-stationarity is determined, additional analysis will likely be required, including release and transport evaluation, further sampling strategies, or different ways of combining and analyzing data (USEPA 2002)[153]. Additional analysis may be as simple as excluding a portion of the dataset that can be determined to be nonrepresentative or to be as complex as using a complex non-stationarity model to determine spatial sampling (Marchant et al. 2009)[432]).

Making unverifiable assumptions regarding data distribution is not advisable (ITRC 2013)[137], particularly regarding data normality. In instances where data distribution is not known, nonparametric methods should be used.

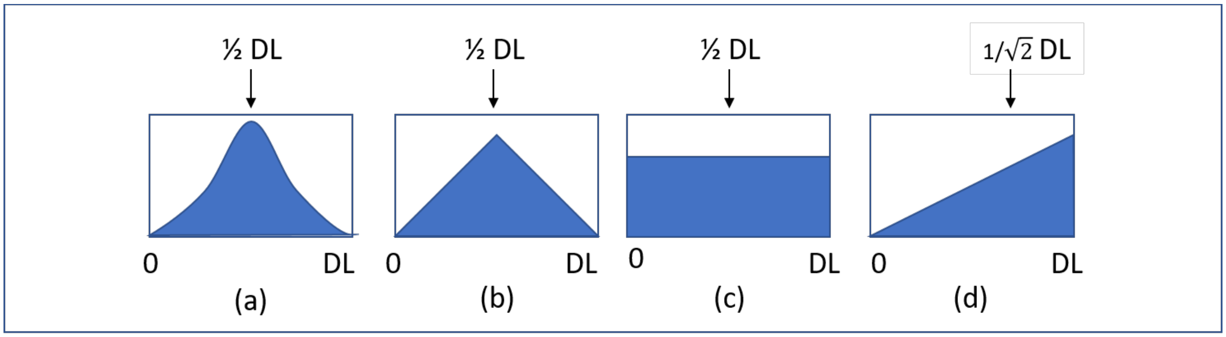
11.3 **Treatment of Nondetects**

Environmental datasets often contain “nondetect” results because of the limited sensitivity of laboratory methods to measure contaminants or because the analyte does not exist in soil (Figure 10-1). This section is intended to cover what to do when the concentration of a given analyte in soil is below the detection limit (nondetect) and how to incorporate such data points in the statistical analysis. It is not intended to discuss laboratory analysis and results, which are discussed in Section 10. Specifically, Section 10.3 provides detailed discussion about detection and quantitation limits.

Nondetects in background datasets can complicate statistical analyses. There are several recommended procedures for treatment of nondetects. In general, nondetect results should be retained when using one of the appropriate methods described below. However, nondetects with detection limit (DL) values larger than the highest detected value in a dataset are usually excluded from further consideration to avoid ensuing complications in statistical analyses.

11.3.1 Substitution methods

Substitution or imputation methods replace nondetects with numerical values that are then treated as equivalent to detected values in statistical analyses. Typical substituted values include 0, various fractions of the DL, the full DL, or randomly assigned values between 0 and the DL. The most common form of imputation is ½ DL substitution, but there are also other proposed fractions of the DL such as DL/√2 (Antweiler 2015)[12]. Each of these alternatives is based on speculations or educated guesses about the shape of probability distribution functions of the nondetects. For example, ½ DL substitution assumes that the distribution of nondetects is symmetric between 0 and DL, while 1/√2 DL is based on the general assumption that the distribution of nondetects can be approximated by a right-triangular distribution with the DL as its mode and zero as its lower limit (see Figure 11-1)



Symmetric distributions = (a), (b), and (c); right-triangular distribution = (d).

Figure 11-1. Examples of assumed nondetect distributions and their corresponding substitution values.

Source: Leyla Shams, NewFields.

Most substitution methods are easy to understand and implement and, in some cases, produce results that are similar to certain more complex methods (for example, the median semi-variance (SemiV) method (Zoffoli et al. 2013)[39]). In their extensive comparative evaluation of the performance of several methods for analyzing simulated censored datasets, Hewett and Ganser (2007)[22] recognized the importance of ease of calculation/accessibility in dealing with datasets that include nondetects. Hewett and Ganser (2007)[22] indicated that substitution methods are expedient and reasonably accurate, especially when dealing with datasets with small proportions of nondetects. However, substitution methods produce biased low estimates of population variance (Hewett and Ganser 2007)[22]. Such results can affect the reliability of certain statistics, including the estimates of the upper and lower confidence limits that rely on variance. For example, as discussed in ITRC (2013)[137], using ½ DL substitution artificially reduces the variance of concentration data, resulting in a confidence interval that is smaller than expected.

In 2006, USEPA guidance supported the use of 0, ½ DL, or DL substitution in datasets with less than 15% nondetects (USEPA 2006)[192]. USEPA's 2015 ProUCL Technical Guide restricted the use of ½ DL substitution only to datasets that are mildly skewed with less than 5% nondetects (USEPA 2015)[197]. Some investigators ((Hornung and Reed 1990)[23], (She 1997)[34], (Antweiler and Taylor 2008)[11]) report the performance of ½ DL as being reasonable.

11.3.2 Kaplan-Meier method

The Kaplan-Meier (KM) method (Kaplan and Meier, 1958) is a nonparametric approach for construction of the cumulative distribution function of a dataset that contains censored data. The constructed CDF in turn is used to estimate the summary statistics of interest. The KM method sorts the dataset by detected and DL values ascendingly and relies on the number of records at and below each detected value to compute its cumulative probability.

The application of the KM method to environmental studies is relatively recent, when algorithms were developed to reformulate the method for left-censored environmental data (nondetects that are reported as less than the DL). Helsel (2005)[20] proposed to transform censored data from left to right by subtracting each detected and DL value by a “large” number (also referred to as “flipping the data”), whereas Popovic et al. (2007)[30] adjusted the algorithm formulas for left-censored data. This latter method has been adopted in USEPA ProUCL ((USEPA 2013)[419], (USEPA 2015)[197]).

KM does not require an assumption of data distribution or any substitution for nondetects, and thus can be applied to cases where the distribution of the data is not known or discernible ((Hewett and Ganser 2007)[22], (USEPA 2015)[197]), and is insensitive to outliers (Antweiler and Taylor 2008)[11]. KM results, however, are reliable only if the pattern of censoring is random and the probability of censoring is independent of DL values (Schmoyeri et al. 1996)[33], (She 1997)[34]. This assumption means that the DL values associated with nondetects in a dataset must occur without displaying any preference to any particular range of concentrations. DL values in typical environmental datasets, however, are often associated with unique and/or low values that are lower than most, if not all, of the detected values. Older datasets with highly variable DLs may be more likely to meet the requirement of having a widespread distribution of DLs relative to measured values. In other words, nondetects are often concentrated along the lower end of the distribution. For such datasets, KM mean and upper confidence limit (UCL) results will be biased high.

Similar to other nondetect treatments, KM results become less reliable when the proportion of nondetects increases. PROPHET Stat Guide (Hayden et al. 1985)[476] warned against the use of KM in cases of heavy censoring or small sample sizes. Helsel (2005)[20] recommended use of the KM method on datasets with no more than 50% censored data, while ITRC (2013)[137] recommended “no more than 50–70% nondetects.” Antweiler and Taylor (2008)[11] recommended KM for summary statistics when datasets include less than 70% censored data.

KM has been promoted by many authors (including (Helsel 2005)[20], (Antweiler and Taylor 2008)[11], (USEPA 2015)[197]), resulting in recommendations for its use within the environmental community. For example, (ITRC (2013)[137], Section 5.7) recommended use of a “censored estimation technique to estimate sample statistics such as the KM method for calculating a UCL on the mean.” Hewett and Ganser (2007)[22] tested KM against substitution, MLE, and regression on order statistics and found that KM “did not perform well for either the 95th percentile or mean.”

11.3.3 Regression on order statistics (ROS)

ROS is a semiparametric imputation technique (Helsel and Cohn 1988)[21]. To apply ROS, the results are sorted in an ascending order in accordance with their detected or DL values, as applicable. This step is followed by producing a plot of observed versus theoretical quantiles, sometime referred to as the Q-Q plot. Among the plotted quantiles, those associated with detected values are then subjected to linear regression. Each nondetect is then substituted with the predicted values based on the interpolated or extrapolated regression line using the order of their corresponding DL values ((Helsel 2005)[20], page 68). A pictorial example is provided in Figure 11-2.

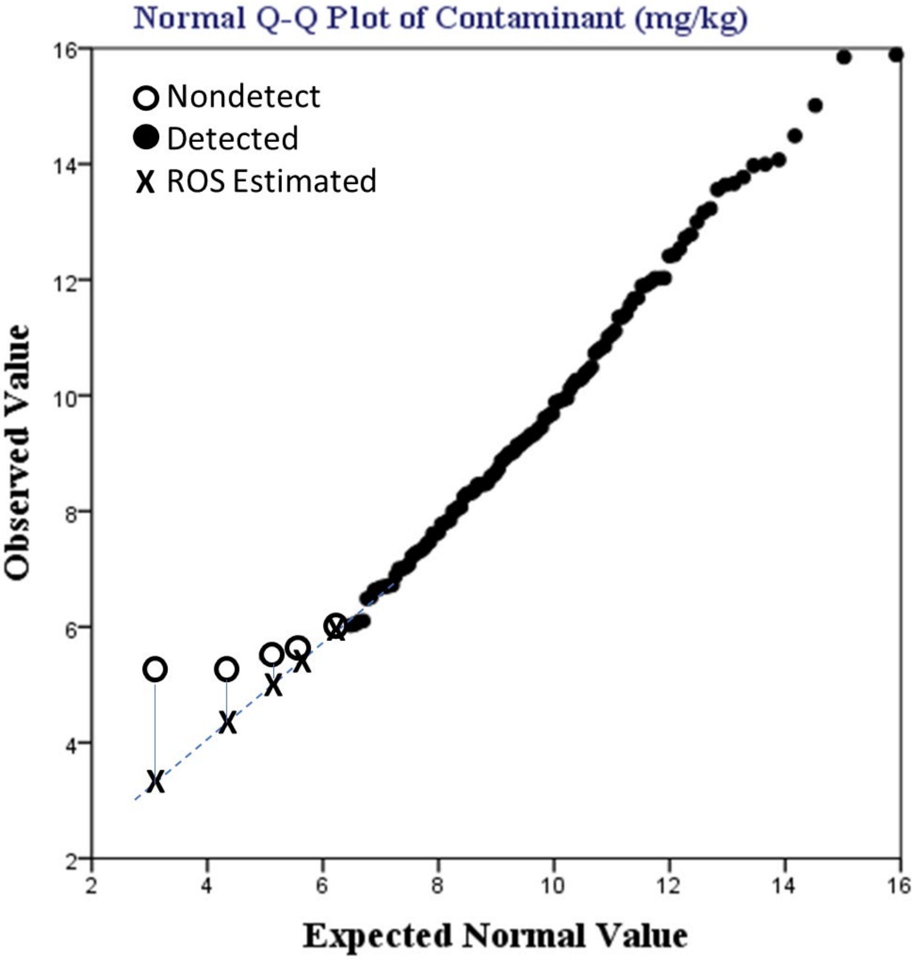


Figure 11-2. ROS application example.

Source: Leyla Shams, NewFields (Generated by IBM SPSS Version 23 plus added graphics).

ROS can be applied to cases where statistical testing has confirmed that the measured data are derived from single populations. ROS is especially suitable for cases with 15–50% nondetects with either one or multiple detection limits ((ITRC 2013)[137], (Sinha, Lambert, and Trumbull 2006)[35]). However, ROS requires an a priori assumption about the distribution of the censored values: in particular, typical ROS applications assume that the distribution of the investigated data are approximately normal or lognormal (ITRC 2013)[137]. If the type of distribution is incorrectly assumed, the resulting ROS estimates (for mean, standard deviation, UCL, and upper tolerance limit (UTL) values) could be inaccurate. The presence of outliers can also distort the regression estimates of slope and intercept that are used to impute values for nondetects (USEPA 2015)[197].

Hewett and Ganser (2007)[22] found that overall, the log-probit regression-based methods "performed well across all scenarios." Overall, ROS is most applicable to datasets that are derived from single populations, have limited skew, lack outliers, and have <50% nondetects (USEPA 2015)[197]. Datasets with 50% or more nondetects should not be subjected to ROS calculations ((ITRC 2013)[137], (USEPA 2015)[197]). ITRC (2013)[137] further recommended that the approach should be applied with datasets representing at least eight to ten detected measurements, while USEPA (2015)[197] recommended a minimum of four to six detected measurements.

11.3.4 Maximum likelihood estimation (MLE)

Maximum likelihood estimation refers to a family of parametric methods that, in essence, estimate parameters of assumed distributions by maximizing the likelihood of the occurrence of the actual detected and nondetect values in the dataset. (Helsel (2005)[20], page 13) stated, "MLE uses three pieces of information to perform computations: a) numerical values above detection limits, b) the proportion of data below each detection limit, and c) the mathematical formula of an assumed distribution. Parameters are computed that best match a fitted distribution to the observed values above each detection limit and to the percentage of data below each limit." The MLE-estimated distribution parameters are then used to calculate the summary statistics of the investigated data.

MLE can be used under a variety of assumed symmetric and asymmetric distributions. The vast majority of MLE applications in the literature have been developed for normal or lognormal distributions, which in some instances may reasonably match observed distribution of background datasets ((Akritas, Ruscitti, and Patil 1994)[10], (Nysen et al. 2015)[29]). Recent articles offer MLE solutions based on new classes of mixed distributions, which can be more representative of typical environmental datasets (Li et al. 2013)[26]).

The primary cases where MLE is applicable are those in which the sample distribution can be reliably determined ((Helsel 2005)[20], (ITRC 2013)[137]). These are often datasets, derived from single populations, with larger sample sizes and/or a small proportion of nondetects (for example, less than 15% nondetects) (USEPA 2006)[192]. ITRC (2013)[137], for example, recommended applying MLE to sample sizes of 50 or above and a detection frequency of more than 50%. If the type of underlying distribution is incorrectly assumed or cannot be identified, the resulting MLE estimates could be misleading.

Similar to datasets with nondetects, when applying MLE on censored datasets to compute upper limits such as UCLs and UTLs (see section 11.6), the use of lognormal distribution should be avoided (Singh, Singh, and Me 1997)[188]. Despite its prevalence in environmental applications, assuming a lognormal distribution may lead to unrealistically elevated UCLs and UTLs, especially when the dataset is highly skewed. This issue has been illustrated throughout the ProUCL Technical Guide (USEPA 2015)[197].

The most recent versions of USEPA's ProUCL program have excluded parametric MLE methods, which USEPA (2015)[197] described as "poor performing," likely due to difficulties in verifying the distribution of left-censored datasets with multiple detection limits. In contrast, Hewett and Ganser (2007)[22] found MLE methods in general to be strong performers when calculating the mean and 95th percentile values for data with known distributions.

11.3.5 Summary of nondetect treatment

The choice of the treatment of nondetects is driven by many factors that may require testing multiple procedures before identifying the appropriate method. Some practical rules for selecting an appropriate method for treatment of nondetects include:

- Substitution, such as 0, ½ DL, or DL, may be used, while recognizing that these methods often yield underestimated standard deviations, as discussed in Section 11.3.1.
- The Kaplan-Meier (KM) method has the advantage of being a nonparametric procedure, as discussed in Section 11.3.2. However, if DLs associated with nondetects are mostly below detected values, the KM mean and UCL would be biased high.
- Regression on order statistics (ROS) is ideal if the distribution of the detected values is clearly evident, as discussed in Section 11.3.3.
- MLE is ideal for datasets for which a reliable distributional assumption can be made, as discussed in Section 11.3.4.
- In cases where the dataset does not meet the above conditions for KM, ROS, or MLE, substitution may be used, while recognizing that these methods often yield underestimated standard deviations, as discussed in Section 11.3.1.

11.4 **Graphical displays**

Graphical displays of data can provide additional insight about a dataset that would not necessarily be revealed by using test statistics or estimates of confidence limits. These displays can be helpful in visualizing differences in means, variance, and distributions of background and site concentrations. Several graphing techniques that are commonly used include Q-Q plots, histograms, box plots, probability plots, and percentile plots. These are described in more detail below.

11.4.1 Quantile-quantile plot (Q-Q plot)

A Q-Q plot is a type of probability scatter plot generated by displaying data quantiles versus the theoretical quantiles of the specified distribution, including normal, gamma, and lognormal distribution. Departures from the linear display of the Q-Q plot suggest that the data do not follow the specified distribution. Note that theoretical quantiles are generated using the specified distribution. The correlation coefficient based upon the Q-Q plot is not a good measure to determine the data distribution. Statisticians have developed goodness of fit tests to determine data distribution. ProUCL has goodness of fit tests for normal, lognormal, and gamma distributions. Example Q-Q plots are shown in Figure 11-3 and Figure 11-4. If the background and site concentrations were very similar (Figure 11-3), the two best fit lines would lie nearly on top of one another. If the concentration of contaminants at the site diverged from the background concentrations, the lines would diverge as shown in Figure 11-4. Q-Q plots can also be used to compare a set of data against a specific distribution in a method called a theoretical quantile-quantile plot.

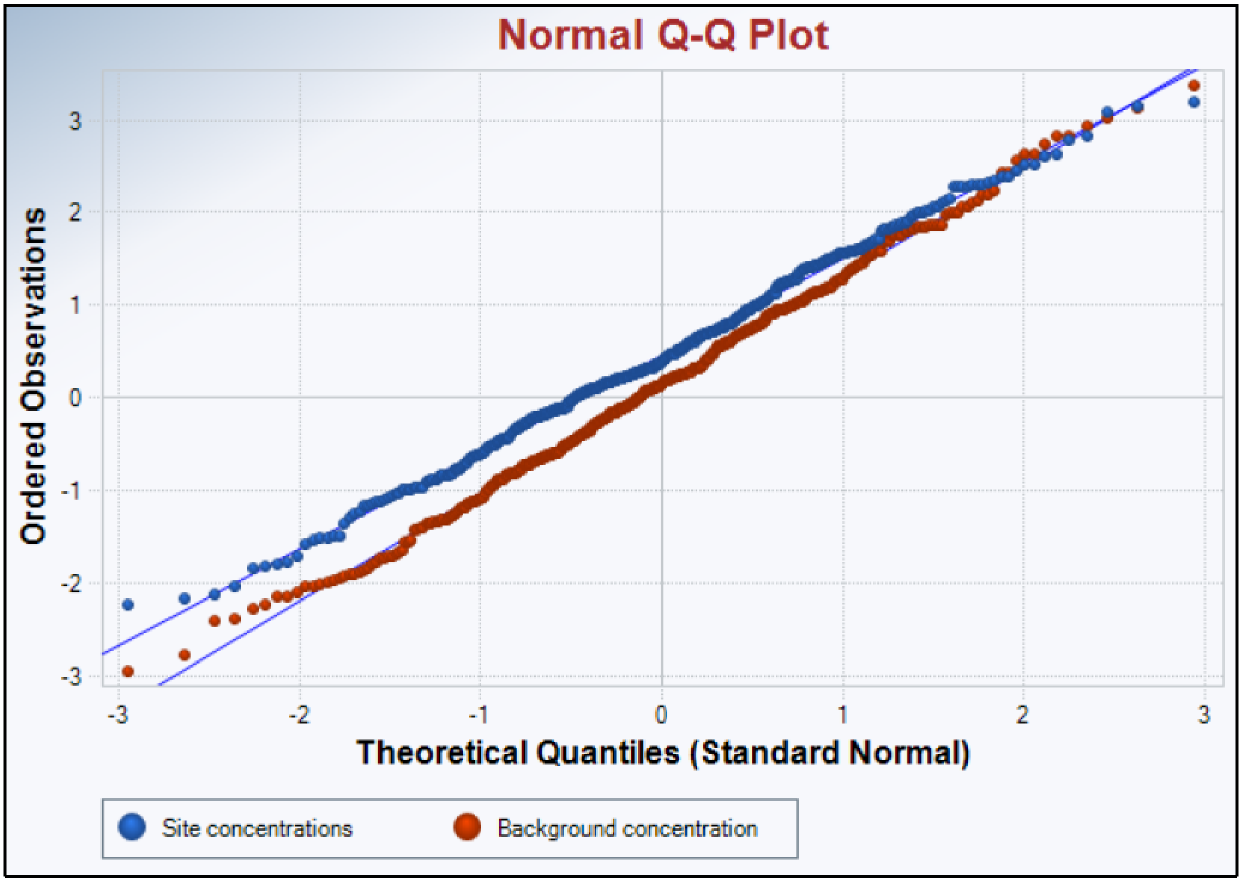


Figure 11-3. Example Q-Q plot—Similar site concentration and background concentration.

Source: Jennifer Weidhaas, University of Utah (Generated by USEPA ProUCL version 5.1).

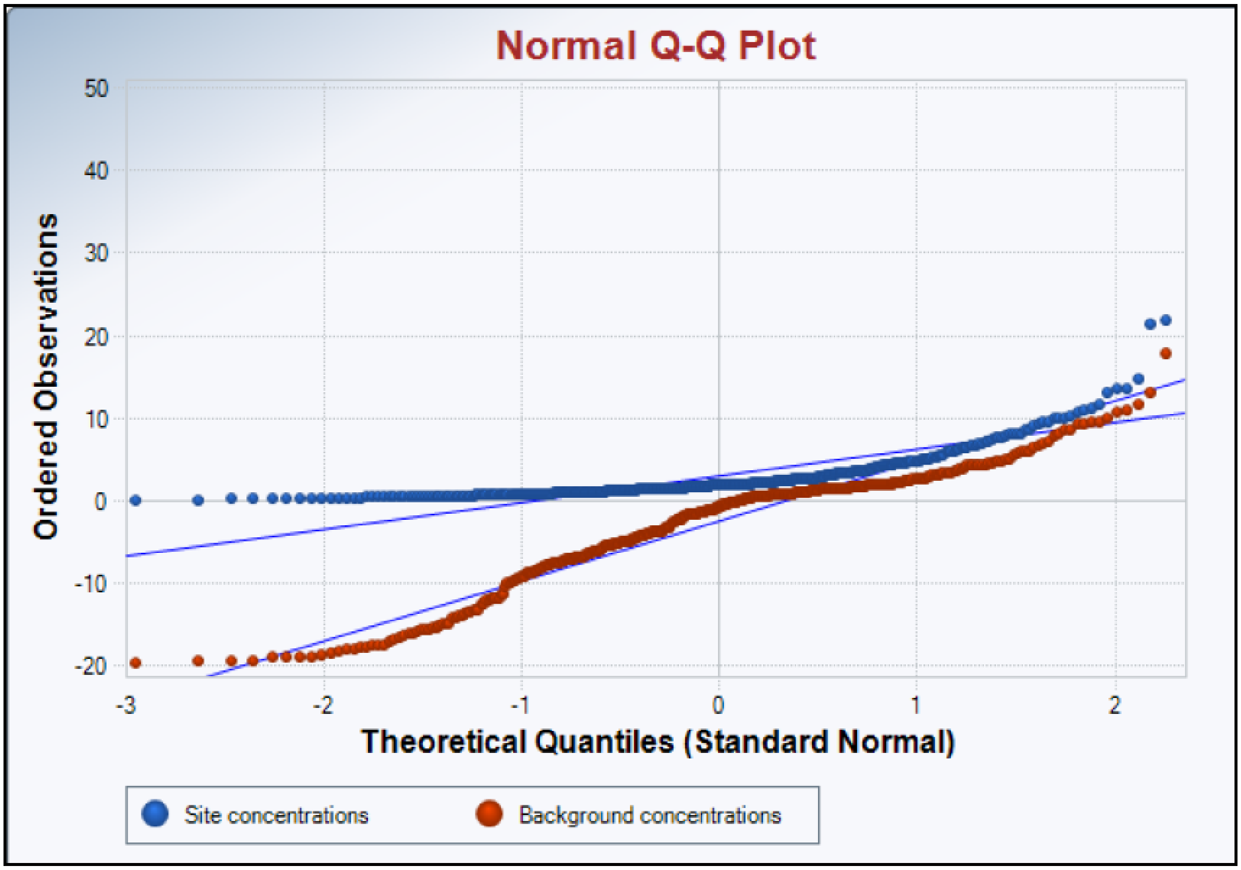


Figure 11-4. Example Q-Q plot—Site concentrations diverging from background concentrations.

Source: Jennifer Weidhaas, University of Utah (Generated by USEPA ProUCL version 5.1).

The combined background and site Q-Q plot with highlighted background data can be used to identify the subpopulations associated with the background and to extract background data from the site data (ASTM E3242-20 (ASTM 2020)[146]). However, the identification of exact break points (or threshold values for different populations) in a Q-Q plot is subjective and requires expertise for accurate evaluation. Guidance on these issues is provided in Section 3 of this document.

11.4.2 Histogram

Histograms are used to display continuous data and to graphically summarize the distribution of the dataset. These histograms can be used to compare the size and shape of the observed data and the potential skewness of the distribution. A histogram showing a lognormally distributed dataset is provided in Figure 11-5 and a histogram showing a normally distributed dataset is provided in Figure 11-6.

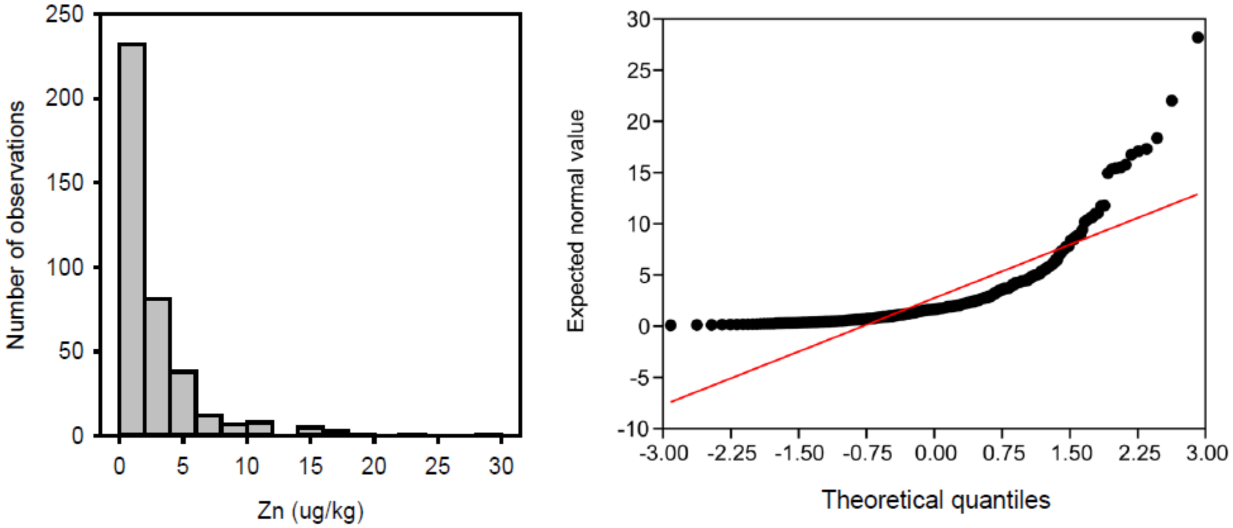


Figure 11-5. Example histogram—Lognormal distributed data versus Q-Q frequency plot.

Source: Jennifer Weidhaas, University of Utah (Generated by Sigma Plot v 14.5).

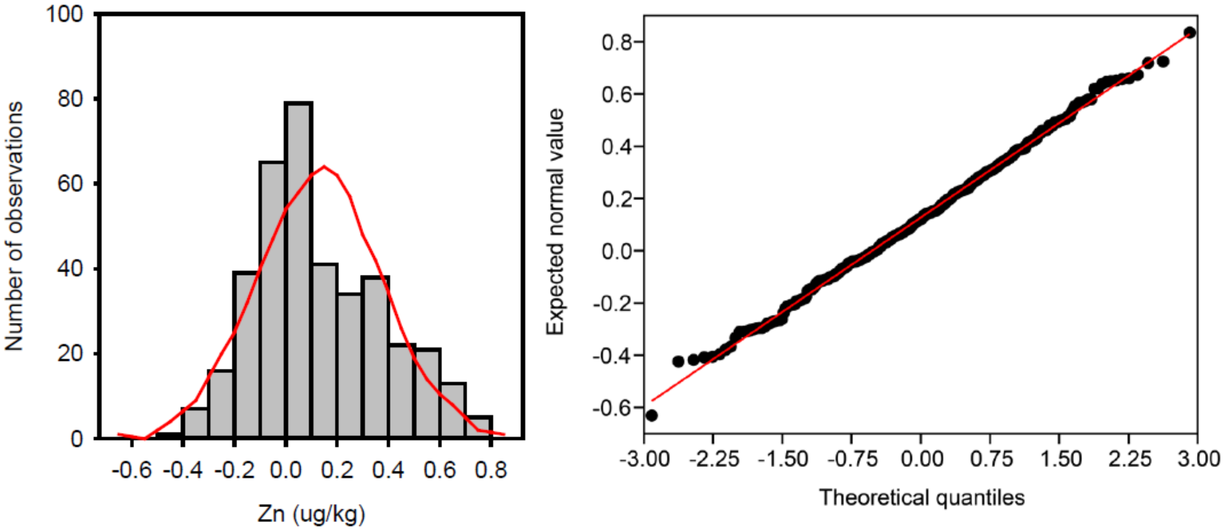


Figure 11-6. Example histogram—Normal distributed data versus Q-Q frequency plot.

Source: Jennifer Weidhaas, University of Utah (Generated by Sigma Plot v 14.5).

The utility of histograms is limited if too few data points or samples are available, the proximity between the population means is too close, or there are unequal standard deviations for the populations. Histogram interpretation can be influenced by the number of bins included in the analysis. If too few bins are used, the shape is lost; conversely, if too many bins are used the shape can be lost as well. However, the histogram can be used to illustrate the conclusions drawn from the Q-Q plot and any outlier tests.

11.4.3 Box plot

Box plots summarize and display the distribution of a set of continuous data, such as the range of metal concentrations in soils at a site. These plots typically highlight five key points of the data: the median, the first and third quartiles (25th and 75th percentile values), and the minimum and maximum. Typically, values that are more than 1.5 times the interquartile range (the 75th percentile value minus the 25th percentile value) are considered outliers. Box plots are valuable as they can be quickly scanned to show the central tendency, range, and presence of outliers. Figure 11-7 illustrates a box plot and depicts the interquartile range. When two box plots are shown side by side it is easy to make comparisons between datasets. Figure 11-8 shows an example comparison of the range of metal concentrations in soils at two sites as depicted in box plots.

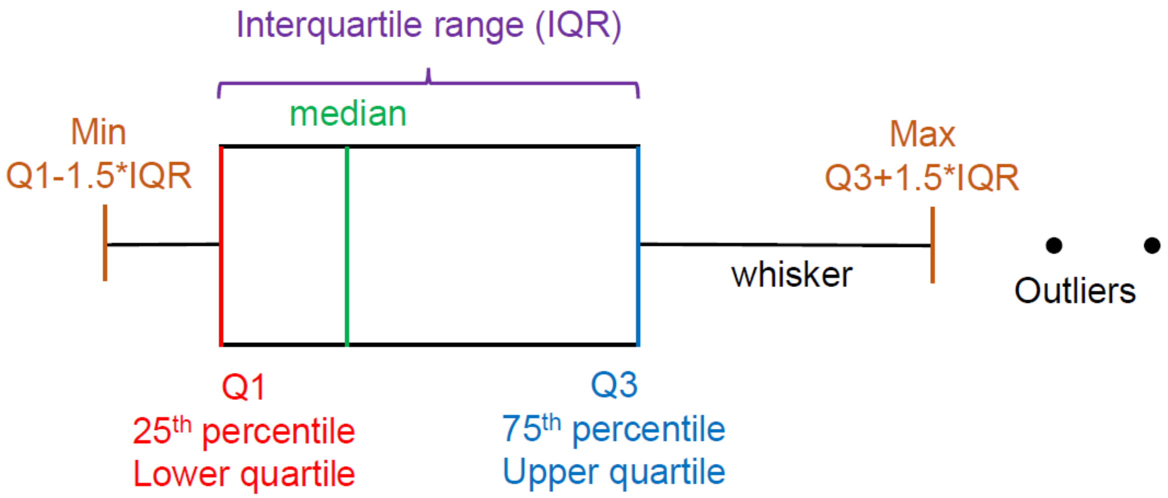


Figure 11-7. Illustration of box plot and key characteristics.

Source: Jennifer Weidhaas, University of Utah.

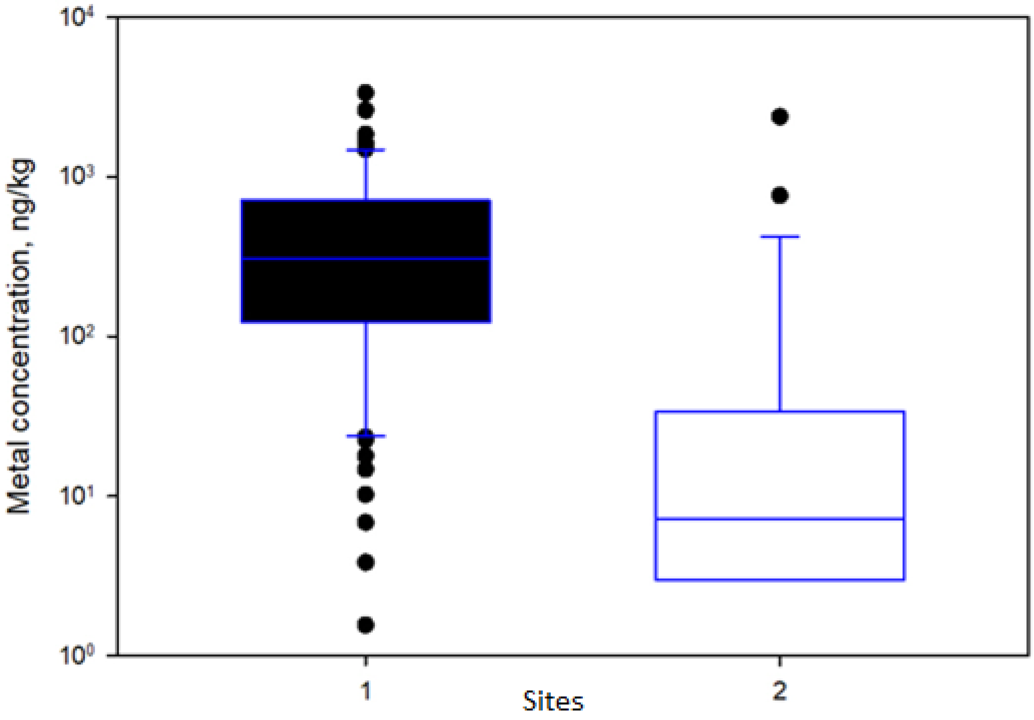


Figure 11-8. Example comparison of the range of metal concentrations in soils at two sites.

