Statistical Analysis of the Environmental Geochemistry of an Unmined Uranium Ore Deposit

Denise Madeline Levitan

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Geosciences

Madeline E. Schreiber, Co-Chair
Robert J. Bodnar, Co-Chair
Patricia M. Dove
J. Donald Rimstidt
Robert R. Seal II

July 18, 2014
Blacksburg, VA

Keywords: Coles Hill, uranium, water, soils, statistics, geochemistry, data analysis

Statistical Analysis of the Environmental Geochemistry of an Unmined Uranium Ore Deposit

Denise Madeline Levitan

ABSTRACT

An evaluation of the geochemistry of the environment prior to large-scale changes enables scientists and other stakeholders to assess both baseline conditions and the potential impact of those changes to the environment. One area in which documentation of pre-development geochemistry is particularly important is in the exploitation of ore deposits. Ore deposits consist of concentrations of elements or minerals that are enriched enough to be of potential economic value. Their unusual geochemistry often leaves a signature on the environment that can both aid in location an economic resource and present environmental management challenges during its lifecycle. Coles Hill, Virginia, represents one such site. The Coles Hill property is the location of uranium-enriched rock, commonly referred to as the Coles Hill uranium deposit.

This dissertation outlines study design, sampling, and statistical analysis methods that can be used in the geochemical characterization of a potential resource extraction site. It presents three studies on geoenvironmental media at Coles Hill. The first study discusses sampling strategies and statistical analysis to address variability in geology, hydrology and climate for baseline assessment and presents an example of such an assessment at Coles Hill. Results suggest a localized environmental impact of the deposit but that differences in bedrock geology within the area surrounding the deposit could also be responsible for some of the variation. This study also emphasizes the importance of consideration of data below analytical detection limits and describes methods for doing so. The second study compares the geochemistry of soil samples collected at Coles Hill with reference data collected by the U.S. Geological Survey using multivariate statistical techniques. Differences are used to suggest potential pathfinder elements such as light rare earth elements to aid in exploration for similar deposits. The third study uses multivariate statistical analysis to examine differences among rocks, soils, and stream sediments to infer important geochemical processes involved in weathering of the deposit. Overall, the results of these studies can aid in the development of future environmental site studies at Coles Hill and elsewhere.

ACKNOWLEDGMENTS

I would like to start by thanking my advisor, Maddy Schreiber, who has done more for me than I could possibly write, especially since I know she'd rather I turn in this dissertation than sing her praises. I would also like to thank my co-advisor, Bob Bodnar, without whom this project would have been impossible. I am grateful for the advice and suggestions of my other committee members, Trish Dove, Don Rimstidt, and Bob Seal.

I would also like to thank my collaborators and former coworkers at the U.S. Geological Survey for encouraging me to come to graduate school and for helping me once I got here. Bob Seal, Nadine Piatak, Jane Hammarstrom, John Jackson, Larry Gough, and Tiffani Westphal assisted in the field, and Natalia Ainsfield did lab work. Mark Engle was extremely helpful in teaching me compositional data analysis. I also appreciate encouragement and suggestions from Robert Ayuso, Nora Foley, Bryn Kimball, Helen Folger, I-Ming Chou, and Larry Drew.

One of the best parts of working at a site like Coles Hill was the diverse group of researchers I got to interact with. I am particularly grateful to John Chermak from Geosciences; Carl Zipper, Patricia Donovan, and Dan Johnson from CSES; Panos Diplas, Billy Kingston, Celso Castro, and Edgardo Zavaleta from CEE; and the many other researchers from Virginia Tech and elsewhere doing concurrent work at Coles Hill. I also appreciate the work of former VT Geosciences graduate students, including Jim Jerden, JP Gannon, and John Wyatt.

Many other students, faculty, and staff from Virginia Tech contributed ideas, field assistance, and moral support during this project. I would like to particularly acknowledge adjuncts Bill Henika and Jim Beard, members of the Hydrogeology and Fluids Research Groups, and the department technical and administrative staff. Field assistants included Carol Johnson, Jeanne Roningen, Rodrigo Prugue, Meijing Zhang, Randy Cosby, Daniel Moncada, Matt MacInnis, Pilar Lecumberri, Liz Krukowski, Yinka Oyewumi, Lindsay Kolbus, James Dale, Sarah Mazza, Kristie Dorfler, and Lindsay Bartz. Anna Hardy and Heather Scott assisted with lab work and data entry, and Zack Munger was always willing to let me use the pH meter first. Due to the diligent efforts of Luke Joyce, Sarah Eagle, and Héctor Lamadrid, I'm still not sure how to make coffee on my own.

William Levitan forwarded me every non-confidential email he received about radionuclide transport and remediation, and Jackie Lewis forwarded me every cheesy science and math joke she heard. These are only the most recent of their innumerable contributions to my education and well-being over the past thirty years. I would also be remiss in not acknowledging booboohead.

Last but not least, I want to thank my many funding sources. This project would not have been possible without funding, support, and access from Virginia Uranium, Inc. Walter Coles Sr. graciously let me tromp through his fields, and Stewart East and Joe Aylor kept me from getting lost in them. Nina Beth Thornton, Frances Haley, and Lorien Koontz Huemoller helped keep things organized, as did VUI's many interns and other staff.

Other funding sources to which I am extremely grateful include the U.S. Geological Survey, a Virginia Water Resources Research Center Student Research Grant, a BP Graduate Scholarship through VT Geosciences, a Geological Society of America Graduate Student Research Grant, the Jay M. McMurray Memorial Grant from the American Association of Petroleum Geologists, the Sand Student Research Presentation Travel Award from the Association for Women Geoscientists, and the Robinson-Holden Summer Fellowship and T.T. Jeffries Geosciences Graduate Research Award from VT Geosciences.

TABLE OF CONTENTS

Abstract ii
Acknowledgments iii
Table of Contents iv
Table of Figures vii
Table of Tables ix
Chapter 1. Introduction 1
References 7
Chapter 2. Developing protocols for geochemical baseline studies 10
Abstract 10
2.1 Introduction 11
2.1.1 The problem of non-detect data 13
2.1.2 The Coles Hill uranium deposit 16
2.2 Methods 19
2.2.1 Sampling strategy 19
2.2.1.1 Spatial site selection 19
2.2.1.2 Temporal sampling 21
2.2.2 Sample collection and analysis 22
2.2.2.1 Stream sediment 22
2.2.2.2 Surface water 22
2.2.3 Geochemical dataset 23
2.2.4 Statistical analysis 24
2.2.4.1 Evaluation of non-detects 24
2.3 Results 28
2.3.1 Variations among sites 28
2.3.1.1 Stream sediments 28
2.3.1.2 Stream water 31
2.3.2 Temporal Variations 33
2.3.3 Comparison of non-detect methods 35
2.3.4 Baseline values 36
2.4 Discussion 40
2.4.1 Sampling protocol and accounting for variability 40
2.4.2 Effect of including non-detect methods in statistical analysis 43

2.4.3 Geochemical baseline for elements of interest in surface water and stream sediment at the Coles Hill deposit 45
2.5 Conclusions 46
Acknowledgments 48
References 48
Chapter 3. Statistical analysis of soil geochemical data to identify pathfinders associated with mineral deposits 53
Abstract 53
3.1 Introduction 54
3.1.1 The Coles Hill uranium deposit 56
3.2 Methods 61
3.2.1 Site selection 61
3.2.2 Sample collection and analysis 62
3.2.3 Selection of reference soil data 63
3.2.4 Statistical analysis 65
3.2.4.1 Multivariate statistical methods 68
3.3 Results and discussion 70
3.3.1 Soil mineralogy at Coles Hill 72
3.3.2 Enrichment factors 72
3.3.3 Correlation and variation between elements 74
3.3.4 Multivariate analysis 78
3.3.5 Geochemical trends 83
3.3.6 Pathfinder relationships 85
3.3.7 Radioisotopes 88
3.4 Conclusions 89
Acknowledgments 91
References 91
Chapter 4. Multivariate exploratory analysis to identify processes controlling the geochemistry of a rock-soil-sediment system at an undeveloped ore deposit 95
Abstract 95
4.1 Introduction 96
4.2 Methods 99
4.2.1 Data sources 99
4.2.2 Data processing 100
4.2.3 Statistical analysis 101

4.3 Results and discussion 103
4.3.1 Are the groups geochemically different? 103
4.3.2 What elements most strongly differentiate AMONG groups? 104
4.3.3 How do these group differences relate to geochemical processes occurring at the site? 108
4.3.4 The role of water 115
4.4 Implications for future research 117
4.5 Conclusions 118
References 120
Appendix A. Supplementary information for Chapter 3 123
Appendix B. Kinetics experiments: Apatite dissolution and autunite precipitation 129
B.1 Introduction 129
B.1.1 Uranium immobilization by phosphate 129
B.1.2 Apatite dissolution 130
B.2 Experimental methods 132
B.2.1 Apatite preparation 132
B.2.2 Solutions 133
B.2.3 Reactor setup 133
B.2.4 Solution analysis 134
B.2.5 Calculations 134
B.3 Results 135
References 139
Appendix C. Data Analysis 141
C.1 Compositional data 141
C.1.1 Log-ratio transformations 141
C.2 Common multivariate methods and figures 143
C.2.1 Hierarchical cluster analysis (HCA) 144
C.2.2 Principal component analysis (PCA) 145
C.2.3 Linear discriminant analysis (LDA) 147
References 148

TABLE OF FIGURES

Figure 1-1: Surface water draining a historic sulfide mine site 1
Figure 1-2: Aerial view of the Coles Hill area 4
Figure 2-1: Geologic map of the Coles Hill region 16
Figure 2-2: Stream sampling sites for the Coles Hill baseline study 18
Figure 2-3: Cumulative distribution functions showing different approaches to censored data .. 27
Figure 2-4: Maps showing concentrations of selected elements in stream sediments in March 2011 30
Figure 2-5: Maps showing dissolved (< 0.45 μm) concentrations in μg/L of selected elements in stream waters in March 2011 32
Figure 2-6: Time-series plots of filtered concentrations of selected elements and discharge at Site 23 throughout the course of this study 34
Figure 2-7: Logarithmic plot of discharge vs. Ba and Ca concentrations at all sites sampled monthly during the course of this study. 34
Figure 2-8: Boxplots showing statistics for uranium stream sediment artificially censored to 20%, 40%, 60%, and 80% of values below detection by three different methods 36
Figure 2-9: Boxplots showing concentrations of geochemical parameters in (a) stream sediment (n=40) and (b) stream water (n=98) during this study 39
Figure 3-1: Geologic map of the Coles Hill region modified from Henika (2002) 57
Figure 3-2: Soil sampling sites 60
Figure 3-3: Selection of reference soil sites from soil sampling sites for USGS national soil survey 65
Figure 3-4: Sample clr PCA biplot 70
Figure 3-5: Tukey boxplots of median enrichment factors [$(X/\mathrm{Al})_{CH}$/median$(X/\mathrm{Al})_{ref}$] in the a) A horizon and b) C horizon 73
Figure 3-6: Dendrograms from hierarchical variable clustering for A and C horizons at Coles Hill and in reference soils 79
Figure 3-7: PCA biplots showing all a) A and b) C horizon samples 82
Figure 3-8: Tukey boxplots showing suggested pathfinder ratios in soils for deposits similar to Coles Hill 86
Figure 3-9: a) Log-log plot of Th vs. U concentrations (mg/kg) and b) Plot of isometric log-ratio balances for K, U, and Th. 88
Figure 3-10: Radionuclide concentrations in the A and C horizons across transects over the North and South ore bodies 89
Figure 4-1: Schematic cross-section of the geoenvironmental media at Coles Hill 100
Figure 4-2: Scatterplot matrix of linear discriminant scores 104
Figure 4-3: Ranked plot of Kruskal-Wallis test statistics by element 105

Figure 4-4: Plot of absolute values of scaled linear discriminant scalings by element 106
Figure 4-5: Element plot of scaled group centers 107
Figure 4-6: Principal component analysis biplot showing scaled results for the first two PCs.. 109
Figure 4-7: Principal component analysis screeplot 110
Figure 4-8: Scatterplot matrix of ilr balances for solid samples 114
Figure 4-9: Scatterplot matrix of isometric log-ratio (ilr) balances for solid and aqueous samples 117
Figure B-1: Published apatite dissolution rates as a function of pH. 132
Figure B-2: Experimental setup. 134
Figure B-3: Rates of Ca release and apatite dissolution as a function of pH 136
Figure B-4: Plot of uranium concentration (U) relative to average feed uranium concentration (U_O) over the course of experiments 138
Figure B-5: Calculated rates (r) of uranium removal and apatite dissolution over the course of experiments 138
Figure C-1: Pie charts showing potential hours per day spent on different tasks. 141
Figure C-2: Sample balance dendrogram showing sequential binary partitioning of four elements 143
Figure C-3: HCA dendrogram showing hexadecimal color codes displayed in their corresponding colors. 144
Figure C-4: PCA biplot of time spent per day on various activities. 146
Figure C-5: Example screeplot of ten PCs by variance (eigenvalue) 147

TABLE OF TABLES

Table 2-1: Non-detect statistical methods 26

Table 2-2: Table of p-values from Wilcoxon signed-rank tests comparing data collected from different bedrock provinces (Piedmont and Triassic basin), from upstream (U/S) and reference sites vs. downstream (D/S) sites, and from Triassic basin reference sites vs. sites downstream of the deposit 29

Table 2-3: Kendall's tau values for selected elements (n=98) in stream water vs. temperature (T) and flow (Q) 33

Table 2-4: Summary statistics for selected elements in stream sediments (n=40) in this study ... 37

Table 2-5: Summary statistics for selected elements in stream waters from monthly sampling (n=98) in this study 38

Table 3-1: Summary statistics for untransformed data 71

Table 3-2: Correlation coefficients (Pearson's r) between U alr values [ln(U/Al)] and other constituents' alr values [ln(*X*/Al)] in A and C horizons of Coles Hill (CH) and reference (ref) soils. 75

Table 3-3: U log-ratio variance values from variation matrices for Coles Hill (CH) and reference (ref) samples, A and C horizons 77

Table 4-1: Sequential binary partition matrix for ilr balances 114

Table A-1: *p*-values for enrichment factor (EF) median comparisons between numerator (*X*/Al at Coles Hill) and denominator (*X*/Al in reference soils) by the Mann-Whitney-Wilcoxon test 123

Table A-2: log-ratio variance matrices for all elements 123

Table A-3: Statistics for principal component analysis 127

Table A-4: Eigenvectors for the first two principal components (see Figure 3-7) for the A and C horizons 128

Table B-1: Calculated rates (r) from apatite dissolution experiments 136

Table B-2: ICP-OES results for lower U experiment. 137

Table B-3: ICP-OES results for higher U experiment 137

Table C-1: Table showing sequential binary partitioning of four elements as illustrated in Figure C-2. 143

CHAPTER 1. INTRODUCTION

The extraction of mineral resources has a long history of environmental impacts. Even Georgius Agricola, in *De Re Metallica*, his 16th-century seminal text on mining, notes these effects. Examples can be seen all over the world, particularly at historic sulfide and coal mines; well-known sites include the Iberian Pyrite Belt in Spain (Espana et al., 2005) and Superfund sites such as Butte, MT (Davis and Ashenberg, 1989), and Iron Mountain, CA (Nordstrom and Alpers, 1999), in the United States (Figure 1-1). Common geochemical impacts include acid mine/rock drainage and the mobilization of metals and other compounds (Langmuir, 1997; Lottermoser, 2010).

Figure 1-1: Surface water draining a historic sulfide mine site. This water is acidic and concentrated in many metals associated with the deposit, which may result in adverse impacts to downstream ecosystems.

An immense amount of research has been done in documenting and characterizing the geochemistry and geochemical processes, such as mineral dissolution, formation of secondary minerals, oxidation-reduction reactions, and element mobility, among others, occurring at surface, near-surface, and exposed mineral deposits. From a practical standpoint, much of this

work is driven by concerns about impacts on nearby ecological and human health, which subsequently drive the implementation of remediation or mitigation to minimize these impacts (Jambor et al., 2003; Lottermoser, 2010; Plumlee et al., 1999).

Mineral deposits represent geochemical anomalies in that they are regions where a particular mineral or element is present in concentrations greater than found elsewhere. These elevated concentrations can alter or drive surrounding environmental processes regardless of whether or not extraction has occurred (Kelepertzis et al., 2012; Levitan et al., 2014; Nordstrom, 2008); detection of unusual geochemistry in water, soil, sediment, and vegetation has long been used as a tool in mineral exploration (Bradshaw and Lett, 1980; Coker, 2010; Leybourne and Cameron, 2010).

Because mineral extraction has such an expansive, millennia-spanning history, few deposits exist that have not seen anthropogenic effects, including mining, construction, agriculture, etc. This is particularly true of near-surface deposits, which generally have the greatest potential for interaction with surface geological media such as water, soil, and alluvium. As a result, it can be difficult to determine what environmental anomalies result from extraction operations as opposed to natural phenomena or earlier human activity.

A number of methods have been proposed to determine "background" or "baseline" conditions, which are a description of the state of a site's environment prior to one or more anthropogenic activities. For example, Runnells et al. (1992) proposed three methods for retrospectively estimating pre-mining conditions: examination of historical records; comparison to concentrations in unmined analogues; and geochemical modeling. Alpers and Nordstrom (2000) expanded this list to also include sediment sampling, stable isotopes, statistical analyses, and mass balance calculations. These retrospective methods are associated with high uncertainty.

However, in the case of a deposit that has not been developed and is not located in the vicinity of other developed sites, the straightforward approach of direct measurement of baseline conditions is possible and, in most cases, legally required. The importance of documenting baseline is frequently seen in the modern discourse on socioscientific issues such as discerning the patterns of climate change (Stocker et al., 2013) or understanding the influence of hydraulic fracturing on aquifers (Osborn et al., 2011).

In addition to addressing legal issues, liability issues, or both, a baseline survey presents the opportunity to study the hydrogeochemical processes occurring at a particular site and, in the case of anomalous geology and geochemistry such as at an ore deposit, to contrast these processes with those occurring elsewhere. Comparisons not only between the site and elsewhere, but among different hydrogeochemical media at the site itself, allows for the interpretation of the relationship and interaction between these reservoirs.

Traditional site characterization has generally involved measurement and reporting of geochemical values, including concentrations of specific elements (particularly those that are regulated or have the potential to cause or impact ecotoxicity) of interest in different media (e.g., rock, soil, water, etc.). However, as discussed in Chapter 2, there is no clear consensus or widely accepted best practices on the methodology for doing so. Once data are collected, their usage varies. In some cases, the data are archived (Anderson et al., 2011; Smith et al., 2005) for potential future use (Drew et al., 2010; Eppinger et al., 2012). The uses can include mineral exploration (Eppinger et al., 2013), setting environmental benchmarks (Kelepertzis et al., 2012), modeling element mobility (Nordstrom, 2011), identifying areas impacted by a variety of natural and anthropogenic processes, such as sea salt deposition, ore deposits, smelting, and urbanization (Reimann et al., 2008; Zhang et al., 2007), and others. The methods used for those purposes

vary, but many involved the use of geochemical modeling, statistical assessment, or a combination. This dissertation presents a systematic approach to the collection and analysis of geochemical baseline data. Analysis focuses on study design and statistical methods, with suggestions for further investigation by geochemical modeling.

This dissertation focuses on geochemical data from the Coles Hill uranium property in Pittsylvania County, VA, USA (Figure 1-2). At Coles Hill, a significant quantity of uranium-mineralized rock is present. Specific geological details of this site are contained within subsequent chapters of this dissertation and can also be found in other works from Virginia Tech (Gannon, 2009; Gannon et al., 2011; Jerden, 2001; Jerden and Sinha, 2003, 2006; Jerden et al., 2003; Tappa et al., 2014; Whitney, 2009; Wyatt, 2009). Though the presence of this deposit has been known since the late 1970s, it has not been developed for extraction or processing of the mineralized rock, and its current status presents an opportunity to study a near-surface mineralized body that has not been impacted by mining.

Figure 1-2: Aerial view of the Coles Hill area. Photo courtesy of Virginia Uranium, Inc.

The remainder of this dissertation is divided into three chapters. Chapter 2 discusses the design and purpose of baseline studies and how to report data from these studies, focusing particularly on data below analytical detection limits. It also summarizes the geochemistry of stream water and sediment in the vicinity of Coles Hill. A version of this manuscript was published as *Levitan, D.M., Schreiber, M.E., Seal, R.R., Bodnar, R.J., Aylor, J.G., 2014. Developing protocols for geochemical baseline studies: An example from the Coles Hill uranium deposit, Virginia, USA. Applied Geochemistry 43, 88-100* (Levitan et al., 2014).