Source: Jennifer Weidhaas, University of Utah (Generated by Sigma Plot v 14.5).

Box plots are typically most useful when datasets contain five or more values and nondetect values need to have a value assigned. Box plots are a quick, simple, and graphical method of displaying data. Further, they are easy to interpret and to compare two or more sets of data side by side. However, identification of outliers is fairly arbitrary and may not be conclusive. Identification of outliers in a box plot is user-defined, as it is possible to set the upper whisker at the maximum concentration, and the lower whisker at the minimum concentration. These plots are also limited in that they show only a single variable in each plot and cannot necessarily show relationships between variables or sites. Further, depending on how the data are plotted, all individual data points may not be shown. The box plot is not intended to be the only graphic used to evaluate site and background.

11.4.4 Probability plots

Probability plots are a graphical technique for assessing whether a dataset (for example, concentrations of a specific element) follows a given distribution (for example, normal or Gaussian distribution). The site data are plotted against a theoretical distribution so that if the data fit the distribution plotted, they fall along a straight line. The Q-Q plot is a type of probability plot (Figure 11-5 and Figure 11-6), namely the normal probability plot. Departures from the theoretical straight line indicate a departure from the specified normal distribution for the data. To determine the goodness of fit of the data to a distribution, the correlation coefficient is used. Typically, probability plots for several competing distributions are generated and the one with the highest correlation coefficient is likely the best choice.

11.4.5 Percentile plots

Percentile plots are nonparametric with concentration on the *y*-axis versus percentile on the *x*-axis (ASTM E3242-20 (ASTM 2020)[146]). To construct these plots, first the concentration data are rank-ordered and then the corresponding percentile is applied to each concentration. An example of this kind of plot is provided in Figure 11-9. Although no distributional assumptions are required, normally distributed data (if applicable) will appear as a straight line when a linear concentration scale is used. Lognormally distributed data (if applicable) will appear as a straight line when a logarithmic concentration scale is used (see Figure 11-9). One application of percentile plots is in the identification of statistical outliers in a candidate background dataset. Such outliers will lie above or below the trend formed by the other data points. Breaks in slope may indicate bimodal or multimodal distributions or may indicate the presence of multiple samples with identical values. As with other graphical techniques, the data should always be inspected when evaluating the plots to identify the reasons for observed breaks in slopes and apparent outliers.

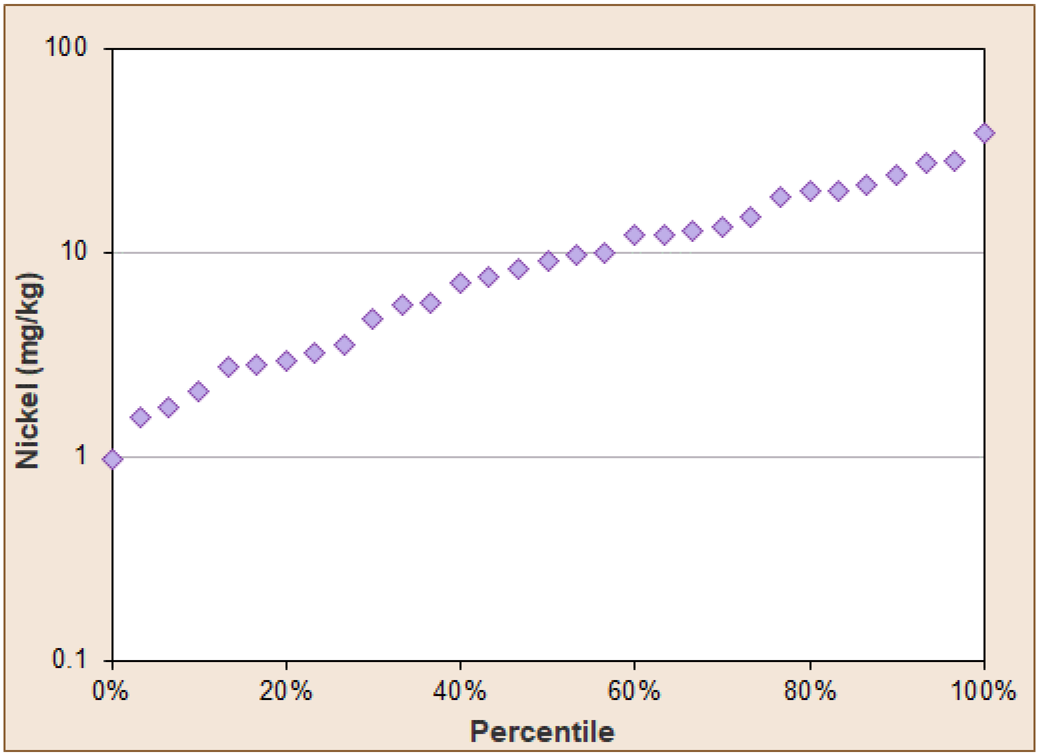


Figure 11-9. Example percentile plot.

Source: Karen Thorbjornsen, APTIM.

11.5 **Outliers**

An outlier or an outlying observation refers to an extreme observation in either direction that appears to deviate markedly in value from other measurements of the dataset in which it appears. Outliers in environmental datasets can be classified into two broad groups (ASTM E178 2016 (ASTM 2016)[171]):

- True outlier—"An outlying observation might be the result of gross deviation from prescribed experimental procedure or an error in calculating or recording the numerical value. When the experimenter is clearly aware that a deviation from prescribed experimental procedure has taken place, the resultant observation should be discarded, whether or not it agrees with the rest of the data and without recourse to statistical tests for outliers. If a reliable correction procedure is available, the observation may sometimes be corrected and retained."

- False outlier—"An outlying observation might be merely an extreme manifestation of the random variability inherent in the data. If this is true, the value should be retained and processed in the same manner as the other observations in the sample. Transformation of data or using methods of data analysis designed for a non-normal distribution might be appropriate."

USEPA (2006)[192] stated that "failure to remove true outliers or the removal of false outliers both lead to a distortion of estimates of population parameters." In practice, only outliers that are demonstrably erroneous or belonging to populations not representative of background conditions should be excluded from the background dataset. In background investigations, typical sources of error that can result in outliers include: (a) transcription error, (b) sampling error, (c) laboratory error, and (d) sampling of media not representative of background conditions as determined by forensic and geochemical analyses. In some cases, soil concentrations associated with known releases or isolated extreme values, such as the outlier displayed in Figure 11-10, may be considered as not representative of background conditions. All other identified outliers should be retained and processed in the same manner as the other observations in the sample. A comprehensive review of outlier removal issues and recommended tests is presented in Grubbs (1969)[16].

11.5.1 Graphical plots

Perceived high and/or low outliers can be visually identified using a probability plot. For this purpose, the measurements in the dataset are sorted in an ascending order and plotted against their corresponding cumulative probabilities, based on a specified distribution. The most common type of probability plot is constructed based in the normal distribution, hence referred to as the normal probability plot. Other forms of the probability plot include the Q-Q (quantile-quantile) plot and the P-P (probability-probability or percent-percent) plot. Normal Q-Q or PP plots can be used for any dataset regardless of their distribution. Log-normal probability plots are also used by practitioners, although USEPA (2015)[199] warns against their use. Therefore, the use of log-normal probability plots is appropriate only if assessed and approved by a professional with statistical expertise.

A potential outlier may manifest itself as a high or a low value separated by a large gap from its preceding or ensuing ranked measurement, as displayed in Figure 11-10. A gap is considered as "large" if it is visually much wider than the gaps displayed by other measurements in the probability plot. Authors, including van der Loo (2010)[38], have proposed parametric procedures to analyze probability plot results to determine outliers.

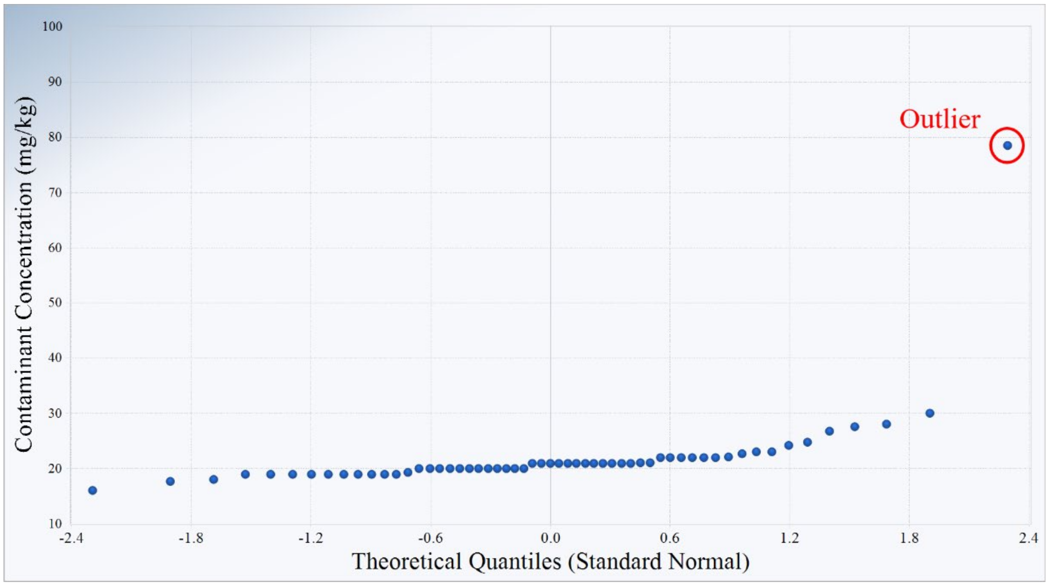


Figure 11-10. Outlier in a Q-Q plot.

Source: Leyla Shams, NewFields (Plot generated by USEPA ProUCL 5.1 (USEPA 2015)[199]).

11.5.2 Statistical tests

There are well-established procedures to test for statistical outliers. Each statistical test requires a pre-identified number of perceived high and/or low outliers and an assumed distribution of the background population. Typical occurrences include a single outlier, or two or more outliers on the same or opposite sides of the dataset.

As listed in Table 11-2, common statistical outlier tests, including those presented in ASTM E178-16a (ASTM 2016)[171], assume the normality or symmetry of the measurements after the removal of outliers. As noted in ASTM E3242-20 (ASTM 2020)[146], for many environmental datasets, the normality assumption is incorrect, which can lead to erroneous outlier identifications. This problem is further exacerbated when the above statistical tests are applied repeatedly, through the iterative removal of perceived outliers. Under such procedures, shrinking standard deviations caused by continuous exclusion of perceived outliers produces an increasing likelihood of incorrect identification of additional outliers. The result is the calculation of biased, unrepresentative results from an incorrectly truncated dataset. In many instances, graphical techniques, such as normal Q-Q plots or box plots, may be more appropriate for identifying outliers. In these cases, a normal Q-Q plot (on untransformed data) may be used as an exploratory tool to identify outliers, and not to determine normality of data (Hoaglin, Mosteller, and Tukey 1983)[434].

Table 11-2. Outlier tests

Source: Developed from (USDON 2002)[36], Table B-3, and ASTM E3242-20 (ASTM 2020)[146], Table X4.1.

Statistical Test	Assumptions	Advantages	Disadvantages
Dixon's test	• Sample size ≤ 25. • Measurements without outliers are normally distributed. • There is only a single suspected outlier, which can be either the smallest or largest value.	• Available in commonly used software, including USEPA's ProUCL. • Recommended in many guidance documents, including USEPA's ProUCL.	• The user must consider the highest or the lowest measured value as a potential outlier prior to the test. • Assumption of normality is rarely applicable to environmental field data. • In cases of skewed or asymmetric data distributions, the test has a tendency to falsely flag the tail value as an outlier.
Discordance test	• 3 < Sample size ≤ 50. • Measurements without outliers are normally distributed. • There is only a single suspected outlier, which can be either the smallest or largest value.	• Available in commonly used software.	• The user must consider the highest or the lowest measured value as a potential outlier before the test. • Assumption of normality is rarely applicable to environmental field data. • In cases of skewed or asymmetric data distributions, the test has a tendency to falsely flag the tail value as an outlier. • Not included in ProUCL.
Rosner's test	• Sample size ≥ 25. • Measurements without outliers are normally distributed.	• Can test for up to 10 suspected outliers. • Available in commonly used software, including USEPA's ProUCL. • Recommended in many guidance documents, including USEPA's ProUCL.	• The user must consider the highest or the lowest measured value as a potential outlier before the test. • Assumption of normality is rarely applicable to environmental field data. • In cases of skewed or asymmetric data distributions, the test has a tendency to falsely flag tail values as outliers.

Statistical Test	Assumptions	Advantages	Disadvantages
Walsh's test	• Sample sizes >60 (for 10% significance level). • Sample sizes > 220 (for 5% significance level). • Can be used to test if the largest *r* measurements or the smallest *r* measurements are outliers. • The measurements without outliers need to be symmetrically distributed, although not necessarily normally distributed.	• Can test for one or more outliers.	• The user must consider the highest or the lowest measured value as a potential outlier prior to the test. • Test can be performed only for certain significant levels. • The significance level depends on the sample size (for example, the significance level of 5% can be used only for sample sizes >220). • The suspected outliers are accepted or rejected as a group, rather than one at a time. • Not included in ProUCL.

The above outlier tests are commonly cited in environmental guidance documents and were developed when environmental datasets were small and the computing power was limited. Newer outlier identification methods that include robust methods are available, such as Hoaglin, Mosteller, and Tukey (1983)[434], Rousseeuw and Leroy (1987)[184], Rousseeuw and Zomeren (1990)[186], Iglewicz and Hoaglin (1993)[178], Maronna, Martin, and Yohai (2006)[181], Singh and Nocerino (1995)[416], and Singh (1996)[187]. Several effective and robust outlier tests are available in the commercial software packages, including SAS and STRATA.

11.6 Confidence Interval Limit, Coefficient, and Limit

A confidence interval is a range having an upper limit value and lower limit value. The true value of the parameter of interest lies within this range/interval. The confidence interval is determined based on the dataset that has been sampled and is only as accurate as the data one has obtained. A confidence interval places a lower limit or an upper limit on the value of a parameter, such as soil metal concentrations.

A confidence coefficient displays how confident the user is the true value will be within the confidence interval. In other words, how much confidence there is that the true value of the parameter of interest does indeed lie between the upper and lower limits of the confidence interval. It is generally recommended that a confidence coefficient of 95% be used, meaning there is a 95% chance that the true value lies between the upper and lower confidence interval values. Confidence coefficients are expressed as values between 0 and 1. A 95% confidence coefficient is expressed as 0.95.

A confidence limit is the lower or upper boundary of a confidence interval (the range of possible values that includes the true value of the parameter of interest). For risk assessment we are usually interested in the upper boundary since we need to be reasonably conservative to ensure we are being protective. When we are choosing an exposure concentration, we usually choose the upper confidence limit of the mean since this gives us a reasonably conservative estimate of what the true mean might be. When we are establishing soil background values (BTVs) we also usually consider using the upper confidence limit of the range of background concentrations to be reasonable to compare to the site concentrations. It is important to carefully consider what site concentrations will be used to compare to this upper confidence limit of soil background to ensure we are being reasonably protective, as discussed in Section 3.

11.7 Statistical Values Used to Represent Background

Although soil background is properly expressed as a range of values, regulators generally need to express it as one single value to be able to use it in risk assessment. Often, for this purpose regulators use soil background threshold values or BTVs. BTV is defined as a measure of the upper threshold of a representative background population, such that only a small portion of background concentrations exceed the threshold value. BTVs are usually used for site delineation purposes, or point-by-point comparison to individual site data to identify localized contamination ((Geiselbrecht et al. 2019)[15], ASTM E3242-20 (ASTM 2020)[146]). Appendix A contains additional information on upper limits commonly used to represent background.

Regardless of the chosen BTV, point-by-point comparisons are prone to produce false positive errors. That is, as the number of comparisons increases, the chance of incorrectly concluding at least one erroneous above-background exceedance approaches 100%, even when the site data are within background ranges (Gibbons 1994)[18]. As stated in ASTM D6312-17 (ASTM 2017)[170], "Even if the false positive rate for a single [comparison] is small (for example, 1 %), the possibility of failing at least one test on any [site dataset] is virtually guaranteed." Due to this limitation, caution should be used when conducting point-by-point comparisons. The U.S. Department of the Navy (USDON 2003)[37] recommended against point-by-point comparisons, except when coupled with reverification sampling ((Gibbons 1994)[18], ASTM D6312-17 (ASTM 2017)[170]). In practice, point-by-point comparisons to BTV are very useful and efficient screening procedures. If exceedances above BTV are detected, then more robust site-to-background tests can be conducted to determine presence or absence of concentrations above background conditions.

Applications of BTV yield reliable results only if the representative background dataset contains an adequate number of measurements. Adequacy of the background sample size is dependent on the intended application, assumptions about the underlying distribution of the chemical concentrations, and the tolerable error rates. These error rates include falsely declaring a site clean (false negative) and falsely declaring a site contaminated (false positive). Falsely declaring a site clean is often of more concern to regulators, and the error rate is set at 5% by convention. The total sum of error rates should not exceed 25% (Gibbons 1994)[18]. As a result, the rate of falsely declaring a site contaminated is often set at a value less than or equal to 20%. Inadequate background sample sizes can lead to unreliable or inappropriate conclusions. A more comprehensive review of this topic and the associated literature is presented by Cochran (1997)[14].

Values commonly used to represent BTVs include the upper percentile, the upper prediction limit (UPL), the upper tolerance limit (UTL), and the upper simultaneous limit (USL). A summary of these values is presented in Table 11-3. USEPA (2015)[197], Section 3 provides detailed discussions and recommendations about the applications of these BTVs. For metals, natural exceedances of BTVs are common and can be verified through geochemical evaluation (Section 5). These exceedances are expected because trace element concentrations naturally span wide ranges, and their distributions are typically right skewed; it is difficult for any background dataset to fully capture these ranges of concentrations.

Table 11-3. Summary of values used to represent BTV

Value	Acronym	Description
Upper percentile	Not applicable	Value below which a specified percentage of observed background concentrations would fall
Upper confidence limit	UCL	Upper limit of 95% confidence interval
Upper prediction limit	UPL	The value below which a specified number of future independent measurements (k) will fall, with a specified confidence level
Upper tolerance limit	UTL	The UCL of an upper percentile of the observed values
Upper simultaneous limit	USL	Value below which the largest value of background observations falls with a specified level of confidence

11.7.1 Upper percentile

An upper percentile is the value below which a specified percentage of observed background concentrations would fall. For example, the 95th percentile is the value below which 95% of observations may be found. Upper percentiles, when used for point-by-point comparison, can yield excessive false positive rates approaching 100%, which are exacerbated when dealing with small background datasets or background datasets consisting of multiple subpopulations.

Estimates of upper percentiles are reliable (not prone to over- or underestimation) if the background dataset is adequately large and representative of a single population. As noted in (USEPA (1989)[130], Section 4.6.1, p. 4–17), the adequacy of the number of measurements depends on statistical and practical considerations. For example, when calculating a nonparametric BTV, if the risk assessor wants to be 90% confident that at least one of the measurements falls within the upper 5th percentile of the background population, then 45 measurements are needed, ($90\% = 1 - (1 - .05)^{45}$). If for practical reasons, only 27 background samples are collected, the resulting confidence would be 75%, ($75\% = 1 - (1 - .05)^{27}$). This example illustrates that the adequate number of measurements is driven by statistical and practical considerations. Other confidence limits are shown in Table 11-4 and calculated using the following equation:

$$\text{Upper confidence limit} = [1 - (1 - (upper\ percentile\ needed)^{number\ of\ samples})] * 100$$

When planning a background sampling event, the assistance of a professional with statistical expertise is recommended.

Table 11-4. Relationship between number of samples tested and the resulting UCL for an upper 5th percentile of the background population

Number of Samples Tested	UCL	UCL Calculation
10	40%	[1-(1-0.05)^10)]*100=40%
20	64%	[1-(1-0.05)^20)]*100=64%
30	79%	[1-(1-0.05)^30)]*100=79%
40	87%	[1-(1-0.05)^40)]*100=87%
50	92%	[1-(1-0.05)^50)]*100=92%
60	95%	[1-(1-0.05)^60)]*100=95%
70	97%	[1-(1-0.05)^70)]*100=97%
80	98%	[1-(1-0.05)^80)]*100=98%
90	99%	[1-(1-0.05)^90)]*100=90%

11.7.2 Upper confidence limit

The accuracy of any statistical estimate is often quantified by its confidence interval, which is the range of values around the estimate that contains, with certain probability (for example, 95%), the true value of that estimate. The upper limit of this interval is referred to as the upper confidence limit or UCL. The UCL of the mean is the common measure of point exposure concentration in risk assessments. However, since the mean is a measure of the central tendency of a dataset, UCL of the mean should not, under all but select circumstances, be used as a BTV because the result would be excessive false positive results.

11.7.3 Upper prediction limit

The UPL is the value below which a specified number of future independent measurements (k) will fall, with a specified confidence level. For example, the 95% UPL of a single observation (k=1) is the concentration that theoretically will not be exceeded in a new or future measured background concentration with a 95% confidence level. Similar to the upper percentile, the use of UPL based on small background datasets (<50 measurements) with multiple subpopulations for point-by-point comparisons can lead to excessive false positive error rates. If the UPL is calculated based on only one future measurement (k = 1) but more (>1) future measurements are obtained, then the resulting UPL will be unrealistically conservative, and the false positive error rate will be even higher.

11.7.4 Upper tolerance limit (UTL)

The UTL is the UCL of an upper percentile of the observed values. A UTL is designated by its confidence and coverage. Coverage defines the targeted upper percentile estimate (for example, 95% coverage implies that the targeted upper percentile is the 95th percentile). Confidence, on the other hand, is related to the interval in which the true value of the upper percentile should occur. For example, the 99-95 UTL represents the 95% upper confidence level (95% UCL) of the 99th percentile value. This means that 95% of future sampling events generate 99th percentiles that will be less than or equal to 99-95 UTL.

When conducting point-by-point comparisons using the UTL, the false positive error rates stay the same, irrespective of the number of comparisons (USEPA 2015)[197]. The 95-95 UTL has become the most common measure of BTV in practice. However, when dealing with site datasets with numerous measurements, even a reasonable false positive error rate would yield a number of erroneous above-background classifications; this needs to be taken into account when interpreting the data.

11.7.5 Upper simultaneous limit (USL)

The USL is the value below which the largest value of background observations falls with a specified level of confidence (USEPA 2015)[197]. USL is specifically used to mitigate the issue of excessive false positive error rate in point-by-point comparisons. However, due to the sensitivity of the USL to outliers, USEPA (2015)[197] recommended its use only in background datasets devoid of outliers or multiple subpopulations.

11.8 Statistical Tests to Compare Site and Background Datasets

There are many statistical tests that can be used to compare a site dataset to a soil background dataset. The test that is most appropriate in each scenario will depend on the site-specific situation. Statistical tests that can be used are listed in the table below along with the associated advantages and disadvantages of each.

In practice, point-to-point comparisons of individual site data to the BTV are used for either delineation purposes or identifying localized contamination (Geiselbrecht et al. 2019)[15]. More comprehensive comparisons of site and background datasets involve the use of two-sample tests (USEPA 2006)[192]. As noted in USDON (2002)[36], in contrast to point-to-point comparisons, two-sample tests are less prone to falsely identifying above-background conditions.

ASTM E3242-20 (ASTM 2020)[146] provides a list of common two-sample tests used in environmental applications, as reproduced in Table 11-5. Each test is based on a specific statistical hypothesis test. Some of these tests, such as the parametric *t*-test and the nonparametric Mann–Whitney U test, are geared toward the comparison of central tendencies of two datasets, to identify widespread contamination. Other tests focus on the comparison of the upper tails of the two datasets to identify localized contaminations (Geiselbrecht et al. 2019)[15].

Table 11-5. Assumptions, advantages, and disadvantages of common two-sample tests

Source: ASTM E3242-20 (ASTM 2020)[146], Table X4.2.

Test Statistic	Objectives/Assumptions	Advantages	Disadvantages
Quantile test	• Objective is to test for differences in the right tail (largest values) of the site and background concentration distributions. • Nondetects are not among the right tail (largest values) in the pooled set of site and background concentrations.	• Nonparametric: No assumption is required regarding the distributions of the site and background concentrations. • Relatively simple to conduct the test. • No distribution assumptions are necessary. • May be used in conjunction with tests that focus on detecting differences in the mean or median of site and background concentrations.	• May require a large number of measurements to have adequate power to detect differences in site and background concentrations. • Test may be inconclusive if nondetects are present among the largest data values in the pooled set of site and background data.
Wilcoxon rank sum (WRS) test also referred to as the "Wilcoxon-Mann-Whitney test" or "Mann Whitney U test")	• Objective is to test for differences in the medians of the site and background concentration data. • All nondetects are associated with a single detection limit. • The detection limit is less than the smallest detected concentration. • At least 50% of both the site and background concentrations are detected values. • The site and background concentration distributions have the same variance.	• Nonparametric: No assumption is required regarding the type of distributions of the site and background concentrations. • Can be applied to datasets with less than 50% nondetects. • More robust with respect to outliers than parametric two-sample tests, such as Student's *t*-test. • May be used in conjunction with tests that focus on detecting differences in the right tails of site and background distributions.	• Not applicable to cases with less than 50% of detected values. • May conclude that site and background concentrations are derived from the same population when concentrations in right tail differ significantly, so it is important to complement the test with tests that focus on detecting differences in the right tails of site and background distributions.
Gehan test	• A generalized form of WRS test. • Objective is to test for differences in the medians of the site and background concentrations. • Nondetects do not need to have the same detection limits. • The censoring mechanism that generated the nondetects is the same for the site and background concentrations.	• Nonparametric: No assumption is required regarding the distributions of the site and background concentrations. • Can be used with nondetects when multiple detection limits are present. • May be used in conjunction with tests that focus on detecting differences in the right tails of site and background distributions.	• The test is not known as well as the WRS test. • Must assume the same censoring mechanisms apply to the site and background data—difference in censoring mechanism can lead to false results. • May conclude that site and background concentrations are derived from the same population when concentrations in right tail differ significantly, so it is important to complement the test with tests that focus on detecting differences in the right tails of site and background distributions.

Test Statistic	Objectives/Assumptions	Advantages	Disadvantages
Tarone-Ware test	• Objective is to test for differences in the medians of the site and background concentrations. • Site and background concentrations are derived from populations with identical distributions; equality of variances of two datasets should be assessed using Levene's test. • Nondetects do not need to have the same detection limits. • The censoring mechanism that generated the nondetects is the same for the site and background concentrations.	• Nonparametric: No assumption is required regarding the distributions of the site and background concentrations. • Suitable for cases of small datasets with large fractions of nondetects. • Can be used when multiple detection limits are present. • May be used in conjunction with tests that focus on detecting differences in the right tails of site and background distributions.	• The test is not known as well as the WRS test. • Must assume the same censoring mechanisms apply to the site and background data—difference in censoring mechanism can lead to false results. • May conclude that site and background concentrations are derived from the same population when concentrations in right tail differ significantly, so it is important to complement the test with tests that focus on detecting differences in the right tails of site and background distributions.
Student's *t*-test	• Objective is to test for differences in the means of the site and background concentrations. • Nondetects have no significant impact on computed means (for example, less than 15% of measurements are nondetects). • Both site and background mean concentrations are normally distributed. • The site and background data concentrations have the same variance.	• Most powerful test for detecting a shift in the site mean from the background mean, if the site and background data are normally distributed. • To ensure the normality of mean concentrations, certain transformations (for example, Box-Cox) may be performed to normalize the data before the test.	• Nondetects must be substituted before implementing the test. • Requires a statistical evaluation of the equality of variances, such as the Levene's test. • In general, the power will be less than that of the WRS test if the mean values are not normally distributed (when site and background datasets are small and/or have large variances). • Outliers can affect the test results. • Not well suited for datasets that contain a large number of nondetects. • May conclude that site and background concentrations are derived from the same population when concentrations in right tail differ significantly, so it is important to complement the test with tests that focus on detecting differences in the right tails of site and background distributions.

Test Statistic	Objectives/Assumptions	Advantages	Disadvantages
Satterthwaite or Welch's two-sample *t*-test	• Objective is to test for differences in the means of the site and background concentrations. • Nondetects have no significant impact on computed means (for example, less than 15% of measurements are nondetects). • Both site and background mean concentrations are normally distributed. • Site and background data distributions are expected or known to have unequal variances.	• Test can be used when the site and background distributions have unequal variances.	• Nondetects must be substituted before implementing the test. • In general, the power will be less than that of the WRS test, if the mean values are not normally distributed (when site and background datasets are small and/or have large variances). • Outliers can affect the test results. • Not well suited for datasets that contain a large number of nondetects. • May conclude that site and background concentrations are derived from the same population when concentrations in right tail differ significantly, so it is important to complement the test with tests that focus on detecting differences in the right tails of site and background distributions.
Two-sample test of proportions	• Objective is to test for differences in the proportions of the site and background concentrations above a given cutoff level, such as an applicable screening or action level. • Test may be used when more than 50% of the site or background datasets are nondetects.	• Nonparametric: No assumption is required regarding the distributions of the site and background concentrations. • Can be used when many nondetects are present.	• Use of inappropriate cutoff values can produce misleading test results.
Levene's test	• Objective is to test the equality of variances of site and background concentrations. • Site and background concentrations are derived from approximately normal distributions.	• Reasonably robust to non-normality of the investigated datasets.	• Nondetects must be substituted before implementing the test. • Not well suited for datasets that contain a large number of nondetects.
Shapiro-Wilk test	• Objective is to test the normality of site and background concentrations.	• Well-known normality test. • A variant of the test, the Shapiro-Francia test, is useful for samples sizes greater than 50.	• Nondetects must be substituted before implementing the test. • Not well suited for datasets that contain a large number of nondetects.

Reprinted, with permission, from ASTM E3242−20 Standard Guide for Determination of Representative Sediment Background, copyright ASTM International, 100 Barr Harbor Drive, West Conshohocken, PA 19428. A copy of the complete standard may be obtained from ASTM International, www.astm.org.

In many instances, both widespread contamination and localized contamination should be tested concurrently. Selection of the appropriate test is contingent on the specific conditions, including the target statistics of interest and the type of distributions displayed by the investigated site and background datasets, as well as their variance equivalency. Geiselbrecht et al. (2019)[15] developed a decision flow diagram for selecting the appropriate types of tests, as reproduced in

Figure 11-11. In practice, nonparametric tests are often preferred because they do not require any specific distributional assumption about the investigated site and background data.

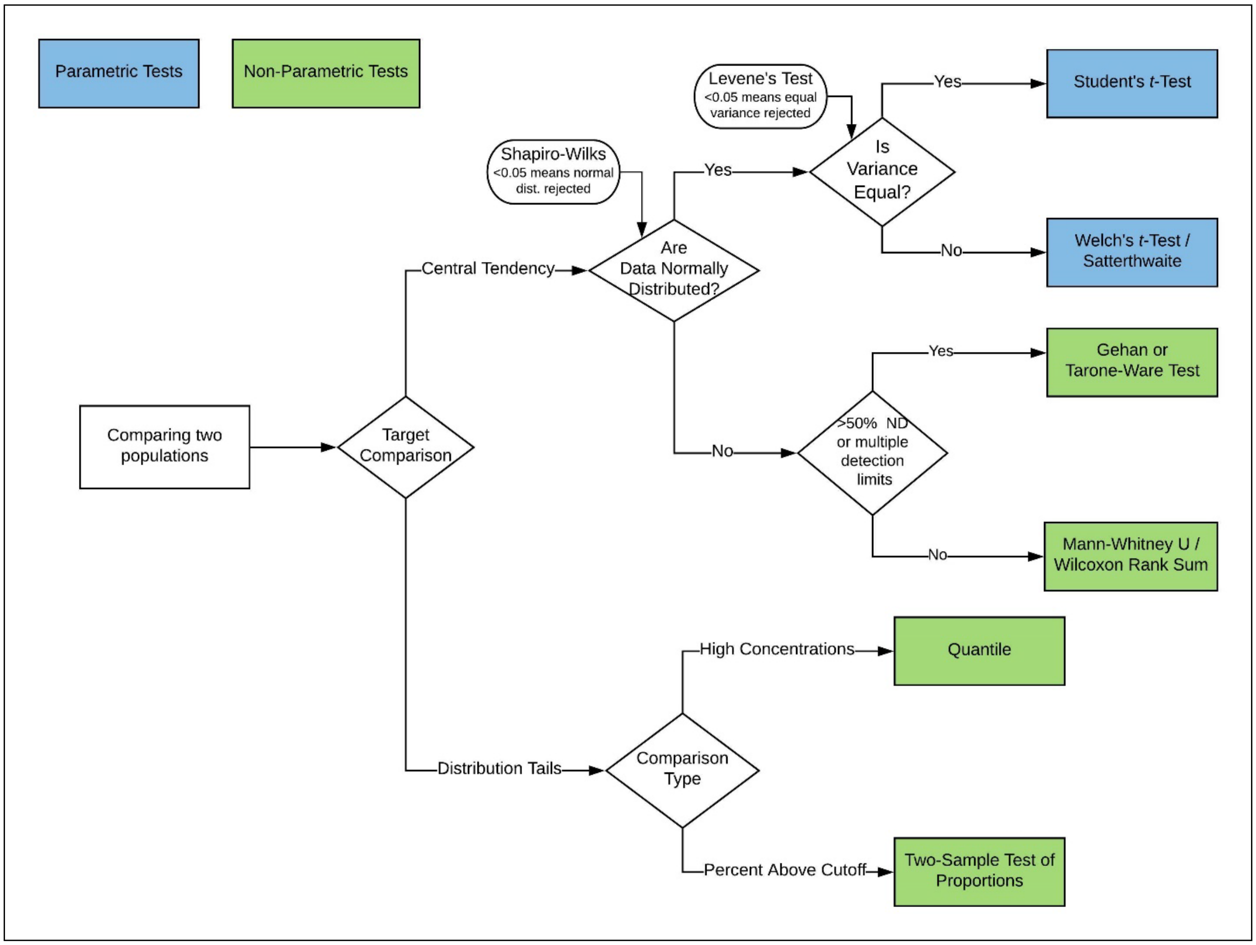


Figure 11-11. Statistical tests for comparison of two datasets.

Source: Adapted from Geiselbrecht et al. (2019)[15]

11.9 Statistical Software

There are many readily available software packages that can be useful for background data analysis (Table 11-6). The following discusses some commonly used software but is far from comprehensive. Other software may be equally pertinent, such as ArcGIS or Visual Sample Plan for specific project needs, but this summary is meant only to make the reader aware of the existence of broadly useful statistical software and to highlight features that may be most useful for their situation.