Chapter 3 uses multivariate statistical methods to determine geochemical patterns within soils overlying the mineralization at Coles Hill. These patterns are evaluated using mineralogical analyses from previous studies (Jerden, 2001). Site soils are also compared to a regional publicly available soil geochemistry database (Smith et al., 2013) to identify differences between the site and the surrounding region. A version of this manuscript was submitted as *Levitan, D.M., Zipper, C.E., Donovan, P., Schreiber, M.E., Seal, R.R. II, Engle, M.A., Chermak, J.A., Bodnar, R.J., Johnson, D.K., and Aylor, J.G. Jr., Statistical analysis of soil geochemical data to identify pathfinders associated with mineral deposits: An example from the Coles Hill uranium deposit, Virginia, USA* to the Journal of Geochemical Exploration.

Chapter 4 uses statistical classification methods to determine differences between geoenvironmental media at the Coles Hill site. Statistical data patterns were used to infer the geochemical processes responsible for these differences. Based on these initial results, related future studies are recommended.

The data set compiled during this study include radionuclide data that are the property of Virginia Uranium, Inc., analyses of core sample that are property of Jim Beard and the Virginia

Museum of Natural History, and other geochemical data that will be published in a forthcoming U.S. Geological Survey Digital Data Series.

REFERENCES

Alpers, C., Nordstrom, D., 2000. Estimation of pre-mining conditions for trace metal mobility in mineralized areas: An overview, ICARD 2000, Denver, CO, pp. 463-472.

Anderson, E.D., Smith, S.M., Giles, S.A., Granitto, M., Eppinger, R.G., Bedrosian, P.A., Shah, A.K., Kelley, K.D., Fey, D.L., Minsley, B.J., Brown, P.J., 2011. Geophysical, Geochemical, and Mineralogical Data from the Pebble Cu-Au-Mo Porphyry Deposit Area, Southwest Alaska: Contributions to Assessment Techniques for Concealed Mineral Resources, Data Series 608.

Bradshaw, P.M.D., Lett, R.E.W., 1980. Geochemical exploration for uranium using soils. Journal of Geochemical Exploration 13, 305-319.

Coker, W.B., 2010. Future research directions in exploration geochemistry. Geochemistry: Exploration, Environment, Analysis 10, 75-80.

Davis, A., Ashenberg, D., 1989. The aqueous geochemistry of the Berkeley Pit, Butte, Montana, U.S.A. Applied Geochemistry 4, 23-36.

Drew, L.J., Grunsky, E.C., Sutphin, D.M., Woodruff, L.G., 2010. Multivariate analysis of the geochemistry and mineralogy of soils along two continental-scale transects in North America. Science of the Total Environment 409, 218-227.

Eppinger, R.G., Fey, D.L., Giles, S.A., Grunsky, E.C., Kelley, K.D., Minsley, B.J., Munk, L., Smith, S.M., 2013. Summary of Exploration Geochemical and Mineralogical Studies at the Giant Pebble Porphyry Cu-Au-Mo Deposit, Alaska: Implications for Exploration Under Cover. Economic Geology 108, 495-527.

Eppinger, R.G., Fey, D.L., Giles, S.A., Kelley, K.D., Smith, S.M., 2012. An exploration hydrogeochemical study at the giant Pebble porphyry Cu-Au-Mo deposit, Alaska, USA, using high resolution ICP-MS. Geochemistry: Exploration, Environment, Analysis 12, 211-226.

Espana, J.S., Pamo, E.L., Santofimia, E., Aduvire, O., Reyes, J., Barettino, D., 2005. Acid mine drainage in the Iberian Pyrite Belt (Odiel river watershed, Huelva, SW Spain): Geochemistry, mineralogy and environmental implications. Applied Geochemistry 20, 1320-1356.

Gannon, J.P., 2009. Evaluation of fracture flow at the Coles Hill uranium deposit in Pittsylvania County, VA using electrical resistivity, bore hole logging, pumping tests, and age dating methods,M.S. Thesis.

Gannon, J.P., Burbey, T.J., Bodnar, R.J., Aylor, J.G., 2011. Geophysical and geochemical characterization of the groundwater system and the role of Chatham Fault in groundwater movement at the Coles Hill uranium deposit, Virginia, USA. Hydrogeology Journal 20, 1-16.

Jambor, J.L., Blowes, D.W., Ritchie, A.I.M., 2003. Environmental aspects of mine wastes, Short Course Series. Mineralogical Association of Canada, Ottawa, ON.

Jerden, J.L., 2001. Origin of uranium mineralization at Coles Hill Virginia (USA) and its natural attenuation within an oxidizing rock-soil-ground water system,PhD Thesis.

Jerden, J.L., Sinha, A.K., 2003. Phosphate based immobilization of uranium in an oxidizing bedrock aquifer. Applied Geochemistry 18, 823-843.

Jerden, J.L., Sinha, A.K., 2006. Geochemical coupling of uranium and phosphorous in soils overlying an unmined uranium deposit: Coles Hill, Virginia. Journal of Geochemical Exploration 91, 56-70.

Jerden, J.L., Sinha, A.K., Zelazny, L., 2003. Natural immobilization of uranium by phosphate mineralization in an oxidizing saprolite-soil profile: chemical weathering of the Coles Hill uranium deposit, Virginia. Chemical Geology 199, 129-157.
Kelepertzis, E., Argyraki, A., Daftsis, E., 2012. Geochemical signature of surface water and stream sediments of a mineralized drainage basin at NE Chalkidiki, Greece: A pre-mining survey. Journal of Geochemical Exploration 114, 70-81.
Langmuir, D., 1997. Aqueous environmental geochemistry. Prentice Hall, Upper Saddle River, N.J.
Levitan, D.M., Schreiber, M.E., Seal II, R.R., Bodnar, R.J., Aylor Jr, J.G., 2014. Developing protocols for geochemical baseline studies: An example from the Coles Hill uranium deposit, Virginia, USA. Applied Geochemistry 43, 88-100.
Leybourne, M.I., Cameron, E.M., 2010. Groundwater in geochemical exploration. Geochemistry: Exploration, Environment, Analysis 10, 99-118.
Lottermoser, B., 2010. Mine wastes: characterization, treatment and environmental impacts. Springer.
Nordstrom, D., 2008. What was the groundwater quality before mining in a mineralized region? Lessons from the Questa project. Geosciences Journal 12, 139-149.
Nordstrom, D.K., 2011. Hydrogeochemical processes governing the origin, transport and fate of major and trace elements from mine wastes and mineralized rock to surface waters. Applied Geochemistry 26, 1777-1791.
Nordstrom, D.K., Alpers, C.N., 1999. Negative pH, efflorescent mineralogy, and consequences for environmental restoration at the Iron Mountain Superfund site, California. Proceedings of the National Academy of Sciences 96, 3455-3462.
Osborn, S.G., Vengosh, A., Warner, N.R., Jackson, R.B., 2011. Methane contamination of drinking water accompanying gas-well drilling and hydraulic fracturing. Proceedings of the National Academy of Sciences 108, 8172-8176.
Plumlee, G.S., Logsdon, M.J., Filipek, L.H., 1999. The environmental geochemistry of mineral deposits, Reviews in Economic Geology. Society of Economic Geologists, Littleton, CO.
Reimann, C., Filzmoser, P., Garrett, R.G., Dutter, R., 2008. Statistical data analysis explained : Applied environmental statistics with R. John Wiley & Sons, Chichester, England; Hoboken, NJ.
Runnells, D.D., Shepherd, T.A., Angino, E.E., 1992. Metals in water. Determining natural background concentrations in mineralized areas. Environmental Science & Technology 26, 2316-2323.
Smith, D., Cannon, W., Woodruff, L., Garrett, R., Klassen, R., Kilburn, J., Horton, J., 2005. Major- and trace-element concentrations in soils from two continental-scale transects of the United States and Canada, Open-File Report 2005-1253.
Smith, D.B., Cannon, W.F., Woodruff, L.G., Solano, F., Kilburn, J.E., Fey, D.L., 2013. Geochemical and mineralogical data for soils of the conterminous United States, Data Series 801.
Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M., 2013. Climate change 2013: The physical science basis. Intergovernmental Panel on Climate Change.
Tappa, M.J., Ayuso, R.A., Bodnar, R.J., Aylor, J.G., Beard, J., Henika, W.S., Vazquez, J.A., Wooden, J.L., 2014. Age of host rocks at the Coles Hill uranium deposit, Pittsylvania

County, Virginia, based on zircon U-Pb geochronology. Economic Geology 109, 513-530.
Whitney, J.A., 2009. Application of electromagnetic methods to identify and characterize sub-surface structures associated with the Coles Hill Uranium Deposit,Mining and Minerals Engineering. M.S. Thesis.
Wyatt, J., 2009. The relationship between structural and tectonic evolution and mineralization at the Coles Hill Uranium Deposit, Pittsylvania County, Virginia,Department of Geosciences. M.S. Thesis.
Zhang, H., Luo, Y., Wong, M., Zhao, Q., Zhang, G., 2007. Defining the geochemical baseline: a case of Hong Kong soils. Environmental Geology 52, 843-851.

CHAPTER 2. DEVELOPING PROTOCOLS FOR GEOCHEMICAL BASELINE STUDIES

Denise M. Levitan, Madeline E. Schreiber, Robert R. Seal II, Robert J. Bodnar, and Joseph G. Aylor Jr.
A version of this chapter was published in Applied Geochemistry 43 (2014) 88-100.

ABSTRACT

In this study, we determined baseline geochemical conditions in stream sediments and surface waters surrounding an undeveloped uranium deposit. Emphasis was placed on study design, including site selection to encompass geological variability and temporal sampling to encompass hydrological and climatic variability, in addition to statistical methods for baseline data analysis. The concentrations of most elements in stream sediments were above analytical detection limits, making them amenable to standard statistical analysis. In contrast, some trace elements in surface water had concentrations that were below the respective detection limits, making statistical analysis more challenging. We describe and compare statistical methods appropriate for concentrations that are below detection limits (non-detect data) and conclude that regression on order statistics provided the most rigorous analysis of our results, particularly for trace elements. Elevated concentrations of U and deposit-associated elements (e.g. Ba, Pb, and V) were observed in stream sediments and surface waters downstream of the deposit, but concentrations were below regulatory guidelines for the protection of aquatic ecosystems and for drinking water. Analysis of temporal trends indicated that concentrations of major and trace elements were most strongly related to stream discharge. These findings highlight the need for sampling protocols that will identify and evaluate the temporal and spatial variations in a thorough baseline study.

2.1 INTRODUCTION

An early step in the development of a mining project is environmental baseline characterization to document the pre-existing conditions at and surrounding the planned operations. Because an ore body, by its nature, contains elements that occur at anomalous concentrations relative to non-mineralized rocks in the adjacent area, such studies provide an opportunity to characterize the geochemical distribution of those elements prior to development and, in some cases, their mobility under pre-disturbance conditions. From a regulatory perspective, baseline studies provide an environmental record prior to development that can be used to locate monitoring sites, to set benchmarks for use during mining, and to establish environmental goals for closure. For example, if the concentration of an element in the soil or water exceeds a regulatory limit prior to site disturbance, it may be necessary to establish a regulatory variance for that element during active mining and later reclamation. Baseline data might also be used to develop generalized geochemical process models for the site. For example, base-flow stream water chemistry could be used along with a geochemical model to describe the chemical weathering reactions prior to site disturbance so that it is possible to recognize how these reactions may be affected by site development. The determination of pre-existing environmental conditions is an important tool in designing mitigation and preventing costly, Sisyphean remediation projects.

The term "baseline" does not have a universally accepted definition (Gałuszka, 2007; Lee and Helsel, 2005; Reimann and Garrett, 2005; Rodrigues and Nalini Júnior, 2009; Salminen and Gregorauskiene, 2000). In some works, "baseline" is used interchangeably with "background," while in others one or both terms are used to describe only "natural" conditions, meaning those that are a result only of local geology, climate, and hydrology, and not the result of any

anthropogenic activity. Salminen and Gregorauskiene (2000) and Reimann and Garrett (2005) presented thorough descriptions of the many definitions of background and baseline.

Here, we define "pre-operational baseline" as a summary of important geochemical factors (including element concentrations and other parameters) that characterizes conditions of one or more media (e.g. soil, sediment, ground water, surface water) prior to mining or other development. Thus, the geochemical baseline reflects the combined effects of the geological history of the area, including hydrothermal activity associated with ore formation, weathering processes that have affected the rocks since original formation of the ore, and historical and current anthropogenic influences (e.g. agriculture or industrial activities). Pre-mining baseline studies are distinct from studies conducted to distinguish natural geochemical variation from anthropogenic effects (e.g. Salminen and Tarvainen, 1997; Zgłobicki et al., 2011) or those intended to identify geochemical anomalies related to hydrothermal activity as a guide for mineral exploration (e.g. Caron et al., 2008; Leybourne and Cameron, 2010). Although similar in process to baseline studies, natural background and exploration surveys seek to identify outliers that can be attributed to the influence of anthropogenic activity or ore deposits.

Due to the applied nature of baseline studies, results are commonly not published in the peer-reviewed literature and instead are integrated into mining or discharge permit applications or environmental impact statements (Pebble Partnership, 2011; US EPA, 2003; US NRC, 1980). There are some exceptions, however, of baseline studies that give rise to analyses and scientific insights that receive broader distribution in the scientific community. Baseline studies reported in the literature have used both exploration and monitoring geochemical data to examine the influence of geology, including the presence of sulfide ore deposits, on stream sediment and water (Claridge and Downing, 1993; Eppinger et al., 2007; Gustavsson et al., 2012; Kelley and

Taylor, 1997). Results of these studies documented higher metal concentrations at and downstream of mineralized areas compared to upstream reaches. Other baseline studies have modeled water-rock interactions at sulfide deposits (Leybourne, 1998; Leybourne et al., 2006; Leybourne et al., 2002). All of the above studies highlighted variations in local geology, particularly in the vicinity of mineral deposits, and included samples from both mineralized and unmineralized areas when evaluating baseline. However, because the goal for most of these studies was to document baseline at a specific site, the studies were not explicit about their rationale for selecting sampling sites, and discussion of temporal variation in geochemical characteristics that are inherent in water samples is sparse.

Although the importance of conducting rigorous geochemical baseline studies prior to mining is widely recognized, there does not exist a consensus on protocols, including rationale of sample site selection, methods for evaluating temporal trends, and approaches for statistical analysis and reporting of baseline data. This paper aims to outline protocols for designing and conducting a geochemical baseline study. The choices that affect these "best practices" are illustrated using an evaluation of the stream sediments and stream water baseline geochemistry at the unmined Coles Hill uranium deposit in south-central Virginia, USA. We place particular emphasis on handling non-detect data. Our results are a first step toward developing a set of standardized best practices for designing baseline surveys. The geochemical interpretation of these data and other data collected from Coles Hill is not presented here but will be included in future papers.

2.1.1 THE PROBLEM OF NON-DETECT DATA

Statistical analysis of data is critical for baseline studies, as these statistics distill large datasets into more manageable values representing prevailing conditions that enable calculation

of constituent fluxes and loads. Statistical summaries and/or values calculated from them are commonly required when applying for discharge and other environmental regulatory permits. Summary descriptive statistics are values representing measures of central tendency (e.g. mean and median), spread (e.g. standard deviation and interquartile range), and extremes (e.g. minimum and maximum). A number of procedures have been used to define baseline but no one method has been universally accepted. Commonly, a single value is used to represent the baseline value for a given element (Salminen and Tarvainen, 1997). However, in a geologically diverse region, using a mean, median, or threshold value may be misleading, as it does not give an indication of the spread of the data. Thus, some studies include measures of spread, such as standard deviation or confidence intervals about the mean (Matschullat et al., 2000) or median (Gustavsson et al., 2012). Other studies present a statistically-determined range, such as using 95% of all observations (Tidball et al., 1974) or all observed data from the 5th to 95th percentiles (Lee and Helsel, 2005; Tidball and Ebens, 1976).

A particular challenge for computing baseline summary statistics is the presence of non-detect or censored values, synonymous terms which represent data that fall below analytical detection or reporting limits. For example, in dilute surface waters, many trace elements of interest have concentrations that consistently fall below analytical detection limits. The majority of the studies cited above (Claridge and Downing, 1993; Eppinger et al., 2007; Gustavsson et al., 2012; Kelley and Taylor, 1997; Leybourne et al., 1998) do not discuss methods for incorporating non-detect data into the statistical analysis, or they apply the widely utilized method of substituting values (generally the detection limit, half the detection limit, or zero) for non-detect measurements. Previous works, including Helsel (2005, 2012); Hron et al. (2010); Martín-Fernández et al. (2012); Martin-Fernandez et al. (2011), have presented methods for estimating

summary statistics, plotting data, and statistically comparing censored data, which have been applied in several recent environmental studies (e.g. Baccarelli et al., 2005; Carranza, 2011; Fiévet and Della Vedova, 2010; Lee and Helsel, 2005). Regulatory entities now accept and even recommend using non-detect methods (US EPA, 2003).

Methods for handling censored data fall into four categories: 1. simple substitution/replacement, 2. parametric functions, 3. regression on order statistics, and 4. survival analysis methods. Substitution or replacement is the replacement of censored data by zero, the detection limit, or a fraction of the detection limit (commonly half). More complex replacement procedures have also been suggested (Martín-Fernández et al., 2003). These methods replace all censored observations within a given variable with an identical value. There are several types of parametric functions that can be used to calculate statistics for a dataset containing non-detects. Maximum likelihood estimation (MLE) uses parameters drawn from the dataset to calculate a distribution. This method can be used on a single variable to calculate statistics (Helsel, 2012). A similar method, using an expectation-maximization algorithm (EM), can use observations from other variables to infer values for the variable with censored observations (Martín-Fernández et al., 2012). Regression on order statistics (ROS) is a semi-parametric approach that calculates a distribution for observations below the detection limit and uses observed values above the detection limit. Survival analysis methods are non-parametric procedures based on data ranking (Helsel, 2012). As the choice of method for handling censored data can influence baseline results, additional guidance for how to select an appropriate method for a specific application is warranted.

2.1.2 THE COLES HILL URANIUM DEPOSIT

The Coles Hill uranium deposit is the largest undeveloped uranium deposit in the United States (Lyntek Inc., 2010; Pincock Allen & Holt, 1982) and is located in Pittsylvania County, Virginia (Figure 2-1), in the Virginia Piedmont physiographic province. Historical land use in the area was for tobacco and cotton farming; currently, cattle grazing is the predominant use of land (Aaron, 2009).