While most of the statistical analysis programs listed below will have the capability to conduct many of the analytical methods required for background statistical analysis, not all programs will be able to easily conduct all methods. Choosing statistical software for the background analysis needs may have other considerations than simply cost or capabilities (for example, limited statistical experience or regulatory agency requirement).

Because software is subject to continual development, the statements made herein regarding specific software programs will become outdated over time. Future versions of programs described herein may feature expanded capabilities, whereas some programs that are currently popular may become less so in the future. Therefore, it is recommended that users research available software options periodically and treat the recommendations presented herein as general guidelines.

It is important that someone using statistical software packages understand the statistics behind the software and the uncertainties. Appropriate use of software requires valid data and someone with sufficient understanding of both statistics and the specific situation being evaluated to make reasonable professional judgment decisions using multiple lines of evidence (of which software statistical analysis would be only one). Decisions should never be made using the software analysis results alone without other lines of evidence.

Table 11-6. Statistical software commonly used in computations involving background concentrations in risk assessments

Software	Capability	Advantages	Disadvantages
USEPA ProUCL	Parametric and nonparametric two-sample tests Multiple options for parametric and nonparametric BTV calculations; options for distribution tests and datasets with nondetects Outlier identification tests Trend tests	• Free • Comprehensive user guide • Relatively easy to master • Benchmark package in risk assessment • Subjected to QA/QC • Provides consistent method for regulatory agencies • Recommended by USEPA • Focus is on environmental data	• Does not have the ability to evaluate data for multiple populations (for example, a moderate and severe contaminant distribution both in the same dataset) • Inflexible input-output formats • Apart from trend test, limited to univariate modeling • Some minor conflicts in presented results, such as nonreproducible results for UCLs based on bootstrap methods • Graphical outputs limited to histograms, box plots, trend graphs, and Q-Q plots • Only compatible with Windows PCs

Software	Capability	Advantages	Disadvantages
Microsoft Excel	Point-by-point comparisons; parametric two-sample tests; limited graphical capabilities	• Easy to use • Often already owned • Subject to QA/QC • Easy data transfer to other statistics programs • Simple tools for basic data visualization	• Limited built-in statistical capabilities • Add in packages or difficult coding needed to calculate BTVs, advanced parametric, and nonparametric tests, and treatment of nondetects • One-time fee or annual license for Excel; add-ins may have a separate one-time or annual fee • Limited advanced graphical capabilities
R code	Open-source code offering a wide variety of statistical and data visualization tools, including all those offered by ProUCL	• Open source, free-to-use software • Large online community support • Customizable statistical analysis capabilities • Customizable graphical outputs • Extensive statistical analyses available	• Not subject to rigorous QA/QC • Requires specialized training, including coding capabilities • User required to manually evaluate dataset for multiple populations (for example, a moderate and severe contaminant distribution both in the same dataset) • Statistical training requirements are significantly higher than for other software • Requires coding capability by the user
Python	Open-source code offering a wide variety of statistical and data visualization tools, including all those offered by ProUCL	• Open source, free-to-use software • Large online community support • Easy to integrate with ArcGIS • Customizable statistical analysis capabilities • Customizable graphical outputs	• Not subject to rigorous QA/QC • Requires specialized training • Requires coding capability by the user
Integrated Statistics Programs (for example, SAS, SPSS, Stata, Statistica, Minitab, MATLAB, Wolfram Mathematica)	Commercial packages offering a wide variety of statistical and data visualization tools, including all offered by ProUCL	• Subject to QA/QC • Graphic capabilities can be somewhat flexible • Require less, or zero, coding knowledge to conduct most analysis • Likely to include software beyond the scope of only background evaluations	• Annual or one-time licensing fees • May require specialized training • Inflexible input-output formats • Slightly limited in analysis capabilities compared to languages such as R or Python

11.9.1 USEPA's ProUCL

USEPA's ProUCL software, currently version 5.1 released May 2016, includes a thorough user guide as well as a technical guide. Version 5.2 is under development and is poised to be released soon. The overwhelming majority of user functionality will be the same between the two versions. The user and technical guides have been updated for version 5,2 as well. The guides walk users through an extensive suite of statistical procedures, including two-sample hypothesis tests, trend analysis methods, BTV calculation, and UCL estimation. The BTV and UCL modules provide several different methods for calculating BTVs and UCLs of the mean. Additionally, there are multiple training presentations available for those just getting started or looking to master the use of ProUCL. Both the ProUCL software and the trainings can be found at USEPA's ProUCL website.

USEPA and state regulatory agencies often use ProUCL because it is recommended by USEPA, it's free and easy to use, and it provides a consistent way to regulate all parties fairly. State regulatory staff often do not have statisticians and do not always have time to learn statistics as thoroughly as they may like to. This software allows users without formal statistical training to obtain accurate results for common statistical procedures. The user interacts with the software via a graphical user interface (GUI) that does not require coding or statistical knowledge (Figure 11-12). As stated earlier, it is important that one using this software understands the statistics being used and the uncertainties involved so they can make an informed decision based on multiple lines of evidence.

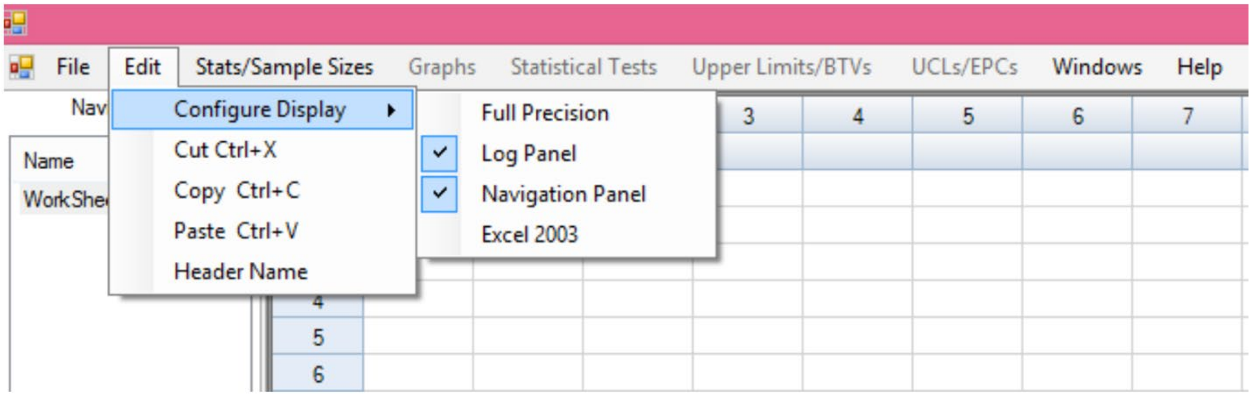


Figure 11-12. Example USEPA ProUCL interface.

Source: (USEPA 2015)[199].

Additionally, users can create graphical outputs, though these are limited to histograms, box plots, Q-Q plots, and trend analysis plots (Figure 11-13). ProUCL does not handle other graphical outputs discussed in later sections, such as index plots, ratio plots, and bar plots.

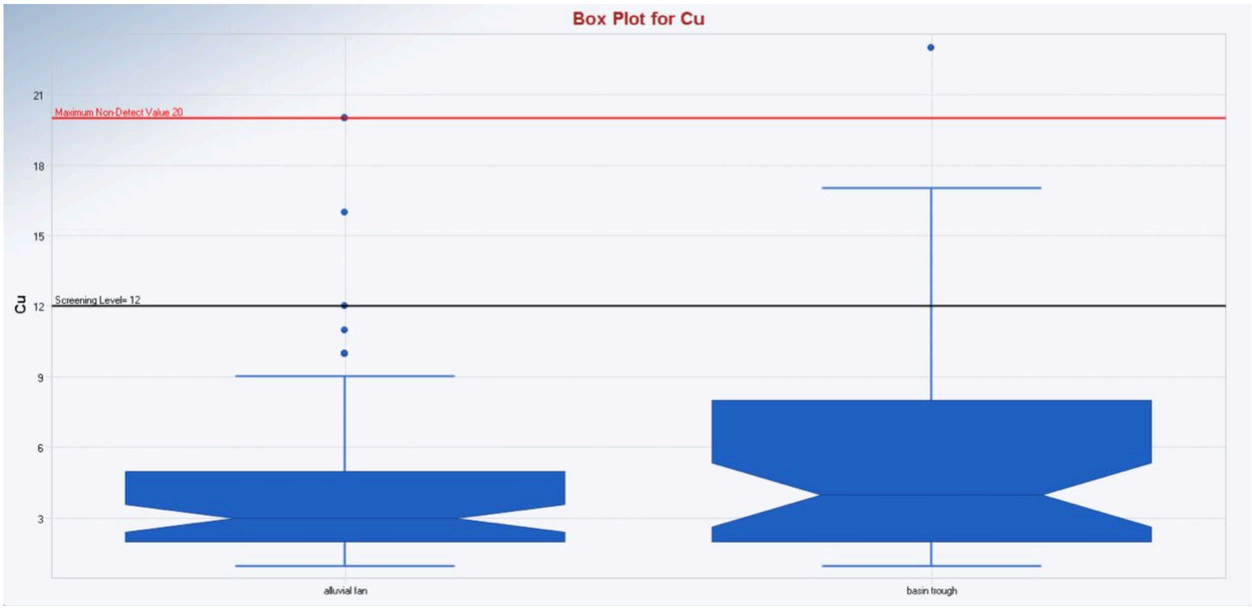


Figure 11-13. Example USEPA ProUCL box plots.

Source: (USEPA 2015)[199].

11.9.2 Microsoft Excel

Microsoft Excel is likely the most well-known software for data entry and tabular manipulation. This software comes with the backing of Microsoft and is directly integrated as part of the ProUCL software, as using Excel formats for input and output within ProUCL is standard. Within Excel itself, there are basic tools for data visualization, a free Excel Analysis Toolpak plugin, and a programming language (Visual Basic for Applications or VBA) to conduct more advanced analyses ("macros") than those provided in Excel's built-in functions. There is a limited free VBA version and a more robust paid version for both Windows and Mac OSX. There are also paid and free statistical analysis plugins for both Windows (for example, XLSTAT and Analyse-it) and Mac OS X (Real Statistics Using Excel). With this in mind, Excel was certainly not intended as a full-fledged background statistical analysis tool and is most suited to simple exploratory data analysis.

11.9.3 R

R is an open-source language and coding environment for statistical computing and graphics. Due to this classification as a language, as opposed to simply a software package, the background statistical analysis possibilities are basically limitless. If a user has the resources and know-how to create an analysis method for the problem in question, it can be tailored to virtually any site-specific problem or issue complexity level. R is also capable of producing all the same statistical analysis that can be achieved in ProUCL.

Because R is a coding language, it does require significantly more statistical understanding to use over USEPA's ProUCL software. To help with this challenge, there are numerous packages available through R's Comprehensive R Archive Network (CRAN). While those packages do not necessarily have any guarantees to their reliability, many are heavily used and regularly maintained by skilled statisticians. Additionally, all code used in building functions included in those packages is visible to R users should they desire to inspect the underlying code further.

One of the biggest advantages that R provides is its ability to produce high-quality graphical deliverables. In addition to the graphic capabilities that come with the core R software, freely available packages, such as *ggplot2* and *plotly,* can provide customizable and interactive (*plotly*) graphical results. This can be taken a step further using R's Shiny app suite, which allows for web hosting of interactive tools. This app suite allows those receiving the deliverable to visualize data or statistical results in flexible ways. Bearing all of that in mind, the background statistical analysis conducted within R is only as reliable as the individual or team that conducted the analysis, so careful quality control and assessment are advised.

R is free and is compatible with Microsoft Windows, Mac OS X, and Linux. Some existing packages within R may function only on a subset of these operating systems. R can be obtained from the R-Project website.

11.9.4 Python

Python is a freely available coding language capable of many of the same things as R. However, this language and its supporting packages tend to focus more on programming and integration with other software (for example, GIS) as opposed to the statistician-oriented approach within R. Python users have developed many freely available packages that provide for a widely customizable approach to statistical analysis. However, at the time this guide was written, there are significantly fewer purpose-built statistical analysis packages available for Python then there are for R.

11.9.5 Integrated statistical programs

There are many commercially available statistical software programs that effectively combine database, graphical, and statistical capabilities with a relatively accessible user interface. The key benefit of using such programs is that they are intuitive to use and can quickly generate decently attractive graphics and accurate analytical results based on a range of typical assumptions and inputs. The downsides of such programs can be lack of flexibility in graphics and analysis relative to programs such as R or Python, and relatively high cost considering most other presented options are free to the user. Integrated statistical programs also lack the focus of a program such as ProUCL, which includes functions specific to the analysis of environmental data. Examples of integrated statistical programs include, but are not limited to SAS, IBM SPSS, Stata, TIBCO Statistica, SYSTAT, MYSTAT, and Minitab. The most popular programs require payment to access, typically through an annual subscription service. GNU PSPP is a free program intended as an alternative to SPSS and similar programs. Currently, PSPP provides only a small number of statistical methods relative to SPSS or the other programs listed above.

While limited relative to R or Python, the integrated statistical programs noted above have sufficient capabilities to support the estimation of background conditions or to compare site conditions to background. These capabilities include summary statistic calculation, hypothesis testing tools (for example, *t*-test and Wilcoxon Rank Sum test), graphical comparison tools, and other tools, depending on the specific program. For background analysis, integrated statistical programs may be more advanced than necessary; the statistical methods and graphics capabilities in ProUCL or potentially less expensive software, such as Excel add-ins like XLStat, or data visualization programs like SigmaPlot, may be sufficient to characterize soil background conditions.

12 REGULATORY FRAMEWORK FROM STATE SURVEY

On March 9, 2020, a survey on use of soil background for risk assessment was sent to designated individuals who provide responses to ITRC surveys for each U.S. state and the District of Columbia. The purpose of the survey was to elicit current state practices for addressing soil background through the risk assessment processes.

This survey focused on finding out how individual states[1] currently handle soil background with respect to definitions, state guidance and regulations, use in risk assessment, and other factors. The survey did not target federal agencies or guidance available from other countries. Also, because in most cases the responses came from one person representing the state, they may not be complete, current, or represent the range of soil background uses in risk assessment from multiple jurisdictions within a state. Readers should consult with the appropriate state agencies for the most current approaches relevant to specific sites.

12.1 **Description of State Survey**

The survey included a mix of multiple choice and open-ended questions. The key questions were:

- Does the state regulatory agency have a definition for natural and/or anthropogenic ambient soil background? If so, please provide the definition(s) and a link to the document where it is presented.
- Does the state regulatory agency recommend using natural and/or anthropogenic ambient soil background values in risk assessment?
- Has the state established statewide or regional natural and/or anthropogenic ambient soil background values? If so, please provide a link.
- Have studies to establish statewide or regional natural and/or anthropogenic ambient soil background been conducted in the state with or without agency approval/endorsement? If so, please provide the reference or URL.
- Will the regulatory agency allow geochemical evaluations to be used to support risk assessment?

12.2 **Overview of State Survey Results**

A total of 36 states, including the District of Columbia (71%), responded to the state survey, leaving 15 states (29%) with no responses, as shown in Table 12-1 Summary of the State Survey

---

[1] For purposes of this section, the word “state” or “states” includes the District of Columbia.

Summary. Figures 12-1 through 12-5 summarize the findings of the survey based on the above questions.

In general, of the 36 responses total, most address natural and/or anthropogenic ambient soil background in some way:

- 61% indicated that states have definitions for natural soil background or for both natural and anthropogenic ambient soil background; no states responded that they have a definition for only anthropogenic ambient soil background.
- 92% indicated that state agencies recommend using soil background values (natural, anthropogenic ambient, or both) sometimes in risk assessment.
- 54% indicated that states have established soil background values, either natural and/or anthropogenic ambient.
- 59% indicated that states have statewide or regional soil background studies.
- 94% indicated that state agencies either allow geochemical evaluations or sometimes allow them, depending on site-specific conditions, to support risk assessments.
- It should be noted that states that did not submit survey responses could have some level of guidance regarding natural or anthropogenic ambient soil background and its use in risk assessment. Thus, the percentages shown above and in Figure 12-1 through Figure 12-5 refer only to the 36 responses received.

12.3 **State Survey Results**

Figure 12-1 shows the breakdown of responses to the first question, whether the states that responded to the survey have definitions for both natural and/or anthropogenic ambient soil background. For reference, Table 12-2 Summary of the State Survey Summary includes any definitions with comments or links that were provided in the survey responses.

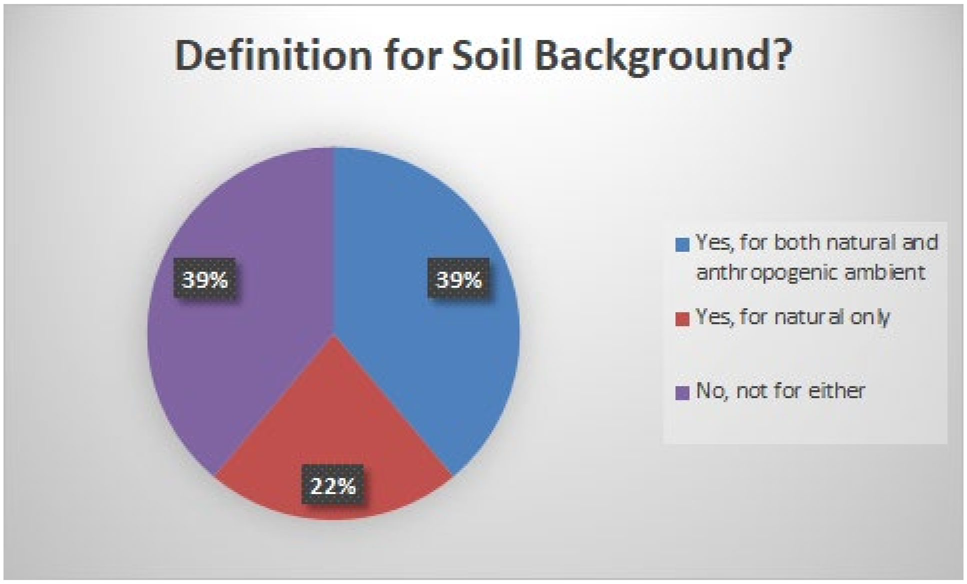


Figure 12-1. Definitions for natural and/or anthropogenic ambient soil background.

Figure 12-2 summarizes the state survey responses regarding the use of natural and/or anthropogenic ambient soil background values in risk assessment. The survey question asked if state agencies "recommend" the use of soil background values in risk assessment. Although 8% of the responding states reported" no," it is possible they may allow it while not actively recommending it. The survey did not address whether those states might "allow" use of soil background in risk assessment.

For further reference, the state guidance summary table (Table 13-1 Summary of the State Survey Summary) identifies specific guidance documents, some identified through the state survey and others researched independently, that address soil background use in risk assessments.

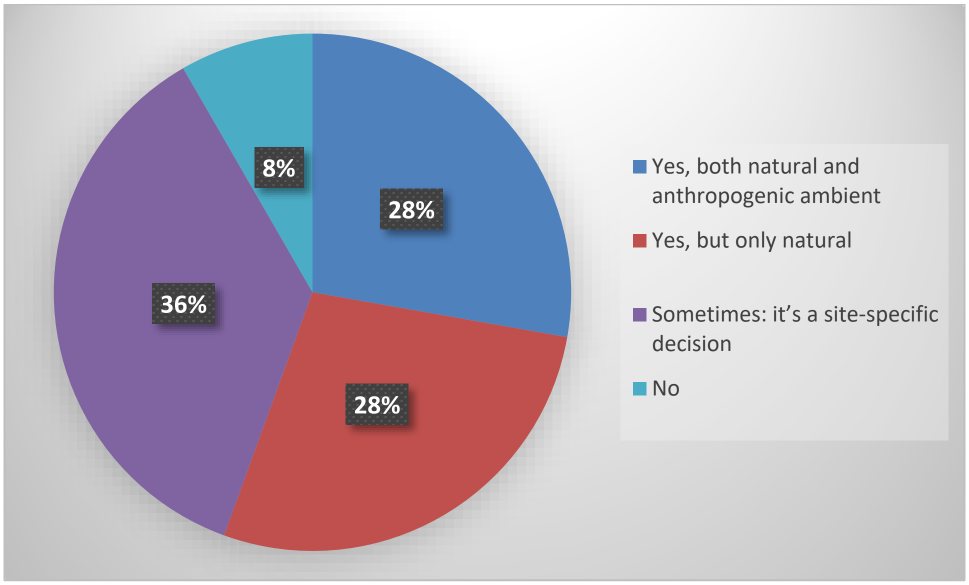


Figure 12-2. Recommend using natural and/or anthropogenic ambient soil background values in risk assessment.

Figure 12-3 shows that more than half the survey responses indicated that the states have established statewide or regional soil background values.

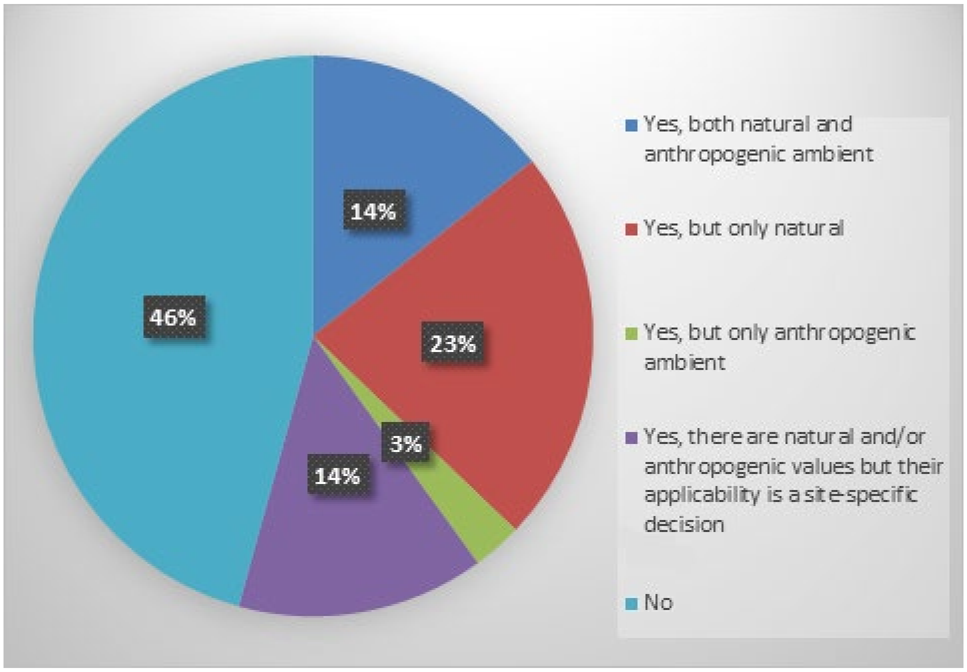


Figure 12-3. Established statewide or regional background soil values.

More than half of the survey responses (Figure 12-4) indicated that statewide or regional natural and/or anthropogenic ambient soil background studies have been conducted within the state. These studies may or may not have been prepared with agency approval or endorsement but rather by independent organizations or researchers. For example, some states will consider soil background surveys prepared by the U.S. Geological Survey, with or without state involvement, to be valid for use in risk assessments.

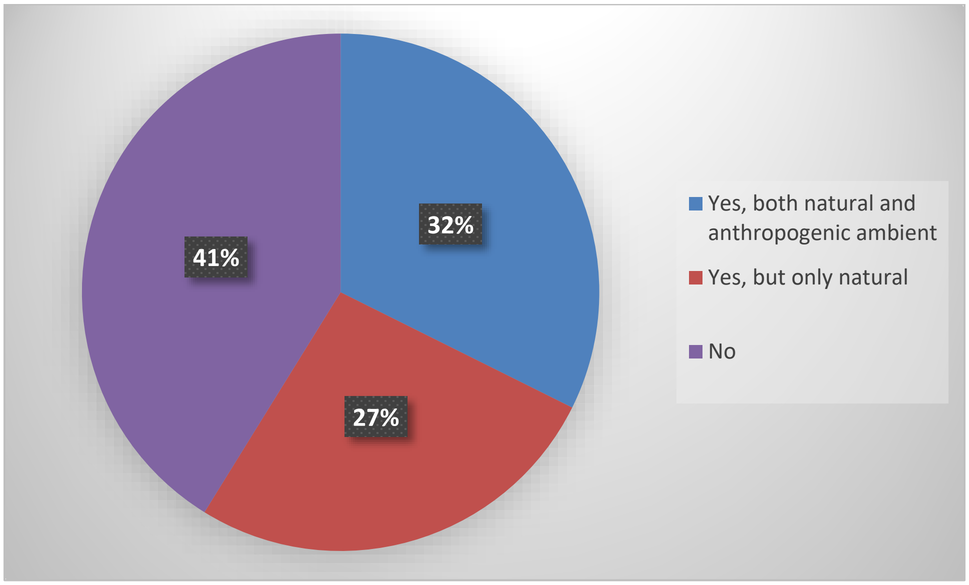


Figure 12-4. Studies to establish statewide or regional soil background with or without agency approval/endorsement.

Figure 12-5 shows that most of the states that responded to the survey (94%) allowed the use of geochemical evaluations to support risk assessments in some way.

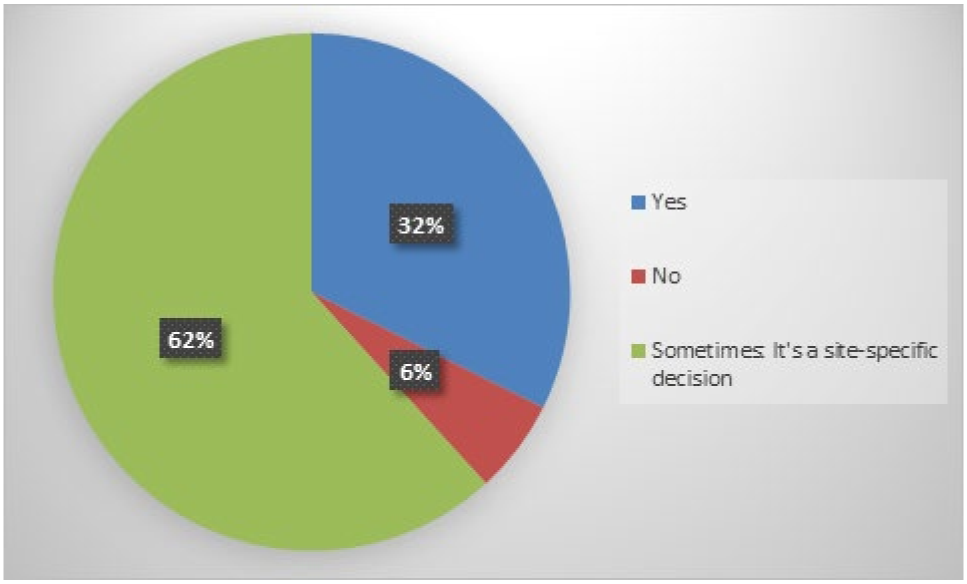


Figure 12-5. Allow geochemical evaluations to support risk assessment.

13 EXISTING GUIDANCE AND STUDIES

This section presents the results of research conducted by the SBR team to identify existing guidance and soil studies that are available regarding soil background and its use in risk assessment. The information is presented in two tables, one listing existing guidance (Table 13-1 Summary of the State Survey Summary) and the other listing existing studies (Table 13-2 Summary of the State Survey Summary); hyperlinks are provided when possible. The information on both tables is categorized by author as follows: federal agencies, state agencies, tribal agencies, private agencies, and other. *(Note: This information was thought to be accurate at the time the document was published in 2021. Please contact the appropriate agency for updated information.)*

The guidance listed on Table 13-1 Summary of the State Survey Summary includes a variety of types of guidance, such as guidance documents, rules, and regulations, some including background values. Table 13-1 Summary of the State Survey Summary also indicates some basic soil background topics that are discussed in this document and whether the topic is addressed in a specific guidance. The following is a list of these topics, along with their locations in this document:

- Soil background values (Section 3)
- Soil background (default and/or site-specific) in risk assessment (Section 4)
- Geochemical evaluations (Sections 5 and 6)
- Environmental forensics (Section 7)
- Sampling (Section 9)
- Analytical methods (Section 10)
- Statistics (Section 11)

14 CASE STUDIES

14.1 Minnesota Pollution Control Agency (MPCA) Soil Background Case Study

This case study presents an example of a state regulatory agency establishing statewide default natural soil background values using data from an existing soil natural background study. The MPCA established their first set of statewide soil background values using the U.S. Geological Survey's (USGS) Geochemical and Mineralogical Data for Soils of the conterminous United States ((MPCA 2016)[459], (Smith et al. 2013)[40]). The objective of USGS's study was to sample for natural soil background, avoiding anthropogenic sources.

Soil screening values used to evaluate cleanup sites are generally derived by state regulatory agencies using the USEPA methodology (USEPA 1996)[427], toxicity values, exposure parameters, and chemical-specific information. These values are purely risk based, intended to protect human health or the environment, and do not take natural soil background into consideration. MPCA's soil screening values are called soil reference values (SRVs). When MPCA revised their SRVs to incorporate more recent toxicity and exposure parameters, they recognized that eight of the revised SRVs for inorganics could potentially be below natural background concentrations in soil. Those eight inorganics were aluminum, arsenic, barium, chromium, cobalt, iron, thallium, and vanadium. This case study describes the process MPCA used to determine whether the SRVs were below natural soil background concentrations and, if so, derive default soil background threshold values (BTVs).

14.1.1 Existing data

Since funding could not be obtained for MPCA to conduct a statewide soil background study, the agency evaluated existing studies that could be used to determine statewide background concentrations for the inorganics of interest.

A report by the Minnesota Geological Survey (Thorleifson et al. 2007)[417] was considered, but the purpose of obtaining that data was for mineral exploration rather than to determine natural background concentrations of inorganics. Samples were obtained from areas thought to have higher concentrations of minerals rather than areas that were not impacted by anthropogenic sources. In addition, the study included both soil and sediment data that would not be easy to identify and separate (Thorleifson et al. 2007)[417].

Two studies often used to determine soil background concentrations of metals, the USGS Boerngen and Shacklette studies ((Boerngen and Shacklette 1981)[420], (Shacklette and Boerngen 1984)[421]), were considered. These studies include concentrations of metals and other trace elements in soil of the conterminous United States, with samples collected in the 1960s and 1970s. However, the USGS more recently conducted a national-scale geochemical and mineralogical survey of soils of the conterminous United States from 2007 to 2013 ((Smith et al. 2013)[40] and (Smith et al. 2014)[213]).

USGS Soil Geochemical and Mineralogical Background Study

The most recent and comprehensive U.S. soil geochemical and mineralogical background study was conducted by the USGS from 2007 to 2013. The USGS developed protocols, performed pilot studies, and collected soil samples across the conterminous United States to determine natural background concentrations for constituents, including inorganics (Smith et al. 2013)[40]. Sites were selected to represent a spatially balanced density of approximately one site per 1,600 km^2 from areas unlikely to be contaminated. In total, 14,434 samples from nearly 5,000 sites were collected across the United States. Samples at each site included a surface sample collected from 0 to 5 centimeters, a composite of the soil A horizon, and a deeper sample called the soil C horizon, not exceeding a depth of approximately 80–100 cm.

Samples were prepared for analyses using standard USGS analytical protocols, which included air drying at ambient temperatures, sieving to less than 2 mm, followed by grinding to less than 150 µm. Prepared samples were aggressively digested using a four-acid mixture (hydrochloric, nitric, hydrofluoric, and perchloric) heated to temperatures between 125 and 150°C. This digestion method dissolves most silicate, sulfide, oxide, and carbonate minerals. Total element concentrations for 41 elements were measured by inductively coupled plasma-atomic emission spectrometry (ICP-AES) and by inductively coupled plasma-mass spectrometry (ICP-MS). Mercury, selenium, and arsenic were analyzed by single-element analytical methods. Organic carbon concentrations for the 0 to 5 cm soil and soil A horizon were calculated from the difference between analyzed inorganic carbon and total carbon. Mineralogy was determined for splits of the soil A and C horizons by a quantitative X-ray diffraction method. Smith et al. (2013)[40] provides details of sample site selection and analytical methods, as well as all geochemical and mineralogical data. Interpolated maps displaying the data are available in Smith et al. (2014)[213] and were also integrated into an interactive website (Smith et al. 2019)[423].

USEPA methods (SW846 3050B or 3051A) used at cleanup sites are similar to analytical methods used by the USGS, but distinctions are evident for preparation methods and, particularly, for sample digestion. The USEPA method includes sample homogenization, but it is up to each individual lab to decide how to do this. USEPA methods also use a less aggressive aqua regia digestion, although there can be differences for digestion of some inorganics. Since USEPA digestions are not intended to yield total or near-total values, geochemical results are not directly comparable to results obtained using USGS's analytical methods. It is likely that concentrations determined for the same soil sample using the more aggressive USGS methods will generally be higher than concentrations determined by USEPA's methods. This, however, may not always be the case and will vary depending on specific inorganic constituents as well as local geology. Sample mineralogy also may influence the effectiveness of the two digestion methods. Overall, inconsistencies between the results of each analytical approach are due to the differences in preparation and digestion methods, and caution must be used when comparing data.

MPCA chose to use this more recent USGS survey of soils to establish state background values. Soil from a total of 137 sites was collected in Minnesota as part of this study. Laboratory analysis of the soil samples was completed in 2013. USGS took the following measures to ensure samples were not collected from areas impacted by anthropogenic sources: no samples were collected within 200 m of a major highway, within 50 m of a rural road, within 100 m of a building or structure, or within 5 km downwind of active major industrial activities such as power plants or smelters. Samples at each site were obtained from three different intervals below the ground surface: 0 to 5 cm, A horizon, and C horizon (Smith et al. 2013)[40]. Splits of all samples collected by the USGS are currently stored in a sample archive in Denver, Colorado.

14.1.2 USGS and USEPA analytical methods

Soil preparation and analytical methods used by the USGS differ from the USEPA methods typically used for cleanup sites in state and federal remediation programs. Differences in concentrations resulting from the two methods depend on the geological characteristics of the area where the sample was taken, and also on the chemical nature of each analyte. As a result, concentrations obtained using these two methods are not generally comparable.

Sample preparation using the USGS method begins with air drying at ambient temperature, sieving to less than 2 mm and grinding to less than 150 µm. The USEPA method (SW846 3050B or 3051A) begins with suggested sample homogenization but leaves the final decision on how to prepare an individual sample up to the laboratory. A consequence is that the USGS preparation method results in finer grained samples compared to samples prepared using USEPA's method, thus leading to more surface area available for digestion.

In addition, a near-total four-acid digestion is used in the USGS method. Sieved and ground samples are digested with a mixture of hot nitric, hydrochloric, hydrofluoric, and perchloric acid. Samples for arsenic analyses are first fused in a mixture of sodium peroxide and sodium hydroxide and then dissolved in hydrochloric acid. The USEPA method uses a less aggressive aqua regia digestion. For arsenic and thallium analyses, digestion is accomplished using hydrogen peroxide and nitric and hydrochloric acid. Digestion for the other inorganics is accomplished using hydrogen peroxide and nitric acid.

Methods of detection used for the eight inorganics in the USGS analytical method are:

- inductively coupled plasma–atomic emission spectrometry (ICP-AES) for aluminum, chromium, iron, and vanadium
- ICP-MS for cobalt and thallium
- hydride-generation atomic absorption spectrometry for arsenic

Methods of detection used for the eight inorganics using the USEPA analytical method are:

- inductively coupled plasma–atomic emission spectrometry (ICP-AES) for aluminum, barium, chromium, cobalt, iron, and vanadium
- ICP-MS for arsenic and thallium

The key distinctions between USGS and USEPAs analytical methods are:

- The USGS preparation method results in more surface area available to be digested, whereas the USEPA preparation method does not require a specific procedure and may result in less surface area available for digestion.

- The USGS digestion is intended to result in a true total concentration of all metals in the soil sample, while USEPA's digestion is not intended to be a total digestion and thus results in lower metals concentrations of some inorganics in the soil sample.

In most cases the USGS method results in higher concentrations than the USEPA method. In these cases, soil background values based on the USGS data would be higher than those based on USEPA's method. Since USEPA's method is used to characterize site concentrations at cleanup sites, this would result in an inappropriate comparison. Higher site concentrations would be incorrectly characterized as being representative of background since USGSs background value would be higher than if the sample were analyzed using USEPA's method.

14.1.3 USEPA method reanalysis

Since it was not appropriate to directly use USGS's data to establish a soil background value for a cleanup site, MPCA reanalyzed a subset of USGS soil samples using the USEPA's method to evaluate the differences in results between the two methods.

The archived USGS soil samples were outside of their holding time when they were reanalyzed for the MPCA study. Samples had been stored at the USGS Denver, Colorado, facility at room temperature since collection. Although the longer than normal storage time under room temperature conditions is not expected to significantly impact the results of the analyses of the eight specific inorganics that were reanalyzed, this has not been confirmed by laboratory testing.

MPCA could not reanalyze more than one sample depth due to cost constraints. Even though human and ecological receptors could be exposed to both the 0 to 5 cm below the ground surface and A horizon, MPCA chose to reanalyze the 0 to 5 cm depth based on the following rationale:

- The 0 to 5 cm depth was most relevant to human receptor exposure.

- Results from the 0 to 5 cm and the A horizon appeared to be similar.

A subset of 48 USGS 0 to 5 cm depth soil samples (out of 137 total available Minnesota samples) was chosen to be reanalyzed using the USEPA method. Samples were selected based on the following factors:

- 24 samples were chosen in a stratified random fashion to represent differences in geology across the state. This was accomplished by choosing samples with varying abundances of quartz and clay, as proportions of these two silicate minerals are a good proxy for soil parent materials.

- 14 samples were chosen in a stratified random fashion to represent the different major land resource areas and land cover across the state, which incorporated variations in vegetation and climate.

- 10 nonrandom samples were chosen based on their proximity to major cities and the Minnesota River to capture areas with potentially higher populations. They were not chosen to capture anthropogenic ambient soil background; the USGS study was aimed at collecting natural soil background.

Three of the archived samples originally chosen for reanalysis using USEPA methods could not be located in the USGS archive so they were not analyzed (two chosen based on differences in geology and one chosen based on proximity to major cities and the Minnesota River). A total of 45 USGS samples were reanalyzed using USEPA's soil analytical method. Thirty-six of these were chosen in a stratified random fashion, and nine were chosen based on their proximity to major cities and the Minnesota River. Randomly chosen datasets are recommended to be used when performing statistical calculations.

Samples were analyzed using USEPA Method 6010C (ICP-AES) for aluminum, barium, chromium, cobalt, iron, and vanadium and USEPA Method 6020A (ICP-MS) for arsenic and thallium.

14.1.4 Statistics

USEPA's ProUCL software and associated technical guide were used to evaluate for potential outliers, determine dataset distributions, and calculate a soil background value (BTV) from the Minnesota sample subset of the USGS study data (USEPA 2013)[419].

Outliers were determined using the Rosner outlier test and further evaluated using Q-Q plots, histograms, and box plots generated using ProUCL, version 5.0. BTV calculations were performed both with and without outliers using ProUCL's "Upper Limit-BTVs" module.

ProUCL' s "Upper Limit-BTVs" calculation was used to determine the BTV for each inorganic. This ProUCL function can be used for datasets with or without nondetect data. It determines whether the dataset is parametric (fits a normal, gamma, or lognormal distribution) or nonparametric and provides the following types of BTVs:

- upper percentiles
- upper prediction limits (UPL)
- upper simultaneous limit (USL)
- upper tolerance limit (UTL)

According to USEPA's ProUCL guidance, upper percentiles and UPLs are not appropriate to use when establishing a default soil BTV where there will be a large number of comparisons to the BTV (expected for statewide default values). Using UPLs, which are designed to minimize false positives, will in this case actually result in high false positive error rates, meaning a larger number of soil concentrations reflecting real background levels would exceed the BTV and appear as if they are contaminated. Another concern with the UPL is that it is calculated based on an expected number of future samples. With a statewide dataset, it is unknown how many future samples will be compared to the BTV.

USLs are also not appropriate to use when establishing a default BTV where there will be a large number of comparisons to the BTV (expected for statewide default values). A USL represents an upper limit on the largest value of the dataset and would result in high false negative error rates, meaning a large number of soil concentrations that represent actual contamination would not exceed the BTV and appear to represent background levels.

UTLs were chosen to establish the BTVs because they are appropriate to use when establishing a default BTV where there will be a large number of comparisons to the BTV (expected for statewide default values). Using UTLs is not likely to result in large false positive or false negative error rates. MPCA chose to calculate 95-95 UTL representing the 95% upper confidence limit on the 95th percentile.

14.1.5 Results

Differences in results from the USGS and USEPA's analytical methods varied widely among the inorganics. For some parameters the differences were very little, but for others they were significant.

MPCA investigated the possibility of using the USEPA method by reanalyzing 45 samples to establish soil BTVs. Since the 45 samples that were reanalyzed using the USEPA method were not all chosen randomly (only 36 were), a comparison was performed between the random (36 samples) and nonrandom (45 samples) datasets to determine whether they were similar. MPCA compared the distributions, graphical displays, and outliers of the five datasets listed below using USEPA's ProUCL.

- USGS method analyzed 137 samples (all samples from Minnesota)
- USGS method analyzed 45 samples (subset including nine nonrandom data points)
- USEPA method reanalyzed 45 samples (subset including nine nonrandom data points)
- USGS method analyzed 36 samples (subset of only randomly selected data points)
- USEPA method reanalyzed 36 samples (subset of only randomly selected data points)

Since these datasets appeared to be similar, MPCA determined that it was appropriate to derive BTVs based on the results of the USEPA method reanalyzed subset of 45 samples despite the fact that only 36 were chosen randomly.

The datasets were evaluated using USEPA's ProUCL with and without outlier removal to determine the following:

- BTVs using the complete 137 sample USGS analyzed dataset for Minnesota
- BTVs using the results of the 45 samples reanalyzed using USEPA methods

Minnesota-specific maps of inorganic concentrations similar to those USGS created for the United States were also created. These maps can be used by risk assessors and project managers on a site-specific basis since the maps show the difference in concentrations across the state of Minnesota. Soil background concentrations of inorganics are heterogeneous. Although a range more accurately describes soil background, one value is often selected due to program needs.

BTVs and the percent difference from the results from the reanalysis using the USEPA analytical method on the 45 samples compared to results from the USGS analytical method using the 137 samples are shown in Table 14-1. This comparison included descriptive statistics and graphical displays. There was very little difference in the results of the two methods for arsenic. The BTV using the entire 137-sample USGS-analyzed dataset and the 45-sample USEPA-reanalyzed dataset was identical (9 mg/kg). Cobalt results were also similar, with the BTVs calculated from the two datasets having an 8% difference. Differences between the aluminum (68%), barium (70%), and chromium (62%) BTVs were much more significant. A BTV for thallium could not be calculated due to the number of nondetects in the dataset. In six cases the BTV was greater than the SRVs. In those cases, the BTV replaced the SRV (MPCA 2016a (MPCA 2016)[460]).

Table 14-1. USEPA analytical method results

Inorganic	**MPCA Residential SRV (mg/kg)**	**USGS BTV (UTL) of 137 Original Samples—USGS Analytical Method (mg/kg)**	**USEPA BTV (UTL) of 45 Reanalyzed Samples—USEPA Analytical Method (mg/kg)**	**Percent Difference USEPA and USGS BTVs**
Aluminum	**15,000**	60,000	**19,000**	-68%
Arsenic	**0.051**	9	**9**	0%
Barium	3,000	710	210	-70%
Chromium	23,000 (III) **11 (VI)**	71	**27**	-62%
Cobalt	**4.6**	13	**12**	-8%
Iron	**11,000**	38,000	**29,000**	-24%
Thallium	**0.15**	None	**0.29**	NA
Vanadium	**1.1**	119	**62**	-48%

Bold, shaded values highlight background concentrations above MPCA SRVs. All BTVs are 95-95 UTL, representing the 95% upper confidence limit on the 95th percentile.

Percent difference USEPA and USGS BTVS defined as (((USEPA BTV – USGS BTV)/USGS BTV))*100)

14.2 Former Firearms Training Range Soil Background Case Study

This site is a former firearms training range used for small firearms training. The COPC associated with historical firearms training activities at the Site identified during environmental investigations included antimony, arsenic, and lead. Remedial soil excavation activities were conducted in 2020. Following the soil excavation activities, only elevated soil concentrations of arsenic were detected in the confirmation samples.

This study presents an example of establishing a soil background level higher than the regulatory default ambient arsenic screening level of 12 mg/kg used as a cleanup goal for arsenic at the Site.

14.2.1 Development of a site-specific arsenic cleanup goal

The removal action confirmation sample results included only the post-removal soil confirmation samples from the margins of the remedial excavation and not the complete soil dataset prior to the soil removal action. Figure 14-1 presents the probability plot of the 30 soil samples.

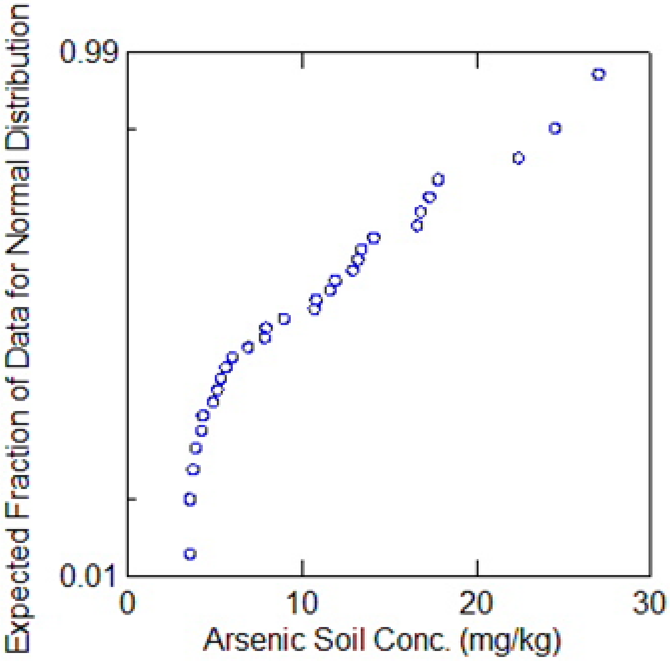


Figure 14-1. Probability plot of 30 samples.

Figure 14-2 summarizes the confirmation dataset in a box and whisker plot.

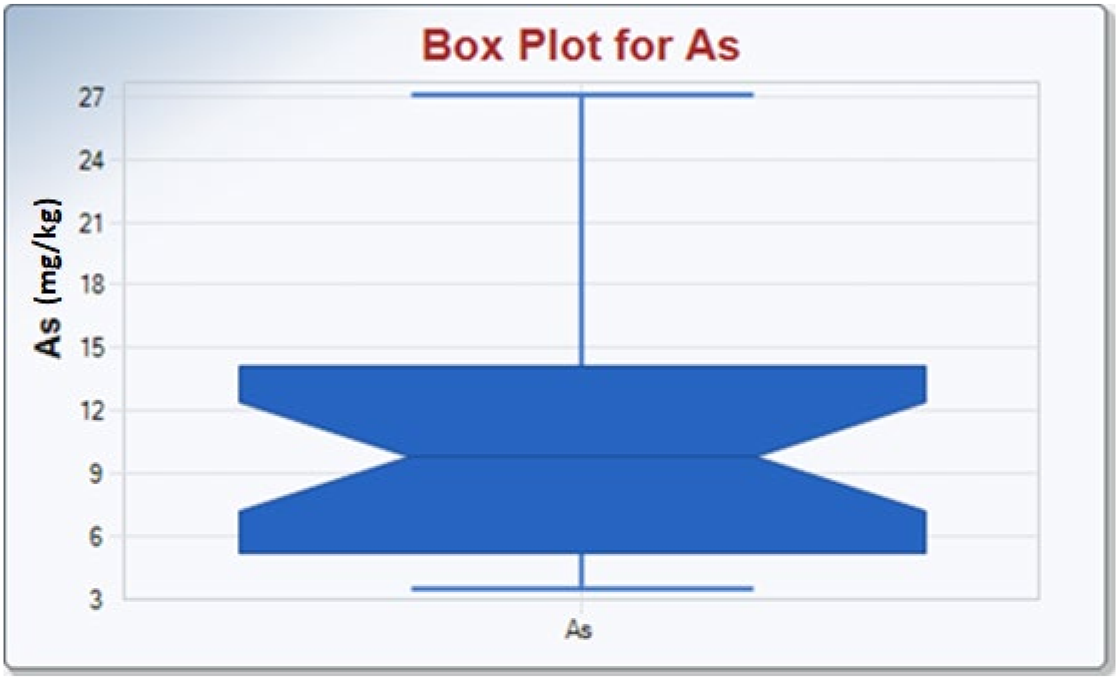


Figure 14-2. Box and whisker plot of confirmation data.

As can be seen from these two plots, the majority of arsenic concentrations are between 5 mg/kg and 14 mg/kg.

Ninety-eight soil samples were analyzed for metals at a nearby active firearms training range. Arsenic concentrations in soil ranged from 0.62 mg/kg to 21 mg/kg. Only four samples exceeded 12 mg/kg, specifically 13, 14, 16, and 21 mg/kg. Six off-site background samples were also collected. Off-site arsenic concentrations ranged from 4.2 mg/kg to 7.0 mg/kg.

For the Site, the two highest concentrations, 24.5 and 27 mg/kg, were removed as potential outliers based on multiple lines of evidence, including 1) graphical evaluation based on probability and box plots; 2) outlier removal based on ProUCL; and 3) arsenic data from a nearby site. Ultimately, professional judgment was used in evaluating the multiple lines of evidence and selecting the final dataset. The Site arsenic dataset was evaluated using the USEPA ProUCL, version 5.1, statistical software. Specifically, the descriptive statistics and BTVs were evaluated for the Site dataset. The ProUCL outputs included general statistics, goodness of fit, and estimated upper tolerance limits (background threshold values). Given the small sample size—only 30 arsenic samples—the 90th- or 95th-percentile estimates were selected as the more appropriate estimators of an upper-bound arsenic background value. Based on the ProUCL outputs, the data may be either approximate normal or lognormal in terms of the data distribution.

Assuming a normal distribution, the 90th- and 95th-percentile arsenic concentrations were 16.5 and 18.4 mg/kg, respectively. Assuming a lognormal distribution, the 90th- and 95th-percentile arsenic concentrations were 17.4 and 21.4 mg/kg, respectively. The nonparametric analysis, which does not assume any specific data distribution, resulted in 90th- and 95th-percentile arsenic concentrations of 17.0 and 17.6 mg/kg, respectively.

14.2.2 Recommendations

Given the small sample size and no discernible distribution of the arsenic dataset, the nonparametric 95th-percentile arsenic concentration of 18 mg/kg was selected as the most appropriate upper bound arsenic screening level for the site. This value is also consistent with (1) the site-specific off-site background data collected, where the maximum arsenic concentration was 18.3 mg/kg, and (2) the arsenic data from the nearby site.

14.3 Region 4 RARE Urban Background Study

In 2017, USEPA Region 4 was awarded a Regional Applied Research Efforts (RARE) Grant to work in collaboration with USEPA Office of Research and Development and southeastern regional states to develop a robust dataset to characterize contaminants present in urban soils deriving from natural and anthropogenic background, unrelated to chemical spills or industrial waste. A sampling and analysis plan (SAP) and a quality assurance project plan (QAPP) were developed and applied to eight cities within Region 4. The SAP and QAPP for the project were developed to ensure data collected would be defensible and reproducible. Soils were analyzed for RCRA 8 metals and PAHs.

14.3.1 Sample Design

A 7-mile by 7-mile grid was applied over each city center and divided into ½-mile by ½-mile grids, totaling 196 cells per city. Using a random number generator, 50 cells were selected for sample collection. Sample location criteria required soils to be representative of an urban setting, undisturbed by recent activity, not potentially impacted by nearby industry, and in public areas such as rights-of-way. Sample locations not acceptable included private/residential yards, industrial properties, areas of recent development, and low-lying areas subject to flooding or inundation. Samples were collected within the selected grid that met the criteria. If no areas within a sample grid met the criteria, then the next randomly selected grid was chosen.

A discrete sample was collected at each location using a turf profiler. The turf profiler sampling tool was modified to be 8.9 cm (3.5 in) wide, 2.5 cm (1 in) thick, and 10.2 cm (4 in) deep. The top 5 cm (2 in) of soil were retained for analysis with top organic material and remaining soils discarded.

Each sample was mixed and rigorously subsampled in the field. Samples designated for PAH analysis were sent in 4-ounce amber jars to an USEPA Contract Laboratory Program laboratory. Samples designated for metals analysis were sent in 4-ounce Whirl-Paks to USEPA's Las Vegas lab to be dried, sieved, and prepared for analysis. Once prepared, samples were shipped to USEPA Region 4 Science and Ecosystem Support Division (SESD) laboratory for analysis. Soil samples were analyzed for metals using USEPA Method 6010C and PAHs using USEPA Method 8270D. Three equipment blanks and three field blanks were collected per city.

14.3.2 Results

Metals analysis included arsenic, barium, cadmium, chromium, lead, mercury, selenium, and silver. Summary statistics were compiled for each contaminant by city and calculated using USEPA ProUCL software. The full dataset of metals and PAHs can be found at the USEPA Regional Urban Background Study website. Selected data are summarized below.

Mean arsenic concentrations in seven of the eight cities were above the USEPA regional screening level (RSL) of 0.68 mg/kg. The mean concentrations ranged from 0.55 mg/kg to 19.71 mg/kg. Noting that anthropogenic background of arsenic in some cities is above USEPA's RSL can be useful not only in determining whether the presence of arsenic in a given location is site-related contamination but in developing appropriate cleanup goals and guiding risk assessment.

Mean lead results in all eight cities were less than the current USEPA screening level of 400 mg/kg. Mean lead concentrations ranged from 15.06 mg/kg to 216.1 mg/kg. Summary statistics for lead are shown in Table 14-2. Box and whisker plots for lead are shown by city in Figure 14-3.

Table 14-2. Lead summary statistics

City	Num Obs	Minimum	Maximum	Mean	Geo-Mean	SD
Chattanooga	50	14	580	96.22	60.1	122
Columbia	50	1.7	200	40.79	25.58	38.41
Gainesville	50	2	110	15.06	9.187	18.62
Lexington	50	18	420	81.51	56.33	88.63
Louisville	50	25	1100	155.7	99.24	192.3
Memphis	50	13	1000	121	63.43	201.7
Raleigh	50	7.6	180	33.84	22.72	37.1
Winston-Salem	50	20	1400	216.1	128.8	247.2

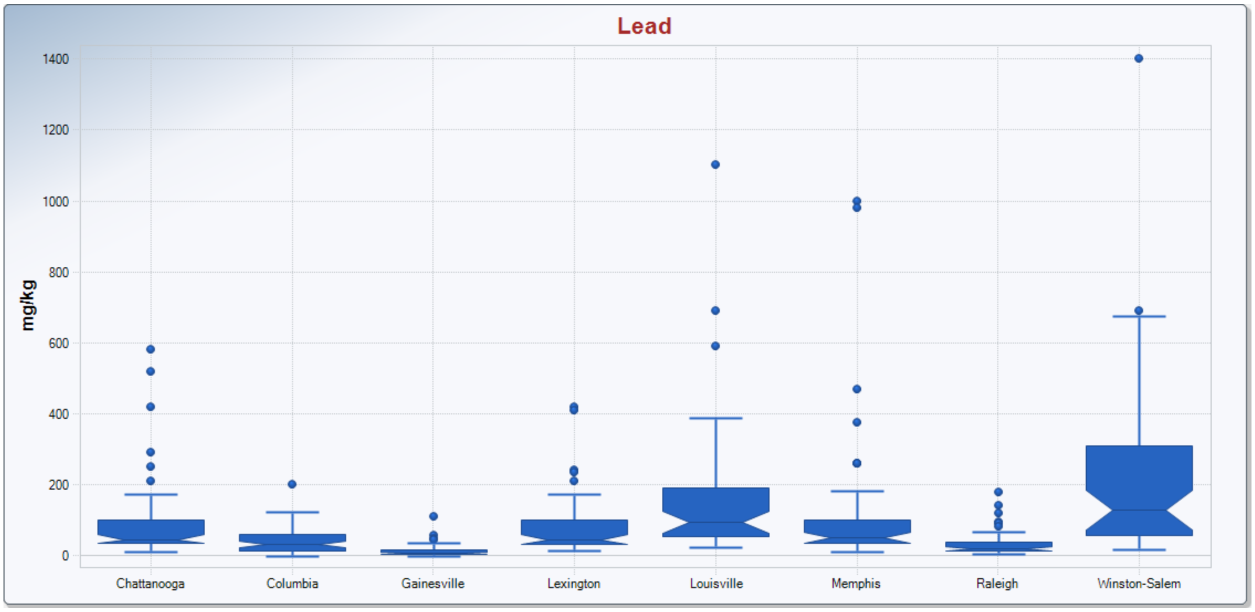


Figure 14-3. Box and whisker plots for lead.

14.3.3 Uses of Methods and Data

The robust dataset collected throughout the southeastern states could be used in developing cleanup goals, assessing risk, evaluating sample results, designing remedies, and evaluating the impact of natural disasters. The Region 4 Urban Background Study results highlight the difference in natural versus anthropogenic background. It is important to establish the difference between background concentrations and elevated contamination due to a spill or release in a particular area.

The methodology has been used by other USEPA regions along with universities to develop their own dataset of urban background soils. The data collected in Region 4 have been used at multiple sites and by two universities. The methods are easily adaptable to fit the area one would test for urban background, as was done for the City of Chattanooga.

The Urban Background Study Sampling and Analysis Plan and Quality Assurance Project Plan is publicly available: (https://www.epa.gov/research-states/regional-urban-background-study-webinar-archive).

The background dataset is also publicly available: https://www.epa.gov/risk/regional-urban-background-study.

14.4 Geochemical Evaluation Case Study—Statistical Outlier is an Uncontaminated Soil Sample

This case study involves a former military munitions site located in mountainous terrain in Colorado. A total of 12 candidate background subsurface soil samples (limited number due to budget constraints) were collected from an undeveloped adjacent property hydraulically upgradient of a former military munitions site that had soils geologically similar to the site. The discrete samples were collected from 0.3 to 0.5 m below ground surface and analyzed for metals; these results formed a candidate subsurface soil background dataset.

Dixon's statistical outlier test was performed for every trace element of interest for this site, including vanadium (a description of Dixon's test is provided in Table 11-2 in Section 11.5). The maximum vanadium (V) concentration (51.1 mg/kg) failed Dixon's outlier test at a significance level (α) of 0.05. Geochemical evaluation was performed to determine whether the elevated vanadium concentration was due to contamination or just inherent variability in soil concentrations. First, soil descriptions in the soil boring logs were studied, along with historical aerial photographs and site history, to identify possible sources for the elevated vanadium at that location; no sources were found. The sample with the outlier concentration was collected near an unpaved forest road, distant from the site.

As discussed in Section 5.3, the next step was to evaluate the concentrations of aluminum versus iron, because they are qualitative grain-size indicators, as well as primary reference elements (Figure 14-4). The concentrations of iron and aluminum covary (as the concentration of iron increases, the concentration of aluminum increases) and the samples exhibit similar Al/Fe ratios, indicating there is no aluminum or iron contamination in any of the candidate background samples. This confirms that these elements can serve as reference elements for the evaluation of trace elements of concern. The samples with higher aluminum and iron concentrations have more fine-grained minerals (such as clays and iron oxides), so they are expected to have naturally higher trace element concentrations.

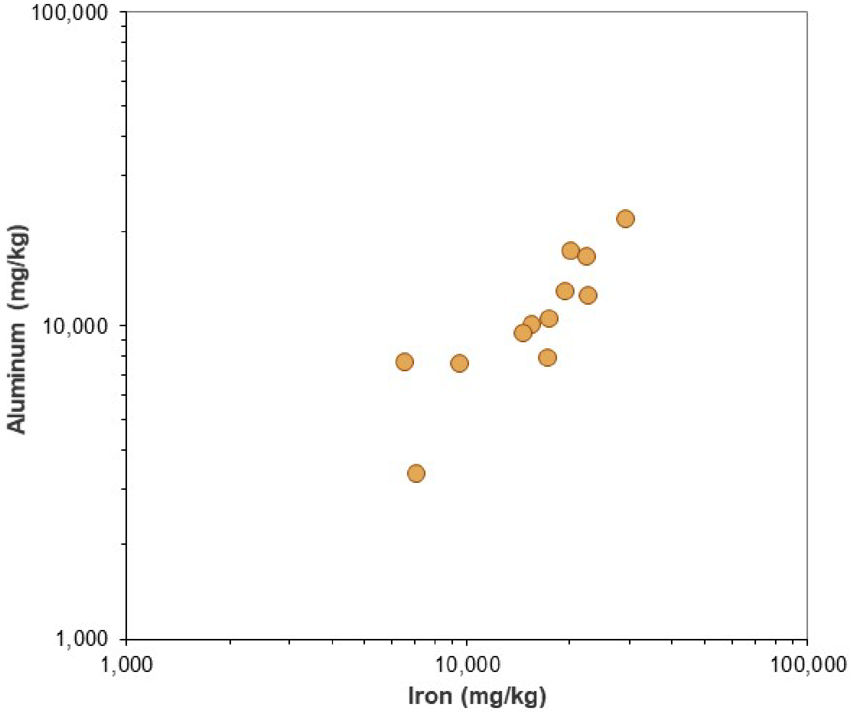


Figure 14-4. Scatter plot of aluminum versus iron for background subsurface soil.

Source: Karen Thorbjornsen, APTIM.

Scatter plots were prepared of vanadium versus aluminum (not depicted here, since the vanadium concentrations did not covary with aluminum concentrations in the scatter plot for this dataset) and vanadium versus iron (Figure 14-5). As discussed in Section 5.5.2, vanadium has a strong affinity for adsorption on iron oxide minerals in soil (Kabata-Pendias 2010)[151]. Vanadium was observed to covary with iron (Figure 14-5). The sample with the highest vanadium concentration also had the highest iron concentration (29,500 mg/kg).

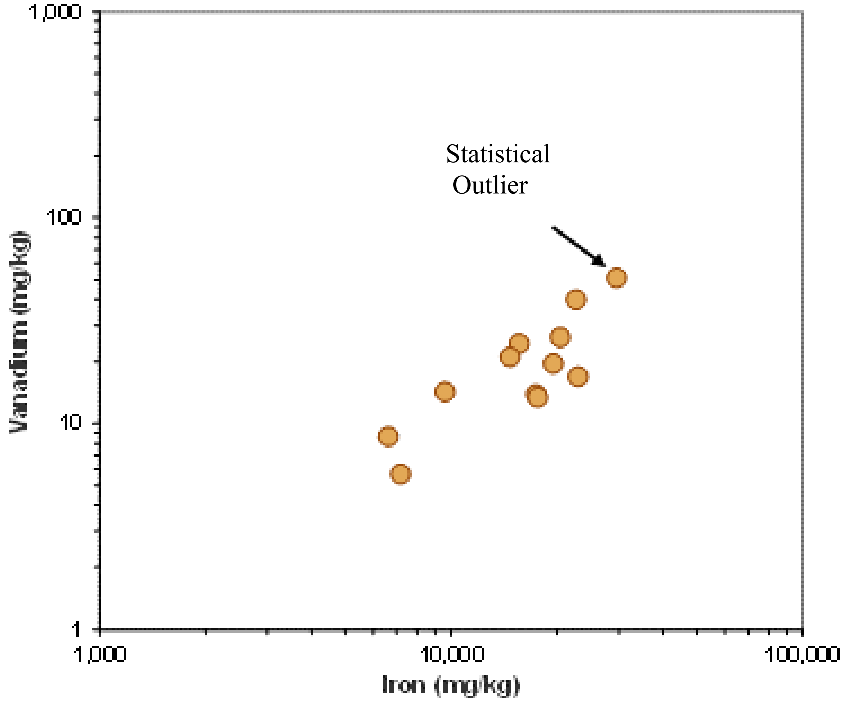


Figure 14-5. Scatter plot of vanadium versus iron for background subsurface soil.

Source: Karen Thorbjornsen, APTIM.

In addition to the scatter plot (Figure 14-5), a ratio plot (Figure 14-6) of vanadium concentrations versus V/Fe ratios was prepared.

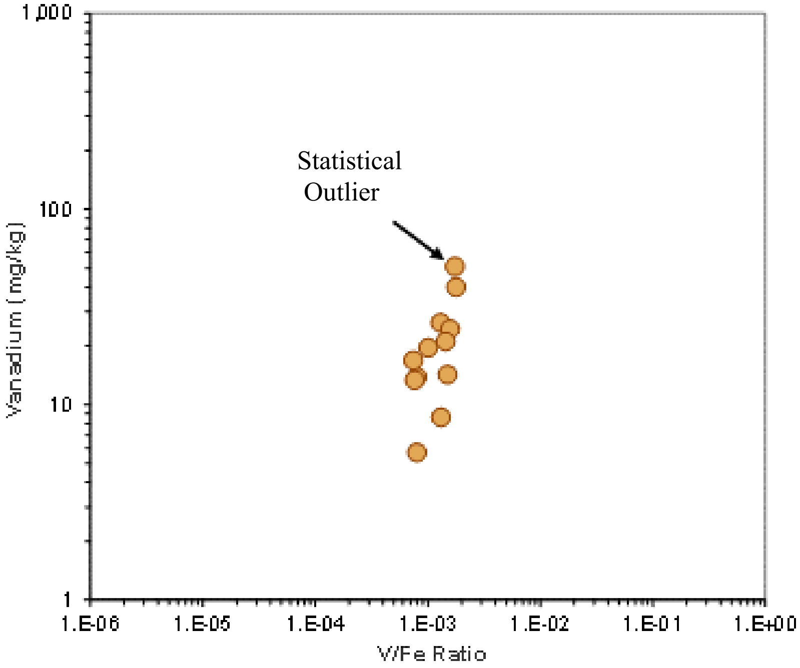


Figure 14-6. Ratio plot of vanadium versus V/Fe for background subsurface soil.

Source: Karen Thorbjornsen, APTIM.

In Figure 14-5, all the background samples (including the statistical outlier) form a common trend with a positive slope. In Figure 14-6, all samples (including the statistical outlier) have similar V/Fe ratios. This evaluation indicates the elevated vanadium concentration is due to preferential enrichment of iron oxide minerals in soil at that location and is not due to contamination by vanadium. Thus, the maximum vanadium concentration (the statistical outlier) was retained in the candidate subsurface soil background dataset.

14.5 Geochemical Evaluation Case Study—Statistical Outlier Is a Contaminated Soil Sample

Twenty-nine candidate background subsurface soil samples were collected from an undeveloped and heavily wooded property adjacent to (but hydraulically upgradient of) a former military facility located in glacial terrain in New Hampshire, with similar soils as the site. The discrete samples were collected from 0.5 to 0.6 m feet below ground surface and analyzed for metals; these results then formed a candidate subsurface soil background dataset.

Rosner's statistical outlier test was performed for every trace element of interest for this site, which included arsenic (a description of Rosner's test is provided in Table 11-2 in Section 11.5). The maximum arsenic (As) concentration in subsurface soil (61.4 mg/kg) failed Rosner's outlier test at a significance level (α) of 0.05. Geochemical evaluation was performed to determine whether this elevated arsenic concentration is due to arsenic contamination of this soil sample or if it is due to inherent variability in soil concentrations. Review of the soil description and the physical setting for that sample location revealed nothing unusual that would indicate a site-related arsenic contaminant source.

Aluminum and iron concentrations covary for this dataset (see scatter plot in Figure 14-7) and the samples have similar Al/Fe ratios. This indicates there is no aluminum or iron contamination in the candidate background samples and confirms that the two elements can serve as reference elements for the evaluation of trace elements of concern.

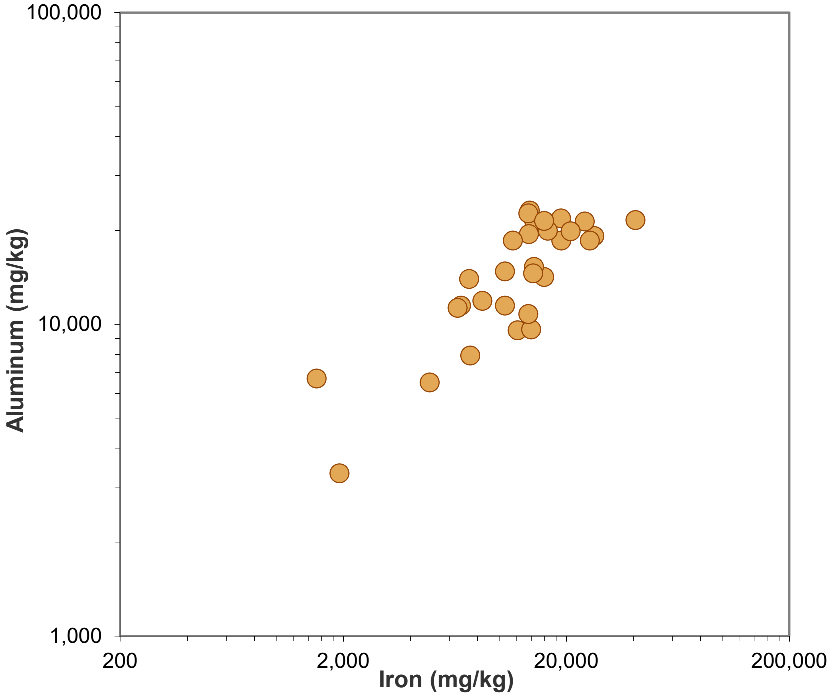


Figure 14-7. Scatter plot of aluminum versus iron for background subsurface soil.