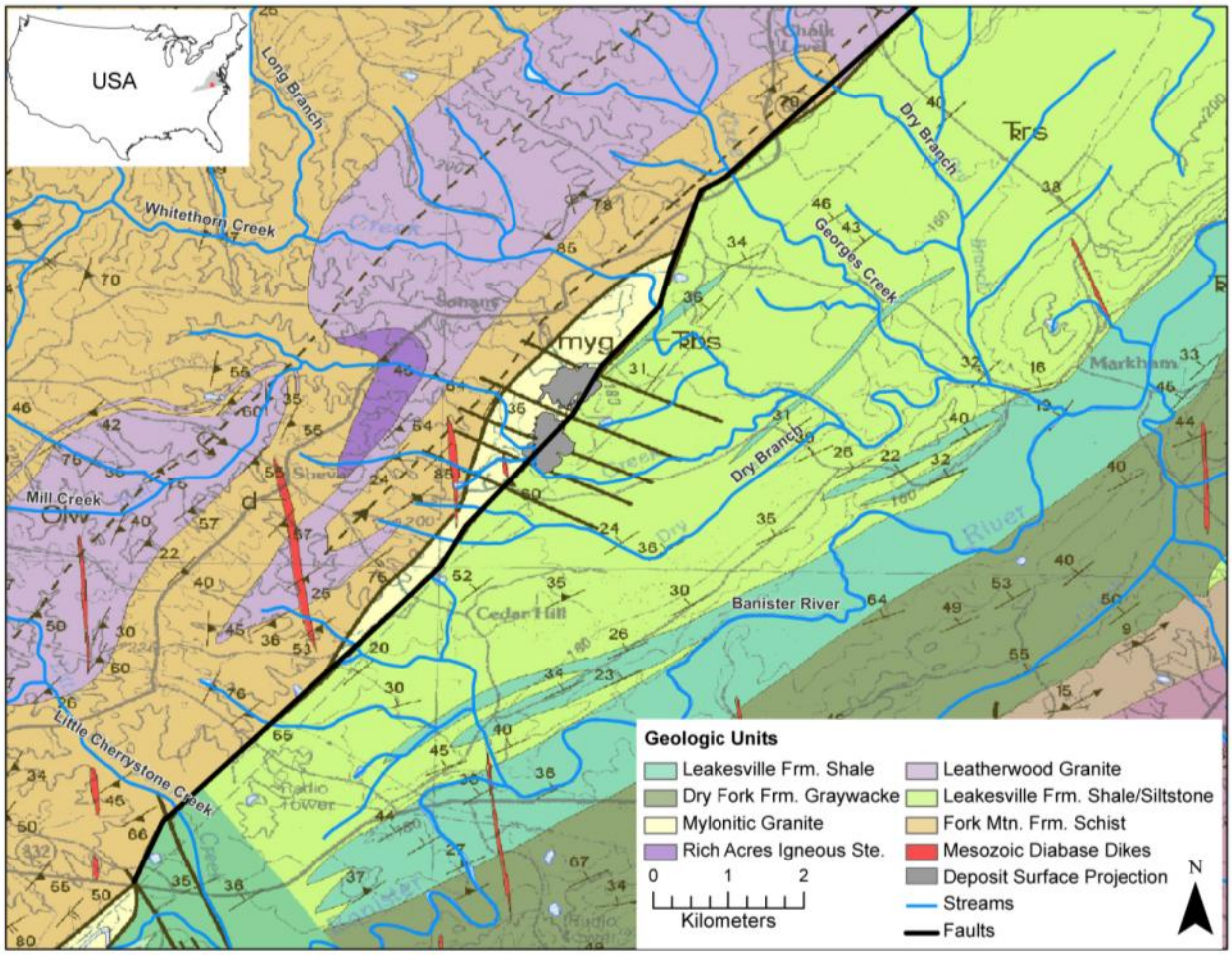

Figure 2-1: Geologic map of the Coles Hill region. A surface projection of the deposit, at a cutoff grade of 0.025 wt. %, is shown in gray. The solid black line oriented NE-SW is the Chatham Fault; the lines perpendicular to the Chatham Fault are cross-faults or lineations (modified from Henika, 2002).

The deposit consists of two ore bodies and is located in the footwall of the Chatham Fault, a normal fault trending northeast-southwest and dipping at ~60° to the southeast (Gannon et al., 2011; Henika, 2002; Meyertons, 1963). The Chatham Fault is cut by a number of high-angle features near the ore bodies that have been referred to as cross faults by Wyatt (2009) and Jerden and Sinha (2003). The Chatham Fault represents the boundary between the Virginia Western Piedmont province to the northwest and the Danville Triassic Basin to the southeast. The Piedmont province in the vicinity of the deposit consists mainly of the Fork Mountain Formation (schist and gneiss) and the Martinsville Igneous Complex, which includes the Leatherwood Granite; the ore is present within mylonitized Leatherwood Granite gneiss and amphibolite. The Danville Triassic Basin to the east of the deposit includes the Leakesville Formation, which consists of shale and siltstone with lesser amounts of sandstone and claystone (Gannon et al., 2011; Henika, 2002; Henika and Thayer, 1983). Figure 1 shows the regional geology of Coles Hill.

The deposit is located in the Whitethorn Creek watershed (Figure 2-2). Whitethorn Creek is an east-flowing stream that discharges into the Banister River approximately 5 km east of the deposit. The Whitethorn Creek watershed covers 170 km^2, mostly upstream of the deposit. Smaller tributaries within this watershed in the vicinity of the deposit include Mill Creek, Dry Branch (southwestern), and several unnamed farm drainages and intermittent streams. Georges Creek and a second Dry Branch (northeastern) flow into Whitethorn Creek downstream of the deposit and within one kilometer of the confluence with the Banister River, but the Coles Hill deposit is not within the surface watersheds of these streams. The water table in the vicinity of the deposit is 0 to 20 m below ground surface (Marline Uranium Corporation, 1983), and much of the groundwater occurs in a fractured rock aquifer and overlying saprolite (Gannon et al.,

2011; Jerden and Sinha, 2003). Small springs are common, and a number of these springs are dammed to create ponds for livestock and irrigation. In the area of the deposit, groundwater flows from west to east, and, although there is some hydrologic connection across the Chatham Fault, the fault is minimally permeable (Gannon et al., 2011).

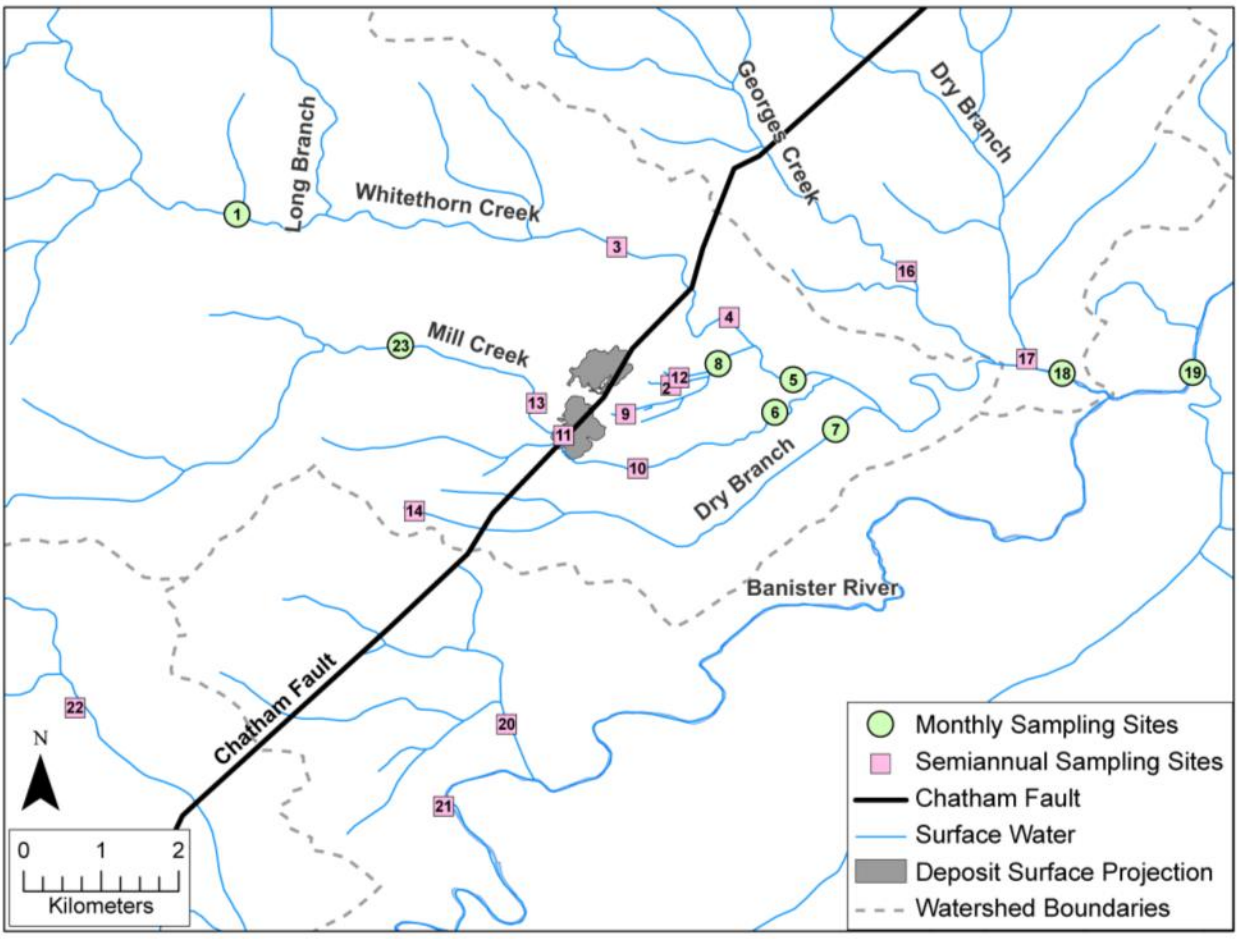

Figure 2-2: Stream sampling sites for the Coles Hill baseline study. Monthly sampling sites are shown as green circles, and additional sites sampled semiannually are shown as pink squares. Surface water features (from the USGS National Hydrography Dataset) are shown in blue, and watershed delineations (from the USDA Watershed Boundary Dataset) are shown as dashed gray lines. A surface projection of the deposit, at a cutoff grade of 0.025 wt. %, is shown in gray.

The study area is located in a temperate humid climate. Average annual temperature is 13°C (55°F) with a minimum in January (average low -5°C or 23°F) and a maximum in July (average high 31°C or 87°F). Annual rainfall is approximately 1150 mm (45 inches) and is

distributed evenly throughout the year (NCDC, 2012). The period of this study (March 2011-March 2012) was warmer and drier than average (NCDC, 2012), and had a mild winter.

2.2 METHODS

2.2.1 SAMPLING STRATEGY

2.2.1.1 SPATIAL SITE SELECTION

One of the study goals was to develop a rationale for sampling strategies for baseline studies. Our strategies are based on the U.S. Environmental Protection Agency (EPA) guidance on hardrock mining (US EPA, 2003) and the U.S. Nuclear Regulatory Commission (NRC) guidance on pre-operational monitoring guidelines for uranium mills (US NRC, 1980), focusing on flowing surface water and stream sediments. A rich literature also exists on general water quality sampling design (e.g. Dobbie et al., 2008; Khalil and Ouarda, 2009; Strobl et al., 2006).

We first identified locations both upstream and downstream of the deposit to allow for comparison of waters that are potentially influenced by the undeveloped deposit. Upstream sites were defined as sites in the Whitethorn Creek watershed upstream of the Chatham Fault and in the Banister River upstream of its confluence with Whitethorn Creek. Downstream sites were defined as any sites downstream of the Chatham Fault within the Whitethorn Creek watershed as well as the Banister River site downstream of the confluence with Whitethorn Creek.

Second, we identified sampling sites that reflected site geological variability to ensure that we could capture and identify geochemical signatures of different geological units and any anomalies associated with the ore body. To cover the geological variability of the site, we

selected sites both in the crystalline Piedmont rocks to the west of the Chatham Fault and Triassic sedimentary rocks to the east of the fault.

Third, we selected sample sites in other watersheds not expected to be affected by the deposit or proposed mining that could be considered reference sites. The relationship of the deposit to the Chatham fault, which separates the Piedmont province from the Triassic basin, is such that the upstream sites are located in the Piedmont and the downstream sites are located in the Triassic. Because the separation between upstream and downstream lies at the intersection of two geologically diverse provinces, the conflation of upstream and downstream differences with the effects of geological heterogeneity is possible. To address the influence of the differing underlying lithologies between upstream and downstream sites, reference watersheds within the Danville Triassic Basin were also included. The reference sites are from nearby streams that do not contain the deposit within their surface watersheds and are located in the Triassic basin, though upstream reaches may, as with the Whitethorn Creek watershed, be located in the Piedmont.

Our final sampling site design included twenty-two sampling sites (Figure 2-2), which include five sites upstream of the deposit (Sites 1, 3, 13, 14, and 23) and ten sites downstream of the deposit (2, 4, 5, 6, 7, 8, 9, 10, 11, 12, and 18) in the Whitethorn Creek watershed, one upstream and one downstream site each in the Banister River (21 and 19, respectively), and four sites in reference watersheds (16, 17, 20, and 22). Monitoring of these sites during any future activity will allow comparison of pre-, syn-, and post-operational conditions, with upstream and reference sites serving as controls for hydrogeochemical changes unrelated to mining or processing.

2.2.1.2 TEMPORAL SAMPLING

As many of the published baseline studies focus on soils and stream sediment, the temporal geochemical variation inherent in water samples is commonly not considered. In baseline studies that address water, sampling is not always conducted during the full range of hydrologic conditions, or in the case of arid regions, surface water is ephemeral and water can only be sampled during runoff events. However, in temperate regions such as our study area, where many streams are perennial and discharge is highly variable, temporal sampling is imperative for evaluating how geochemical characteristics of water fluctuate with temperature, which can relate to sorption and biological activity, and discharge.

The timing of our sampling events is consistent with NRC pre-operational monitoring guidelines (US NRC, 1980). NRC recommends sampling stream sediments twice, once following spring runoff and once in late summer following a period of extended low flow, and sampling flowing surface waters monthly. Using this guidance, we sampled stream sediments twice, in March and September 2011. The entire network of stream water sites were sampled twice (semiannually), in March and September 2011. A subset of eight (sites 1, 5, 6, 7, 8, 18, 19, and 23) was sampled on a monthly basis from March 2011 to March 2012. In addition, because we anticipated that stream discharge could influence the geochemical characteristics of our water samples, we measured stream discharge during each event at each sampling site.

It is important to note that our study was conducted over a one year period, which is the time period required by NRC for pre-operational monitoring. However, we anticipate continued monitoring of these sites to collect additional baseline data under a wider range of climatic and hydrologic conditions.

2.2.2 SAMPLE COLLECTION AND ANALYSIS

2.2.2.1 STREAM SEDIMENT

Stream sediment samples were collected from the top 10 cm of the streambed using a stainless steel trowel and high-density polyethylene (HDPE) bucket. Composite samples of at least 30 increments were collected following procedures outlined in Shelton and Capel (1994). Composited samples were mixed in a stainless steel or HDPE bowl and wet-sieved in the field to <2 mm. Following collection, samples were air-dried and re-sieved in the laboratory for homogenization.

Metal and carbon analyses for stream sediments were conducted by SGS Laboratories in Toronto, Canada, under a contract with the USGS Analytical Chemistry Services Group in Denver, CO. Metals were measured using inductively-coupled plasma mass spectrometry (ICP-MS; As, Pb, Th, U) and inductively-coupled plasma atomic emission spectroscopy (ICP-AES; Al, Ca, Fe, K, Mg, Na, Ti, Ba, Cu, Mn, V, Zn) following a multi-acid (HCl-HNO_3-$HClO_4$-HF) digestion. Data were deemed acceptable if the relative standard deviation of duplicate run samples was within 15%. Carbon species were analyzed using an elemental analyzer. QA/QC samples consisted of field duplicates, which were collected from the same streambed area immediately following collection of the original sample.

2.2.2.2 SURFACE WATER

Grab samples of surface water were collected from each site using a dip method (USGS, 2006). Samples intended for dissolved constituent analysis were filtered in-line through a hydrophilic 0.45 μm disposable capsule filter using a peristaltic pump and platinum-cured silicone tubing. Field parameters including specific conductance (SC), pH, temperature, dissolved oxygen (DO), and oxidation-reduction potential (ORP) were measured *in situ* using a

multiprobe (Hanna Instruments 9828). Stream velocity was measured using a flow meter along a defined transect; discharge (Q) was calculated using the velocity-area method, assuming a rectangular stream channel.

Water chemical analyses were conducted by the USGS Mineral Resources Program laboratory in Denver, CO (Taggart, 2002). Major cations (Ca, K, Mg, Na) were analyzed by ICP-AES. Trace elements were analyzed by ICP-MS. Anions (Cl and sulfate) were analyzed by ion chromatography. Dissolved organic carbon (DOC) was measured by carbon analyzer. Alkalinity was determined by titration and calculated using the inflection point method (Rounds, 2006). Samples were also analyzed for selected radionuclides; those results are not presented here. The QA/QC program consisted of field equipment blanks, duplicate samples, and analysis of USGS standard reference samples for major and trace elements dissolved in water. One blank and at least one duplicate were collected during every sampling event. Sample batches were flagged for reanalysis if reference samples differed by more than 10% of their published values. Additional laboratory QA/QC procedures are described by Taggart (2002).

2.2.3 GEOCHEMICAL DATASET

Major elements that are indicative of the general geochemical character of the media and which were incorporated into the statistical analysis included Ca, K, Mg, and Na in both media, and Cl and sulfate in waters; trace elements associated with the deposit and/or regulated in discharge by federal (40CFR440) and state (9VAC25-151-150) regulations included Al, As, Ba, Cu, Fe, Mn, Pb, Th, Ti, U, V, and Zn. For waters, unless otherwise specified, data for dissolved (filtered) constituents were used, and other relevant parameters include pH, alkalinity, and DOC. Total organic carbon (TOC) in stream sediments was also included. To avoid bias from the greater frequency of monthly sampling sites (compared to semiannual sites), summary and

temporal analyses were conducted on monthly samples (n=98), whereas analyses based on site characteristics were conducted on semiannual samples (n=44), which had greater spatial coverage. We present statistical analyses for the selected elements and parameters listed above, although data exist for other elements in the various media. Data not presented in this paper were excluded because (1) only elements with direct geochemical relevance to mineralization at the Coles Hill deposit (Jerden and Sinha, 2003) are included, and (2) dissolved concentrations of several constituents (e.g. Cd, Cr, Ni, and others) were almost entirely (>90%) below analytical detection limits.

2.2.4 STATISTICAL ANALYSIS

Summary statistics and statistical hypothesis testing were used to evaluate differences in geochemical conditions relative to underlying geology or location relative to the deposit (i.e. upstream or downstream). Summary statistics were calculated, including mean, median, standard deviation, minimum, maximum, and quartiles. When non-detect data are included, ROS functions in R (see below for discussion) were used to calculate statistics. Non-parametric hypothesis tests, including the Mann-Whitney-Wilcoxon rank-sum test (two groups) or the generalized Wilcoxon test (three or more groups), were used to assess site and temporal variations. For all tests, an alpha of 0.05 was used. Sites were compared on the basis of their locations relative to the ore deposit (upstream/reference or downstream; west in Piedmont or east in Triassic Basin) and on underlying geological units. Temporal comparisons were done between each month.

2.2.4.1 EVALUATION OF NON-DETECTS

Many elements of interest (As, Cu, Pb, Mn, Th, U, Ti, V, and Zn) in our water dataset contained non-detect data. As discussed in the introduction, there are four approaches for

handling non-detect data, including ROS, parametric functions, substitution, and survival analysis. Table 1 summarizes methods for evaluating non-detects and in what situations they are best used. Figure 3 shows cumulative distribution functions indicating how each method approaches data below the detection limit, which is shown as a dotted vertical line. A plot of actual values is shown for comparison. Three of the methods, ROS, the MLE function, and survival analysis, do not provide discrete values for observations below the detection limit. Survival analysis does not provide information on data below detection but can be used to estimate a mean. The ROS and MLE function methods use a data distribution based on above-detection data and enable the calculation of summary statistics, including those that fall below the detection limit. Substitution and EM with imputation replace below-detection data with discrete values; in the case of substitution, each replacement value is identical.

Table 2-1: Non-detect statistical methods. Guidance on using each method is given, as are corresponding functions in the R statistical software. Much of the information in this table can also be found in the work by Helsel (2005, 2012).