Source: Karen Thorbjornsen, APTIM.

As discussed in Section 5.5.2, arsenic has a strong affinity to adsorb on iron oxide minerals in soil ((Bowell 1994)[414], (Kabata-Pendias 2010)[151], so the arsenic and iron concentrations are expected to covary in soil samples in the absence of arsenic contamination. In the scatter plot of arsenic versus iron concentrations (Figure 14-8), all but one of the samples lie along the same trend, reflecting the natural association of arsenic with iron oxides in soil. The sample with the statistical outlier concentration lies above and to the left of the trend. Although it has the highest arsenic, it has relatively low iron (as well as relatively low concentrations of the other major elements).

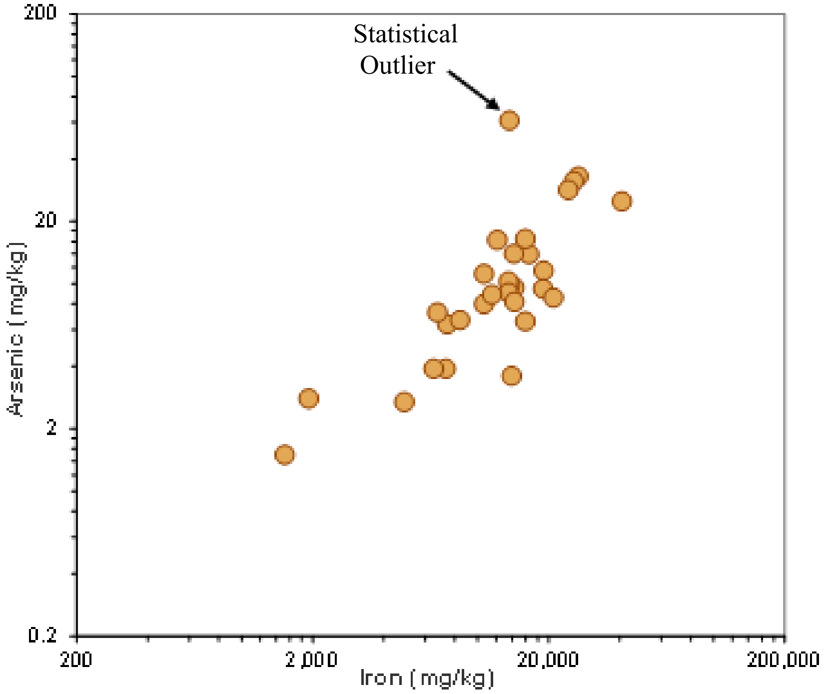


Figure 14-8. Scatter plot of arsenic versus iron for background subsurface soil.

Source: Karen Thorbjornsen, APTIM.

In the ratio plot in Figure 14-9, all samples (except the statistical outlier) have similar As/Fe ratios; the statistical outlier has a higher As/Fe ratio relative to the other 28 subsurface soil samples. The arsenic concentration in this sample is anomalously high, likely due to contamination of the sample by arsenic. Thus, this concentration was excluded from the candidate background dataset for arsenic, leaving this dataset with 28 arsenic concentrations.

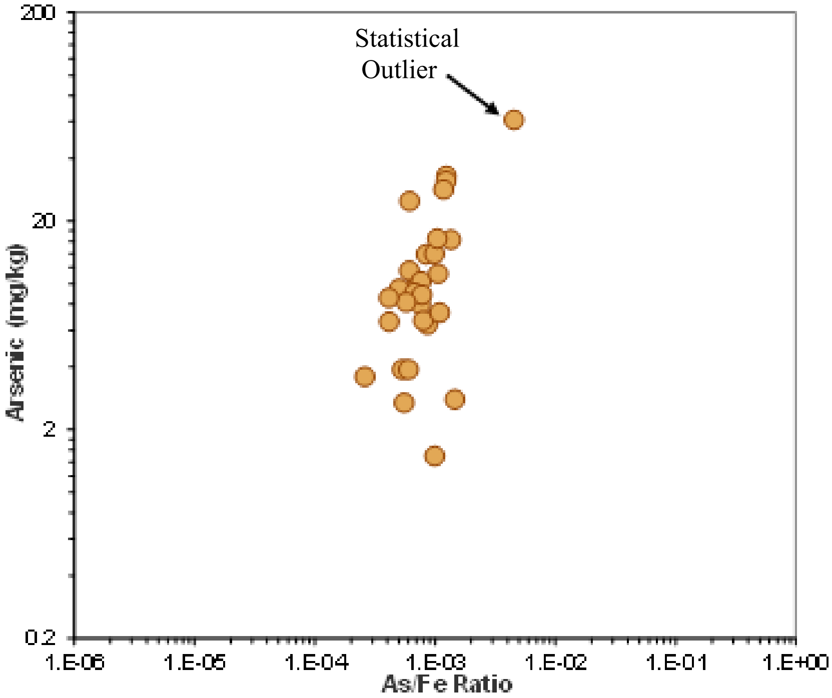


Figure 14-9. Ratio plot of arsenic versus As/Fe for background subsurface soil.

Source: Karen Thorbjornsen, APTIM.

For arsenic at this site, geochemical evaluation confirmed the result of the statistical outlier test.

14.6 Geochemical Evaluation Case Study–Contaminated Soil Sample Is Not a Statistical Outlier

A total of 10 candidate background surface soil samples (limited number due to budget constraints) were collected from a portion of a former military munitions site located in Wyoming that had never been developed and was hydraulically upgradient of the area of interest. The background soils were geologically similar to those in the area of interest. The discrete surface soil samples were collected from 0 to 5 cm below ground surface and analyzed for metals; these results then formed a candidate background dataset.

Dixon’s statistical outlier test was performed for every trace element of interest for this site, which included lead (a description of Dixon’s test is provided in Table 11-2 in Section 11.5). The maximum lead (Pb) concentration in subsurface soil (17.4 mg/kg) passed Dixon’s outlier test at a significance level (α) of 0.05. From visual inspection this concentration appeared to be higher than the other lead concentrations, so a geochemical evaluation was performed to double-check whether the concentration was elevated due to lead contamination of soil at that location or just due to inherent variability in soil concentrations.

Review of the soil description for that sample location revealed nothing unusual that would indicate a site-related lead contaminant source; the sampled soils were the same silty sand as noted for the other background sample locations. The aluminum and iron concentrations covary (see scatter plot in Figure 14-10) and the samples have similar Al/Fe ratios, indicating that there is no aluminum or iron contamination in the candidate background samples and confirming that the two elements can serve as reference elements.

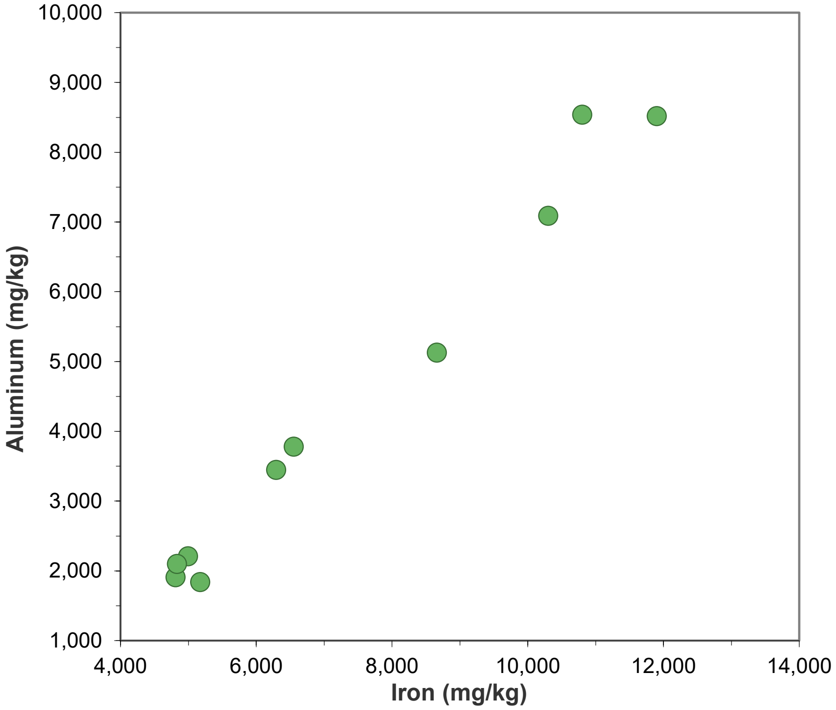


Figure 14-10. Scatter plot of aluminum versus iron for background surface soil.

Source: Karen Thorbjornsen, APTIM.

Although lead can adsorb on iron oxides and manganese oxides, it also has a strong affinity to adsorb on clay minerals in soil ((McLean and Bledsoe 1992)[415], (Kabata-Pendias 2010)[151]). A scatter plot of lead versus aluminum (Figure 14-11) and corresponding ratio plot (Figure 14-12) were prepared.

In Figure 14-11, all but one of the background samples lie along a common trend; the sample with the maximum lead concentration has one of the lowest aluminum concentrations and lies above the trend. This sample also has low concentrations of the other major elements.

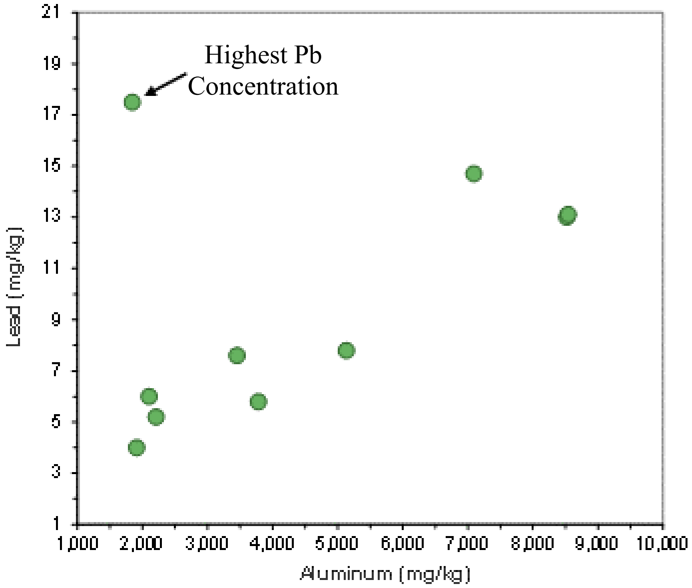


Figure 14-11. Scatter plot of lead versus aluminum for background surface soil.

Source: Karen Thorbjornsen, APTIM.

In Figure 14-12, most samples have similar Pb/Al ratios; the sample with the maximum lead concentration has a higher Pb/Al ratio than the other nine surface soil samples. Geochemical evaluation indicates the maximum lead concentration is anomalously high (likely due to lead contamination of the sample and suspected to be munitions-related) and is not naturally elevated, even though this data point was not identified as a statistical outlier in this dataset.

Based on the results of the geochemical evaluation, the anomalous concentration was excluded from the candidate background dataset for lead, leaving this dataset with nine lead concentrations. This approach resulted in a more representative background dataset.

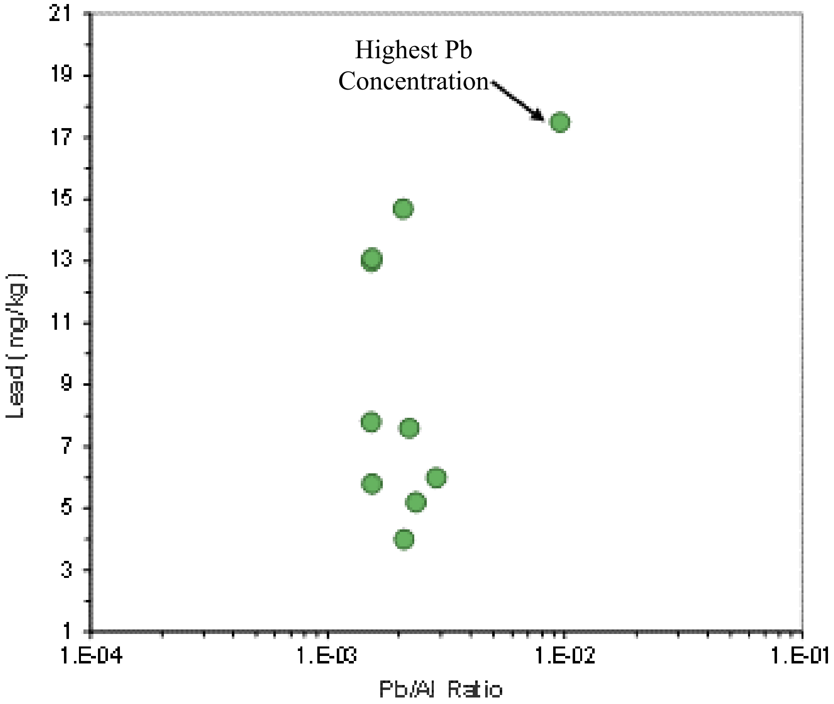


Figure 14-12. Ratio plot of lead versus Pb/Al for background surface soil.

Source: Karen Thorbjornsen, APTIM.

14.7 Environmental Forensics Case Study—PAHs from Leaked Petroleum Versus Contaminated Fill

This site had various commercial and industrial uses since the 1800s, including a city dump, rail yard, and oil products terminal. The site is adjacent to a waterway whose shorelines progressed over many decades due to infilling of marshland associated with the site's historic city dump. The oil terminal operated on only a portion of the site, wherein gasoline, kerosene, and diesel/fuel oil #2 (only) were off-loaded by barge, stored in aboveground tanks, and distributed by trucks from a loading rack. The rail yard was located at the terminus of a major railroad and operated from the mid-1800s until the 1960s.

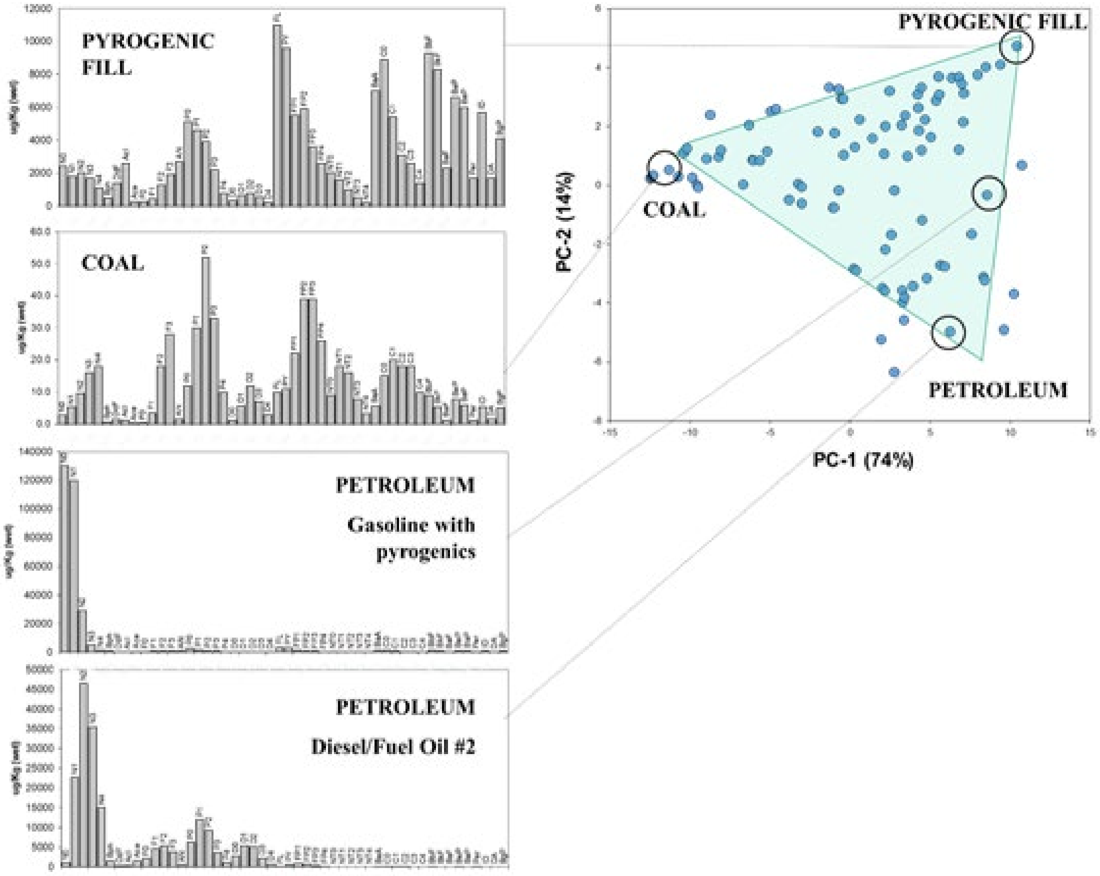
Preliminary site assessments across the 13-acre site measured the concentrations of 16 priority pollutant PAHs (EPAPAH16) for 185 soil samples, which revealed regulatory exceedances of the cleanup standards in 74 soil samples at the site. Only 33 of these exceedances were due to 2- or 3-ring PAHs reasonably associated with gasoline, kerosene, or diesel/fuel oil #2; most exceedances were for high molecular weight, 4- to 6-ring PAHs in deeper soil intervals (>1 m). The study's objective was to determine the sources of all PAHs detected in the soil samples, and specifically to distinguish PAHs derived from petroleum attributable to oil terminal operations (gasoline, kerosene, and diesel/fuel oil #2) versus nonpetroleum sources, potentially attributable to contaminated fill material and historical rail yard operations (for example, coal, coal ash, slag, and cinders).

Of the 185 soil samples, 108 were analyzed for TPH and PAH, the latter using the expanded list of PAH analytes that included 44 parent and alkylated PAHs and aromatic thiaarenes. The GC/FID chromatograms acquired using USEPA Method 8015 and PAH histograms acquired using modified USEPA Method 8270 revealed variable fingerprints for hydrocarbons and PAHs among the site's soils, including apparent mixtures. The PAH concentrations were evaluated using PCA (Section 7.2.1), which yielded the factor score plot shown in Figure 14-13. Three end-member sources of PAHs were recognized: pyrogenic fill, particulate coal, and light petroleum fuels. Based upon the PCA factor scores, many samples exhibited mixtures among these end-members.

Pyrogenic fill was dominated by 4- to 6-ring PAHs with homologues exhibiting a skewed pattern (Figure 14-13), which is typical of pyrogenic sources (Section 7.3.1.2; Figure 7-2; (Blumer 1976)[279]). Soils containing particulate coal were recognized by a broad range of PAHs with bell-shaped homologue profiles typical of petrogenic sources (Figure 7-2; (Blumer 1976)[279]). However, because no such broad-boiling petroleum was handled at the site (only light fuels), this pattern was attributed to coal particles, not petroleum. This conclusion was supported by visual and microscopic examination of the soil samples. Finally, some soil samples exhibited a strong dominance of 2-ring PAHs attributed to weathered gasoline, while others exhibited 2- and 3-ring PAHs, along with sulfur-containing dibenzothiophenes, each exhibiting bell-shaped homologue profiles typical of petroleum, specifically, in this case, diesel/fuel oil #2 (Figure 14-13).

Subsequent spatial analysis of the PCA-based results revealed that the soil samples identified as containing petroleum were mostly located proximal to former oil terminal operations (truck rack and pipelines). Those soil samples containing particulate coal were widely distributed on the site but were most common in the shallower soil intervals (<0.6 m). Finally, those soil samples containing pyrogenic fill also were widely distributed at the site, but these were found predominantly within deeper soil intervals (>1 m).

The chemical fingerprinting results, when interpreted in light of the site's operational history, were used to constrain the spatial extent of soil at the site that was impacted by spilled petroleum attributable to the former oil terminal operations. Coal was attributed to historic rail yard operations wherein coal was stockpiled and used as a locomotive fuel. Its impact was limited to shallower soils. PAHs contained within the pyrogenic fill were found to be the dominant source of PAHs detected in the site's soils and owing to the abundance of high molecular weight PAHs present in the fill (Figure 14-13). These were largely driving the regulatory exceedances across the site and attributed to the infilling associated with the site's historical dump operations.



Principal component analysis (PCA) factor score plot (right) identifying three end-member PAH sources in the site soils, represented by PAH histograms (left). Note petroleum sources contain little 4- to 6-ring PAHs, which largely were responsible for regulatory exceedances at the site. Contaminated fill associated with historical landfilling resulted in an anthropogenic ambient background for this site.

Figure 14-13. Principal component analysis (PCA) factor score plot.

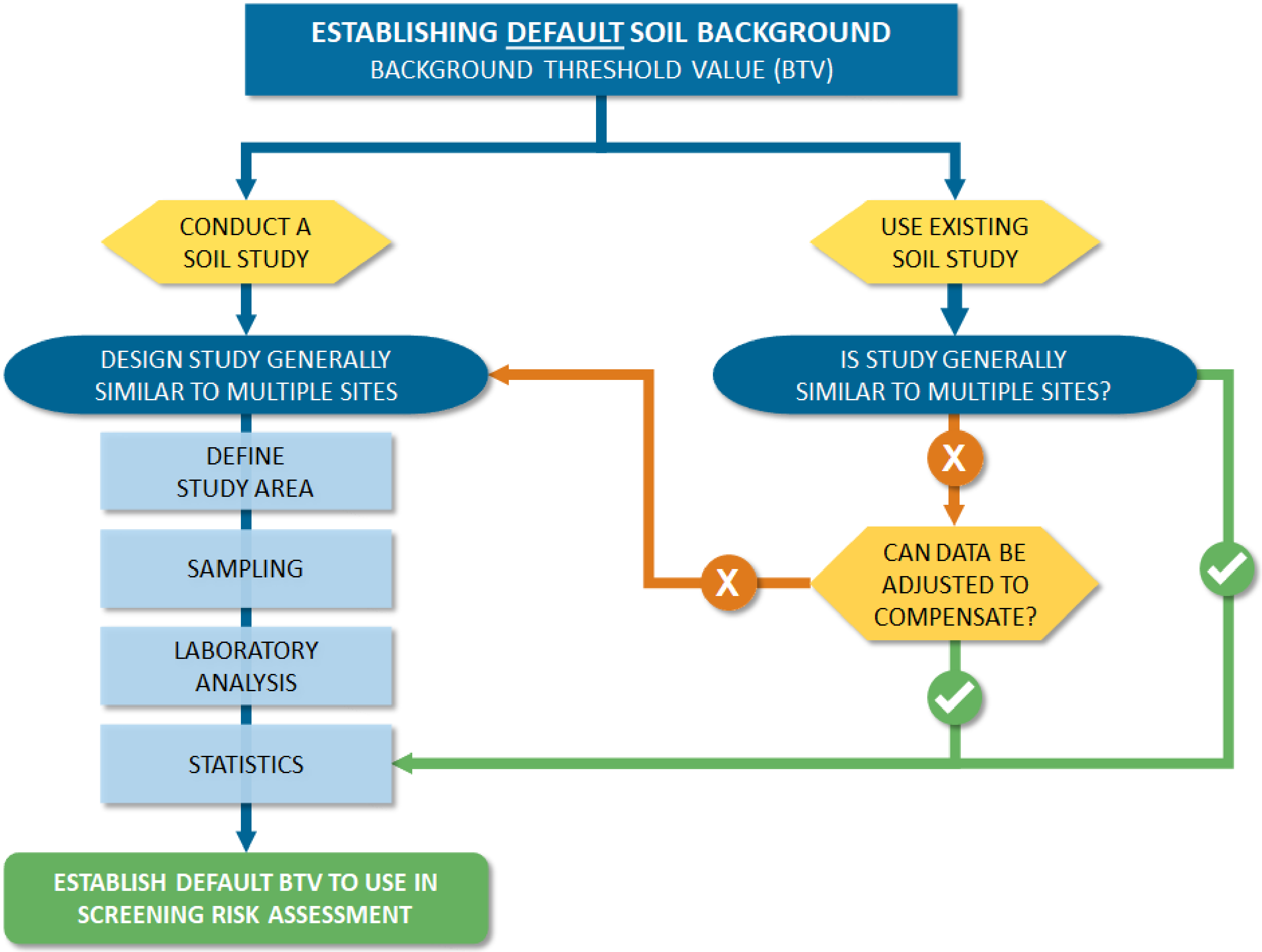
Source: Stout (unpublished).

FRAMEWORKS

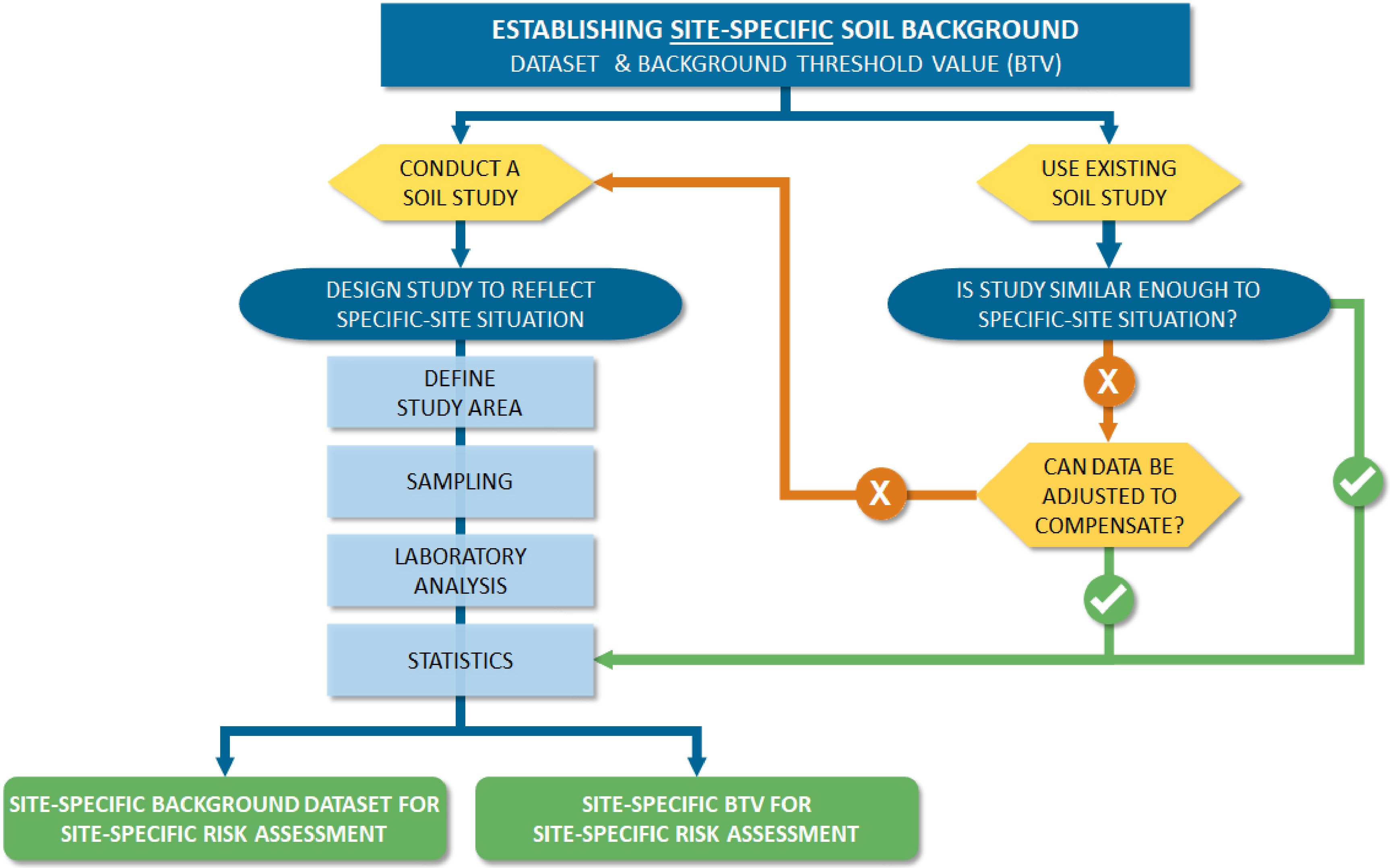
Three frameworks are provided to depict the process generally discussed in this document for establishing default and site-specific soil background and using background in risk assessment:

- Framework 1 depicts the process for establishing default soil background
- Framework 2 depicts the process for establishing site-specific soil background
- Framework 3 depicts the process for using soil background in risk assessment

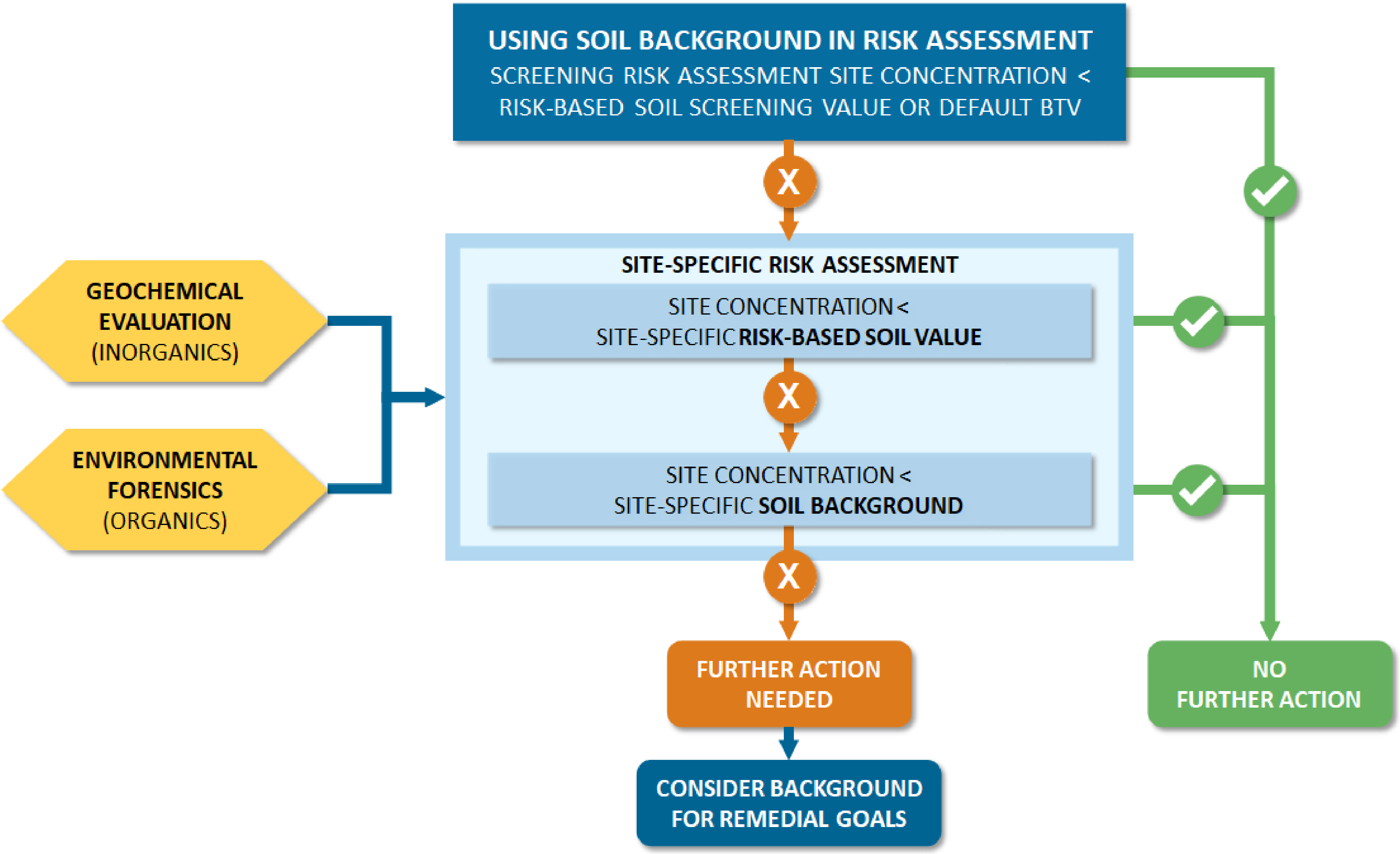
The process outlined in the frameworks can vary based on a specific situation and the regulatory agency providing oversight. Geochemical evaluations and environmental forensics can be used in any step of the process. Users are encouraged to review the referenced sections of this document for details on the elements of the process described in the frameworks.



Framework 1. Establishing Default Soil Background



Framework 2. Establishing Site-Specific Soil Background



USING SOIL BACKGROUND IN RISK ASSESSMENT

SCREENING RISK ASSESSMENT SITE CONCENTRATION < RISK-BASED SOIL SCREENING VALUE OR DEFAULT BTV

SITE-SPECIFIC RISK ASSESSMENT

SITE CONCENTRATION < SITE-SPECIFIC **RISK-BASED SOIL VALUE**

SITE CONCENTRATION < SITE-SPECIFIC **SOIL BACKGROUND**

GEOCHEMICAL EVALUATION (INORGANICS)

ENVIRONMENTAL FORENSICS (ORGANICS)

FURTHER ACTION NEEDED

CONSIDER BACKGROUND FOR REMEDIAL GOALS

NO FURTHER ACTION

Framework 3 Using Soil Background in Risk Assessment

Appendix A. UPPER LIMITS USED TO ESTIMATE BACKGROUND THRESHOLD VALUES

The USEPA has issued several comprehensive documents dealing with calculating upper limits to compute BTV estimates ((USEPA 2002)[131], (USEPA 2002)[154], (USEPA 2006)[192], (USEPA 2009)[198], (USEPA 2015)[197]) based upon a background dataset. However, there is no general consensus among practitioners about the statistic (upper limits) that should be used to estimate a BTV. For the sake of a typical reader and completeness, a brief description of upper limits (with interpretation), including upper percentiles, UPLs, UTLs, and USLs, used to estimate BTVs is provided in this appendix. Mathematical details and formulae to compute these statistics for datasets with and without NDs are given in ProUCL 5.1 Technical Guide. In the following the background population is also referred as the target population.

Upper percentile: A value, based on the data, below which a selected percentage (for example, 95%) of the data points will fall. In comparison with other BTV estimates (for example, UTL 95-95), the use of the 95th percentile yields a higher number of false positives, resulting in potentially unnecessary cleanup decisions, especially when many observations coming from a population (comparable to the background population) are compared with the sample percentile, p0.95.

Upper prediction limits (UPL): Let UPL95 represent a 95% UPL for a future/next observation. One is 95% sure that a *single* future value from the target (background) population will be less than or equal to the UPL95 with a confidence coefficient of 0.95 (95%). If an on-site value, x_onsite, is less than the UPL95, it may be concluded that x_onsite comes from the background population with a CC of 0.95. A UPL95 is not meant to be used to perform more than one future comparison.

However, in practice, users tend to use a UPL95 for many future comparisons, which results in a higher number of false positives (locations declared contaminated when in fact they are clean). When k (>1) future comparisons are made with a UPL95, some of those values will exceed UPL95 just by chance, each with a probability of 0.05. For proper comparison, UPLs need to be computed according to the number, k, of comparisons that will be performed. For details refer to ProUCL 5.1 Technical Guide.

Upper tolerance limit (UTL): A UTL (1-α)-p (for example, UTL 95-95) based upon an established background dataset represents that limit such that p% (for example, = 95%) of the sampled data will be less than or equal to that limit with a CC equal to (1-α) * 100% (for example, =95%). A UTL (1-α)-p represents a (1–α) * 100% upper confidence limit for the pth percentile of the underlying background population. It is expected that at least p% of the observations coming from the background population will be less than or equal to the UTL (1-α)-p with a CC equal to (1-α) * 100%. Specifically, a UTL 95-90 represents an upper tolerance limit providing coverage to at least 90% of the observations of the target population with CC=95%.

A UTL 95-90 assumes that as much as 10% of the observations can exceed the background UTL 95-90 when site concentrations are not different from the background population, and a UTL 95-90 can declare 10% of the observations coming from the background population as not coming from the background population just by chance, with a probability of 0.95. A UTL is used when many comparisons are planned. For a UTL 95-90, 10 exceedances per 100 comparisons (of background values) can result just by chance for an overall CC of 0.95. The number (and not percentage) of false positives can become large when many values from the target population are compared with a UTL.

Upper simultaneous limit (USL): A (1 – α) * 100% USL based upon an established background dataset is meant to provide simultaneous coverage for all sample observations in the background dataset (Singh and Nocerino 1995)[416] with probability (1-α). It is expected that all observations (present and future) belonging to the target (background, comparable to background) population will be less than or equal to a 95% USL (USL95) with a CC=0.95.

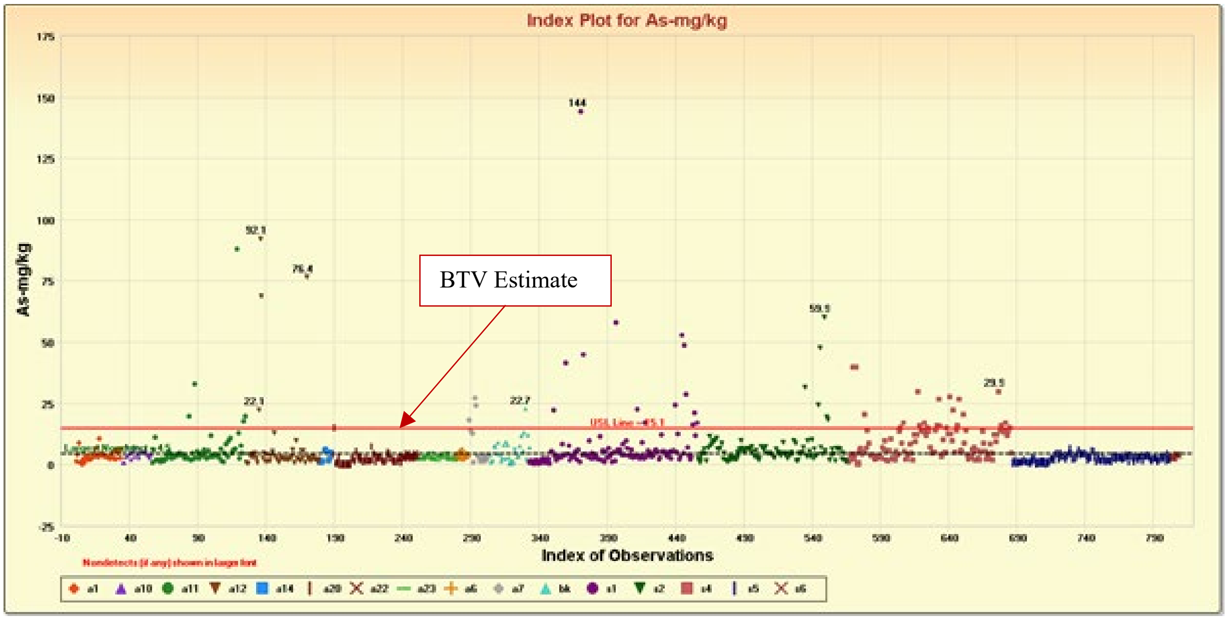
Like a UTL, a USL is used when any number (small or large) of on-site observations are compared with a BTV estimate. Unlike a UTL, a USL does not assume a priori that a certain percentage of background observations may not belong to the background population. Depending upon the variability of the background data, some of these statistics (for example, USL95, UTL 95-95) may exceed the largest value in the background dataset. To account for data variability of all sampled and unsampled locations, critical values associated with a USL95 increase as the sample size increases.

Appendix B. INDEX PLOTS

An index plot based upon a dataset, x1, x2, ..., xn of size n represents a scatter plot obtained by plotting n pairs. An index plot represents a visual way of identifying intermediate and extreme outliers potentially present in a mixture on-site dataset. The data values need not be ordered in any way. The generation of index plots does not require any distributional assumptions and the availability of spatial data (x, y coordinates; latitude and longitude) associated with sampled locations. Since an index plot represents a scatter plot, any software equipped to generate a scatter plot can be used to generate an index plot.

For clarity and usefulness of an index plot, one may want to order data by their population identification (ID) codes (for example, AOC1, AOC2...) (where AOC=area of concern) in the generation of an index plot as shown in Figure B1. On an index plot, observations from the various groups, including AOCs and an extracted background (for example, labeled as *bk-extrct*) data, are color-coded, where different colors represent different groups (AOCs). To avoid the distortion of the scale of an index plot, ND observations with elevated detection limits (DLs) may not be displayed on the index plot. An index plot can be formalized by drawing horizontal lines at one or more BTV estimates (for example, USL95, UTL 95-95), or at a prespecified regional cleanup level meant to compare individual on-site concentrations. The use of such index plots (Figures B1 and B2) comparing background (original or extracted) and the remaining on-site data (not included in extracted background data) may help the site managers in quickly identifying site AOCs and in making cleanup and remediation decisions at those AOCs.

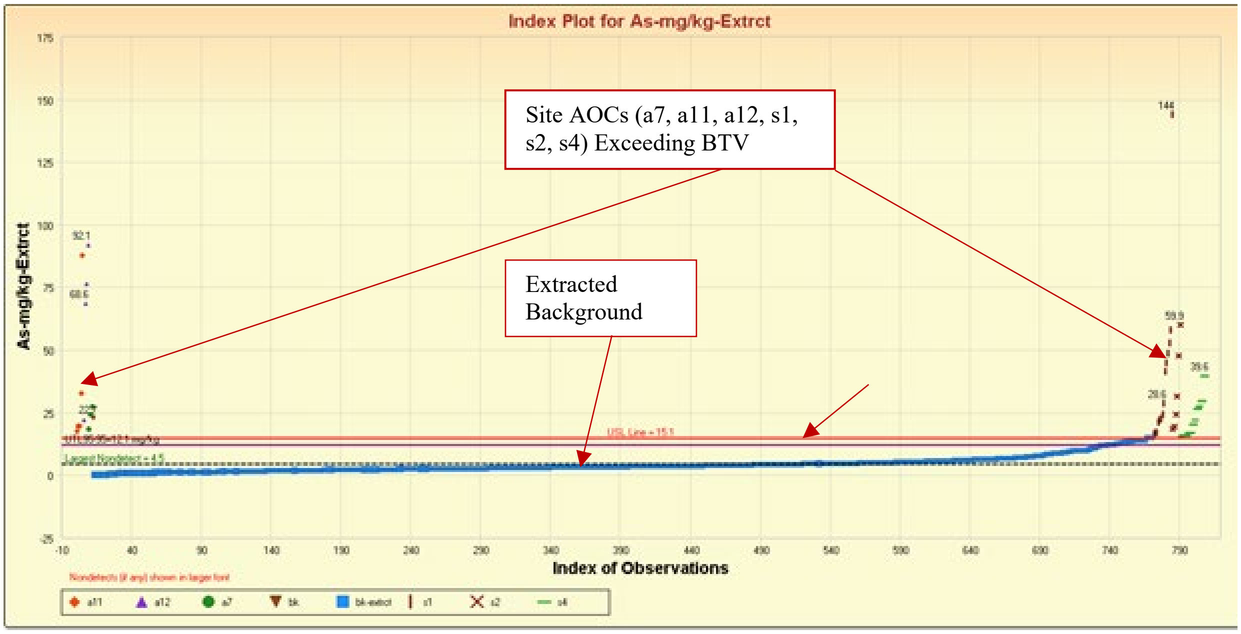
Figure B1 contains a color-coded index plot of the surface soil arsenic (prior to background extraction) dataset used in Section 3.9.5. It exhibits data of the AOCs and the existing background (labeled as bk) with a horizontal line displayed at a BTV estimate. Figure B2 has the similar index plot comparing the extracted background dataset (shown in blue) to the remaining contaminated AOCs concentrations.



Horizontal line displayed at the BTV estimate =15.1 mg/kg.

Figure B1. Index plot comparing surface soil arsenic concentrations of the existing background (*bk*) and various AOCs (prior to background extraction).

Source: Anita Singh, ADI-NV Inc.



Horizontal lines displayed at the BTV estimate, 15.1 mg/kg, and the largest nondetect value, 4.5 mg/kg.

Figure B2. Index plot comparing surface soil arsenic concentrations of the extracted background (*bk-extrct*) and the remaining AOCs (not included in *bk-extrct*).

Source: Anita Singh, ADI-NV Inc.

Appendix C. ADDITIONAL SOURCES OF INFORMATION FOR PAHS IN SOIL

Basic Methods Used in the Forensic Analysis of PAHs in Soils

Achten, C. and Hoffman, T. 2009. "Native Polycyclic Aromatic Hydrocarbons (PAH) in Coals – A Hardly Recognized Source of Environmental Contamination," *Sci. Total Environ.* 407: 2461-2473.

Blumer, M. 1975. "Polycyclic Aromatic Compounds in Nature," *Scientific American* 234: 35-45.

Colombo, J.C., Pelletier, E., Brochu, C., Khalil, M., Catoggio, J.A. 1989. "Determination of Hydrocarbons Sources Using N-alkane and Polycyclic Aromatic Hydrocarbon Distributions Indexes." Case study: Rio de La Plata estuary, Argentina. *Environ. Sci. Technol.* 23: 888-894.

Costa, H.J. and Sauer, J.T.C. 2005. "Forensic Approaches and Considerations in Identifying PAH Background." *Environ. Forensics* 6: 9-16.

Gonzalez-Vila, F.J., Lopez, J.L., Martin, F., del Rio, J.C. 1991. "Determination in Soils of PAH Produced by Combustion of Biomass Under Different Conditions." *Fresenius J. Anal. Chem.* 339: 750-753.

Lima, A.L.C., Farrington, J.W., Reddy, C.M. 2005. "Combustion-Derived Polycyclic Aromatic Hydrocarbons in The Environment – A Review," *Environ. Forensics* 6: 109-131.

Mandalakis, M., Gustafsson, O, Alsberg, T., Egeback, A.L., Reddy, C.M., Xu, L., Klanova, I., Holoubek, I., Stephanou, E.G. 2005. "Contribution of Biomass Burning to Atmospheric Polycyclic Aromatic Hydrocarbons at Three European Background Sites." *Environ. Sci. Technol.* 39: 2976-2982.

Stout, S.A. and Emsbo-Mattingly, S.D. 2008. "Concentration and Character of PAHs and Other Hydrocarbons in Coals of Varying Rank – Implications for Environmental Studies of Soils and Sediments Containing Particulate Coal." *Org. Geochem.* 39: 801-819.

Stout, S.A. and Uhler, A.D. 2003. "Distinguishing 'Background' Hydrocarbons from Contamination Using Chemical Fingerprinting." *Environ. Claims. J.* 15(2): 241-256.

Stout, S.A., Uhler, A.D., Emsbo-Mattingly, S.D. 2004. "Comparative Evaluation of Background Anthropogenic Hydrocarbons in Surficial Sediments from Nine Urban Waterways." *Environ. Sci. Technol.* 38: 2987-2994.

Stout, S.A., Emsbo-Mattingly, S.D., Douglas, G.S., Uhler, A.D., McCarthy, K.J. 2015. "Beyond 16 Priority Pollutants: A Review of PACs Used in Environmental Forensics." *Polycyclic Aromatic Compounds*, 35:285–315.

Vogt, N.B., Brakstad, F., Thrane, K., Nordenson, S., Krane, J., Aamot, E., Kolset, K., Esbensen, K., Steinnes, E., 1987. "Polycyclic Aromatic Hydrocarbons in Soils and Air: Statistical Analysis and Classification by the SIMCA Method." *Environ. Sci. Technol.* 21, 35-44.

Wakeham, S., Schaffner, C., Giger, W. 1980. "Polycyclic Aromatic Hydrocarbons in Recent Lake Sediments – II. Compounds Derived from Biogenic Precursors During Early Diagenesis." *Geochim. Cosmochim. Acta* 44: 415-429.

Wang, D., Tian, F., Yang, M., Liu, C., Li, Y.F. 2009. “Application of Positive Matrix Factorization to Identify Potential Sources of PAHs in Soil of Dalian, China.” *Environ. Pollut.* 157(5): 1559-1564.

Youngblood, W.W. and Blumer, M. 1975. “Polycyclic Aromatic Hydrocarbons in the Environment: Homologous Series in Soils and Recent Marine Sediments.” *Geochim. Cosmochim. Acta* 39: 1303-1314.

Zuo, Q., Duan, Y. H., Yang, Y., Wang, X. J., Tao, S. 2007. “Source Apportionment of Polycyclic Aromatic Hydrocarbons in Surface Soil in Tianjin, China.” *Environ. Pollut.* 147, 303-310.

PAH Background Concentrations in Soils

ATSDR (Agency for Toxic Substance and Disease Registry). 1995. “Toxicological Profile for Polycyclic Aromatic Hydrocarbons (PAHs),” Available at http://www.atsdr.cdc.gov/toxprofiles/tp69.html.

Azzolina, N.A., Kreitinger, J.P., Skorobogatov, Y., Shaw, R.K. 2016. “Background Concentrations of PAHs and Metals in Surface and Subsurface Soils Collected Throughout Manhattan, New York,” *Environ. Forensics* 17(4): 294-310.

Bradley, L.J.N. Magee, B.H., Allen, S.L. 1994. “Background Levels of Polycyclic Aromatic Hydrocarbons (PAH) and Selected Metals in New England Urban Soils,” *J. Soil Contamin.* 3(4): 349-361.

EPRI. 2000. “Literature Review of Polycyclic Aromatic Hydrocarbons.” Electric Power Research Inst., Palo Alto, CA, Publ. No. TR-114755, March 2000.

EPRI. 2003. “Polycyclic Aromatic Hydrocarbons (PAHs) in Surface Soil in Illinois.” Electric Power Research Inst., Palo Alto, CA, Publ. No. 1011376, Jan. 2003.

IARC (International Agency for Research on Cancer). 1983. “IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans.” Vol 32. Polynuclear aromatic compounds: Part I. Chemical, environmental and experimental data. Lyons, France: World Health Organization.

Jones, K.C., Stratford, J.A., Waterhouse, K.S., Vogt, N.B. 1989. Organic Contaminants in Welsh Soils: Polynuclear Aromatic Hydrocarbons.” *Environ. Sci. Technol.* 23: 540-550.

Kay, R.T., Arnold, T.L., Cannon, W.F., Graham, D., Morton, E., Bienert, R. 2003. “Concentration of Polynuclear Aromatic Hydrocarbons and Inorganic Constituents in Ambient Surface Soils,” Chicago, Illinois: 2001-02. U.S. Geol. Survey, Water Resources Investigation Report 03-4105.

LaGoy, P.K. and Quirk, T.C. 1994. “Establishing Generic Remediation Goals for the Polycyclic Aromatic Hydrocarbons: Critical Issues.” *Environ Health Perspect.* 102(4): 348-352.

Liu, Y., Gao, P., Su, J., da Silva, E.B., de Oliveria, L.M., Townsend, T., Xiang, P, Ma, L.Q. 2019. “PAHs in Urban Soils of Two Florida Cities: Background Concentrations, Distribution, and Sources.” *Chemosphere* 214: 270-277.

MADEP. 2002. “Technical Update: Background Levels of Polycyclic Aromatic Hydrocarbons and Metals in Soil. 310 CMR 40.0006.” Massachusetts Department of Environmental Protection, Boston, MA, USA.

Menzie, C.A., Potocki, B.B., Santodonato, J. 1992. "Exposure to Carcinogenic PAHs in the Environment." *Environ. Sci. Technol.* 26(7): 1278-1284.

Montgomery, J.H. 2007. *Groundwater Chemicals Desk Reference*, 4th ed., CRC Press, Boca Raton, Florida, 1752 p.

Morillo, E., Romero, A.S., Maqueda, C., Madrid, L., Ajmone-Marsan, F., Grcman, H., Davidson, C.M., Hursthouse, A.S., Villaverde, J. 2007. "Soil Pollution by PAHs in Urban Soils: A Comparison of Three European Cities." *J. Environ. Monitor*. 9: 1001-1008.

Nam, J.J., Thomas, G.L, Jaward, F.M., Steimmes, E., Gustafsson, O., Jones, K.C. 2008. "PAHs in Background Soils from Western Europe: Influence of Atmospheric Deposition and Soil Organic Matter." *Chemosphere* 70: 1596-1602.

Nam, J.J., Sweetman, A.J., Jones, K.C., 2009. "Polynuclear Aromatic Hydrocarbons (PAHs) in Global Background Soils." *J. Environ. Monit*. 11: 45-48.

Rabideau, A.J., Bronner, C., Milewski, D., Golubski, J., Weber, A.S. 2007. "Background Concentrations of Polycyclic Aromatic Hydrocarbon (PAH) Compounds in New York State Soils." *Environ. Forensics* 8: 221-230.

Sugihara, T., Mueller, R., Boyer, J., Evenson, J., Froehlich, D., Giles, G., Lester, L., Motter, A. Neumann, G., Schik, K. 2020. "Concentrations of Polycyclic Aromatic Hydrocarbons in New Jersey Soils." NJ DEP, Feb. 2020, 157 p.

Teaf, C.M., Convert, D.J., Kothur, S.R. 2008. "Polycyclic Aromatic Hydrocarbons (PAHs) in Urban Soil: A Florida Risk Assessment Perspective." *Int. J. Soil Sed. Water* 1(2): 12 p.

Van Metre, P.C., Mahler, B.J., Furlong, E.T. 2000. "Urban Sprawl Leaves its PAH Signature." *Environ. Sci. Technol*. 34: 4064-4070.

Wang, G., Zhang, Q., Ma, P. Rowden, J., Mielke, H.W., Gonzales, C., Powell, E. 2008. "Sources and Distribution of Polycyclic Aromatic Hydrocarbons in Urban Soils: Case Studies of Detroit and New Orleans." *Soil Sed. Contamin*. 17(6): 547-563.

Wilcke, W. 2007. "Global Patterns of Polycyclic Aromatic Hydrocarbons (PAHs) in Soil." *Geoderma* 141(3-4): 157-166.

Comparisons Among the PAH Fingerprints Through the Use of Diagnostic Ratios

Azzolina, N.A., Kreitinger, J.P., Skorobogatov, Y., Shaw, R.K. 2016. "Background Concentrations of PAHs and Metals in Surface and Subsurface Soils Collected Throughout Manhattan, New York," *Environ. Forensics* 17(4): 294-310.

Budzinski, H. Raoux, C., Baumard, P., Bellocq, J., Garrigues, P. 1995. "Differentiation of Contamination Sources in Recent Sediments Through PAH Distributions: An Overview," pp. 616-618 in *Organic Geochemistry Developments and Applications to Energy, Climate, Environment and Human History: Selected Papers from the 17th International Meeting on Organic Geochemistry*, J. Grimalt and C. Dorronsoro, eds., European Assoc. Organic Geochemists.

Budzinski, H. Jones, I., Bellocq, J., Pierard, C., Garrigues, P. 1997. “Evaluation of Sediment Contamination by Polycyclic Aromatic Hydrocarbons in the Gironde Estuary.” *Mar. Chem.* 58: 85-97.

Colombo, J.C., Pelletier, E., Brochu, C., Khalil, M., Catoggio, J.A. 1989. “Determination of Hydrocarbons Sources Using N-alkane and Polycyclic Aromatic Hydrocarbon Distributions Indexes.” Case study: Rio de La Plata estuary, Argentina. *Environ. Sci. Technol.* 23: 888-894.

Costa, H.J. and Sauer, J.T.C. 2005. “Forensic Approaches and Considerations in Identifying PAH Background.” *Environ. Forensics* 6: 9-16.

Costa. H.J., White, K.A., Ruspantini, J.J. 2004. “Distinguishing PAH Background and MGP Residues in Sediments of a Freshwater Creek.” *Environ. Forensics* 5(3): 171-182.

Dickhut, R.M., Canuel, E.A., Gustafson, K.E., Liu, K., Arzayus, K.M., Walker, S.E., Edgecombe, G., Gaylor, M.O., MacDonald, E.H., 2000. “Automotive Sources of Carcinogenic Polycyclic Aromatic Hydrocarbons Associated with Particulate Matter in the Chesapeake Bay Region.” *Environ. Sci. Technol*. 34: 4635-4640.

Emsbo-Mattingly, S.D., Litman, E. 2016. “Polycyclic Aromatic Hydrocarbon Homolog and Isomer Fingerprinting.” In: S.A. Stout, Z. Wang, eds., *Standard Handbook Oil Spill Environmental Forensics,* Elsevier Inc., Boston, pp. 254-312.

Herman, K., Wannamaker, E.J., Jegadeesan, G.B. 2012. “Sediment PAH Allocation Using Parent PAH Proportions and a Least Root Mean Squares Mixing Model.” *Environ. Forensics* 13: 225-237.

Stout, S.A. and Emsbo-Mattingly, S.D. 2008. “Concentration and Character of PAHs and Other Hydrocarbons in Coals of Varying Rank – Implications for Environmental Studies of Soils and Sediments Containing Particulate Coal.” *Org. Geochem.* 39: 801-819.

Yunker, M.B., Macdonald, R.W., Vingarzan, R., Mitchell, R.H., Goyette, D., Sylvestre, S. 2002. “PAHs in the Fraser River Basin: A Critical Appraisal of PAH Ratios as Indicators of PAH Source and Composition.” *Org. Geochem*. 33: 489-515.

Appendix D. TEAM CONTACTS

Team Leaders:

Bonnie Brooks
Department of Ecology—Washington
360-584-8856
bonnie.brooks@ecy.wa.gov

Claudio Sorrentino
Department of Toxic Substances Control—California
916-255-6656
Claudio.Sorrentino@DTSC.ca.gov

Program Advisors:
Jim Rocco
Sage Risk Solutions LLC
jrocco@sagerisk.com

Lesley Hay Wilson
Sage Risk Solutions LLC
lhay_wilson@sagerisk.com

Writing Group Leaders:

Introduction, Sampling & Laboratory Analysis (ISL)
Harold Templin
Indiana Department of Environmental Management
htemplin@idem.in.gov

Jay Peters
Haley & Aldrich
jpeters@haleyaldrich.com

Establishing Soil Background (ESB)
Erin Endsley
Wisconsin Department of Natural Resources
erin.endsley@wisconsin.gov

Geoff Siemering
University of Wisconsin
gsiemering@wisc.edu

Using Soil Background in Risk Assessment (SBRA)

Susan Turnblom
Oregon Department of Environmental Quality
turnblom.susan@deq.state.or.us

Barrie Selcoe
Jacobs
Barrie.Selcoe@jacobs.com

Geochemical Evaluations and Environmental Forensics (GEEF)

Brady Johnson
Idaho Department of Environmental Quality
brady.johnson@deq.idaho.gov

Karen Thorbjornsen
APTIM
karen.thorbjornsen@aptim.com

Appendix E. GLOSSARY

A

Anthropogenic ambient soil background—Amount of a substance, or family of closely related substances (for example, similar element species or similar compounds), present in soil due to anthropogenic nonpoint sources or due to their ability to be transported long distances. This is potentially most relevant to some recalcitrant organics, such as polycyclic aromatic hydrocarbons (PAHs) and 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) that can be present in soil at low concentrations not because of local anthropogenic sources, but because of their persistence, their ubiquity, or their ability to be transported long distances. Additionally, this could be relevant to metals with a widespread urban source, such as the historical association of lead with vehicle emissions or long-distance transport of mercury.

B

Background breakpoint (BP)—A point in the pooled dataset determined by using iterative normal Q-Q plots in untransformed raw scale. A background BP distinguishes uncontaminated and contaminated site locations.

Background reference area - The area identified as appropriate for collection of samples used to ultimately determine a soil background concentration or range.

Background threshold value (BTV)—A background population parameter representing some upper threshold, such as a 95th percentile or a 99th percentile of the distribution of the background population. Though this background parameter is not specifically defined in USEPA documents, USEPA guidance documents ((USEPA 2002)[131], (USEPA 2002)[154], (USEPA 2006)[192]) describe several methods to estimate this background parameter. For example, in computing a 95-95 UTL based upon a background dataset, you are estimating the 95th percentile of the background population. Some details about BTVs are described in Appendix A.

BTV estimate—Statisical limits such as UPLs, UTLs, and USLs represent estimates of a BTV. BTV estimates are computed based upon a background dataset.

C

Chemicals of potential concern (COPC)—In a risk assessment, a substance detected at a hazardous waste site that has the potential to affect receptors adversely due to its concentration, distribution, and mode of toxicity (USEPA 1997)[272]. COPC are generally categorized operationally, based on how they are measured in the analytical laboratory. “Inorganic” COPC generally address metals, elements, and unique inorganic compounds such as perchlorate. “Organic” COPC include VOCs (such as acetone, benzene, and trichloroethylene), semivolatile organic compounds (such as chlorophenols, chlorobenzenes, and phthalate esters), pesticides (such as atrazine, DDT, toxaphene), PCBs, and polychlorinated dibenzodioxin and dibenzofurans.

Conceptual site model (CSM)—Iterative representation of the site that summarizes and helps project planners visualize and understand available information. The CSM is the primary planning and decision-making tool used to identify the key issues and the data necessary to transition a project from characterization through post-remedy.

D

Default soil background—Established by regulatory agencies for a larger area (for example, a state, a region, or a unique geological area) that generally shares similar physical, chemical, geological, and biological characteristics. Since default soil background is intended to be used to evaluate a large number of sites, it is generally established to be on the conservative side. Default soil background can be established for both natural and anthropogenic ambient soil background concentrations.

E

Environmental forensics—A well-established discipline that considers scientific and nonscientific information to interpret the potential sources and ages of certain chemical compounds detected at a site, typically at anomalous concentrations.

Established background dataset—In this report, a dataset (traditional existing or extracted), agreed upon by all parties to be used as a background dataset.

Existing/traditional background dataset—A dataset containing COPC concentrations of samples collected from uncontaminated off-site locations collected using USEPA policy guidelines described in USEPA documents (e.g., (USEPA 1989)[130], (USEPA 1992)[249], (USEPA 1992)[250], (USEPA 2002)[131], (USEPA 2002)[154]).

Extracted background dataset/sitewide anthropogenic extracted background dataset—Dataset consisting of all observations (detects and NDs) less than or equal to the background BP of the pooled dataset. Note that this definition is specific to this document; background datasets derived per agency guidelines are not required to have background BPs or be strictly nonoverlapping with contaminated data.

F

G

Graphical display—A graphical representation of data displaying data points with their group identification as well as statistics (for example, UTL, USL) of interest.

H

I

Index plot—An index plot displaying data points with group identification as well as horizontal lines displayed at BTV estimates of interest, such as UPL, UTL, or USL.

J

K

L

Limit of detection (LOD)—The measure of an analytical method to detect the presence of an analyte with a 99% level of confidence.

M

Method detection limit (MDL)—The minimum measured concentration that can be reported with 99% confidence that the measured concentration is distinguishable from method blank results (USEPA 2015)[197].

N

Natural soil background—The concentration of a substance, or family of closely related substances (for example, similar element species or similar compounds) present in soil due to geological characteristics, natural processes, or releases from nonanthropogenic sources (for example, wildfires, volcanic activity). Natural soil background does not include releases from local anthropogenic sources, releases from distant anthropogenic sources of persistent chemicals (due to their ability to be transported long distances), or other anthropogenic sources of ubiquitous or widespread contamination. Natural soil background is potentially one of the broadest applicable definitions, in that natural background for certain chemicals has been established statewide or over a large geographic area provided the geological formations are conserved.

Normal Q-Q plot/probability plot—A graphical method for comparing an empirical dataset (*y*-axis) to a theoretical normal distribution (*x*-axis). Each point is the cartesian coordinates of the theoretical quantile and the ordered (in ascending order) empirical data. A normal probability plot (Chambers et al. 1983)[46] is generated similarly. The two plots—normal Q-Q plot and probability plot—are used to address the same objectives (for example, (Chambers et al. 1983)[46], (Reimann, Filzmoser, and Garrett 2005)[62]). These graphs are often used as exploratory tools to determine: data skewness, data variability, data distribution tails, outliers, and the presence of more than one population in the dataset.

O

Oxyanion—An ion that is negatively charged and contains oxygen.

P

Petrogenic—Chemicals derived directly from fossil fuels (crude oil, refined petroleum, or coal). Most common use is in reference to petrogenic PAHs derived from natural heating of fossil fuels in the earth over geologic time.

Pooled/mixture/broader dataset—A pooled dataset consists of observations collected from on-site AOCs and operating units. The pooled dataset may also consist of observations of a traditional background dataset (when available).

Pyrogenic—Chemicals produced during the processes of heating, such as combustion and carbonization. Most common use is in reference to pyrogenic PAHs derived from the anthropogenic thermal decomposition of organic matter (oil, coal, or biomass).

Q-Q plot—A quantile-quantile (Q-Q) plot displaying goodness of fit test statistics and other descriptive statistics, including the correlation coefficient, mean, and standard deviation.

R

S

Site-specific background population—A population consisting of areas/locations 1) not impacted by site-related activities and releases; 2) influenced by non-site-related anthropogenic releases (if any) comparable to those of the site; and 3) with natural variability and geological properties comparable to the site.

Site-specific soil background—The concentration of a substance established for a specific site to compare to that one site's data to estimate if site concentrations may reflect background. This requires collection of data representative of the site's natural and/or anthropogenic ambient background that has similar physical, chemical, geological, biological, and ecological characteristics and land use as the site being evaluated. Depending on the project goal, it may also be necessary to include other anthropogenic sources, such as municipal or industrial outfalls and urban runoff, that are influencing chemical concentrations at the site being evaluated.

Skewness—A measure of the lack of radial symmetry in a distribution about the mean.

Soil—For purposes of this document, the unconsolidated material present at ground surface and extending down to solid rock.

Soil background—Includes both natural soil background and anthropogenic ambient soil background. At some locations, the unconsolidated material may have formed in place from the underlying bedrock or associated with material that was transported to the location.

T

U

V

W

X

Y

Z

Appendix F. ACRONYMS

AES	atomic emission spectroscopy
AFFF	aqueous film forming foam
AOC	area of concern
ATV	all-terrain vehicle
BOC	biogenic organic compounds
BP	breakpoint
BTV	background threshold value
CERCLA	Comprehensive Environmental Response, Compensation, and Liability Act
CMB	chemical mass balance
COPC	chemicals of potential concern
CRAN	Comprehensive R Archive Network
CSM	conceptual site model
CVAA	cold vapor atomic absorption
DL	detection limit
DQO	data quality objectives
DU	decision units
ECR	excess lifetime cancer risk
EPC	exposure point concentration
FID	flame ionization detection
FSAP	field sampling and analysis plan
GC/ECD	gas chromatography/electron capture detector
GC/FID	gas chromatography/flame ionization detector
GC/MS	gas chromatography/mass spectrometry
GIS	geographic information system
HI	hazard index
HQ	hazard quotient
HRGC/HRMS	high-resolution gas chromatography/high-resolution mass spectrometry
ICP	inductively coupled plasma
ICP/AES	inductively coupled plasma/atomic emission spectroscopy
ICP/MS	inductively coupled plasma/mass spectrometry
IELCR	incremental excess lifetime cancer risk
ISM	incremental sampling methodology
ISO	International Organization for Standardization
KM	Kaplan-Meier method
LOC	levels of chlorination
LOD	limit of detection
MDL	method detection limit
MLE	maximum likelihood estimation
MPCA	Minnesota Pollution Control Agency
MS	mass spectrometry
ND	nondetect
NELAC	National Environmental Laboratory Accreditation Conference
NOM	naturally occurring materials
OCP	organochlorine pesticides

254	OSWER	Office of Solid Waste and Emergency Response
255	PAH	polycyclic aromatic hydrocarbons
256	PCA	principal component analysis
257	PCB	polychlorinated biphenyls
258	PCDD	polychlorinated dibenzodioxins
259	PCDD/F	polychlorinated dibenzo-p-dioxins and dibenzofurans
260	PCDF	polychlorinated dibenzofurans
261	PMF	positive matrix factorization
262	QA	quality assurance
263	QA/QC	quality assurance/quality control
264	QAPP	quality assurance project plan
265	QC	quality control
266	QL	quantitation limit
267	RARE	Regional Applied Research Efforts
268	RCRA	Resource Conservation and Recovery Act
269	RL	reporting limit
270	ROS	regression on order statistics
271	RSL	regional screening level
272	SAP	sampling and analytical plan
273	SBR	soil background and risk assessment
274	SIM	selected ion monitoring
275	SRV	soil reference values
276	SU	sampling units
277	TCDD	total tetrachlorodibenzo-p-dioxin
278	TEF	toxicity equivalency factor
279	TEQ	toxicity equivalent
280	TPH	total petroleum hydrocarbons
281	UCL	upper confidence limit
282	UPL	upper prediction limit
283	USEPA	United States Environmental Protection Agency
284	USGS	United States Geological Survey
285	USL	upper simultaneous limit
286	UTL	upper tolerance limit
287	VOC	volatile organic compound
288	VSP	Visual Sample Plan
289	XRF	x-ray fluorescence

REFERENCES

93. Achten, Christine, and Thilo Hofmann. 2009. "Native polycyclic aromatic hydrocarbons (PAH) in coals-A hardly recognized source of environmental contamination." *The Science of the total environment* 407:2461-73. doi: https://doi.org/10.1016/j.scitotenv.2008.12.008.