Method	Use with	R function	Advantages	Disadvantages
ROS	Small datasets (n<50)	cenros (NADA)	• Can be used for highly censored data	• Semi-parametric (relies on assumption of log-normal or normal distribution
			• Provides statistical summaries without data values for censored data	• Limited for more in-depth analyses due to lack of discrete values
EM (imputation)	Large datasets (n>50)	impRZilr (robCompositions)	• Can be used for highly censored data	• Imputed values are estimates; interpretation using imputed values could be misleading
			• Provides estimates for more in-depth analysis	• Requires accompanying detectable data from other elements • Parametric (assumes distribution)
MLE (function)	Large datasets (n>50)	cenmle (NADA)	• Can be used for highly censored data	• Inaccurate at data extremes • Parametric (assumes distribution)
Substitution (e.g. 1/2 DL)	Small fraction BDL, multiple DL	n/a	• Simple	• Not accurate for high proportions of non-detects • Statistical measures of spread highly biased
Survival Analysis	<50% BDL, multiple DL	cenfit (NADA)	• Non-parametric	• Does not provide min or max • Cannot provide estimations for percentiles below detection limit

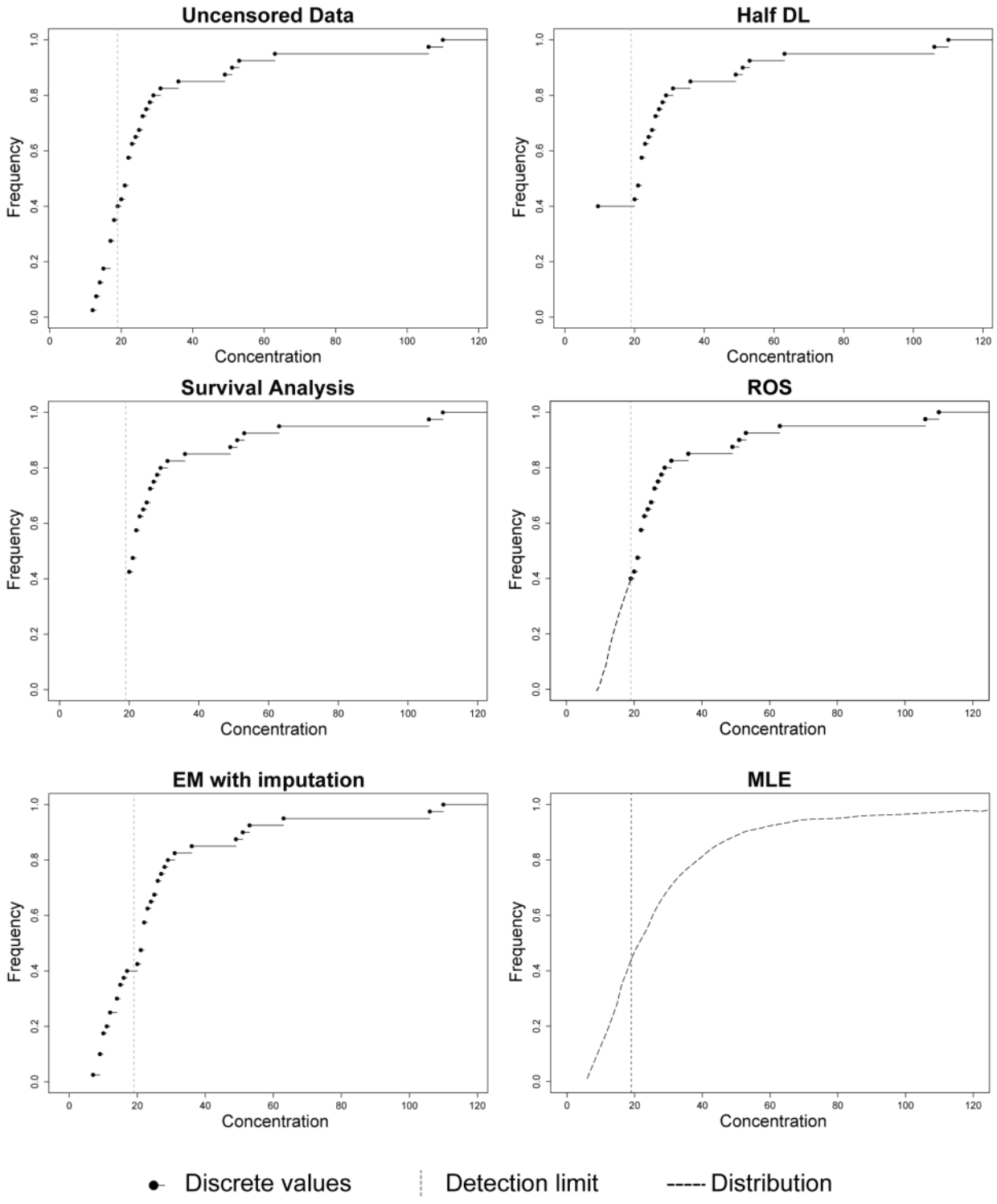

Figure 2-3: Cumulative distribution functions showing different approaches to censored data. Points and solid lines represent empirical data, whereas dashed lines represent functions. The detection limit is shown as a vertical dotted line.

Based on the proportions of censored data and the size of our dataset (n=40 for stream sediments and 98 for waters), the most appropriate methods for handling censored data are ROS, MLE, or EM. To do further evaluation, we applied the three methods described above (ROS, EM, substitution) that can provide the full suite of summary statistics to examine their effect on Th, U, and V concentrations in stream sediments artificially censored at 20%, 40%, 60%, and 80% for comparison with the complete dataset. In stream sediments, concentrations for these three elements are all above detection limits, and artificially censoring these data at different levels shows the effect censoring can have on estimated parameters. Real data were used rather than simulations to assess the variations these methods introduce when data are not random but also do not necessarily conform to a particular distribution, as can be the case in baseline studies with a limited number of data points (Helsel and Gilliom, 1986). Calculations were done using R (R Development Core Team, 2011). Using regression on order statistics (ROS) from the R package NADA (Lee, 2010), rounded zero imputation from the R package robCompositions (Templ et al., 2013), and substituting half the detection limit, summary statistics were calculated for the artificially-censored data to be compared with the complete data.

2.3 RESULTS

2.3.1 VARIATIONS AMONG SITES

2.3.1.1 STREAM SEDIMENTS

The site distribution of selected elements in stream sediments is shown in Figure 2-4. Stream sediments were compared by location relative to the deposit and according to underlying geology. Comparisons of locations relative to the deposit consisted of two analyses, the first comparing upstream and reference sites to downstream sites, and the second comparing

reference Triassic basin sites to downstream Triassic basin sites. Significant differences in median values are seen between upstream/reference and downstream sites for Al, Ba, Ca, K, Th, U, Zn, and organic carbon (Table 2-2). No significant differences were observed between reference and downstream Triassic basin sites. Data distributions for Al, As, Ba, Ca, Cu, K, Na, and organic carbon were significantly different based on underlying bedrock geology (Table 2-2).

Table 2-2: Table of p-values from Wilcoxon signed-rank tests comparing data collected from different bedrock provinces (Piedmont and Triassic basin), from upstream (U/S) and reference sites vs. downstream (D/S) sites, and from Triassic basin reference sites vs. sites downstream of the deposit. Significant (<0.05) p-values are in bold italics. Measurements for As, Th, and Zn in stream water were 100% below detection and are thus not included in this analysis.

	Seds			Water		
Analyte	Bedrock	U/S-D/S	Ref.	Bedrock	U/S-D/S	Ref.
Al	***0.031***	***0.012***	0.481	***0.003***	0.062	0.982
As	***0.044***	0.197	0.886	n/a	n/a	n/a
Ba	***0.000***	***0.021***	0.888	0.118	0.185	0.639
Ca	***0.016***	***0.046***	0.452	0.107	0.425	0.708
Cu	***0.014***	0.066	0.963	***0.001***	***0.011***	0.706
Fe	0.505	0.194	0.963	0.791	0.319	0.270
K	***0.003***	***0.016***	0.452	***0.014***	***0.016***	0.360
Mg	0.522	0.296	0.925	***0.004***	***0.037***	0.796
Mn	0.533	0.570	0.888	0.500	0.343	0.060
Na	***0.029***	0.064	0.604	0.762	0.664	0.336
Pb	0.610	0.074	0.302	***0.001***	***0.009***	0.453
Th	0.160	***0.012***	0.670	n/a	n/a	n/a
Ti	0.385	0.957	1.000	0.084	0.059	0.348
U	0.200	***0.004***	0.370	0.112	***0.007***	0.058
V	0.320	0.250	1.000	***0.006***	***0.014***	0.423
Zn	0.105	***0.049***	0.963	n/a	n/a	n/a
OC	***0.006***	***0.000***	0.279	***0.003***	***0.016***	0.597

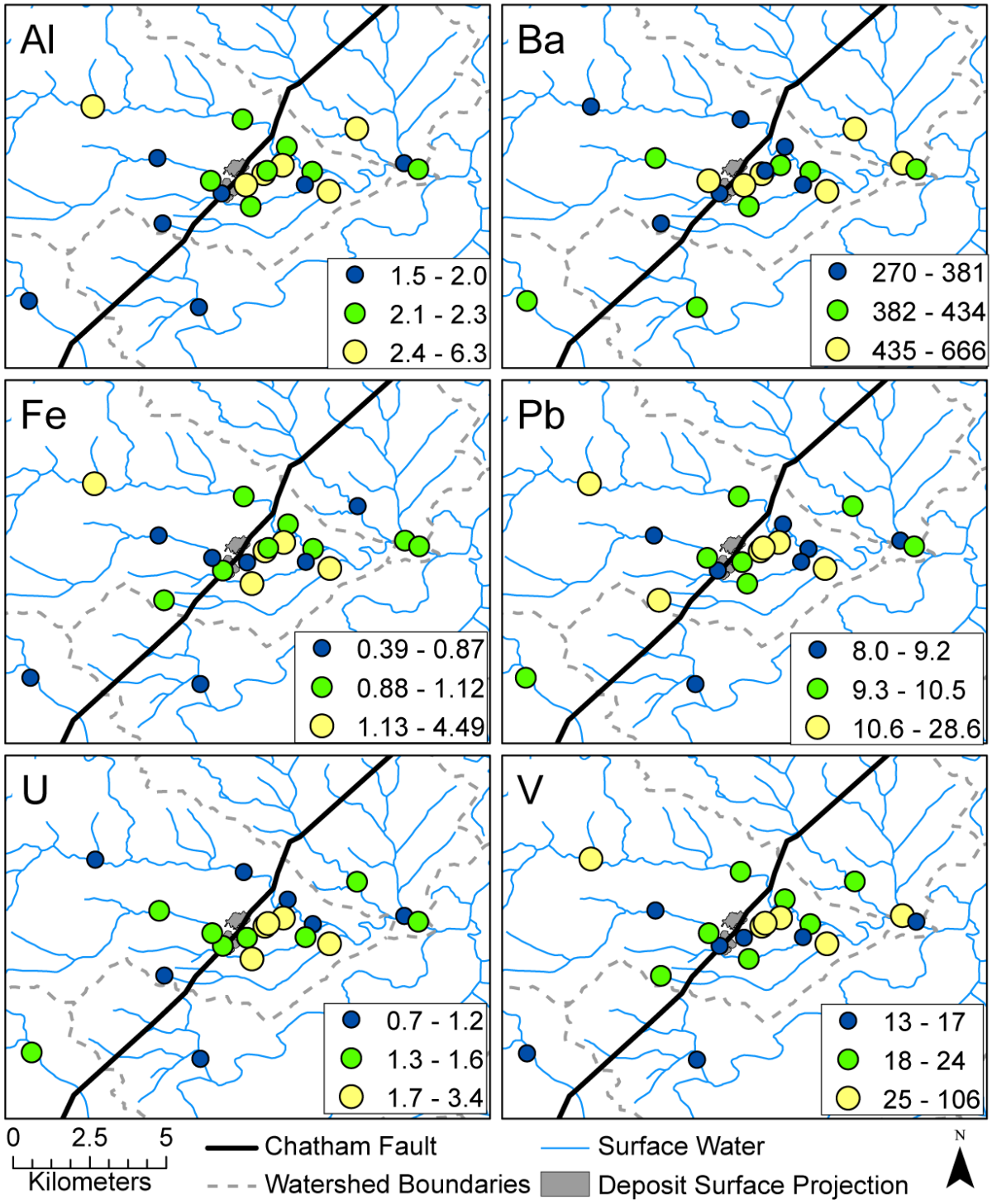

Figure 2-4: Maps showing concentrations of selected elements in stream sediments in March 2011. Al and Fe are in %; all others are in mg/kg. Larger and lighter circles indicate higher concentrations. Symbol bins are divided into terciles (three evenly-sized groups).

2.3.1.2 STREAM WATER

The site distribution of selected dissolved elements in streams is shown in Figure 2-5. Element concentrations in the dissolved fractions at the eleven downstream sites were compared to concentrations from the eleven upstream/reference sites, and concentrations at reference Triassic basin sites were compared to downstream Triassic sites (Table 2-2). Downstream sites were significantly higher in Cu, K, Mg, Pb, U, V, and DOC than upstream sites. No significant differences were found between reference and downstream Triassic basin sites. Arsenic, Th, and Zn measurements in the spatial dataset were 100% BDL; no comparison could be made for these elements. The highest concentrations for all elements were found in farm drainages east of the deposit that drain the area where the ore bodies outcrop (Sites 2 and 12 in Figure 2-2). Significant differences in data distributions for Al, Cu, Fe, K, Pb, V, and DOC were seen as a function of underlying geology (Table 2-2).

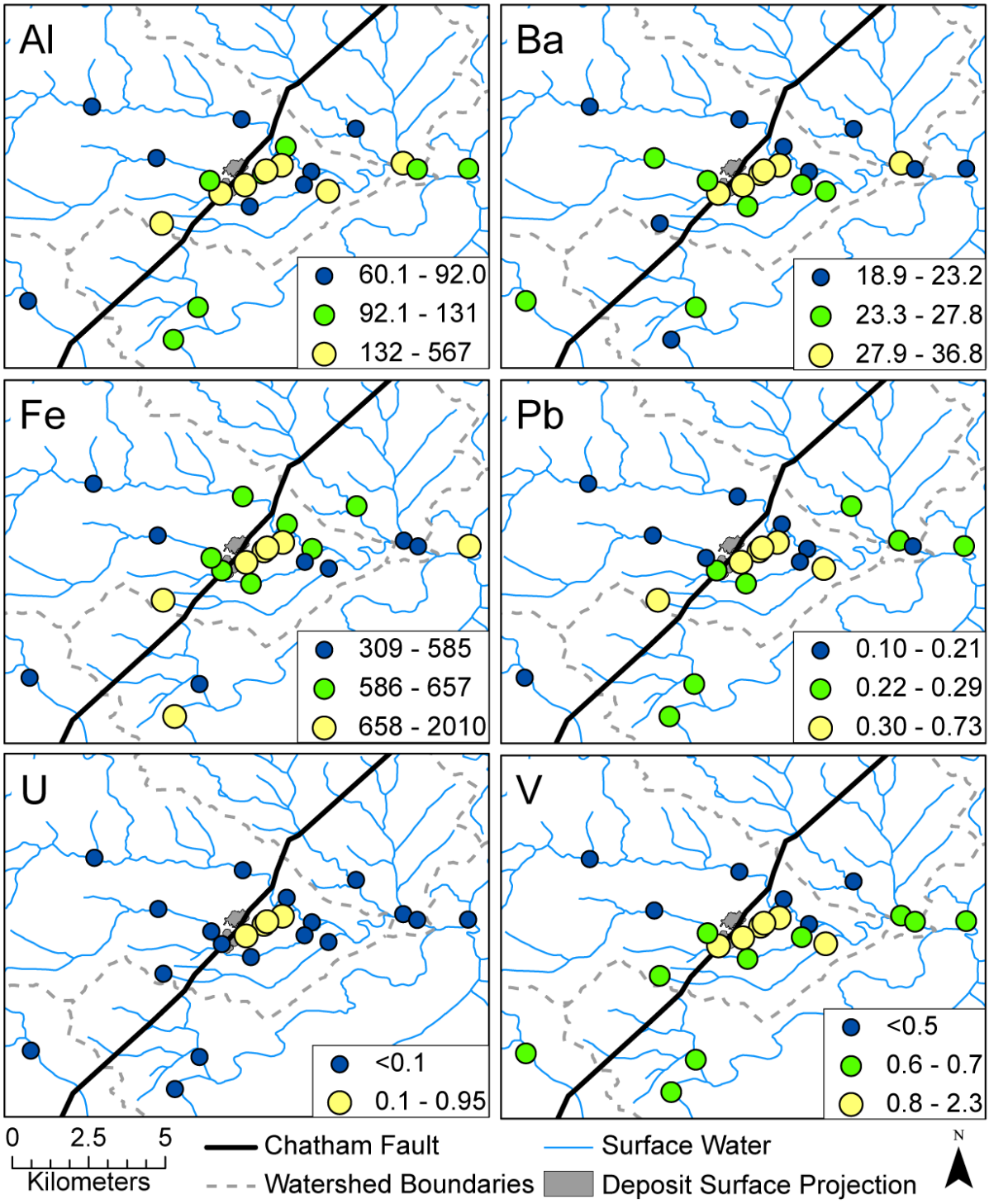

Figure 2-5: Maps showing dissolved (< 0.45 µm) concentrations in µg/L of selected elements in stream waters in March 2011. Larger and lighter circles indicate higher concentrations. Symbol bins are divided into terciles for Al, Ba, Fe, and Pb. For elements with non-detect observations (U, V), the break between the first and second tercile was adjusted to the detection limit so that the lowest concentration bin contains only those observations below detection.

2.3.2 TEMPORAL VARIATIONS

Temporal variations were assessed for stream waters at sites sampled monthly (Figure 2-6). Significant differences (paired Wilcoxon test) in concentrations were observed during the thirteen months of this study, and most of these differences appear to relate to changes in discharge. Seasonal changes were assessed by measuring correlations of dissolved elements with temperature and flow (discharge) by Kendall's tau (Table 2-3). Temperature correlated negatively with changes in Na, sulfate, and Pb, and positively with Th and V concentrations, but other elements analyzed showed no correlation with temperature. Concentrations for Ca, Mg, Na, Ba, Cu, Mn, U, V, DOC, and alkalinity showed a negative correlation with flow (Figure 2-7). Increases in Fe appear to correspond with increases in flow (Figure 2-6), but this trend is not sufficiently consistent to be statistically significant.

Table 2-3: Kendall's tau values for selected elements (n=98) in stream water vs. temperature (T) and flow (Q). Values corresponding with significant ($p < 0.05$) correlation are in bold.

Analyte	T	Q
Ca	0.09	***-0.29***
K	0.13	-0.09
Mg	-0.04	***-0.25***
Na	***-0.14***	***-0.22***
Cl	0.10	0.01
SO4	***-0.33***	-0.05
Al	0.01	-0.13
Ba	-0.09	***-0.35***
Cu	0.04	***-0.14***
Fe	0.02	-0.10
Mn	0.11	***-0.22***
Pb	***-0.15***	-0.10
Th	***0.38***	-0.04
Ti	-0.07	-0.13
U	0.04	***-0.43***
V	***0.18***	***-0.24***
DOC	-0.05	***-0.17***
pH	0.04	0.03
ORP	-0.01	0.08
Alk.	***0.21***	***-0.30***

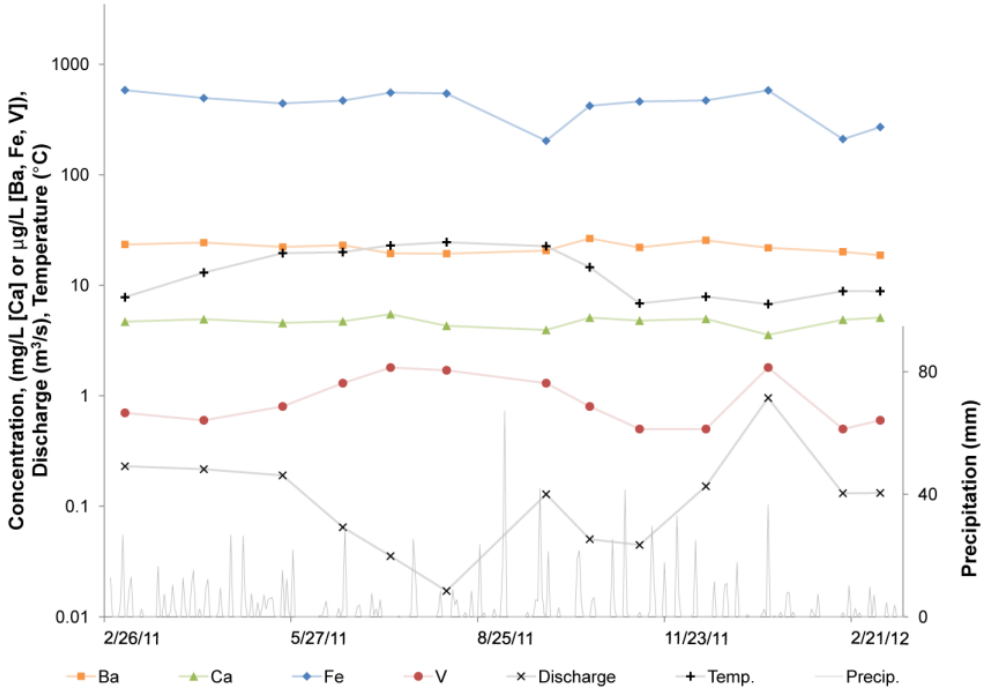

Figure 2-6: Time-series plots of filtered concentrations of selected elements and discharge at Site 23 throughout the course of this study. BDL V concentrations were plotted as the detection limit (0.5 µg/L). Precipitation data are from the Chatham, VA, station (NCDC, 2012).

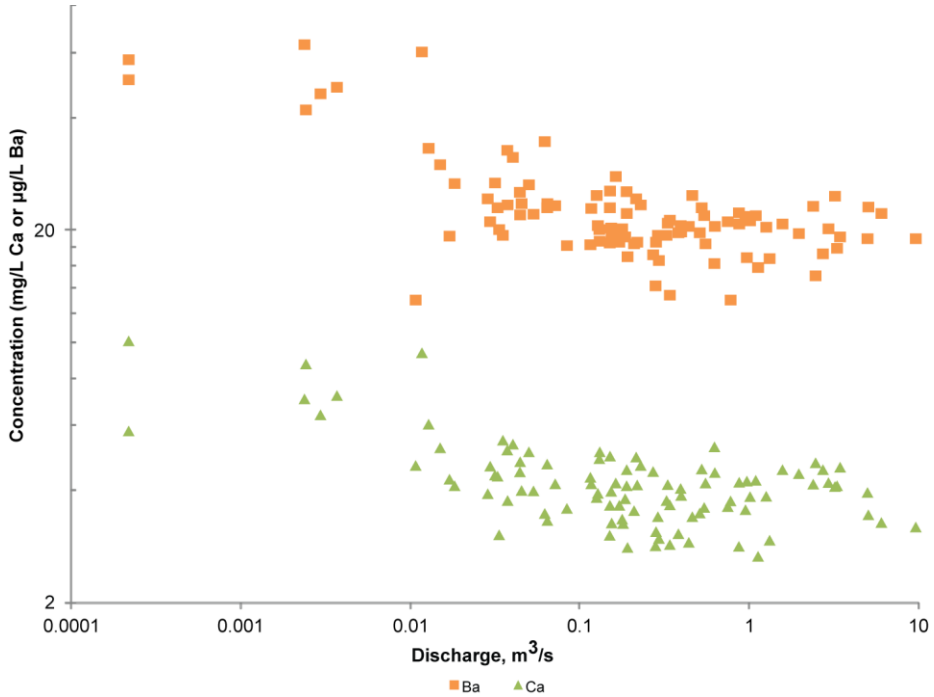

Figure 2-7: Logarithmic plot of discharge vs. Ba and Ca concentrations at all sites sampled monthly during the course of this study.

2.3.3 COMPARISON OF NON-DETECT METHODS

When non-detect methods (ROS, EM, substitution) were compared on artificially-censored stream sediment data, statistical estimates became less accurate relative to original data as the proportion of non-detects increased (Figure 2-8). By definition, percentiles greater than the percentage of non-detects remained unchanged by censoring. Percentiles lower than the percentage of non-detects showed wide variation. In all cases, the mean decreased as the censored proportion increased and, in most cases, so did the median. Overall, estimated statistics were similar between the three methods used, except for the estimated medians at higher percentages of censoring and data spread in the substitution method. Estimated median values for analyses with greater than half the observations censored were also lower than the actual median, though in all but one case, ROS medians were most similar to the actual median. Data spread, as shown by the interquartile range (the extent of the box in Figure 2-8, or the 25^{th} to 75^{th} percentiles), was generally greater in the case of non-detects relative to uncensored data, with the exception of 80% censored substitution, where all data in the range are replaced by an identical value, collapsing the box. Based on the results of this analysis, we chose to use the ROS method for censored data in this dataset.

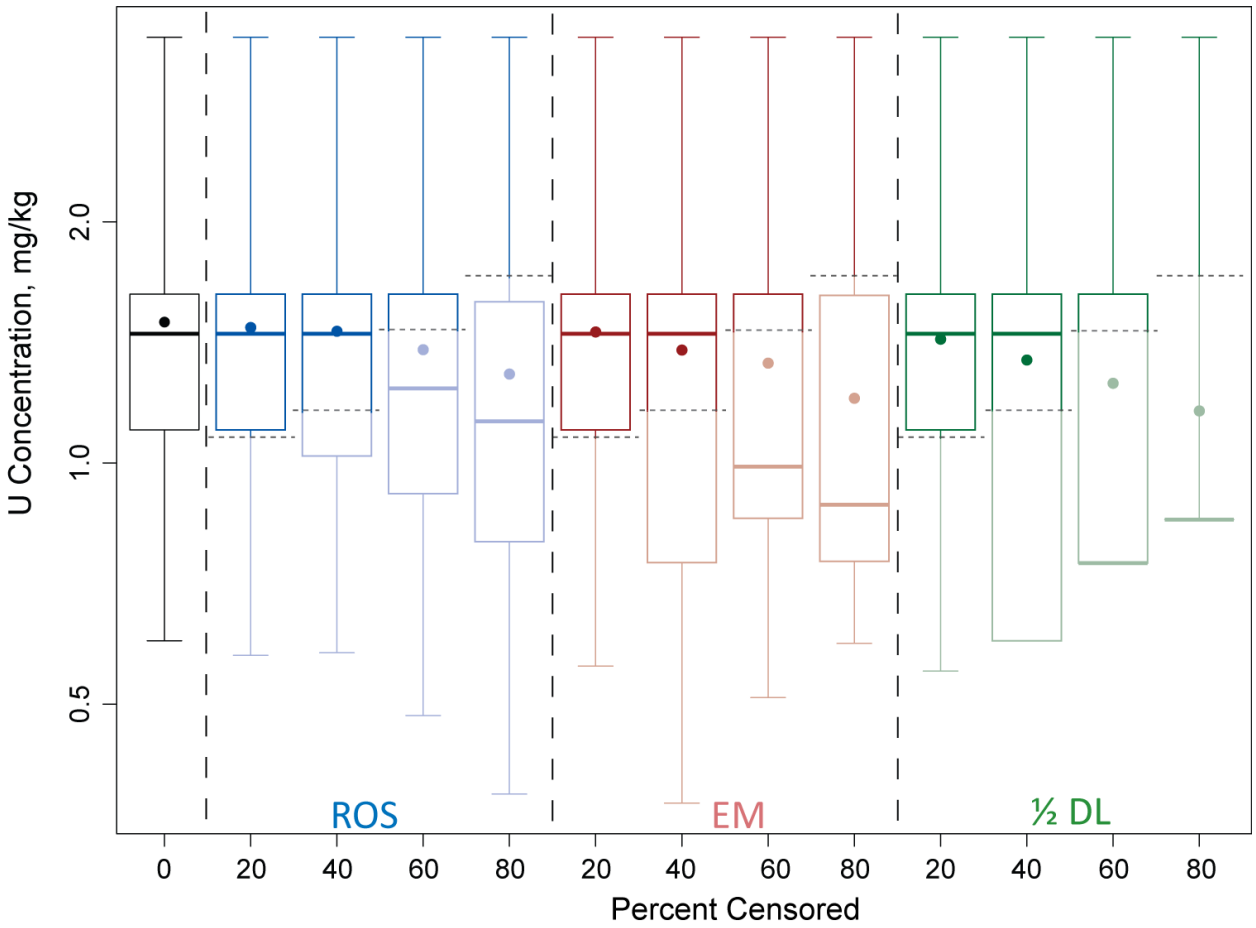

Figure 2-8: Boxplots showing statistics for uranium stream sediment artificially censored to 20%, 40%, 60%, and 80% of values below detection by three different methods. Methods used were ROS, imputation based on EM, and substitution of half the detection limit (Table 1). The plot on the left shows the original data distribution. Whiskers extend to maximum and minimum values, whereas the inner box represents, from bottom to top, the first quartile, median, and third quartile. The dashed horizontal line extending beyond the width of the box and whiskers represents the detection limit; data shown in gray below that line are simulated. Data means are shown as points.

2.3.4 BASELINE VALUES

Ultimately, baseline for each element is given as a range of values, including the minimum, maximum, and median of the data collected in this study (Table 2 and Table 3, for sediment and water, respectively). These statistics lump both upstream and downstream sites, and samples from all flow conditions as described above. Accompanying boxplots provide a

visualization of baseline ranges (Figure 2-9). For parameters with non-detect data, censored boxplots are shown. Elements with non-detect data include As in stream sediments and Cu, Mn, Pb, Th, Ti, U, and V in water.

Table 2-4: Summary statistics for selected elements in stream sediments (n=40) in this study. Italicized values contain below detection limit (BDL) values and were calculated by regression on order statistics (ROS). One measurement was removed from the Pb dataset for data quality concerns. Table includes detection limits (DL), percentage of data below detection limits (%BDL), minimum (Min), first quartile (1Q), third quartile (3Q), maximum (Max), and standard deviation (SD).

Analyte	Units	DL	BDL	Min	1Q	Median	Mean	3Q	Max	SD
Al	%	0.01	0%	1.49	1.95	2.24	2.37	2.53	6.34	0.82
As	mg/kg	1	38%	*0.13*	*0.80*	2	4.3	4	21	5.6
Ba	mg/kg	5	0%	269	363	393	412	447	666	85
Ca	%	0.01	0%	0.04	0.10	0.20	0.20	0.27	0.37	0.10
Cu	mg/kg	0.5	0%	0.6	1.8	2.7	3.7	4.1	16.9	3.5
Fe	%	0.01	0%	0.32	0.80	0.93	1.32	1.18	6.02	1.14
K	%	0.01	0%	0.65	1.01	1.15	1.21	1.25	2.21	0.33
Mg	%	0.01	0%	0.02	0.08	0.11	0.14	0.14	0.53	0.11
Mn	mg/kg	5	0%	55	177	222	274	282	1140	212
Na	%	0.01	0%	0.14	0.26	0.36	0.35	0.41	0.98	0.15
Pb	mg/kg	5	0%	7.8	8.7	9.6	10.8	10.8	28.6	3.9
Th	mg/kg	0.2	0%	1.9	3.7	4.1	4.2	4.6	10.8	1.4
Ti	%	0.01	0%	0.11	0.21	0.24	0.25	0.28	0.46	0.06
U	mg/kg	0.1	0%	0.6	1.1	1.5	1.5	1.6	3.4	0.6
V	mg/kg	1	0%	12	17	22	28	27	110	22
Zn	mg/kg	1	0%	4	10	13	16	17	69	12
TOC	%	0.01	0%	0.02	0.07	0.12	0.27	0.35	1.79	0.37

Table 2-5: Summary statistics for selected elements in stream waters from monthly sampling (n=98) in this study. Italicized values contain below detection limit (BDL) values and were calculated by regression on order statistics (ROS). Mean and standard deviation (SD) for Cu, Mn, Pb, Th, Ti, U, and V were calculated by ROS. For Th and U, with %BDL>80, results are tenuous. Alk.=Alkalinity as mg $CaCO_3$/L Table includes detection limits (DL), percentage of data below detection limits (%BDL), minimum (Min), first quartile (1Q), third quartile (3Q), maximum (Max), and standard deviation (SD).

Analyte	Units	DL	% BDL	Min	Q1	Median	Mean	Q3	Max	SD
Al	**µg/L**	2	0	2.2	19.5	36.9	130.1	146	1190	201
Ba	**µg/L**	0.2	0	13	19	21	23	24	63	9.2
Ca	**mg/L**	0.1	0	2.67	3.57	4.13	4.30	4.62	10.1	1.26
Cl	**mg/L**	0.06	0	2.6	3.8	4.5	4.5	4.9	8.9	1.2
Cu	**µg/L**	0.5	57.1	*0.04*	*0.21*	*0.42*	*0.68*	1.0	4.2	*0.74*
Fe	**µg/L**	50	0	130	326	465	466	555	1220	198
K	**mg/L**	0.1	0	1.05	1.71	2.03	2.36	2.50	7.14	1.22
Mg	**mg/L**	0.1	0	1.25	1.59	1.75	2.05	2.08	4.86	0.76
Mn	**µg/L**	0.2	13.3	*5.3*	*27.7*	*41.1*	*87.8*	71.2	793	*147*
Na	**mg/L**	0.1	0	2.77	3.51	4.18	4.28	4.84	7.82	0.98
Pb	**µg/L**	0.05	43.9	*0.0054*	*0.045*	0.090	*0.19*	0.28	1.2	*0.21*
SO4	**mg/L**	0.08	0	2.0	2.5	3.2	4.1	4.5	13.2	2.6
Th	**µg/L**	0.2	80.6	*0.0016*	*0.023*	*0.065*	*0.21*	*0.19*	1.6	*0.39*
Ti	**µg/L**	0.5	26.5	*0.020*	*0.33*	0.80	*3.2*	3.2	28.2	*5.3*
U	**µg/L**	0.1	84.7	*0.0019*	*0.013*	*0.029*	*0.055*	*0.063*	0.34	*0.07*
V	**µg/L**	0.5	32.7	*0.14*	*0.45*	0.70	*0.85*	1.1	2.9	*0.56*
Alk.	**mg/L**	n/a	0	7.04	18.5	22.8	22.8	26.5	54.2	7.6
DOC	**mg/L**	0.2	0	1.20	1.98	2.45	4.17	4.81	17.00	3.61
pH	**units**	n/a	n/a	3.88	6.395	6.765	6.67	7.09	7.87	0.65
ORP	**mV**	n/a	n/a	-133.2	50.7	68.7	91.9	117.0	391.0	79.5

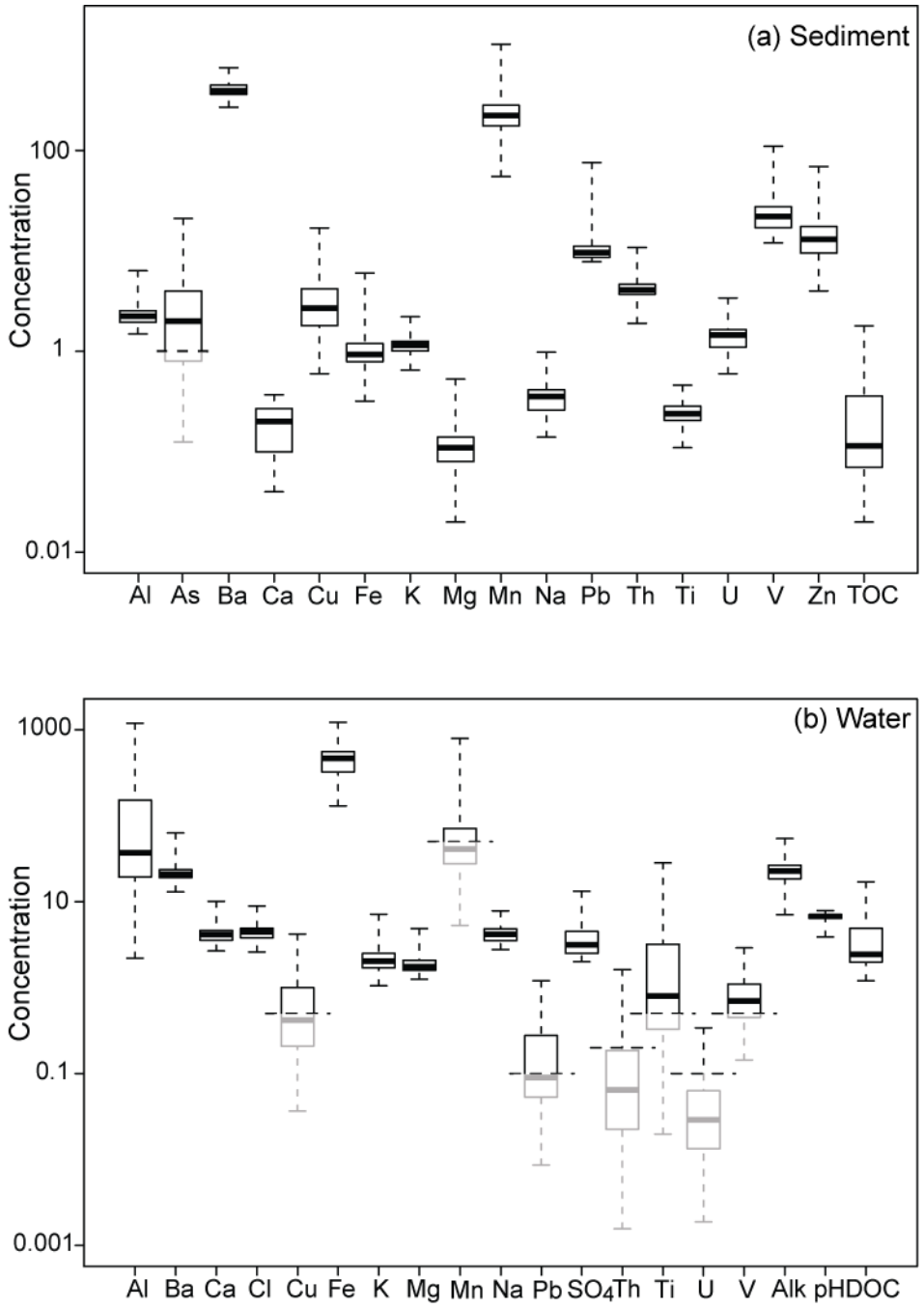

Figure 2-9: Boxplots showing concentrations of geochemical parameters in (a) stream sediment (n=40) and (b) stream water (n=98) during this study. Units for measurements, which include (a) mg/kg, % and (b) mg/L and µg/L, correspond to those in (a) Table 2 and (b) Table 3.Whiskers extend to maximum and minimum values, whereas the inner box represents, from bottom to top, the first quartile, median, and third quartile. For elements with non-detect data (e.g. As in stream sediment), the dashed horizontal line extending beyond the width of the box and whiskers represents the detection limit; data shown in gray below that line are simulated using ROS.

2.4 DISCUSSION

2.4.1 SAMPLING PROTOCOL AND ACCOUNTING FOR VARIABILITY

For stream water, in order to most efficiently and cost-effectively conduct both spatial and temporal comparisons, a number of sites were identified for sampling twice: once during a high-flow season (early spring), and once during base-flow conditions (late summer), coinciding with collection of stream sediment, whereas a subset of sites was sampled every month. By sampling both the spatially diverse (semiannual) sites and the temporally diverse (monthly) sites, it was feasible to collect samples to assess the effect of both factors on surface water constituents (Figs. 5 and 6). Due to weather variability, it was not possible to align sampling events with optimal peak flow and base flow periods. However, the sampling of a subset of sites monthly during a range of weather and flow conditions enabled analysis of surface water quality to encompass meteorological and subsequent hydrological variations (Figure 2-6). These factors are associated with significant differences in the concentrations of many relevant parameters in this study, including those commonly associated with uranium deposits, such as Pb and Th concentrations, which showed variation with temperature, and U and V concentrations, which showed variation with flow. A factor commonly adding variability to the concentration-flow relationship is runoff from spring snowmelt (e.g. Nagorski et al., 2003; Shafer et al., 1997). Due to the climate of southern Virginia and the mildness of the winter during this study, this effect was not observed.

In addition to constituents with regulatory guidelines, parameters with the potential to impact element mobility and toxicity should also be included in geochemical baseline studies. These parameters are relevant throughout a mining lifecycle as, for instance, in the development of discharge levels and design of treatment systems. Such parameters include pH and redox

potential, which can control element charge, speciation, sorption, and stability (Langmuir, 1997); organic and inorganic carbon (i.e. DOC and alkalinity in water or TOC and carbonate in solids), which can complex with trace elements or compete for exchange sites to impact solubility (Benedetti et al., 1995; Reuter and Perdue, 1977; Shafer et al., 1997); concentrations of elements in suspended particulate or dissolved components of water; and mineralogy and grain size of solid phases, which factor into dissolution/precipitation and sorption processes (Langmuir, 1997). At a site such as Coles Hill, where pH- and redox-sensitive trace metals (e.g. U, V, Mn) are of interest, the inclusion of pH and redox potential are of particular importance. The grain sizes present can also be relevant, as trace elements are often associated with smaller particles; some of these particles may be lost during wet sieving in the field (Horowitz and Elrick, 1987). Discussion of processes affecting element transport at Coles Hill will be addressed in a later study.

Concentrations of selected constituents by location relative to the deposit suggested that the deposit locally influences the geochemistry of surrounding environmental media (Table 2-2). Stream sediments and stream water integrate signatures from throughout their catchments. Ideally, sampling sites would be identified probabilistically (Dobbie et al., 2008), but access issues commonly make random site selection impractical. The number and density of sampling sites is dependent on site-specific factors, including the variability of rock types and the coverage of stream reaches in the vicinity of the deposit (Figure 2-1).