472. ADEM, 2017. Alabama Environmental Investigation and Remediation Guidance (AEIRG), Montgomery, AL. https://adem.alabama.gov/programs/land/landforms/AEIRGInvestigation.pdf.

10. Akritas, Michael G., Thomas F. Ruscitti, and G. P. Patil. 1994. "7 Statistical analysis of censored environmental data." In *Handbook of Statistics*, 221-242. Elsevier. doi:https://doi.org/10.1016/S0169-7161(05)80009-4.

1. Ames, K. C., and E. A. Prych. 1995. Background concentrations of metals in soils from selected regions in the State of Washington. U. S. Geological Survey and Open-File Reports Section Earth Science Information Center. http://pubs.er.usgs.gov/publication/wri954018.

230. Anderson, T.W. 2003. *An Introduction to Multivariate Statistical Analysis*: Wiley, https://books.google.com/books?id=Cmm9QgAACAAJ.

12. Antweiler, Ronald C. 2015. "Evaluation of Statistical Treatments of Left-Censored Environmental Data Using Coincident Uncensored Data Sets. II. Group Comparisons." *Environmental Science & Technology* 49 (22):13439-13446. doi: https://doi.org/10.1021/acs.est.5b02385.

11. Antweiler, Ronald C., and Howard E. Taylor. 2008. "Evaluation of Statistical Treatments of Left-Censored Environmental Data using Coincident Uncensored Data Sets: I. Summary Statistics." *Environmental Science & Technology* 42 (10):3732-3738. doi: https://doi.org/10.1021/es071301c.

200. Arche Consulting. 2020. "Threshold calculator for metals in soil (v3.0) ". https://www.arche-consulting.be/tools/threshold-calculator-for-metals-in-soil/.

480. Asmaryan, Shushanik, Vahagn Muradyan, Lilit Sahakyan, Armen Saghatelyan, and Timothy Warner. 2014. "Development of remote sensing methods for assessing and mapping soil pollution with heavy metals." In *GlobalSoilMap: Basis of the Global Spatial Soil Information System - Proceedings of the 1st GlobalSoilMap Conference*, 429-432. doi:https://doi.org/10.1201/b16500-77.

139. ASTM. 2009. ASTM D1452. Practice for Soil Exploration and Sampling by Auger Borings. ASTM International, West Conshohocken, PA. https://www.astm.org/DATABASE.CART/HISTORICAL/D1452-09.htm.

141. ASTM. 2013. ASTM D6169 / D6169M-13, Standard Guide for Selection of Soil and Rock Sampling Devices Used With Drill Rigs for Environmental Investigations. ASTM International, West Conshohocken, PA. https://www.astm.org/Standards/D6169.htm.

142. ASTM. 2014. ASTM D6282 / D6282M-14, Standard Guide for Direct Push Soil Sampling for Environmental Site Characterizations. ASTM International, West Conshohocken, PA. https://www.astm.org/Standards/D6282.htm.

135. ASTM. 2014. ASTM E2616-09(2014): Standard Guide for Remedy Selection Integrating Risk-Based Corrective Action and Non-Risk Considerations. ASTM International, West Conshohocken, PA. https://www.astm.org/search/fullsite-search.html?query=E2616&.

144. ASTM. 2015. ASTM D4700-15, Standard Guide for Soil Sampling from the Vadose Zone. ASTM International, West Conshohocken, PA. https://www.astm.org/Standards/D4700.htm.

140. ASTM. 2015. ASTM D6151 / D6151M-15, Standard Practice for Using Hollow-Stem Augers for Geotechnical Exploration and Soil Sampling. ASTM International, West Conshohocken, PA. https://www.astm.org/Standards/D6151.htm.

145. ASTM. 2016. ASTM D6914 / D6914M-16, Standard Practice for Sonic Drilling for Site Characterization and the Installation of Subsurface Monitoring Devices. ASTM International, West Conshohocken, PA. https://www.astm.org/Standards/D6914.htm.

171. ASTM. 2016. ASTM E178-16a, Standard Practice for Dealing With Outlying Observations. ASTM International, West Conshohocken, PA. https://www.astm.org/Standards/E178.htm.

170. ASTM. 2017. ASTM D6312-17, Standard Guide for Developing Appropriate Statistical Approaches for Groundwater Detection Monitoring Programs at Waste Disposal Facilities. ASTM International, West Conshohocken, PA. https://www.astm.org/Standards/D6312.htm.

128. ASTM. 2017. ASTM E456-13A(2017)e4, Standard Terminology Relating to Quality and Statistics. ASTM International, West Conshohocken, PA. https://www.astm.org/Standards/E456.htm.

127. ASTM. 2019. ASTM E177-19, Standard Practice for Use of the Terms Precision and Bias in ASTM Test Methods. ASTM International, West Conshohocken, PA. https://www.astm.org/Standards/E177.htm.

143. ASTM. 2020. ASTM D6286 / D6286M-20, Standard Guide for Selection of Drilling and Direct Push Methods for Geotechnical and Environmental Subsurface Site Characterization. ASTM International, West Conshohocken, PA. https://www.astm.org/Standards/D6286.htm.

136. ASTM. 2020. ASTM E1689-20, Standard Guide for Developing Conceptual Site Models for Contaminated Sites. ASTM International, West Conshohocken, PA. https://www.astm.org/Standards/E1689.htm.

146. ASTM. 2020. ASTM E3242-20, Standard Guide for Determination of Representative Sediment Background Concentrations. ASTM International, West Conshohocken, PA. https://www.astm.org/Standards/E3242.htm.

275. ATSDR, 2000. Toxicological Profile for Polychlorinated Biphenyls PCBs, Division of Toxicology/Toxicology Information Branch, Atlanta, Georgia. https://www.atsdr.cdc.gov/toxprofiles/tp17.pdf.

232. BC Environment, 2001. Identifying Populations - A guide for data analysts and interpreters on the identification of statistical populations. Contaminated Sites Statistical Applications Guidance Document No. 12-7, British Columbia Environment, British Columbia, CA. https://www2.gov.bc.ca/assets/gov/environment/air-land-water/site-remediation/docs/guidance-documents/gd07.pdf.

147. BC Environment, 2020. The British Columbia Field Sampling Manual - Part D1 - Soil Sampling and Investigations, British Columbia Ministry of the Environment and Climate Change Strategy, Victoria, B.C. https://www2.gov.bc.ca/gov/content/environment/research-monitoring-reporting/monitoring/laboratory-standards-quality-assurance/bc-field-sampling-manual.

277. Benotti, Mark, Loretta Fernandez, Graham Peaslee, Gregory Douglas, Allen Uhler, and Stephen Emsbo-Mattingly. 2020. "A forensic approach for distinguishing PFAS materials." *Environmental Forensics* 21:1-15. doi: https://doi.org/10.1080/15275922.2020.1771631.

279. Blumer, Max. 1976. "Polycyclic Aromatic Compounds in Nature." *Scientific American* 234:34-45. doi: https://doi.org/10.1038/scientificamerican0376-34.

99. Boehm, Paul, Jaana Pietari, Linda Cook, and Tarek Saba. 2018. "Improving rigor in polycyclic aromatic hydrocarbon source fingerprinting." *Environmental Forensics*:1-13. doi: https://doi.org/10.1080/15275922.2018.1474287.

420. Boerngen, Josephine G., and Hansford T. Shacklette. 1981. Chemical analyses of soils and other surficial materials of the conterminous United States. U. S. Geological Survey. http://pubs.er.usgs.gov/publication/ofr81197.

414. Bowell, R. J. 1994. "Sorption of arsenic by iron oxides and oxyhydroxides in soils." *Applied Geochemistry* 9 (3):279-286. doi: https://doi.org/10.1016/0883-2927(94)90038-8.

255. Boyle, R. W. 1974. "The use of major elemental ratios in detailed geochemical prospecting utilizing primary halos." *Journal of Geochemical Exploration* 3 (4):345-369. doi: https://doi.org/10.1016/0375-6742(74)90004-1.

4. Brisson, M. R., and D. Popp. 2017. Detected or Not – Defining Laboratory Analytical Limits. *The Synergist*. https://synergist.aiha.org/201711-detected-or-not.

285. Bro, Rasmus, and Age K. Smilde. 2014. "Principal component analysis." *Analytical Methods* 6 (9):2812-2831. doi: https://doi.org/10.1039/C3AY41907J.

160. Brooks, B., 2020. Soil Background Threshold Evaluation, Minnesota Pollution Control Agency, Environmental Analysis and Outcomes Division, St. Paul, MN.

286. Brusseau, Mark, R. Anderson, and Bo Guo. 2020. "PFAS concentrations in soils: Background levels versus contaminated sites." *Science of The Total Environment* 740:140017. doi: https://doi.org/10.1016/j.scitotenv.2020.140017.

288. Buck, Robert, James Franklin, Urs Berger, Jason Conder, Ian Cousins, Pim De Voogt, Allan Jensen, Scott Mabury, and Stefan van Leeuwen. 2011. "Perfluoroalkyl and Polyfluoroalkyl Substances in the Environment: Terminology, Classification, and Origins." *Integrated Environmental Assessment and Management* 7:513-41. doi: https://doi.org/10.1002/ieam.258.

464. Cal DTSC, 2008. Proven Technologies and Remedies Guidance – Remediation of Metals in Soils: Appendix B Strategies for Establishing and Using Background Estimates of Metals in Soils, Department of Toxic Substances Control, Sacramento, CA. https://dtsc.ca.gov/wp-content/uploads/sites/31/2018/11/Appdx_B_083108.pdf.

52. Cal DTSC, 2009. Arsenic Strategies Determination of Arsenic Remediation Development of Arsenic Cleanup Goals, California Department of Toxic Substances Control, Sacramento, California, January 16, 2009. https://dtsc.ca.gov/wp-content/uploads/sites/31/2018/01/Arsenic-Cleanup-Goals-Jan09-A.pdf.

201. Cal DTSC, 2009. Use of northern and southern California polycyclic aromatic hydrocarbon (PAH) studies in the manufactured gas plant site cleanup process, California Department of Toxic Substances Control, Sacramento, California, JUly 1, 2009. https://dtsc.ca.gov/wp-content/uploads/sites/31/2018/01/MGP_PAH_Advisory_070109.pdf.

444. CCME, 2007. Reference Method for the Canada Wide Standard for Petroleum Hydrocarbons in Soil - Tier 1 Method, Canadian Council of Ministers of the Environment, Winnipeg, Manitoba. http://www.ccme.ca.

46. Chambers, J.M., W.S. Cleveland, B. Kleiner, and P. Tukey. 1983. *Graphical Methods for Data Analysis*. New York: Chapman and Hall/CRC. doi:https://doi.org/10.1201/9781351072304.

202. Checkai, Ron, Eric Van Genderen, José Paulo Sousa, Gladys Stephenson, and Erik Smolders. 2014. "Deriving site-specific clean-up criteria to protect ecological receptors (plants and

soil invertebrates) exposed to metal or metalloid soil contaminants via the direct contact exposure pathway." *Integrated Environmental Assessment and Management* 10 (3):346-357. doi: https://doi.org/10.1002/ieam.1528.

173. Chen, Ming, Gerco Hoogeweg, and W. Harris. 2001. "Arsenic Background Concentrations in Florida, U.S.A. Surface Soils: Determination and Interpretation." *Environmental Forensics - ENVIRON FORENSICS* 2:117-126. doi: https://doi.org/10.1006/enfo.2001.0050.

14. Cochran, William G. . 1997. *Sampling Techniques, 3rd Edition*, *Wiley Series in Probability and Mathmatical Statistics-Applied*. New York: John Wiley & Sons, https://www.wiley.com/en-us/Sampling+Techniques%2C+3rd+Edition-p-9780471162407.

47. Cook, P. D. 1998. "Estimating Background Concentrations of Inorganic Analytes from On-Site Soil Sample Data." In, edited by K. Hoddinott, 3-22. West Conshohocken, PA: ASTM International. doi:https://doi.org/10.1520/STP13283S.

256. Cornell, R.M., and U Schwertmann. 2003. "The Iron Oxides: Structure, Properties, Reactions, Occurences and Uses, Second Edition." In.: Wiley-VCH Verlag GmbH & Co. KGaA. doi:https://doi.org/10.1002/3527602097.

296. Creaser, C. S., A. R. Fernandes, S. J. Harrad, T. Hurst, and E. A. Cox. 1989. "Background levels of polychlorinated biphenyls in British soils - II." *Chemosphere* 19 (8):1457-1466. doi: https://doi.org/10.1016/0045-6535(89)90094-5.

233. De Oliveira, Vinicius, Cleide Abreu, Ricardo Coelho, and Leônidas Melo. 2014. "Cadmium background concentrations to establish reference quality values for soils of Sao Paulo State, Brazil." *Environmental Monitoring and Assessment* 186. doi: 10.1007/s10661-013-3462-2.

485. Douglas, Gregory S., Stephen D. Emsbo-Mattingly, Scott A. Stout, Allen D. Uhler, and Kevin J. McCarthy. 2007. "CHAPTER 9 - CHEMICAL FINGERPRINTING METHODS." In *Introduction to Environmental Forensics (Second Edition)*, edited by Brian L. Murphy and Robert D. Morrison, 311-454. Burlington: Academic Press. doi:https://doi.org/10.1016/B978-012369522-2/50010-5.

114. Douglas, Gregory S., Stephen D. Emsbo-Mattingly, Scott A. Stout, Allen D. Uhler, and Kevin J. McCarthy. 2015. "Chapter 8 - Hydrocarbon Fingerprinting Methods." In *Introduction to Environmental Forensics (Third Edition)*, edited by Brian L. Murphy and Robert D. Morrison, 201-309. San Diego: Academic Press. doi:https://doi.org/10.1016/B978-0-12-404696-2.00008-4.

301. Dunn, W.J., D.L. Stalling, and S. Wold. 1984. "The use of Simca Pattern Recognition in the Analysis of Complex Chromatographic Data." In *QSAR in Environmental Toxicology*, edited by Kaiser K.L.E.: Springer, Dordrecht. doi:https://doi.org/10.1007/978-94-009-6415-0_6.

115. Dworian, Paul. 1996. "Interference by Natural Organics in Diesel Analyses." Eighth International Conference on Cold Regions Engineering, Fairbanks, Alaska, August 12-16.https://cedb.asce.org/CEDBsearch/record.jsp?dockey=0102157.

303. Elmendorf, D. L., C. E. Haith, G. S. Douglas, and Roger Prince. 1994. "Relative rates of biodegradation of substituted polycyclic aromatic hydrocarbons." In *Bioremediation of Chlorinated and Polycyclic Aromatic Hydrocarbons Compounds*, 188-202. Lewis Publisher.

152. EPA, Ohio, 2013. Discrete Soil Sampling, Ohio EPA Division of Environmental Response and Revitalization, Columbus, Ohio.

https://epa.ohio.gov/portals/30/rules/FSOP%202.1.1,%20Discrete%20Soil%20Sampling,%202-12-13%20FINAL.pdf.

53. Fleischhauer, Henry L., and Nic Korte. 1990. "Formulation of cleanup standards for trace elements with probability plots." *Environmental Management* 14 (1):95-105. doi: https://doi.org/10.1007/BF02394023.

119. Galarneau, Elisabeth. 2008. "Source specificity and atmospheric processing of airborne PAHs: Implications for source apportionment." *Atmospheric Environment* 42 (35):8139-8149. doi: https://doi.org/10.1016/j.atmosenv.2008.07.025.

15. Geiselbrecht, Allison, Shahrokh Rouhani, Karen Thorbjornsen, Douglas Blue, Steven Nadeau, Tessa Gardner-Brown, and Steven Brown. 2019. "Important Considerations in the Derivation of Background at Sediment Sites." *Integrated Environmental Assessment and Management* 15. doi: https://doi.org/10.1002/ieam.4124.

479. Gholizadeh, Asa, Mehdi Saberioon, Eyal Ben-Dor, and Lubos Boruvka. 2018. "Monitoring of Selected Soil Contaminants using Proximal and Remote Sensing Techniques: Background, State-of-the-Art and Future Perspectives." *Critical Reviews in Environmental Science and Technology* 48:243-278. doi: https://doi.org/10.1080/10643389.2018.1447717.

18. Gibbons, Robert D. 1994. *Statistical Methods for Groundwater Monitoring*: John Wiley & Sons. doi:https://doi.org/10.1002/9780470172940.

16. Grubbs, Frank E. 1969. "Procedures for Detecting Outlying Observations in Samples."

54. Halil, Aral, and C. Sarac. 1988. "Partitioning of Geochemical Populations by Sinclair's Method. An Application on a Geochemical Stream Sediment Data from the Belgium Ardennes." *Communications, University of Ankara Faculty of Sciences* Series C, Vol 6:313 - 323, https://dspace.ankara.edu.tr/xmlui/bitstream/handle/20.500.12575/62120/16191.pdf?sequence=1&isAllowed=y.

258. Hawkes, H.E., and J.S. Webb. 1962. *Geochemistry in Mineral Exploration*. New York, NY: Harper & Row.

476. Hayden, R. E., P.J. Remington, M.A Theobals, and J.F. Wiley. 1985. Methods for designing treatments to reduce interior noise of predominant sources and paths in a single engine light aircraft. Bolt Beranek and Newman Inc. Cambridge, MA: National Aeronautics and Space Administration, Langley Research Center. https://core.ac.uk/download/pdf/42837965.pdf.

20. Helsel, Dennis R. 2005. *Nondetects and data analysis: Statistics for censored environmental data*: John Wiley & Sons, http://pubs.er.usgs.gov/publication/70180734.

21. Helsel, Dennis R., and Timothy A. Cohn. 1988. "Estimation of descriptive statistics for multiply censored water quality data." *Water Resources Research* 24 (12):1997-2004. doi: https://doi.org/10.1029/WR024i012p01997.

22. Hewett, P., and G. H. Ganser. 2007. "A comparison of several methods for analyzing censored data." *Annals of Occupational Hygiene* 51 (7):611-32. doi: https://doi.org/10.1093/annhyg/mem045.

439. HI DOH, 2011. Update to Soil Action Levels for Inorganic Arsenic and Recommended Soil Management Practices, Updated September 2012, Department of Health, Honolulu, Hawaii. http://hawaiidoh.org/references/HDOH%202011c.pdf.

235. HI DOH, 2012. Hawaiian Islands Soil Metal Background Evaluation Report, Department of Health Hazard Evaluation and Emergency Response, Honolulu, Hawaii. https://health.hawaii.gov/heer/files/2012/05/Hawaiian-Islands-Soil-Metal-Background-Evaluation-Report-May-2012.pdf.

125. Hindersmann, Benjamin, and Christine Achten. 2018. "Urban soils impacted by tailings from coal mining: PAH source identification by 59 PAHs, BPCA and alkylated PAHs." *Environmental Pollution* 242:1217-1225. doi: https://doi.org/10.1016/j.envpol.2018.08.014.

434. Hoaglin, David, Frederick Mosteller, and John W. Tukey. 1983. *Understanding Robust and Exploratory Data Analysis*. Vol. 96: John Wiley and Sons. doi:https://doi.org/10.2307/2670272.

23. Hornung, Richard W., and Laurence D. Reed. 1990. "Estimation of Average Concentration in the Presence of Nondetectable Values." *Applied Occupational and Environmental Hygiene* 5 (1):46-51. doi: https://doi.org/10.1080/1047322X.1990.10389587.

312. Horstmann, M., M. S. McLachlan, and M. Reissinger. 1993. "Investigations of the origin of PCDD/F in municipal sewage sludge." *Chemosphere* 27 (1):113-120. doi: https://doi.org/10.1016/0045-6535(93)90283-B.

176. IDEM, 2012. Remediation Closure Guide, Indiana Department of Environmental Management, Office of Land Quality, Indianapolis, IN. https://www.in.gov/idem/cleanups/files/remediation_closure_guide.pdf.

177. IDEQ, 2014. Statistical Guidance for Determining Background and Ground Water Quality and Degradation, Idaho Department of Environmental Quality, Boise, ID. https://www.deq.idaho.gov/media/1226/guidance-statistical-degradation.pdf.

475. IDEQ, 2018. Idaho Risk Evaluation Manual for Petroleum Releases, Idaho Department of Environmental Quality, Boise, ID. file:///C:/Users/jrocc/AppData/Local/Temp/MicrosoftEdgeDownloads/f3705aa2-8af1-468d-880a-73711105aee5/idaho-risk-evaluation-manual-for-petroleum-releases-2018.pdf.

178. Iglewicz, B., and D.C. Hoaglin. 1993. *How to Detect and Handle Outliers*: ASQC Quality Press, https://books.google.com/books?id=siInAQAAIAAJ.

6. ISO. 2017. Standard 17025 - General requirements for the competence of testing and calibration laboratories. International Organization for Standardization. https://www.iso.org/standard/66912.html.

488. ITRC, 2007. Triad Implementation Guide, Interstate Technology Regulatory Council, Characterization Sampling, and Monitoring Team, Washington, D.C. http://www.itrcweb.org/GuidanceDocuments/SCM-3.pdf.

148. ITRC, 2012. ISM-1 - Incremental Sampling Methodology, Interstate Technology Regulatory Council, Remediation Process Optimization Team, Washington, D.C. https://www.itrcweb.org/GuidanceDocuments/ISM-1_2012_with_Clarifications.pdf.

137. ITRC, 2013. GSMC-1 - Groundwater Statistics and Monitoring Compliance: Statistical Tools for the Project Life Cycle, Interstate Technology Regulatory Council, Groundwater Statistics and Monitoring Compliance Team, Washington, D.C. https://projects.itrcweb.org/gsmc-1/.

237. ITRC, 2015. Risk-3 - Decision Making at Contaminated Sites: Issues and Options in Human Health Risk Assessment, Interstate Technology Regulatory Council, Risk Assessment Team, Washington, D.C. www.itrcweb.org/risk-3.

206. ITRC, 2017. BCS-1 - Bioavailability of Contaminants in Soil, Interstate Technology Regulatory Council, Bioavailability in Contaminated Soil Team, Washington, D.C. https://bcs-1.itrcweb.org/1-introduction/.

316. ITRC, 2018. TPH Risk Evaluation at Petroleum- Contaminated Sites, Interstate Technology Regulatory Council, TPH Risk Evaluation Team, Washington, D.C. https://tphrisk-1.itrcweb.org/.

317. ITRC, 2019. ASCT-1 Implementing Advanced Site Characterization Tools, Interstate Technology Regulatory Council, Advanced Site Characterization Tools Team, Washington, D.C. https://asct-1.itrcweb.org.

318. ITRC, 2020. History and Use of Per- and Polyfluoroalkyl Substances (PFAS), Interstate Technology Regulatory Council, PFAS Team, Washington, D.C. https://pfas-1.itrcweb.org/fact-sheets/.

149. ITRC, 2020. ISM-1 Clarifications - Clarifications to ITRC 2012 ISM-1 Guidance, Interstate Technology Regulatory Council, Remediation Process Optimization Team, Washington, D.C. https://www.itrcweb.org/GuidanceDocuments/ISM-1_Clarifications.pdf.

431. ITRC, 2020. ISM-2 - Incremental Sampling Methodology Update, Interstate Technology Regulatory Council, Remediation Process Optimization Team, Washington, D.C. https://ism-2.itrcweb.org/.

319. ITRC, 2020. Naming Conventions and Physical and Chemical Properties of Per- and Polyfluoroalkyl Substances (PFAS), Interstate Technology Regulatory Council, PFAS Team, Washington, D.C. https://pfas-1.itrcweb.org/fact-sheets/.

447. ITRC, 2021. PFAS Technical and Regulatory Guidance Document, Interstate Technology Regulatory Council, PFAS Team, Washington, D.C. https://pfas-1.itrcweb.org/.

323. Johnson, Glenn W., Robert Ehrlich, William Full, and Scott Ramos. 2007. "Principal Components Analysis and Receptor Models in Environmental Forensics." In *Introduction to Environmental Forensics*, edited by Brian L. Murphy and Robert D. Morrison. New York, New York: Elsevier.

448. Johnson, Glenn W., I. I. I. John F. Quensen, Jeffrey R. Chiarenzelli, and M. Coreen Hamilton. 2006. "Chapter 10 - Polychlorinated Biphenyls." In *Environmental Forensics Contaminant Specific Guide*, edited by Robert D. Morrison and Brian L. Murphy, 187-225. Burlington: Academic Press. doi:https://doi.org/10.1016/B978-012507751-4/50032-X.

239. Johnson, Richard A., and Dean W. Wichren. 2015. *Applied Multivariate Statistical Analysis (5th Edition)*. Englewood Cliffs: Prentice Hall.

151. Kabata-Pendias, Alina. 2010. *Trace elements in soils and plants: Fourth edition*. doi:https://doi.org/10.1201/b10158.

325. Katsoyiannis, Athanasios, and Knut Breivik. 2014. "Model-based evaluation of the use of polycyclic aromatic hydrocarbons molecular diagnostic ratios as a source identification tool." *Environmental Pollution* 184:488-494. doi: https://doi.org/10.1016/j.envpol.2013.09.028.

179. KDHE, 2010. Determining Background Levels for Chemicals of Concern, Kansas Department of Health and Environment, Bureau of Environmental Remediation/Remedial Section, (BER Policy # BER-RS-038 (Revised)). https://www.kdheks.gov/ber/policies/BER_RS_038.pdf.

180. KEEC, 2004. Kentucky Guidance for Ambient Background Assessment, Kentucky Energy & Environment Cabinet, Natural Resources and Environmental Protection Cabinet. https://eec.ky.gov/Environmental-Protection/Waste/superfund/Documents/AmbientBackgroundAssessment.pdf.

56. Kim, Kyoung-Ho, Seong-Taek Yun, Hyun-Koo Kim, and Ji-Wook Kim. 2015. "Determination of natural backgrounds and thresholds of nitrate in South Korean groundwater using

model-based statistical approaches." *Journal of Geochemical Exploration* 148:196-205. doi: https://doi.org/10.1016/j.gexplo.2014.10.001.

330. Larsen, Randolph K., and Joel E. Baker. 2003. "Source Apportionment of Polycyclic Aromatic Hydrocarbons in the Urban Atmosphere: A Comparison of Three Methods." *Environmental Science & Technology* 37 (9):1873-1881. doi: https://doi.org/10.1021/es0206184.

259. Levinson, A.A. 1974. *Introduction to Exploration Geochemistry*. Calgary: Applied Publishing.

333. Li, An, Jae-Kil Jang, and Peter A. Scheff. 2003. "Application of EPA CMB8.2 Model for Source Apportionment of Sediment PAHs in Lake Calumet, Chicago." *Environmental Science & Technology* 37 (13):2958-2965. doi: https://doi.org/10.1021/es026309v.

26. Li, Shi, Stuart Batterman, Feng-Chiao Su, and Bhramar Mukherjee. 2013. "Addressing extrema and censoring in pollutant and exposure data using mixture of normal distributions." *Atmospheric environment (Oxford, England : 1994)* 77:10.1016/j.atmosenv.2013.05.004. doi: https://doi.org/10.1016/j.atmosenv.2013.05.004.

334. Lima, Ana, John Farrington, and Christopher Reddy. 2005. "Combustion-Derived Polycyclic Aromatic Hydrocarbons in the Environment—A Review." *Environmental Forensics - ENVIRON FORENSICS* 6:109-131. doi: https://doi.org/10.1080/15275920590952739.

38. Loo, Mark van der. 2010. "Distribution based outlier detection in univariate data."

337. MADEP, 2002. Technical Update: Background Levels of Polycyclic Aromatic Hydrocarbons and Metals in Soil. 310 CMR 40.0006, Massachusetts Department of Environmental Protection, Boston, MA. https://www.mass.gov/doc/technical-update-background-levels-of-polycyclic-aromatic-hydrocarbons-and-metals-in-soil-0/download.

432. Marchant, B. P., S. Newman, R. Corstanje, K. R. Reddy, T. Z. Osborne, and R. M. Lark. 2009. "Spatial monitoring of a non-stationary soil property: phosphorus in a Florida water conservation area." *European Journal of Soil Science* 60 (5):757-769. doi: https://doi.org/10.1111/j.1365-2389.2009.01158.x.

181. Maronna, Ricardo, Douglas Martin, and Victor Yohai. 2006. "Robust Statistics: Theory and Methods." doi: https://doi.org/10.1002/0470010940.

57. Matschullat, Jörg, R. Ottenstein, and Clemens Reimann. 2000. "Geochemical background - Can we calculate it?" *Environmental Geology* 39:990-1000. doi: https://doi.org/10.1007/s002549900084.

425. Matzke, B. D., Wilson J. E., Newburn L. L., Dowson S. T., Hathaway J. E., Sego L. H., Bramer L. M., and Pulsipher B. A., 2014. Visual Sample Plan Version 7.0 User's Guide (PNNL-23211), Pacific Northwest National Laboratory, Richland, Washington. https://vsp.pnnl.gov/docs/PNNL-23211.pdf.

454. Mayes, B. A., E. McConnell, B. Neal, M. Brunner, S. Hamilton, T. Sullivan, A. C. Peters, M. Ryan, J. Toft, A. W. Singer, J. Brown, R. Menton, and J. A. Moore. 1998. "Comparative Carcinogenicity in Sprague-Dawley Rats of the Polychlorinated Biphenyl Mixtures Aroclors 1016, 1242, 1254, and 1260." *Toxicological sciences : an official journal of the Society of Toxicology* 41:62-76. doi: https://doi.org/10.1006/toxs.1997.2397.

456. McKillup, Steve. 2012. *Statistics Explained: An Introductory Guide for Life Scientists*. Cambridge and New York: Cambridge University Press.

468. McLachlan, Geoffrey J. 2004. *Discriminant Analysis and Statistical Pattern Recognition*: Wiley Interscience.

415. McLean, J., and B. E. Bledsoe, 1992. Ground Water Issue: Behavior of Metals in Soils, Office of Research and Development, Washington, D.C., (EPA/540/S-92/018).

https://nepis.epa.gov/Exe/ZyNET.exe/10002DSF.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1991+Thru+1994&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=&File=D%3A%5Czyfiles%5CIndex%20Data%5C91thru94%5CTxt%5C00000003%5C10002DSF.txt&User=ANONYMOUS&Password=anonymous&SortMethod=h%7C-&MaximumDocuments=1&FuzzyDegree=0&ImageQuality=r75g8/r75g8/x150y150g16/i425&Display=hpfr&DefSeekPage=x&SearchBack=ZyActionL&Back=ZyActionS&BackDesc=Results%20page&MaximumPages=1&ZyEntry=1&SeekPage=x&ZyPURL.

339. Meijer, S. N., W. A. Ockenden, A. Sweetman, K. Breivik, J. O. Grimalt, and K. C. Jones. 2003. "Global Distribution and Budget of PCBs and HCB in Background Surface Soils: Implications for Sources and Environmental Processes." *Environmental Science & Technology* 37 (4):667-672. doi: https://doi.org/10.1021/es025809l.

340. Menzie, Charles A., Bonnie B. Potocki, and Joseph Santodonato. 1992. "Exposure to carcinogenic PAHs in the environment." *Environmental Science & Technology* 26 (7):1278-1284. doi: https://doi.org/10.1021/es00031a002.

260. Morrison, Jean M., Martin B. Goldhaber, Lopaka Lee, JoAnn M. Holloway, Richard B. Wanty, Ruth E. Wolf, and James F. Ranville. 2009. "A regional-scale study of chromium and nickel in soils of northern California, USA." *Applied Geochemistry* 24 (8):1500-1511. doi: https://doi.org/10.1016/j.apgeochem.2009.04.027.

344. Morrison, R.D., and B.L. Murphy. 2006. *Environmental Forensics: Third Edition*. Burlington, MA: Academic Press.

459. MPCA, 2016. Background Threshold Value (BTV) Evaluation, Remediation and Environmental Analysis and Outcomes Divisions, Saint Paul, MN. https://www.pca.state.mn.us/waste/risk-based-site-evaluation-guidance.

460. MPCA, 2016. Soil Reference Value (SRV) Technical Support Document, Remediation and Environmental Analysis and Outcomes Divisions, Saint Paul, MN. https://www.pca.state.mn.us/waste/risk-based-site-evaluation-guidance.

473. MTDEQ, 2005. Action Level for Arsenic in Surface Soil, Montana Department of Environmental Quality Remediation Division, Helena, MT.

469. Myers, Jonathan, and Karen Thorbjornsen. 2004. "Identifying Metals Contamination in Soil: A Geochemical Approach." *Soil and Sediment Contamination: An International Journal* 13 (1):1-16. doi: https://doi.org/10.1080/10588330490269732.

347. Nam, Jae, Andrew Sweetman, and Kevin Jones. 2009. "Polynuclear aromatic hydrocarbons (PAHs) in global background soils." *Journal of environmental monitoring : JEM* 11:45-8. doi: https://doi.org/10.1039/b813841a.

8. NELAC. 2016. Volume 1: Management and Technical Requirements for Laboratories Performing Environmental Analysis The NELAC Institute (TNI). https://www.nelac-institute.org/content/CSDP/standards.php?ap3=1_2.

350. Noémi, Barabás, Pierre Goovaerts, and Peter Adriaens. 2004. "Modified Polytopic Vector Analysis To Identify and Quantify a Dioxin Dechlorination Signature in Sediments. 2. Application to the Passaic River." *Environmental Science & Technology* 38 (6):1821-1827. doi: https://doi.org/10.1021/es026229r.

29. Nysen, Ruth, Christel Faes, Pietro Ferrari, Philippe Verger, and Marc Aerts. 2015. "Parametric and semi-nonparametric model strategies for the estimation of distributions of chemical contaminant data."

262. Pal, Arundhati, and A. K. Paul. 2008. "Microbial extracellular polymeric substances: Central elements in heavy metal bioremediation." *Indian journal of microbiology* 48:49-64. doi: https://doi.org/10.1007/s12088-008-0006-5.

60. Panno, S.V., W.R. Kelly, A.T. Martinsek, and K.C. Hackley. 2007. Database for estimating background and threshold nitrate concentrations using probability graphs. Illinois Department of Natural Resources Illinois State Geological Survey. Illinois State Geological Survey. http://hdl.handle.net/2142/50299.

59. Panno, Samuel, Walton Kelly, Adam Martinsek, and Keith Hackley. 2006. "Estimating Background and Threshold Nitrate Concentrations Using Probability Graphs." *Ground water* 44:697-709. doi: https://doi.org/10.1111/j.1745-6584.2006.00240.x.

58. Papastergios, Georgios, J. L. Fernandez-Turiel, Anestis Filippidis, and Domingo Gimeno. 2011. "Determination of geochemical background for environmental studies of soils via the use of HNO3 extraction and Q–Q plots." *Environmental earth sciences* 64:743-751. doi: https://doi.org/10.1007/s12665-010-0894-7.