In addition to spatial considerations, the residence times for different media require different temporal resolution. The residence time of surface water in the vicinity of the deposit is short relative to that of stream sediments. As a result, sampling frequency was designed to capture as much of the medium's variability as feasible. In low-gradient, low-flow streams, such

as those sampled in this study, a single sampling of stream sediments may be sufficient, as this study did not reveal significant differences in stream sediment composition between sampling events during different flow regimes, but other sites, including streams with marked seasonal differences in flow patterns, may be better served by sampling during high and base flows to capture a range of transport, depositional, and reactive conditions (Davide et al., 2003; Geesey et al., 1984; Jain and Sharma, 2001). Surface waters have the most potential for geochemical variation, particularly from changes in flow and temperature, which can influence partitioning between solid and sorbed phases, including onto biological materials subject to temperature influence (Nimick et al., 2003). Geochemical variation can occur at many temporal scales, including diel cycles from photo- and bio-chemical reactions (Fuller and Davis, 1989; Nimick et al., 2003) and seasonal changes, as seen in this study, with temperature and runoff and their attendant variations in storage, biological activity, and transport processes, which can also impact sediment trace element concentrations (Brick and Moore, 1996). Capturing as much of this variation as possible is essential to a thorough baseline study, and thus multiple sampling events were incorporated to properly assess the dynamic character of water chemistry (Figure 2-6). Monthly sampling was used in this study, but the correlation of discharge with geochemical results, a trend observed both in the present study and in numerous previous works (e.g. Godsey et al., 2009; Nordstrom, 2011), suggests a regular sampling schedule is less important than a sampling strategy aimed at capturing climatic/hydrologic variability.

The duration of a baseline study is determined by the schedule of the development project. Some regulations (e.g. US NRC, 1980) specify a minimum duration for a baseline study. However, the collection of more baseline data means that more variation can be accounted for. For example, this study was conducted during a year with a warmer than average winter with

little snowfall. An extension of the study to include a less mild winter has the potential to yield different results. In addition, a baseline study should transition into a monitoring program once the site becomes operational.

2.4.2 EFFECT OF INCLUDING NON-DETECT METHODS IN STATISTICAL ANALYSIS

Baseline geochemical surveys have the potential to generate large datasets. In this study, one challenge to using common methods of statistical analysis was the presence of non-detect data. Some jurisdictions regulate contaminants on the basis of load, and a process-oriented understanding of a site can benefit from a rigorous material balance of various reservoirs. Both, especially when concentrations are near detection limits, are impaired when many samples are below detection limits. We examined a number of different approaches for evaluating non-detect data, and the most effective approach can depend on the data size, distribution, and proportion of censored values. Handling censored data is of particular relevance in this study because concentrations of many of the constituents associated with uranium deposits, particularly uranium and its progeny, were below detection limits in a majority of samples. In using censored data methods, datasets can be summarized without discarding or substituting these points. However, a greater number of non-detects leads to greater uncertainty in predicted values (Figure 2-8). When discrete predicted values are used, as in substitution or imputation methods, specific numbers are assigned to specific samples; the assignment of these values can be misinterpreted as actual instead of predicted information. For ROS, as used in this study, the uncertainty includes a higher reliance on the assumption of lognormally-distributed data (Helsel, 2012). When actual values are lognormally distributed, ROS provides a more accurate prediction of statistical parameters than when data are not lognormal. This caveat is particularly relevant in this study, as the concentrations of U and some progeny (e.g. Th) contain a majority of non-

detects. Because of this, ROS results for these elements (U and Th) should be regarded as approximate or "tenuous" (Helsel, 2012; Lee, 2010). If future measurements and/or analytical techniques with lower detection limits, such as high-resolution ICP-MS, preconcentration, or as-yet undeveloped methods, are used at these sites, the assumption of a lognormal distribution for censored data for highly censored parameters can be tested (Reimann et al., 2008).

Presenting non-detect data graphically is another concern when dealing with this type of data. Many common graphical methods commonly used in the scientific literature portray data as points or lines at specific values. When specific values are unknown, as in the case of non-detect data, these graphical methods may be altered (Helsel, 2012). The methods for altering graphics are, broadly, similar to those used in presenting non-detect data numerically. Methods include truncation of data plots to show only above-detection data, plotting substituted values (Figure 2-6), plotting below-detection data according to a distribution (Figure 2-3, Figure 2-8, Figure 2-9), and plotting below-detection data as ranges (Figure 2-5).

Among the methods previously used for characterizing baseline, some utilize mean and standard deviation as essential metrics and thus are based on an assumption of normally- or lognormally-distributed data. This assumption is not always valid (Reimann and Filzmoser, 2000), and actual data for many dissolved major and trace elements in this study are not normally or lognormally distributed; thus more robust methods such as non-parametric statistics are preferable. In this study, we used semi-parametric statistical methods due to the high proportions of non-detect data for some analytes. A number of baseline methods in previous studies do not have straightforward approaches to determining values for non-detect data. In some cases, this can result in reporting a value as less than a detection limit, but it can also result in propagation of error associated with simulating censored data. For example, substituting

values such as half of the detection limit can introduce false patterns into datasets (Helsel, 2006; Helsel, 2012), and computing statistics using only data above the detection limit omits a portion of the meaningful data (Helsel, 2012). From the perspective of trying to gain a more process-oriented understanding of baseline conditions, the next step after a study such as this one, knowledge of elemental fluxes or loads is essential. Calculation of loads necessarily requires an understanding of a range of concentrations, which are impossible to calculate without a quantification or estimation of the total range of concentrations.

2.4.3 GEOCHEMICAL BASELINE FOR ELEMENTS OF INTEREST IN SURFACE WATER AND STREAM SEDIMENT AT THE COLES HILL DEPOSIT

Baseline data are generally used for two purposes. One is as part of a regulatory program, and the other is for generating a rigorous understanding of the geochemical controls on the fluxes of constituents in natural settings. As a result, baseline is more complex than a single data mean. Gustavsson et al. (2012) suggested that baseline levels comprise a range and can best be presented by providing this range as well as a graphical illustration of data distributions. Using a range has value for monitoring during operations, whereas the upper extreme of the range can be considered, along with appropriate environmental standards, for determining reasonable closure goals. If the baseline study includes a well-designed sampling program, this range will encompass meteorological, hydrological, geological, and other variations within the study area. In this study, we found significant differences in element concentrations with the above-listed variations.

At the unmined Coles Hill uranium deposit, concentrations of U and other trace elements in stream sediments fall within typical ranges for these elements in stream sediments (Horowitz and Stephens, 2008) and do not exceed probable effects concentrations (PECs) of elements for which PECs exist [33 mg/kg As, 149 mg/kg Cu, 128 mg/kg Pb, 459 mg/kg Zn; MacDonald et al.

(2000)]; maximum concentrations exceed threshold effect concentrations (TECs) for As (maximum observed concentration of 21 mg/kg exceeded the TEC, 9.79 mg/kg). PECs are freshwater sediment concentration levels above which harmful effects to aquatic biota are likely to be observed, and TECs are levels below which harmful effects are unlikely to be observed (MacDonald et al., 2000).

Dissolved U concentrations in streams around Coles Hill are well below current US EPA drinking water standards (30 μg/L; (US EPA, 2009)) and Tier II screening benchmarks for both acute (33.5 μg/L) and chronic (1.87 μg/L) toxicity to freshwater biota (Suter, 1996). Concentrations of major and other trace elements in streams near the deposit are similar to typical values for surface waters (Langmuir, 1997) and low relative to primary drinking water (EPA maximum contaminant levels are 2000 μg/L Ba, 1300 μg/L Cu, 15 μg/L Pb, and 30 μg/L U) and mining discharge regulations set by the US EPA (40CFR440 Subpart C specifies a daily maximum effluent concentration of 4.0 mg/L for U) and the Commonwealth of Virginia (9VAC25-151-150 sets mining benchmark discharge requirements of 18 μg/L Cu, 1000 μg/L Fe, and 120 μg/L Pb). Project-specific discharge limits have not been set for development at Coles Hill.

2.5 CONCLUSIONS

A baseline study conducted for the purpose of pre-mining characterization must capture the full range of existing geochemical conditions, including any anomalous values from pre-existing anthropogenic activities or from natural geochemical anomalies that can be attributed to the ore deposit, particularly when these anomalies exceed regulatory standards. Use of standard statistical procedures that require exclusion of these values, as a means of achieving data distributions that accommodate those procedures, could mask data from existing pre-mining high

points, as occurs when using only a certain percentile range of the data, or lead to the acceptance of higher values than are actually present, as sometimes occurs with distribution-based baseline methods such as mean intervals and Tukey fences.

In this study, we present a baseline focused on establishing geochemical conditions in stream sediment and water prior to a proposed mining operation. We focused on baseline study design, including sampling site selection and frequency of sampling to encompass geological, hydrologic, and climatic variability. We also addressed issues of statistical analysis, including how to choose a method for handling non-detect data that is appropriate for the study goals and the dataset.

In our study, some elements of interest (U, Th, V) in surface water contained a high proportion of non-detect data. Through comparison of different methods for analyzing non-detect data, we found that regression on order statistics provided the most rigorous treatment of these data. Results of statistical analysis of our dataset show significant increases in concentrations of U and other constituents in stream sediments and water between stream reaches upstream and downstream of the Coles Hill uranium deposit. Relationships were observed between dissolved concentrations and stream discharge, underscoring the need for multiple sampling events in a variety of conditions in order to document the full geochemical variability of a site.

The methods described can be expanded to other types of resource extraction, such as hydraulic fracturing, changes in land use, or even more broadly, changes in climate.

ACKNOWLEDGMENTS

We thank the numerous field and lab assistants from Virginia Tech and the USGS. Virginia Uranium, Inc. provided partial financial support to Virginia Tech. VUI and Walter Coles Sr. allowed property access, and Stewart East provided valuable knowledge about the area. Larry Gough, J. Donald Rimstidt, Matt Leybourne, and an anonymous reviewer provided helpful comments on this manuscript. The use of product or trade names is for descriptive purposes only and does not constitute endorsement by the U.S. government.

REFERENCES

Aaron, L.G., 2009. Pittsylvania County, Virginia: A Brief History. The History Press, Charleston, SC.

Baccarelli, A., Pfeiffer, R., Consonni, D., Pesatori, A.C., Bonzini, M., Patterson Jr, D.G., Bertazzi, P.A., Landi, M.T., 2005. Handling of dioxin measurement data in the presence of non-detectable values: Overview of available methods and their application in the Seveso chloracne study. Chemosphere 60, 898-906.

Benedetti, M.F., Milne, C.J., Kinniburgh, D.G., Van Riemsdijk, W.H., Koopal, L.K., 1995. Metal ion binding to humic substances: Application of the non-ideal competitive adsorption model. Environmental Science & Technology 29, 446-457.

Brick, C.M., Moore, J.N., 1996. Diel Variation of Trace Metals in the Upper Clark Fork River, Montana. Environmental Science & Technology 30, 1953-1960.

Caron, M.-E., Grasby, S.E., Ryan, M.C., 2008. Spring water trace element geochemistry: A tool for resource assessment and reconnaissance mineral exploration. Applied Geochemistry 23, 3561-3578.

Carranza, E.J.M., 2011. Analysis and mapping of geochemical anomalies using logratio-transformed stream sediment data with censored values. Journal of Geochemical Exploration 110, 167-185.

Claridge, P.G., Downing, B.W., 1993. Environmental geology and geochemistry at the Windy Craggy massive sulphide deposit, northwestern British Columbia. CIM Bulletin 86, 50-57.

Davide, V., Pardos, M., Diserens, J., Ugazio, G., Thomas, R., Dominik, J., 2003. Characterisation of bed sediments and suspension of the river Po (Italy) during normal and high flow conditions. Water Research 37, 2847-2864.

Dobbie, M.J., Henderson, B.L., Stevens, J.D.L., 2008. Sparse sampling: Spatial design for monitoring stream networks. 113-153.

Eppinger, R.G., Briggs, P.H., Dusel-Bacon, C., Giles, S.A., Gough, L.P., Hammarstrom, J.M., Hubbard, B.E., 2007. Environmental geochemistry at Red Mountain, an unmined volcanogenic massive sulphide deposit in the Bonnifield district, Alaska Range, east-central Alaska. Geochemistry: Exploration, Environment, Analysis 7, 207-223.

Fiévet, B., Della Vedova, C., 2010. Dealing with non-detect values in time-series measurements of radionuclide concentration in the marine environment. Journal of Environmental Radioactivity 101, 1-7.

Fuller, C.C., Davis, J.A., 1989. Influence of coupling of sorption and photosynthetic processes on trace element cycles in natural waters. Nature 340, 52-54.

Gałuszka, A., 2007. A review of geochemical background concepts and an example using data from Poland. Environmental Geology 52, 861-870.
Gannon, J.P., Burbey, T.J., Bodnar, R.J., Aylor, J.G., 2011. Geophysical and geochemical characterization of the groundwater system and the role of Chatham Fault in groundwater movement at the Coles Hill uranium deposit, Virginia, USA. Hydrogeology Journal 20, 1-16.
Geesey, G.G., Borstad, L., Chapman, P.M., 1984. Influence of flow-related events on concentration and phase distribution of metals in the lower Fraser River and a small tributary stream in British Columbia, Canada. Water Research 18, 233-238.
Godsey, S.E., Kirchner, J.W., Clow, D.W., 2009. Concentration–discharge relationships reflect chemostatic characteristics of US catchments. Hydrological Processes 23, 1844-1864.
Gustavsson, N., Loukola-Ruskeeniemi, K., Tenhola, M., 2012. Evaluation of geochemical background levels around sulfide mines – A new statistical procedure with beanplots. Applied Geochemistry 27, 240-249.
Helsel, D.R., 2005. Nondetects and data analysis: Statistics for censored environmental data. Wiley-Interscience.
Helsel, D.R., 2006. Fabricating data: How substituting values for nondetects can ruin results, and what can be done about it. Chemosphere 65, 2434-2439.
Helsel, D.R., 2012. Statistics for Censored Environmental Data Using Minitab and R. Wiley.
Helsel, D.R., Gilliom, R.J., 1986. Estimation of distributional parameters for censored trace level water-quality data 2. Verification and applications. Water Resources Research 22, 147-155.
Henika, W.S., 2002. Geologic map of the Danville 30 by 60 minute quadrangle, Virginia, Publication 166.
Henika, W.S., Thayer, P.A., 1983. Geologic map of the Spring Garden quadrangle, Virginia, Publication 48.
Horowitz, A.J., Elrick, K.A., 1987. The relation of stream sediment surface area, grain size and composition to trace element chemistry. Applied Geochemistry 2, 437-451.
Horowitz, A.J., Stephens, V.C., 2008. The effects of land use on fluvial sediment chemistry for the conterminous U.S. — Results from the first cycle of the NAWQA Program: Trace and major elements, phosphorus, carbon, and sulfur. Science of the Total Environment 400, 290-314.
Hron, K., Templ, M., Filzmoser, P., 2010. Imputation of missing values for compositional data using classical and robust methods. Computational Statistics & Data Analysis 54, 3095-3107.
Jain, C.K., Sharma, M.K., 2001. Distribution of trace metals in the Hindon River system, India. Journal of Hydrology 253, 81-90.
Jerden, J.L., Sinha, A.K., 2003. Phosphate based immobilization of uranium in an oxidizing bedrock aquifer. Applied Geochemistry 18, 823-843.
Kelley, K.D., Taylor, C.D., 1997. Environmental geochemistry of shale-hosted AgPbZn massive sulfide deposits in northwest Alaska: Natural background concentrations of metals in water from mineralized areas. Applied Geochemistry 12, 397-409.
Khalil, B., Ouarda, T.B.M.J., 2009. Statistical approaches used to assess and redesign surface water-quality-monitoring networks. Journal of Environmental Monitoring 11, 1915-1929.
Langmuir, D., 1997. Aqueous environmental geochemistry. Prentice Hall, Upper Saddle River, N.J.

Lee, L., 2010. NADA: Nondetects And Data Analysis for environmental data, 1.5-3 ed, R package.
Lee, L., Helsel, D., 2005. Baseline models of trace elements in major aquifers of the United States. Applied Geochemistry 20, 1560-1570.
Leybourne, M.I., 1998. Hydrogeochemistry of ground and surface waters associated with massive sulphide deposits, Bathurst Mining Camp, New Brunswick: Halfmile Lake and Restigouche deposits,PhD Thesis.
Leybourne, M.I., Cameron, E.M., 2010. Groundwater in geochemical exploration. Geochemistry: Exploration, Environment, Analysis 10, 99-118.
Leybourne, M.I., Clark, I.D., Goodfellow, W.D., 2006. Stable isotope geochemistry of ground and surface waters associated with undisturbed massive sulfide deposits; constraints on origin of waters and water–rock reactions. Chemical Geology 231, 300-325.
Leybourne, M.I., Goodfellow, W.D., Boyle, D.R., 1998. Hydrogeochemical, isotopic, and rare earth element evidence for contrasting water–rock interactions at two undisturbed Zn–Pb massive sulphide deposits, Bathurst Mining Camp, N.B., Canada. Journal of Geochemical Exploration 64, 237-261.
Leybourne, M.I., Goodfellow, W.D., Boyle, D.R., 2002. Sulphide oxidation and groundwater transport of base metals at the Halfmile Lake and Restigouche Zn–Pb massive sulphide deposits, Bathurst Mining Camp, New Brunswick. Geochemistry: Exploration, Environment, Analysis 2, 37-44.
Lyntek Inc., 2010. Preliminary economic assessment on the Coles Hill Uranium Property, Lakewood, CO.
MacDonald, D.D., Ingersoll, C.G., Berger, T.A., 2000. Development and evaluation of consensus-based sediment quality guidelines for freshwater ecosystems. Archives of Environmental Contamination and Toxicology 39, 20-31.
Marline Uranium Corporation, 1983. An evaluation of uranium development in Pittsylvania County, Virginia. Marline Uranium Corporation, USA.
Martín-Fernández, J.A., Barceló-Vidal, C., Pawlowsky-Glahn, V., 2003. Dealing with zeros and missing values in compositional data sets using nonparametric imputation. Mathematical Geology 35, 253-278.
Martín-Fernández, J.A., Hron, K., Templ, M., Filzmoser, P., Palarea-Albaladejo, J., 2012. Model-based replacement of rounded zeros in compositional data: Classical and robust approaches. Computational Statistics & Data Analysis 56, 2688-2704.
Martin-Fernandez, J.A., Palarea-Albaladejo, J., Olea, R.A., 2011. Dealing with zeros, in: Pawlowsky-Glahn, V., Buccianti, A. (Eds.), Compositional Data Analysis. Wiley, pp. 43-58.
Matschullat, J., Ottenstein, R., Reimann, C., 2000. Geochemical background – can we calculate it? Environmental Geology 39, 990-1000.
Meyertons, C.T., 1963. Triassic formations of the Danville basin, Resources, V.D.o.M., Report of Investigations 6.
Nagorski, S.A., Moore, J.N., McKinnon, T.E., Smith, D.B., 2003. Geochemical response to variable streamflow conditions in contaminated and uncontaminated streams. Water Resources Research 39, 1044.
NCDC, 2012. National Climatic Data Center. National Oceanic and Atmospheric Administration, Asheville, NC.

Nimick, D.A., Gammons, C.H., Cleasby, T.E., Madison, J.P., Skaar, D., Brick, C.M., 2003. Diel cycles in dissolved metal concentrations in streams: Occurrence and possible causes. Water Resources Research 39, 1247.
Nordstrom, D.K., 2011. Hydrogeochemical processes governing the origin, transport and fate of major and trace elements from mine wastes and mineralized rock to surface waters. Applied Geochemistry 26, 1777-1791.
Pebble Partnership, 2011. Pebble Project Environmental Baseline Document, http://www.pebbleresearch.com/ebd/.
Pincock Allen & Holt, 1982. Geologic reserves: Coles Hill South Uranium Deposit, Pittsylvania County, Virginia. Pincock Allen & Holt, Tucson, Arizona.
R Development Core Team, 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, Software Program.
Reimann, C., Filzmoser, P., 2000. Normal and lognormal data distribution in geochemistry: death of a myth. Consequences for the statistical treatment of geochemical and environmental data. Environmental Geology 39, 1001-1014.
Reimann, C., Filzmoser, P., Garrett, R.G., Dutter, R., 2008. Statistical data analysis explained : Applied environmental statistics with R. John Wiley & Sons, Chichester, England; Hoboken, NJ.
Reimann, C., Garrett, R.G., 2005. Geochemical background—concept and reality. Science of the Total Environment 350, 12-27.
Reuter, J.H., Perdue, E.M., 1977. Importance of heavy metal-organic matter interactions in natural waters. Geochimica Et Cosmochimica Acta 41, 325-334.
Rodrigues, A.S.L., Nalini Júnior, H.A., 2009. Geochemical background values and its implications in environmental studies. Rem: Revista Escola de Minas 62, 155-165.
Rounds, S.A., 2006. Alkalinity and acid neutralizing capacity, Techniques of WaterResources Investigations Book 9, Chapter A6.6.
Salminen, R., Gregorauskiene, V., 2000. Considerations regarding the definition of a geochemical baseline of elements in the surficial materials in areas differing in basic geology. Applied Geochemistry 15, 647-653.
Salminen, R., Tarvainen, T., 1997. The problem of defining geochemical baselines. A case study of selected elements and geological materials in Finland. Journal of Geochemical Exploration 60, 91-98.
Shafer, M.M., Overdier, J.T., Hurley, J.P., Armstrong, D., Webb, D., 1997. The influence of dissolved organic carbon, suspended particulates, and hydrology on the concentration, partitioning and variability of trace metals in two contrasting Wisconsin watersheds (U.S.A.). Chemical Geology 136, 71-97.
Shelton, L.R., Capel, P.D., 1994. Guidelines for collecting and processing samples of stream bed sediment for analysis of trace elements and organic contaminants for the National Water-Quality Assessment Program, Open-File Report 94-458.
Strobl, R.O., Robillard, P.D., Shannon, R.D., Day, R.L., McDonnell, A.J., 2006. A Water Quality Monitoring Network Design Methodology for the Selection of Critical Sampling Points: Part I. Environmental Monitoring and Assessment 112, 137-158.
Suter, G.W., 1996. Toxicological benchmarks for screening contaminants of potential concern for effects on freshwater biota. Environmental toxicology and chemistry 15, 1232-1241.
Taggart, J., 2002. Analytical methods for chemical analysis of geologic and other materials. US Geological Survey Open-File Report 2, 223.

Templ, M., Hron, K., Filzmoser, P., 2013. robCompositions: Robust estimation for compositional data, R package.
Tidball, R.R., Ebens, R.J., 1976. Regional geochemical baselines in soils of the Powder River Basin, Montana-Wyoming, in: Laudon, R.B. (Ed.), Geology and Energy Resources of the Powder River, 28th Annual Field Conference Guidebook. Wyoming Geological Association, pp. 299-310.
Tidball, R.R., Erdman, J.A., R.J., E., 1974. Geochemical baselines for sagebrush and soil. Powder River Basin, Montana-Wyoming, Open-File Report 74-250.
US EPA, 2003. EPA and hardrock mining: A source book for industry in the Northwest and Alaska,
US EPA, 2009. National Primary Drinking Water Regulations, EPA 816-F-09-004.
US NRC, 1980. Radiological Effluent and Environmental Monitoring at Uranium Mills, Regulatory Guide 4.14.
USGS, 2006. Collection of water samples, U.S. Geological Survey Techniques of Water-Resources Investigations, book 9.
Wyatt, J., 2009. The relationship between structural and tectonic evolution and mineralization at the Coles Hill Uranium Deposit, Pittsylvania County, Virginia,Department of Geosciences. M.S. Thesis.
Zgłobicki, W., Lata, L., Plak, A., Reszka, M., 2011. Geochemical and statistical approach to evaluate background concentrations of Cd, Cu, Pb and Zn (case study: Eastern Poland). Environmental Earth Sciences 62, 347-355.

CHAPTER 3. STATISTICAL ANALYSIS OF SOIL GEOCHEMICAL DATA TO IDENTIFY PATHFINDERS ASSOCIATED WITH MINERAL DEPOSITS

Denise M. Levitan, Carl E. Zipper, Patricia Donovan, Madeline E. Schreiber, Robert R. Seal II, Mark A. Engle, John A. Chermak, Robert J. Bodnar, Daniel K. Johnson, and Joseph G. Aylor Jr.

A version of this chapter was submitted to the Journal of Geochemical Exploration and is undergoing revisions

ABSTRACT

Soil geochemical anomalies can be used to identify pathfinders for application in exploration for ore deposits. Soil geochemical exploration programs for U are confounded because U deposits can be classified into at least 15 distinct genetic types. In this study, we use compositional data analysis with multivariate statistical methods to analyze soil geochemistry data collected from the Coles Hill uranium deposit, Virginia, USA, to identify pathfinders associated with this deposit. Elemental compositions and relationships were compared between the Coles Hill soils and reference soils extracted from a regional subset of a national-scale geochemical survey. Results show that pathfinders for the Coles Hill deposit include light rare earth elements (La and Ce), which when normalized by their Al content are correlated with U/Al, and elevated Th/Al values, which are not correlated with U/Al. These results can be used in genetic and weathering models of the Coles Hill deposit and can also be applied to future prospecting for similar U deposits in the eastern United States and in regions with similar geologic/climatic conditions.

3.1 INTRODUCTION

The soils and weathered materials overlying ore deposits can serve as a component of geochemical exploration programs to discover new deposits (Grunsky, 2010). Studies that use soil geochemistry in exploration rely on anomalous concentrations and patterns of pathfinder elements associated with the deposit type of interest (Bradshaw and Lett, 1980). While the term "pathfinder" is generally used to describe concentrations of one or more elements other than the economic one(s), here we expand the use to include ratios between elements, as are commonly used in the analysis of compositional data. Pathfinders are identified on the basis of the underlying geological, mineralogical, and geochemical attributes of an ore deposit type. The pathfinders should have a wider distribution in the immediate vicinity of the deposit than the elements of economic interest in order to define a target that can be refined with more expensive exploration techniques such as drilling. In some locations, the element of economic importance in the deposit may not be a good pathfinder in soil because it may be mobile in the weathering environment and therefore not indicative of enrichment in underlying unweathered rock. As such, Snelling (1984) suggested that U may be a poor element to use in exploration for U deposits because it can be easily mobilized by near-surface oxidized waters, leading to low concentrations in soils and other weathering products.

Uranium deposits are among the most diverse types of mineral deposit on Earth (Cuney, 2009). In fact, economic geologists classify them into at least 15 geologically distinct genetic types and 40 subtypes (Fayek, 2013). Economic occurrences of U are found in sedimentary, metamorphic, and igneous rocks and in rocks ranging in age from Archean to Quaternary (Cuney, 2009; Dahlkamp, 1993; IAEA, 2009). The wide range in geologic environments and time periods during which U deposits have formed leads to a highly variable and often diagnostic

geochemical signature for the different U deposit types. Geochemical anomalies are often associated with uranium deposits and have been documented in a variety of media, including altered rock associated with mineralization, sediments, soils, and stream, lake, and groundwater. For example, Power et al. (2012) examined samples of soils and siliciclastic units overlying the Phoenix U deposit in the Athabasca Basin, Canada, and found U, Mo, Co, Ag, and W anomalies in humus, B-horizon, and uppermost sandstone units overlying ore zones. In contract, sedimentary U deposits, including sandstone-hosted (roll-front) U deposits of mostly Phanerozoic age, are characterized by anomalous V, Mo, Se, As, and locally Cu and Ag (Cox and Singer, 1986; Rose and Wright, 1980). Unconformity-related U deposits, such as those that occur in the Athabasca Basin of Canada, are associated with anomalous Ni, Co, As, and Pb, with B and Cu also serving as pathfinders (Jefferson et al., 2007). As these examples show, U deposits are characterized by a widely variable suite of pathfinders, depending on the geological environment of formation, among other factors. Therefore, the use of soil geochemistry to prospect for new U deposits or to refine the location of a deposit discovered by other means, such as radiometric or geophysical surveying, requires that pathfinders specific to that deposit type be identified. Soil geochemistry largely reflects the underlying geology and climate in which weathering is taking place; however, anthropogenic overprinting is also seen on a variety of scales (Cohen et al., 2012).

Here, we describe soil geochemistry using data from transects over the undeveloped Coles Hill U deposit, Virginia, USA, that have been analyzed using multivariate statistical methods. Study goals are to identify anomalous distributions of elements in soils and to identify elements related to U at Coles Hill. We achieve these goals while demonstrating the use of large-scale geochemical survey data in a site-specific study. We expect these data will prove to be

valuable not only in exploration for similar types of deposits and for development of post-mining reclamation, but also contribute to development of a genetic model for the Coles Hill deposit.

3.1.1 THE COLES HILL URANIUM DEPOSIT

The Coles Hill U deposit, the largest undeveloped U deposit in the USA (Lyntek Inc., 2010; Pincock Allen & Holt, 1982), is located in Pittsylvania County, Virginia (Figure 3-1). The deposit was discovered in 1979 by Marline Uranium Corporation as one of several radiometric anomalies along the margin of the Danville-Dan River Triassic Basin (Jerden and Sinha, 2003). The deposit consists of the North and South ore bodies (Figure 3-1), which combined comprise an estimated 54 million kilograms (119 million pounds) of U_3O_8 with a cutoff grade of 0.025 wt.% and an average grade of 0.06 wt.% (Lyntek Inc., 2010; Marline Uranium Corporation, 1983). The ore bodies are exposed at the surface in minor (few-meter scale) outcrops. The North ore body dips slightly to the north, whereas the South ore body plunges to the south (Lyntek Inc., 2010). The host rock for the ore bodies is a mostly mylonitized quartzo-feldspathic granite gneiss interlayered with amphibolites. The sulfide content of the mineralized rock is minimal.

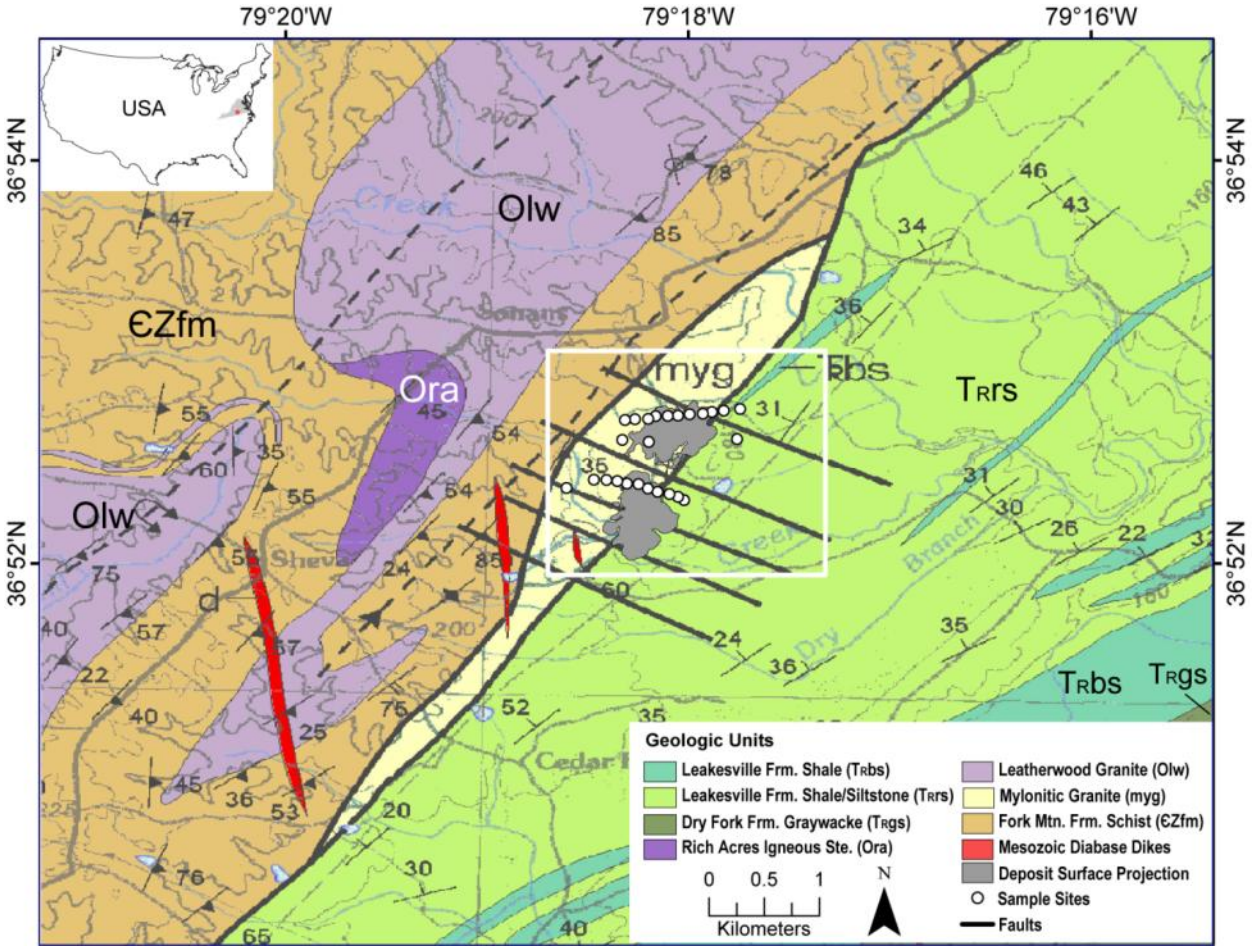

Figure 3-1: Geologic map of the Coles Hill region modified from Henika (2002). A surface projection of the U mineralization, at a cutoff grade of 0.025 wt. %, is shown in grey. Coles Hill sampling sites are shown as white circles. White box indicates extent of soil sampling sites map (Figure 3-2).

The deposit is located in the footwall of the Chatham Fault, a normal fault trending northeast-southwest and dipping at approximately 60° to the southeast (Gannon et al., 2011; Henika, 2002; Meyertons, 1963). The Chatham Fault divides the Virginia Piedmont Province to the west from the Danville Triassic Basin to the east. Both of these physiographic subregions occur within the Piedmont ecoregion of the eastern USA. The Piedmont Province in the vicinity of the deposit consists mainly of the Fork Mountain Formation (muscovite-biotite schist and garnetiferous biotite gneiss) and the Martinsville Igneous Complex, including the mylonitized

Leatherwood Granite in which the ore occurs. The Danville Triassic Basin to the east of the deposit includes the Leakesville Formation, consisting of shale and siltstone with lesser amounts of sandstone and claystone (Gannon et al., 2011; Henika, 2002; Henika and Thayer, 1983).

The Coles Hill deposit exhibits characteristics of several different U deposit types. The Coles Hill deposit is an epigenetic deposit in which U mineralization in the form of pitchblende, coffinite and U-bearing apatite occur in veins within a deformed meta-igneous host rock. Recent age dating (Tappa et al., 2014) indicates that mineralized orthogneiss and mineralized amphibolite are 419 ± 19 and 426 ± 21 Ma, respectively. To date, a confirmed genetic model for Coles Hill has yet to emerge, though some details of the mineralization and alteration history of the deposit are given in Jerden (2001). Dahlkamp (1993) included the Coles Hill deposit among the Vein Type (3) deposits, and a more recent summary by Dahlkamp (2010) states that "the Coles Hill/Swanson deposit is primarily controlled by structure and may be attributed to vein-type uranium deposits." More recently, Aylor et al. (2014) suggested two possible sources for the U. One model invokes remobilization of U from the Triassic basin by westward-flowing meteoric waters that intersected the Chatham Fault at depth and migrated upward and laterally into the Leatherwood Granite. A second U source model involves U mineralization by hydrothermal fluids associated with emplacement and cooling of the Leatherwood Granite. These fluids would have migrated upwards into a structural trap where U minerals were precipitated in hydrothermal fractures. Neither of these models has been tested owing to lack of detailed geochemical and geological data for the Coles Hill deposit.

The near surface in the vicinity of the Coles Hill deposit is characterized by a 15-20 m thick soil/saprolite zone. In the vicinity of the deposit, the water table (0 to 20 m below ground surface) and redox boundary occur within the saprolite/soil (Jerden et al., 2003). Two primary

soil associations predominate in the area surrounding the deposit (Figure 2; NRCS, 2011). The Madison-Cecil soil association, produced as a result of weathering of Western Piedmont province rock, is present to the northwest of the deposit; the Appling, Cecil, Cullen, Helena, and Madison soil series that occur in this association are generally deeply weathered, well-drained (the Helena series is moderately well-drained), and moderately permeable. The Pinkston-Mayodan-Creedmoor soil association, weathered from Triassic basin sedimentary rocks, is present to the southeast; the Mayodan, Creedmoor, and Sheva soil series that occur within this association are deeply weathered and moderately well to well-drained, but are less permeable than Madison-Cecil soils. A band of alluvial soils also extends through the study area, comprising the Chenneby, Mattaponi, and Toccoa soil series, which are somewhat poorly drained (Chenneby) to moderately well-drained (Mattaponi and Toccoa). Glacial sediments are not present in this region. Soils in this area are generally thick (>160 cm) and loamy, with most soil series over and adjacent to the deposit classified as sandy loams.

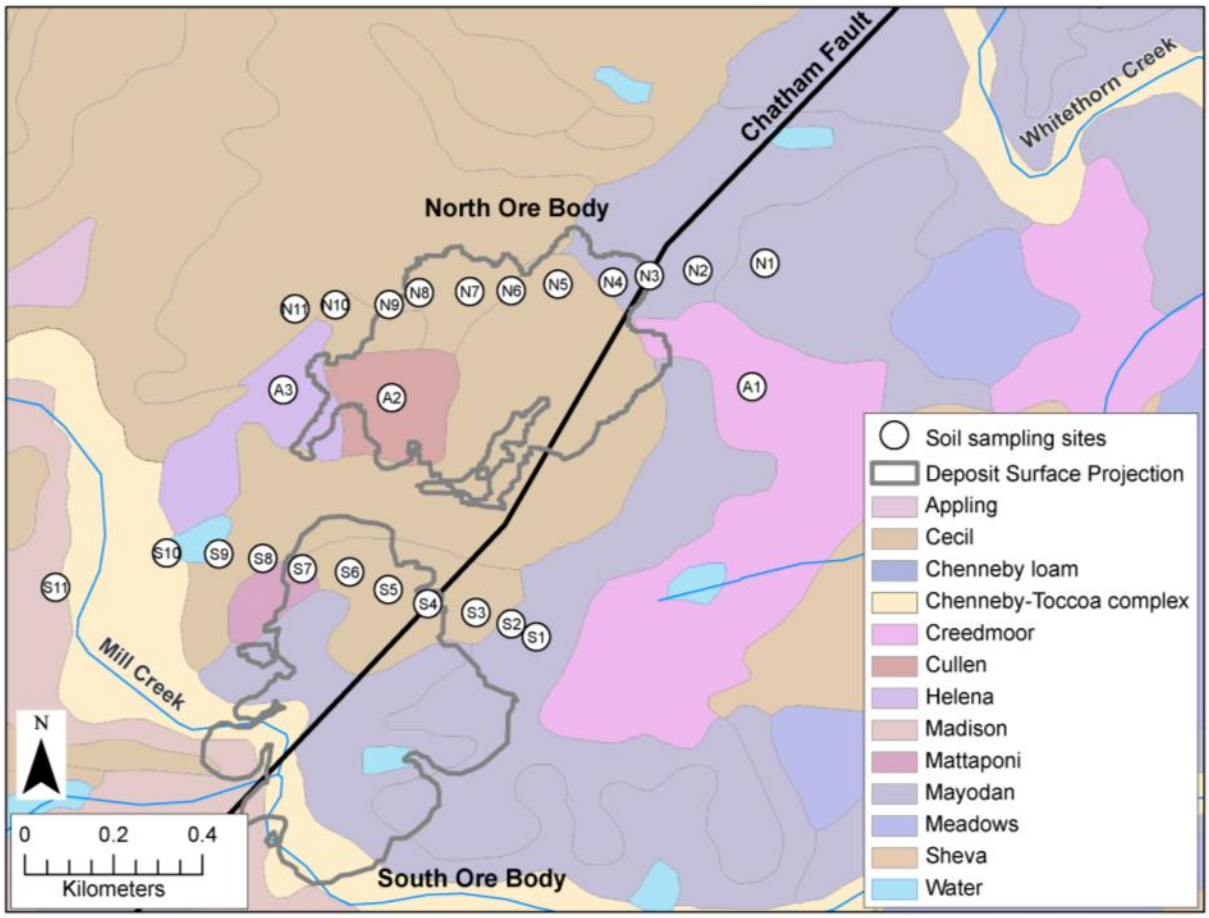

Figure 3-2: Soil sampling sites. White circles indicate sampling sites. Outline of surface projection of the deposit is shown. Background colors indicate soil series designated by NRCS (2011). Area of map extent relative to Figure 3-1 is shown in Figure 3-1.