30. Popovic, M., H. Nie, David Chettle, and F. McNeill. 2007. "Random left censoring: A second look at bone lead concentration measurements." *Physics in medicine and biology* 52:5369-78. doi: https://doi.org/10.1088/0031-9155/52/17/018.

353. Rauert, Cassandra, Mahiba Shoieb, Jasmin K. Schuster, Anita Eng, and Tom Harner. 2018. "Atmospheric concentrations and trends of poly- and perfluoroalkyl substances (PFAS) and volatile methyl siloxanes (VMS) over 7 years of sampling in the Global Atmospheric Passive Sampling (GAPS) network." *Environmental Pollution* 238:94-102. doi: https://doi.org/10.1016/j.envpol.2018.03.017.

62. Reimann, Clemens, Peter Filzmoser, and Robert Garrett. 2005. "Background and threshold: Critical comparison of methods of determination." *The Science of the total environment* 346:1-16. doi: https://doi.org/10.1016/j.scitotenv.2004.11.023.

354. Reimann, Clemens, and Robert G. Garrett. 2005. "Geochemical background—concept and reality." *Science of The Total Environment* 350 (1):12-27. doi: https://doi.org/10.1016/j.scitotenv.2005.01.047.

64. Renez, Andrew N., Robert G. Garrett, Inez M. Kettles, Eric C. Grunsky, and R J McNeil. 2011. "Using soil geochemical data to estimate the range of background element concentrations for ecological and human-health risk assessments." Workshop on the role of geochemical data in environmental and human health risk assessment, Halifax.https://geoscan.nrcan.gc.ca/starweb/geoscan/servlet.starweb?path=geoscan/fulle.web&search1=R=288746.

184. Rousseeuw, Peter J., and Annick M. Leroy. 1987. *Robust Regression and Outlier Detection*. USA: John Wiley & Sons, Inc. doi:https://doi.org/10.1002/0471725382.

186. Rousseeuw, Peter J., and Bert Zomeren. 1990. "Unmasking Multivariate Outliers and Leverage Points." *Journal of The American Statistical Association - J AMER STATIST ASSN* 85:633-639. doi: https://doi.org/10.1080/01621459.1990.10474920.

33. Schmoyeri, R. L., J. J. Beauchamp, C. C. Brandt, and F. O. Hoffman. 1996. "Difficulties with the lognormal model in mean estimation and testing." *Environmental and Ecological Statistics* 3 (1):81-97. doi: https://doi.org/10.1007/BF00577325.

421. Shacklette, Hansford T., and Josephine G. Boerngen. 1984. Element Concentrations in Soils and Other Surficial Materials of the Conterminous United States. U. S. Geological Survey. https://pubs.usgs.gov/pp/1270/pdf/PP1270_508.pdf.

34. She, Nian. 1997. "Analyzing Censored Water Quality Data Using a Non-Parametric Approach." *JAWRA Journal of the American Water Resources Association* 33 (3):615-624. doi: https://doi.org/10.1111/j.1752-1688.1997.tb03536.x.

360. Shields, Walter J., Yves Tondeur, Laurie Benton, and Melanie R. Edwards. 2006. "14 - Dioxins and Furans." In *Environmental Forensics*, edited by Robert D. Morrison and Brian L. Murphy, 293-312. Burlington: Academic Press. doi:https://doi.org/10.1016/B978-012507751-4/50036-7.

212. Siemering, Geoffrey S., and Robert Thiboldeaux. 2021. "Background concentration, risk assessment and regulatory threshold development: Polycyclic aromatic hydrocarbons (PAH) in Milwaukee, Wisconsin surface soils." *Environmental Pollution* 268:115772. doi: https://doi.org/10.1016/j.envpol.2020.115772.

65. Sinclair, A. J. 1974. "Selection of threshold values in geochemical data using probability graphs." *Journal of Geochemical Exploration* 3 (2):129-149. doi: https://doi.org/10.1016/0375-6742(74)90030-2.

67. Sinclair, A. J. 1983. "Chapter 3 - Univariate Analysis." In *Handbook of Exploration Geochemistry*, edited by R. J. Howarth, 59-81. Elsevier Science B.V. doi:https://doi.org/10.1016/B978-0-444-42038-1.50009-5.

90. Sinclair, A. J. 1991. "A fundamental approach to threshold estimation in exploration geochemistry: probability plots revisited." *Journal of Geochemical Exploration* 41 (1):1-22. doi: https://doi.org/10.1016/0375-6742(91)90071-2.

66. Sinclair, A.J. 1976. *Applications of Probability Graphs in Mineral Exploration*: Association of Exploration Geochemists, https://books.google.com/books?id=Ou0DAQAAIAAJ.

188. Singh, A. K., Anita Singh, and Engelhardt Me, 1997. The Lognormal Distribution in Environmental Applications, Office of Research and Development, 01/01. https://nepis.epa.gov/Exe/ZyNET.exe/P10078OC.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1995+Thru+1999&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=&File=D%3A%5Czyfiles%5CIndex%20Data%5C95thru99%5CTxt%5C00000026%5CP10078OC.txt&User=ANONYMOUS&Password=anonymous&SortMethod=h%7C-&MaximumDocuments=1&FuzzyDegree=0&ImageQuality=r75g8/r75g8/x150y150g16/i425&Display=hpfr&DefSeekPage=x&SearchBack=ZyActionL&Back=ZyActionS&BackDesc=Results%20page&MaximumPages=1&ZyEntry=1&SeekPage=x&ZyPURL.

416. Singh, A., and J.M. Nocerino. 1995. "Robust Procedures for the Identification of Multiple Outliers." In *The Handbook of Environmental Chemistry*, edited by Einax J., 229-277. Berlin, Heidelberg: Springer. doi:https://doi.org/10.1007/978-3-540-49148-4_8.

187. Singh, Anita. 1996. "Outliers and robust procedures in some chemometric applications." *Chemometrics and Intelligent Laboratory Systems* 33 (2):75-100. doi: https://doi.org/10.1016/0169-7439(95)00087-9.

189. Singh, Anita, A. K. Singh, and Ross Iaci. 2002. "Estimation of the Exposure Point Concentration Term Using a Gamma Distribution."

91. Singh, Anita, Ashok K. Singh, and George Flatman. 1994. "Estimation of background levels of contaminants." *Mathematical Geology* 26 (3):361-388. doi: https://doi.org/10.1007/BF02089229.

361. Singh, Sunita 2016. "Remote sensing applications in soil survey and mapping: A Review " *International Journal of Geomatics and Geosciences* 7 (2):192:203, http://www.ipublishing.co.in/jggsarticles/volseven/EIJGGS7015.pdf.

35. Sinha, P., M. B. Lambert, and V. L. Trumbull. 2006. "Evaluation of statistical methods for left-censored environmental data with nonuniform detection limits." *Environ Toxicol Chem* 25 (9):2533-40. doi: https://doi.org/10.1897/05-548r.1.

362. Slonecker, E. Terrence, and Gary B. Fisher. 2014. An evaluation of remote sensing technologies for the detection of fugitive contamination at selected Superfund hazardous waste sites in Pennsylvania. U. S. Geological Survey. Reston, VA. http://pubs.er.usgs.gov/publication/ofr20141081.

423. Smith, David B., Federico Solano, Laurel G. Woodruff, William F. Cannon, and Karl J. Ellefsen. 2019. Geochemical and mineralogical maps, with interpretation, for soils of the conterminous United States. U. S. Geological Survey. Reston, VA. http://pubs.er.usgs.gov/publication/sir20175118.

213. Smith, David, William Cannon, Laurel Woodruff, Federico Solano, and K.J Ellefsen. 2014. Geochemical and mineralogical maps for soils of the conterminous United States: U.S. Geological Survey Open-File Report 2014–1082. U.S. Geological Survey. https://pubs.usgs.gov/of/2014/1082/.

40. Smith, David, William Cannon, Laurel Woodruff, Federico Solano, James Kilburn, and David Fey. 2013. Geochemical and mineralogical data for soils of the conterminous United States: U.S. Geological Survey Data Series 801. U.S. Geological Survey. https://pubs.usgs.gov/ds/801/.

263. Sposito, Garrison. 2016. *The Chemistry of Soils. Third Edition*: Oxford University Press.

482. Stensvold, Krista A. 2012. "Distribution and variation of arsenic in Wisconsin surface soils, with data on other trace elements." *Scientific Investigations Report*, https://pubs.usgs.gov/sir/2011/5202/.

365. Stout, Scott A, and Allen Uhler. 2003. "Distinguishing “Background” Hydrocarbons from Contamination Using Chemical Fingerprinting." *Environmental Claims Journal* 15:241-259. doi: https://doi.org/10.1080/713609448.

363. Stout, Scott A., and Stephen D. Emsbo-Mattingly. 2008. "Concentration and character of PAHs and other hydrocarbons in coals of varying rank – Implications for environmental studies of soils and sediments containing particulate coal." *Organic Geochemistry* 39 (7):801-819. doi: https://doi.org/10.1016/j.orggeochem.2008.04.017.

369. Stout, Scott A., Stephen D. Emsbo-Mattingly, Gregory S. Douglas, Allen D. Uhler, and Kevin J. McCarthy. 2015. "Beyond 16 Priority Pollutant PAHs: A Review of PACs used in Environmental Forensic Chemistry." *Polycyclic Aromatic Compounds* 35 (2-4):285-315. doi: https://doi.org/10.1080/10406638.2014.891144.

364. Stout, Scott A., and Thomas P. Graan. 2010. "Quantitative Source Apportionment of PAHs in Sediments of Little Menomonee River, Wisconsin: Weathered Creosote versus Urban Background." *Environmental Science & Technology* 44 (8):2932-2939. doi: https://doi.org/10.1021/es903353z.

370. Stout, Scott A., A. D. Uhler, and S. D. Emsbo-Mattingly. 2004. "Comparative evaluation of background anthropogenic hydrocarbons in surficial sediments from nine urban waterways." *Environ Sci Technol* 38 (11):2987-94. doi: https://doi.org/10.1021/es040327q.

371. Stout, Scott A., A.D. Uhler, K.J. McCarthy, and S.D Emsbo-Mattingly. 2002. "Chemical Fingerprinting of Hydrocarbons." In *Introduction to Environmental Forensics,*, edited by B. Murphy and R. Morrison, 135-260. Boston, MA: Academic Press.

366. Stout, Scott A., and Zhendi Wang. 2008. "Diagnostic compounds for fingerprinting petroleum in the environment." In *Environmental Forensics, Issues in Environmental Science and Technology*, 54-104. The Royal Society of Chemistry.

264. Stumm, W, and J. J. Morgan. 1996. *Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters, 3rd Edition*: Wiley-Interscience doi:https://doi.org/10.1016/s0016-7037(97)81133-7.

463. Taggart, Joseph E., Jr. 2002. *Analytical methods for chemical analysis of geologic and other materials, U.S. Geological Survey [electronic resource] / by Joseph E. Taggart, Jr., editor*. Edited by Survey Geological, *U.S. Geological Survey open-file report ; 02-223.* Denver, Colo: U.S. Dept. of the Interior, U.S. Geological Survey.

269. Thorbjornsen, Karen. 2008. "Building a Better Background Data Set: The Importance of Considering Geochemistry." 2008 Joint Meeting of The Geological Society of America, Soil Science Society of America, American Society of Agronomy, Crop Science Society of America, Gulf Coast Association of Geological Societies with the Gulf Coast Section of SEPM, Pittsburgh, PA.https://gsa.confex.com/gsa/2008AM/webprogram/Paper150295.html.

267. Thorbjornsen, Karen, and Jonathan Myers. 2007. "Identification of Metals Contamination in Firing-Range Soil Using Geochemical Correlation Evaluation." *Soil & Sediment Contamination: An International Journal* 16:337-349. doi: https://doi.org/10.1080/15320380701404391.

268. Thorbjornsen, Karen, and Jonathan Myers. 2007. "Identifying Metals Contamination in Groundwater Using Geochemical Correlation Evaluation." *Environmental Forensics* 8:25-35. doi: https://doi.org/10.1080/15275920601180529.

417. Thorleifson, L., Kenneth L. Harris, H. Hobbs, C. Jennings, A. R. Knaeble, R. Lively, B. A. Lusardi, and G. Meyer. 2007. "OFR07-01, Till geochemical and indicator mineral reconnaissance of Minnesota."

471. Tobiszewski, Marek, and Jacek Namieśnik. 2012. "PAH diagnostic ratios for the identification of pollution emission sources." *Environmental Pollution* 162:110-119. doi: https://doi.org/10.1016/j.envpol.2011.10.025.

374. Townsend, Donald I. 1983. "Change of isomer ratio and fate of polychlorinated-p-dioxins in the environment." *Chemosphere* 12 (4):637-643. doi: https://doi.org/10.1016/0045-6535(83)90221-7.

437. Tukey, John W. 1977. *Exploratory Data Analysis*. Reading, PA: Addison-Wesley Publishing Company.

375. Uhler, A. D., and S. Emsbo-Mattingly. 2006. "Environmental Stability of PAH Source Indices in Pyrogenic Tars." *Bulletin of environmental contamination and toxicology* 76:689-96. doi: https://doi.org/10.1007/s00128-006-0975-1.

36. USDON. 2002. Guidance For Environmental Background Analysis, Volume I: Soil. In *NFESC User's Guide UG-2049-ENV*, edited by U.S. Department of the Navy (US DON), Naval Facilities Engineering Command. Prepared by Battelle Memorial Institute, Earth Tech, Inc., and NewFields, Inc., Washington D.C. https://www.navfac.navy.mil/content/dam/navfac/Specialty%20Centers/Engineering%20and%20Expeditionary%20Warfare%20Center/Environmental/Restoration/er_pdfs/gpr/navfacesc-ev-ug-2049-env-bkgrd-soil-200204.pdf.

37. USDON. 2003. Guidance for Environmental Background Analysis. Vol II: Sediment. In *NFESC User's Guide UG-2054-ENV*, edited by U.S. Department of the Navy (USDON), Naval Facilities Engineering Command. Prepared by Battelle Memorial Institute, Earth

Tech, Inc., and NewFields, Inc., Washington D.C. file:///C:/Users/jrocc/AppData/Local/Temp/MicrosoftEdgeDownloads/dd20f213-ddb9-4be4-ba3c-83278554e076/navfacesc-ev-ug-2054-env-bkgrd-seds-200304.pdf.

404. USEPA, 1979. EPA's Final PCB Ban Rule: Over 100 Questions & Answers To Help You Meet These Requirements, U.S. Environmental Protection Agency, Office of Toxic Substances, Washington, D.C. https://nepis.epa.gov/Exe/ZyNET.exe/910154VW.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1976+Thru+1980&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=&File=D%3A%5Czyfiles%5CIndex%20Data%5C76thru80%5CTxt%5C00000023%5C910154VW.txt&User=ANONYMOUS&Password=anonymous&SortMethod=h%7C-&MaximumDocuments=1&FuzzyDegree=0&ImageQuality=r75g8/r75g8/x150y150g16/i425&Display=hpfr&DefSeekPage=x&SearchBack=ZyActionL&Back=ZyActionS&BackDesc=Results%20page&MaximumPages=1&ZyEntry=1&SeekPage=x&ZyPURL.

130. USEPA, 1989. Risk Assessment Guidance for Superfund Volume I Human Health Evaluation Manual (Part A), U.S. Environmental Protection Agency, Office of Emergency and Remedial Response, Washington, D.C., (EPA/540/1-89/002). https://www.epa.gov/sites/production/files/2015-09/documents/rags_a.pdf.

489. USEPA, 1992. Guidance for Conducting Remedial Investigations and Feasibility Studies Under CERCLA - Interim Final, U.S. Environmental Protection Agency, Office of Emergency and Remedial Response, Washington, D.C., (PE89-184626; EPA/540/G-89/004; OSWER Directive 9355.3-01), October. https://semspub.epa.gov/work/HQ/100001529.pdf.

486. USEPA, 1992. Guidance for Data Useability in Risk Assessment (Part A) Final, U.S. Environmental Protection Agency, Office of Emergency and Remedial Response, Washington, D.C., (Publication: 9285.7-09A, PB92 - 963356), April. https://semspub.epa.gov/work/05/424356.pdf.

250. USEPA, 1992. Guidance for Performing Site Inspections under CERCLA. Interim Final. Chapter 4, U.S. Environmental Protection Agency, Office of Emergency and Remedial Response, Washington, D.C., (EPA/540-R-92-021. Directive 9345.1-05). https://nepis.epa.gov/Exe/ZyNET.exe/2000IRVN.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1991+Thru+1994&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=&File=D%3A%5Czyfiles%5CIndex%20Data%5C91thru94%5CTxt%5C00000014%5C2000IRVN.txt&User=ANONYMOUS&Password=anonymous&SortMethod=h%7C-&MaximumDocuments=1&FuzzyDegree=0&ImageQuality=r75g8/r75g8/x150y150g16/i425&Display=hpfr&DefSeekPage=x&SearchBack=ZyActionL&Back=ZyActionS&BackDesc=Results%20page&MaximumPages=1&ZyEntry=1&SeekPage=x&ZyPURL.

249. USEPA, 1992. Hazard Ranking System Guidance Manual, Interim Final. Chapter 5, U.S. Environmental Protection Agency, Office of Soild Waste and Emergency Response, Washington, D.C., (EPA 540-R-92-026, PB92-963377, Publication 9345.1-07). https://semspub.epa.gov/work/HQ/174020.pdf.

132. USEPA, 1994. Selecting and Using Reference Information in Superfund Ecological Risk Assessments, U.S. Environmental Protection Agency, Office of Solid Waste and

Emergency Response, Washington, D.C., (EPA/540/F-94/050; PB94-963319). https://www.epa.gov/sites/production/files/2015-09/documents/v2no4.pdf.

220. USEPA, 1995. Determination of Background Concentrations of Inorganics in Soils and Sediments at Hazardous Waste Sites. Engineering Forum Issue, U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C., (EPA/540/S-96/500). https://www.epa.gov/sites/production/files/2015-06/documents/determine_background_concentrations.pdf.

251. USEPA, 1995. Establishing Background Levels, OSWER Office of Emergency and Remedial Response Quick Reference Fact Sheet, U.S. Environmental Protection Agency, Office of Emergency and Remedial Response, Washington, D.C., (EPA/540/F-94/031, PB94-963314, Directive 9285.7-20FS). https://semspub.epa.gov/work/11/174014.pdf.

427. USEPA, 1996. Soil Screening Guidance: User's Guide, Second Edition, U.S. Environmental Protection Agency, Office of Emergency and Remedial Response, Washington, D.C., (EPA/540/R-96/018). https://semspub.epa.gov/work/HQ/175238.pdf.

272. USEPA, 1997. Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessment. Interim Final, Office of Solid Waste and Emergency Response, Washington, D.C., (EPA 540-R-97-006, OSWER 9285.7-25, PB97-963211). https://semspub.epa.gov/work/HQ/157941.pdf.

273. USEPA, 1998. Guidelines for Ecological Risk Assessment, U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, D.C., (EPA/630/R-95/002F). https://www.epa.gov/sites/production/files/2014-11/documents/eco_risk_assessment1998.pdf.

490. USEPA, 2000. Data Quality Objectives Process for Hazardous Waste Site Investigations EPA QA/G-4HW - Final, U.S. Environmental Protection Agency, Office of Environmental Information, Washington, D.C., (EPA/600/R-00/007), January. https://www.epa.gov/sites/default/files/2015-07/documents/g4hw-final.pdf.

442. USEPA, 2001. Eco Update - The Role of Screening-Level Risk Assessments and Refining Contaminants of Concern in Baseline Ecological Risk Assessments., U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C., (Publication 9345.0-14. EPA 540/F-01/014). https://clu-in.org/download/contaminantfocus/sediments/eco-update-slera0601.pdf.

441. USEPA, 2001. Risk Assessment Guidance for Superfund Volume I Human Health Evaluation Manual (Part D, Standardized Planning, Reporting, and Review of Superfund Risk Assessments), U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C. https://www.epa.gov/sites/production/files/2018-03/documents/175137.pdf.

153. USEPA, 2002. Guidance for Choosing a Sampling Design for Environmental Data Collection: Use in the Development of a Quality Assurance Project Plan (EPA QA/G-5S), U.S. Environmental Protection Agency, Office of Environmental Information, Washington, D.C., (EPA/240/R-02/005). https://www.epa.gov/sites/production/files/2015-06/documents/g5s-final.pdf.

154. USEPA, 2002. Guidance for Comparing Background and Chemical Concentrations in Soil for CERCLA Sites, U.S. Environmental Protection Agency, Office of Emergency and Remedial Response, Washington, D.C., (EPA 540-R-01-003; OSWER 9285.7-41). https://www.epa.gov/sites/production/files/2015-11/documents/background.pdf.

133. USEPA, 2002. Guidance for Quality Assurance Project Plans, U.S. Environmental Protection Agency, Office of Environmental Information, Washington, D.C., (EPA/240/R-02/009). https://www.epa.gov/sites/production/files/2015-06/documents/g5-final.pdf.

131. USEPA, 2002. Role of Background in the CERCLA Cleanup Program, U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response; Office of Emergency and Remedial Response, Washington, D.C., (OSWER 9285.6-07P). https://www.epa.gov/sites/production/files/2015-11/documents/bkgpol_jan01.pdf.

409. USEPA, 2003. Exposure and Human Health Reassessment Of 2,3,7,8-Tetrachlorodibenzo-P-Dioxin TCDD and related compounds, U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C., (EPA/600/P-00/001). https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=87843.

192. USEPA, 2006. Data Quality Assessment: Statistical Methods for Practitioners, U.S. Environmental Protection Agency, Office of Environmental Information, Washington, DC, (EPA/240/B-06/003). https://nepis.epa.gov/Exe/ZyPDF.cgi/900B0D00.PDF?Dockey=900B0D00.PDF.

134. USEPA, 2006. Guidance on Systematic Planning Using the Data Quality Objective Process, U.S. Environmental Protection Agency, Office of Environmental Information, Washington, D.C., (EPA/240/B-06/001). https://www.epa.gov/quality/guidance-systematic-planning-using-data-quality-objectives-process-epa-qag-4

161. USEPA, 2006. IDL- MDL- PQL; What the "L" is going on?; What does all this alphabet soup really mean?, U.S. Environmental Protection Agency, Region III Quality Assurance, Washington, D.C., (MDL Factsheet - Revision No.:2.5). https://www.epa.gov/sites/production/files/2015-06/documents/whatthel.pdf.

410. USEPA, 2006. An Inventory of Sources and Environmental Releases of Dioxin-Like Compounds in the United States for the Years 1987, 1995 and 2000, U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC, (EPA/600/P-03/002F). https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=159286.

223. USEPA, 2007. Estimation of relative bioavailability of lead in soil and soil-like materials using in vivo and in vitro methods, U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, D.C., (OSWER 9285.7-77). https://semspub.epa.gov/work/HQ/175416.pdf.

406. USEPA, 2007. Pilot Survey of Levels of Polychlorinated Dibenzo-p-dioxins, Polychlorinated Dibenzofurans, Polychlorinated Biphenyls, and Mercury in Rura lSoils of the United States, U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C., (EPA/600/R-05/048F). https://cfpub.epa.gov/si/si_public_record_Report.cfm?Lab=NCEA&dirEntryId=150944&CFID=27805128&CFTOKEN=30961662.

162. USEPA, 2007. Report of the Federal Advisory Committee on Detection and Quantitation Approaches and Uses in Clean Water Act Programs, U.S. Environmental Protection Agency, Federal Advisory Committee on Detection and Quantitation Approaches and Uses in Clean Water Act Programs, Washington, D.C., (MDL Factsheet - Revision No.:2.5). https://www.epa.gov/sites/production/files/2015-10/documents/detection-quant-faca_final-report_2012.pdf.

198. USEPA, 2009. Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities: Unified Guidance, U.S. Environmental Protection Agency, Office of Resource Conservation and Recovery, Washington, DC, (EPA/530/R-09/007). https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P10055GQ.TXT.

462. USEPA, 2011. Environmental Cleanup Best Management Practices: Effective Use of the Project Life Cycle Conceptual Site Model, U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, D.C., (EPA 542-F-11-011). https://www.epa.gov/sites/production/files/2015-04/documents/csm-life-cycle-fact-sheet-final.pdf.

224. USEPA, 2012. Recommendations for Default Value for Relative Bioavailability of Arsenic in Soil, U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, D.C., (OSWER 9200.1-113). https://www.epa.gov/superfund/soil-bioavailability-superfund-sites-guidance#arsenic.

428. USEPA, 2012. Standard Operating Procedures Soil Sampling, U.S. Environmental Protection Agency, U.S. EPA Environmental Response Team, Washington, D.C. https://nepis.epa.gov/Exe/ZyNET.exe/P100C9V4.TXT?ZyActionD=ZyDocument&Client=EPA&Index=2000+Thru+2005&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=&File=D%3A%5Czyfiles%5CIndex%20Data%5C00thru05%5CTxt%5C00000028%5CP100C9V4.txt&User=ANONYMOUS&Password=anonymous&SortMethod=h%7C-&MaximumDocuments=1&FuzzyDegree=0&ImageQuality=r75g8/r75g8/x150y150g16/i425&Display=hpfr&DefSeekPage=x&SearchBack=ZyActionL&Back=ZyActionS&BackDesc=Results%20page&MaximumPages=1&ZyEntry=1&SeekPage=x&ZyPURL.

419. USEPA, 2013. ProUCL Version 5.0 Technical Guide, Statistical Software for Environmental Applications for Data Sets with and without Nondetect Observations, U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC, (EPA/600/R-07/041). https://www.epa.gov/sites/production/files/2015-03/documents/proucl_v5.0_tech.pdf.

429. USEPA, 2014. Soil Sampling, U.S. Environmental Protection Agency, Region 4 Science and Ecosystem Support Division, Washington, D.C., (SESDPROC-300-R3). https://www.epa.gov/sites/production/files/2015-06/documents/Soil-Sampling.pdf.

195. USEPA. 2015. "EPA Contract Laboratory Program Statement of Work for inorganic Superfund Methods Multi-Media, Multi-Concentration SOM02.3." accessed 5/1/2021. https://19january2017snapshot.epa.gov/clp/epa-contract-laboratory-program-statement-work-inorganic-superfund-methods-multi-media-multi-0_.html.

197. USEPA, 2015. ProUCL Version 5.1 Technical Guide, Statistical Software for Environmental Applications for Data Sets with and without Nondetect Observations, U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC, (EPA/600/R-07/041). https://19january2017snapshot.epa.gov/sites/production/files/2016-05/documents/proucl_5.1_tech-guide.pdf.

199. USEPA, 2015. ProUCL Version 5.1 User Guide, Statistical Software for Environmental Applications for Data Sets with and without Nondetect Observations, U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC, (EPA/600/R-07/041). https://www.epa.gov/sites/production/files/2016-05/documents/proucl_5.1_user-guide.pdf.

168. USEPA, 2016. Definition and Procedure for the Determination of the Method Detection Limit, Revision 2, U.S. Environmental Protection Agency, Office of Water, Washington, D.C., (EPA 821-R-16-006). https://www.epa.gov/sites/production/files/2016-12/documents/mdl-procedure_rev2_12-13-2016.pdf.

167. USEPA, 2016. National Functional Guidelines for High Resolution Superfund Methods Data Review, U.S. Environmental Protection Agency, Office of Water, Washington, D.C., (OEM 9200.3-115; EPA 542-B-16-001). https://www.epa.gov/sites/production/files/2016-05/documents/hrsm_nfg.pdf.

165. USEPA, 2017. National Functional Guidelines for Inorganic Superfund Methods Data Review, U.S. Environmental Protection Agency, Office of Superfund Remediation and Technology Innovation (OSRTI), Washington, D.C., (OLEM 9355.0-135; EPA-540-R-201 7-001). https://www.epa.gov/sites/production/files/2017-01/documents/national_functional_guidelines_for_inorganic_superfund_methods_data_review_01302017.pdf.

164. USEPA, 2017. National Functional Guidelines for Organic Superfund Methods Data Review, U.S. Environmental Protection Agency, Office of Superfund Remediation and Technology Innovation (OSRTI), Washington, D.C., (OLEM 9355.0-136; EPA-540-R-2017-002). https://www.epa.gov/sites/production/files/2017-01/documents/national_functional_guidelines_for_organic_superfund_methods_data_review_013072017.pdf.

252. USEPA, 2018. Frequently Asked Questions About the Development and Use of Background Concentrations at Superfund Sites: Part One, General Concepts, Office of Superfund Remediation and Technology Innovation, Washington, D.C., (OLEM Directive 9200.2-141 A). https://semspub.epa.gov/work/HQ/100001657.pdf.

166. USEPA. 2020. "The SW-846 Compendium." U.S. Environmental Protection Agency, accessed October 11, 2020. https://www.epa.gov/hw-sw846/sw-846-compendium.

478. USEPA. 2021. "Documents for Recommended Toxicity Equivalency Factors for Human Health Risk Assessments of Dioxin and Dioxin-Like Compounds." U.S. Environmental Protection Agency, accessed December 28, 2020. https://www.epa.gov/risk/documents-recommended-toxicity-equivalency-factors-human-health-risk-assessments-dioxin-and#:~:text=EPA%20recommends%20the%20use%20of,cancer%20and%20non%2Dcancer%20effects.

9. van Buuren, Beverly H. 2017. *Introduction to Reporting Limits*. Moss Landing, CA: Marine Pollution Studies Laboratory at the Moss Landing Marine Laboratories. Presentation. https://acwi.gov/monitoring/webinars/mpsl_qa_services_intro_rls_012517.pdf.

396. Van den Berg, M., L. S. Birnbaum, M. Denison, M. De Vito, W. Farland, M. Feeley, H. Fiedler, H. Hakansson, A. Hanberg, L. Haws, M. Rose, S. Safe, D. Schrenk, C. Tohyama, A. Tritscher, J. Tuomisto, M. Tysklind, N. Walker, and R. E. Peterson. 2006. "The 2005 World Health Organization reevaluation of human and Mammalian toxic equivalency factors for dioxins and dioxin-like compounds." *Toxicol Sci* 93 (2):223-41. doi: https://doi.org/10.1093/toxsci/kfl055.

430. VSP Development Team, 2020. Visual Sample Plan: A Tool for Design and Analysis of Environmental Sampling; Version 7.13, Pacific Northwest National Laboratory. https://www.pnnl.gov/projects/visual-sample-plan.

384. Wang, Zhen, Jingwen Chen, Fulin Tian, Ping Yang, Xianliang Qiao, and Ziwei Yao. 2010. "Application of Factor Analysis with Nonnegative Constraints for Source Apportionment of Soil Polycyclic Aromatic Hydrocarbons (PAHs) in Liaoning, China." *Environmental Forensics* 11 (1-2):161-167. doi: https://doi.org/10.1080/15275920903558968.

385. Wang, Zhendi, C. Yang, F. Kelly-Hooper, B. P. Hollebone, X. Peng, C. E. Brown, M. Landriault, J. Sun, and Z. Yang. 2009. "Forensic differentiation of biogenic organic compounds from petroleum hydrocarbons in biogenic and petrogenic compounds cross-

contaminated soils and sediments." *Journal of Chromatography A* 1216 (7):1174-1191. doi: https://doi.org/10.1016/j.chroma.2008.12.036.

386. Wang, Zhendi, C. Yang, Z. Yang, B. Hollebone, C. E. Brown, M. Landriault, J. Sun, S. M. Mudge, F. Kelly-Hooper, and D. G. Dixon. 2012. "Fingerprinting of petroleum hydrocarbons (PHC) and other biogenic organic compounds (BOC) in oil-contaminated and background soil samples." *Journal of Environmental Monitoring* 14 (9):2367-2381. doi: https://doi.org/10.1039/C2EM30339F.

389. Wei, Lifei, Yangxi Zhang, Ziran Yuan, Zhengxiang Wang, Feng Yin, and Liqin Cao. 2020. "Development of Visible/Near-Infrared Hyperspectral Imaging for the Prediction of Total Arsenic Concentration in Soil." *Applied Sciences* 10:2941. doi: https://doi.org/10.3390/app10082941.

270. Whitney, Philip R. 1975. "Relationship of manganese-iron oxides and associated heavy metals to grain size in stream sediments." *Journal of Geochemical Exploration* 4 (2):251-263. doi: https://doi.org/10.1016/0375-6742(75)90005-9.

158. Winegardner, Duane. 2019. *An Introduction to Soils for Environmental Professionals*. doi:https://doi.org/10.1201/9781315137322.

483. Wischkaemper, H.K., A.F. Beliveau, and R.W R.W. Henderson, 2013. U.S. EPA Region 4 technical services section issue paper for polychlorinated biphenyl characterization at Region 4 Superfund sites, U.S. Environmental Protection Agency, U.S. EPA Region 4/Technical Services Section, Athens, GA, (EPA 600/R-11/005A). https://www.epa.gov/sites/production/files/2015-09/documents/r4_issue_paper_for_pcbs_5-15-2013.pdf.

477. Wolfe, Douglas A. 2010. "Ranked set sampling." *WIREs Computational Statistics* 2 (4):460-466. doi: https://doi.org/10.1002/wics.92.

271. Xue, Wen, Sae Yun Kwon, Stephan E. Grasby, Elsie M. Sunderland, Xin Pan, Ruiyang Sun, Ting Zhou, Haiyu Yan, and Runsheng Yin. 2019. "Anthropogenic influences on mercury in Chinese soil and sediment revealed by relationships with total organic carbon." *Environmental Pollution* 255:113186. doi: https://doi.org/10.1016/j.envpol.2019.113186.

39. Zoffoli, Hugo José Oliveira, Carlos Alberto Alves Varella, Nelson Moura Brasil do Amaral-Sobrinho, Everaldo Zonta, and Alfredo Tolón-Becerra. 2013. "Method of median semi-variance for the analysis of left-censored data: Comparison with other techniques using environmental data." *Chemosphere* 93 (9):1701-1709. doi: https://doi.org/10.1016/j.chemosphere.2013.05.041.

402. Zou, Yonghong, Lixia Wang, and Erik R. Christensen. 2015. "Problems in the fingerprints based polycyclic aromatic hydrocarbons source apportionment analysis and a practical solution." *Environmental Pollution* 205:394-402. doi: https://doi.org/10.1016/j.envpol.2015.05.029.

403. Zuo, Q., Y. H. Duan, Y. Yang, X. J. Wang, and S. Tao. 2007. "Source apportionment of polycyclic aromatic hydrocarbons in surface soil in Tianjin, China." *Environmental Pollution* 147 (2):303-310. doi: https://doi.org/10.1016/j.envpol.2006.05.029.