In the weathered bedrock overlying the deposit and in the saprolite between the bedrock and soil, barium meta-autunite [meta-uranocircite, $Ba(UO_2)_2(PO_4)_2$•6-8H_2O] has formed in fractures and as rims on apatite grains in minor amounts. The formation of this mineral is believed to immobilize the oxidized U, resulting in surrounding groundwater U concentrations lower than expected (13.9 µg/L and less) in proximity to the deposit (Jerden and Sinha, 2003; Jerden et al., 2003). Uranium concentrations in surface waters draining the site are also low, ranging from below detection (0.1 µg/L) to a maximum of 0.34 µg/L in ephemeral streams

draining the area of the North ore body (Levitan et al., 2014). Other secondary minerals in the saturated saprolite include kaolinite and Mn oxides (Jerden et al., 2003). In the unsaturated saprolite and in the soils overlying the deposit, U is associated with Fe oxy-hydroxides (mainly through sorption), found in the structure of Al-phosphates of the crandallite group (Jerden and Sinha, 2006), and incorporated in light rare-earth element (LREE)-bearing phosphates such as rhabdophane (Jerden, 2001). Other secondary minerals in the unsaturated zone overlying the deposit include vermiculite, montmorillonite, and gibbsite.

The deposit is located in a temperate humid climate. Average annual temperature is 13°C (55°F) with a minimum in January (average low -5°C or 23°F) and a maximum in July (average high 31°C or 87°F). Annual rainfall is approximately 1150 mm (45 inches) and is distributed fairly evenly throughout the year (NCDC, 2012). Historical land use in the area was for tobacco and cotton farming, and cattle grazing is currently the predominant use of land (Aaron, 2009). During historic crop cultivation, various amendments were applied to the land surface, including fertilizers and lead arsenate pesticides.

3.2 METHODS

3.2.1 SITE SELECTION

Twenty-five soil sampling sites (Figure 3-2) were selected along east-west transects, which were based on locations of electrical resistivity profiles conducted by Gannon et al. (2011). Two transects each comprised of eleven sites were sampled across each ore body, with a sampling distance lag of ~100m. Samples from an additional three sites between the two transects were selected to encompass soil types not found along the transects. Whereas nearly three-quarters of the soil sampling sites are underlain by mylonitized granite, six different soil

series are recognized (Figure 3-2). At each sampling site, soil samples were collected from both the A and C horizons, yielding a total of 50 samples for analysis.

3.2.2 SAMPLE COLLECTION AND ANALYSIS

Procedures for collection and analysis of soil samples followed U.S. Geological Survey (USGS) methods for determining trace-metal content (Smith et al., 2005). Soil samples were collected using hand-driven steel augers to a depth of 160 cm or auger refusal. Extracted soils were laid on plastic sheeting, and horizons were identified by changes in soil color, texture, and structure. In the event that a C horizon was not identified in the top 160 cm of soil, the bottom 25 cm of the collected soil (generally B horizon) was sampled. To reduce cross contamination between holes, augers were brushed clean after each sample, and a shallow test hole was augered at each site prior to primary sampling, with the resulting soil discarded. After the A and C horizons were identified, subsamples of soil were collected from each horizon, bagged, and labeled on site. In the laboratory, subsamples were air-dried, then sieved to less than 2 mm using a stainless steel sieve under laboratory hoods. Samples were stored in low-density polyethylene bags.

Major and trace element and carbon analyses were completed by SGS Laboratories in Toronto, Canada, under a contract with the USGS Analytical Chemistry Services Group in Denver, Colorado, USA. Forty-two major and trace elements were measured using inductively-coupled plasma mass spectrometry (ICP-MS) and inductively-coupled plasma atomic emission spectroscopy (ICP-AES) following a multi-acid (HCl-HNO_3-$HClO_4$-HF) digestion. Carbon was measured by a carbon analyzer. Radionuclides were analyzed by Pace Analytical Services in Greensburg, PA. Methods include gamma spectroscopy (EPA Method 901.1 for Ra-226 and -

228), extraction chromatography (RP280m DOE method for Pb-210), and alpha spectrometry (HSL-300m for Th and U species). Quality assurance/quality control (QA/QC) samples consisted of field duplicates (samples collected from adjacent auger holes) and lab duplicates (samples divided following drying and sieving), as well as laboratory internal QA/QC including standard samples. A 20% relative difference threshold was used between duplicate samples for elements for which the analytical precision allowed for relative differences of 20% or less (i.e., if a primary sample measures 0.02 and the duplicate 0.03, these measurements were not included in the QA/QC analysis). Overall, fewer than 10% of elements in five of six duplicate sets failed to meet the threshold, and therefore the dataset was considered of acceptable quality. A comparison of field and lab duplicates suggested that variability was attributable to both field and laboratory procedures.

Mineralogy was determined by X-ray diffraction (XRD). Back-loaded powder mounts were run using CuKα radiation over a range of 3 to 80 °2θ at 0.035 °2θ/s. Minerals were identified using PANalytical X'Pert HighScore Plus v.3.0 software loaded with patterns from the International Centre for Diffraction Data (ICDD) PDF-2.

3.2.3 SELECTION OF REFERENCE SOIL DATA

In 2013, the USGS released data for a national soil geochemical and mineralogical survey designed to document the abundance and spatial distribution of elements and minerals in the soils of the conterminous United States (Smith et al., 2013). National soil survey sampling sites were selected using a spatially randomized design intended to represent 1600 km^2 areas (Smith et al., 2013). Analyses of data from pilot transects of North America (Smith et al., 2005) have been the subject of a number of studies, many of which were included in a special issue of *Applied*

Geochemistry (Vol. 24, Issue 8, 2009), as USGS reports (Smith et al., 2005), and in other publications (Drew et al., 2010).

We used the publicly available data from the USGS national soils survey as a reference for determining anomalous elemental concentrations in soils from the Coles Hill U deposit area. Soils in the present study were collected and analyzed using procedures identical to those of the national soil survey, sampling for which was completed a year prior to the present study. National soil survey samples (A and C horizons) collected at locations within 250 km of Coles Hill and within the Piedmont (45) EPA Level III ecoregion (US EPA, 2013) were selected as the reference dataset (n=48) in order to have similar climate, slope, and bedrock type to the Coles Hill samples, though the variations in these aspects for the reference samples are likely to be of greater magnitude than in the Coles Hill sample sites (Figure 3-3).

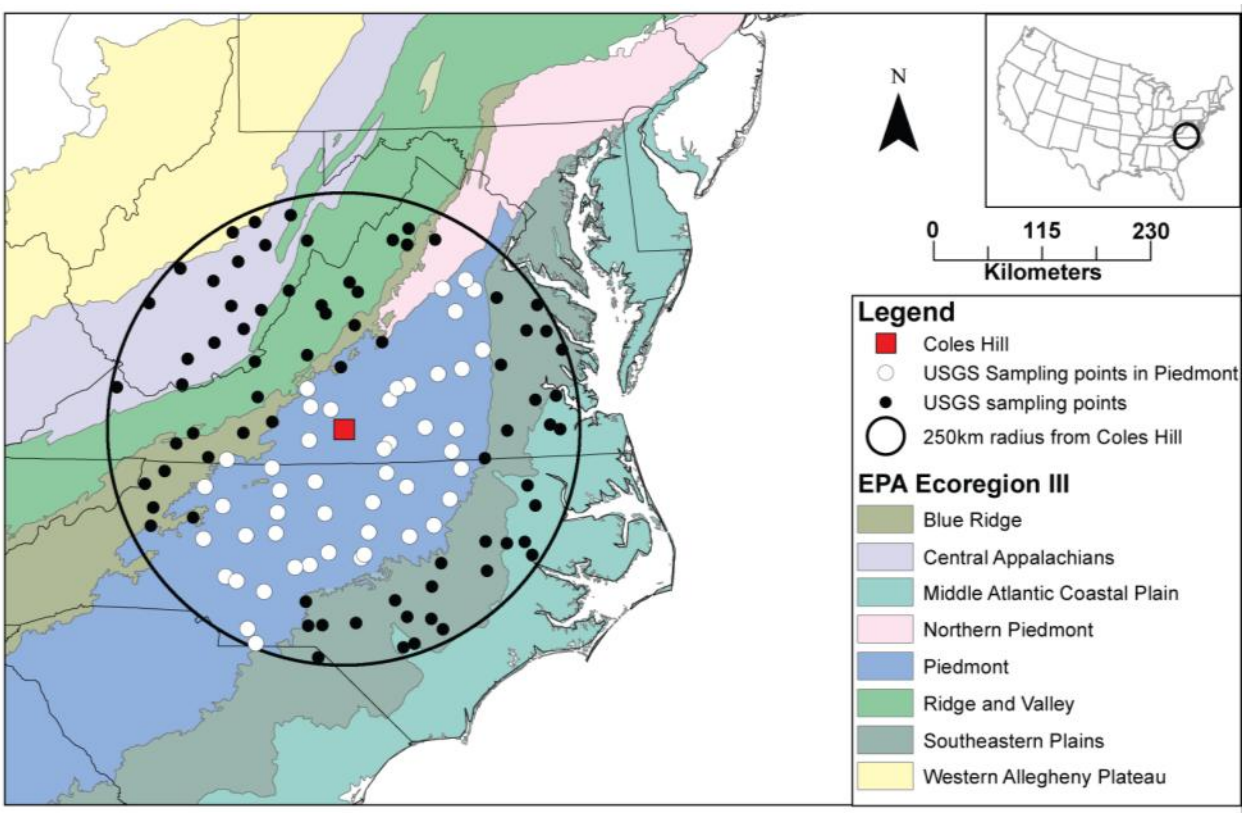

Figure 3-3: Selection of reference soil sites from soil sampling sites for USGS national soil survey (Smith et al., 2013). Coles Hill is indicated by a red box, and a black circle shows a 250 km radius around the center of Coles Hill. Map base colors represent US EPA (2013) Level III ecoregions, with Piedmont shown in blue. White dots represent show USGS sampling sites in the Piedmont ecoregion within 250 km of Coles Hill, this study's reference soils. Black dots represent USGS sampling sites within 250 km of Coles Hill in other ecoregions, which were not used in this study.

3.2.4 STATISTICAL ANALYSIS

Separate analyses were conducted for soil samples from the A and C horizons. These statistical analyses were done using R (R Development Core Team, 2011) and CoDaPack 2.0 (Comas-Cufí and Thió-Henestrosa, 2011), in many cases using code modified from Reimann et al. (2008). Many of the statistical methods used are discussed in Appendix C. Soil samples were classified based on horizon and origin (Coles Hill or reference). All elements present in both datasets, and with >50% of observations recorded with concentrations greater than detection

limits, and organic carbon (C_{org}), were analyzed. Thirty-six elements and C_{org} were considered. Non-detect observations were replaced using a classical parametric approach based on isometric log-ratio transformation (Martín-Fernández et al., 2012; Palarea-Albaladejo et al., 2013; Templ et al., 2013).

In the analysis of geochemical concentration data, the problem of closure arises (Aitchison, 1986; Pawlowsky-Glahn and Egozcue, 2006). Closed data are observations which are expressed as parts of a whole (e.g. mg/kg) and often sum to a constant constraining value (e.g. 10^6 mg/kg). In the analysis of compositional data, the issue of closure can lead to the misinterpretation of statistical results, such as spurious correlations. As a result, methods to "open" such data have been proposed (Aitchison, 1986; Egozcue et al., 2003). These methods have been used successfully in a number of geochemical studies (Buccianti et al., 2006; Carranza, 2011; Engle et al., 2011; Pawlowsky-Glahn and Buccianti, 2011; Zuo et al., 2013).

Because element concentrations in soils represent closed compositional data, concentration data were opened for data analysis by applying a family of log-ratio transformations (Aitchison, 1982). Up-to-date information on compositional data and its analysis are described in several books (Buccianti et al., 2006; Pawlowsky-Glahn and Buccianti, 2011). In our case, data were centered-log-ratio (clr) transformed for most analyses and additive-log-ratio transformed (alr) for analyses (i.e., correlation tests) where clr does not preserve the relevant aspects or structure of the original data. The clr transformation takes the natural log of an element concentration divided by the geometric mean of the other element concentrations for that sample. The alr transformation uses the natural log of an element concentration relative to a selected divisor. The use of an alr transformation requires the "sacrifice" of one variable to serve as the divisor in the transformation. An ideal divisor will have values well above detection and a

degree of precision that provides a range of values. Aluminum was used as the divisor in this study because it satisfies both of these requirements. Aluminum concentrations were nearly identical between Coles Hill and reference soils; the log-ratio of Al to the remainder of elements (10^6-Al) was not statistically significantly different (Mann-Whitney-Wilcoxon tests resulted in p-values of 0.97 and 0.62 and Kolmogorov-Smirnov tests resulted in p-values of 0.76 and 0.48 for the A and C horizons, respectively, suggesting that the distributions and medians of Al between the two sample sets were not significantly different).

To identify differences in trace element distribution between the Coles Hill (CH) and reference (ref) datasets, enrichment factors (EFs) were calculated for the concentration of each element (*X*) relative to the concentration of Al for each Coles Hill sample as:

$$EF = \frac{X_{CH}/Al_{CH}}{median\left(X_{ref}/Al_{ref}\right)}$$

The significance of EFs was assessed by comparing the numerator to the pre-median denominator by the non-parametric Mann-Whitney-Wilcoxon test using a Bonferroni correction (α/x, where x is the number of variables assessed) on a family alpha of 0.05.

Variation tables were created for all elements, divided into four groups by horizon (A or C) and dataset (Coles Hill or reference). A variation table shows the log of the variance of the ratio between two elements among all samples. High values indicate a high degree of variability in the ratio between the two elements suggesting they tend to span over different orders of magnitude, whereas low values indicate the ratio exhibits little variability and thus the two parts

in the ratio are roughly proportional. In addition, correlation analysis using Pearson's r was conducted on alr-transformed data to evaluate relationships between different variables.

3.2.4.1 MULTIVARIATE STATISTICAL METHODS

Modern chemical techniques, such as spectrometry using inductively-coupled plasma (e.g., ICP-MS, ICP-AES), can measure concentrations of dozens of elements from easily obtainable sample sizes at low cost. As a result, geochemical datasets have the potential to consist of a large number of variables. In many studies, such as those focusing on the fate of a single element, only a small number of elements may be of interest, but in exploratory data analysis studies, a large selection of variables can be of interest. Multivariate methods can be used to analyze such datasets, providing results and elucidating patterns that lead to an understanding of the geochemical processes involved in the interactions of elements within a given system (Drew et al., 2010; Grunsky, 2010; Reimann et al., 2008). At the same time, the available statistical methods can be limited by the size of the dataset relative to the number of variables analyzed, so sample size must be considered in the development of a statistical approach and/or the intended statistical approaches must be considered when determining the number of samples to collect.

In this study, multivariate analysis consisted of hierarchical cluster analysis and principal component analysis (PCA). Both analyses were conducted on the combined clr-transformed Coles Hill and reference soil concentration data to evaluate relationships between variables and differences between the datasets. Hierarchical variable clustering was calculated by average linkage for each soil horizon and dataset. The hierarchical clustering method uses an algorithm that groups into clusters the most similar elements by minimizing the Aitchison distance between

individual sample data and then continues to group the clusters until all are combined. Clusters combined at the lowest height correspond to the shortest Aitchison distances and, thusly, are the most similar.

Principal component analysis (PCA) is a technique that is used to reduce the number of dimensions of multivariate data to interpret trends and variations in the data from a reduced number of latent variables. The principal components (PCs) generated in PCA can yield information about the dataset, and they can also be used in further analyses, particularly those that require fewer variables. The results of PCA are typically presented as biplots (Figure 3-4), which are two-dimensional plots depicting one PC on the x-axis and another PC on the y-axis. Loading vectors (generally shown as arrows from the origin, which represents the geometric mean of the dataset) show how each of the original variables (i.e., elements) relates to the PCs plotted. Scores (generally shown as points) show the PC values for each sample. When PCA is done with clr-transformed data, links (segments between vector tips) can be used to interpret processes, rather than the rays (Aitchison and Greenacre, 2002; Grunsky, 2010); in general, shorter links show closer relation (i.e., lower log-ratio variance between the two elements associated with the rays; reflected in the variation matrix), whereas longer links indicate a lack of proportionality between the elements involved. Similarly, multiple links can be compared. Parallel links suggest correlation between the pairs of elements while orthogonal links suggest independence; the cosine of the angle between links provides an estimate of correlation between the associated log-ratios. Figure 3-4 shows an example PCA biplot of a hypothetical dataset including elements A,B,C, and D representing individual elements or parameters with scores as gray dots, loading vectors originating from the geometric mean of the dataset as black arrows, and four links as red dashed lines. The length of link AC is long relative to the other links, which

is interpreted as higher variance in the ratio of A to C. The length of link CD is relatively short, showing a more proportional relationship between C and D. The low angle between links AC and BC suggests a correlation between A/C and B/C, while the near-right angle between links CD and BD suggests independence between C/D and B/D. Through such techniques, relationships between pairs of elements can be identified in clr biplots.

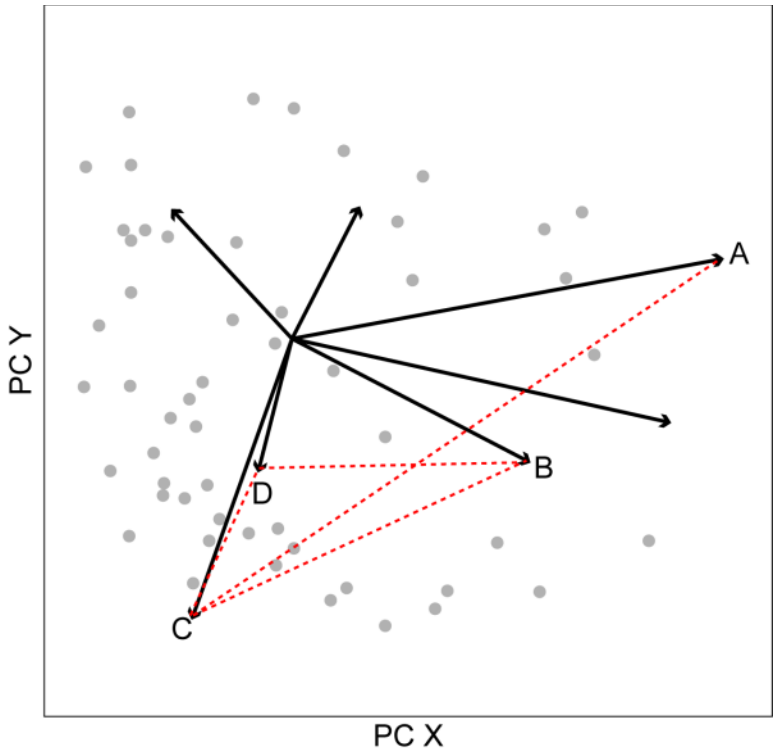

Figure 3-4: Sample clr PCA biplot. Figure includes loading vectors (black arrows,four of which are labeled A, B, C, and D), sample scores (gray circles), and links (red dashed lines connecting labeled vectors) plotted along two principal components (X and Y).

3.3 RESULTS AND DISCUSSION

Minimum, median, and maximum values for each data subset (A and C horizons from Coles Hill and reference datasets) are shown in Table 3-1.

Table 3-1: Summary statistics for untransformed data. Minimum (Min), median (Med), and maximum (Max) concentrations for all elements and organic carbon (C_{org}) by origin dataset (Coles Hill and Reference) and soil horizon (A and C).

		Coles Hill A			Reference A			Coles Hill C			Reference C		
	Units	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**
Al	wt %	1.6	4.08	8.36	0.92	4.29	10	1.53	9.86	15	5.87	10.1	14.8
As	mg/kg	<1	2	6	<1	2.85	24.3	<1	2	10	<1	2.9	51.6
Ba	mg/kg	165	377	855	98	339	1420	150	491	1260	39	351.5	3150
Be	mg/kg	0.3	1	2.3	0.2	1	4	0.3	1.7	2.9	0.5	1.45	6.1
Bi	mg/kg	0.1	0.28	0.86	0.05	0.1	0.32	0.07	0.16	0.4	<0.04	0.165	0.81
Ca	wt %	0.06	0.17	0.43	0.02	0.415	4.04	<0.01	0.02	0.58	<0.01	0.03	4.58
Ce	mg/kg	22.8	67.1	221	6.28	34.65	211	20	95.6	584	9	35.95	572
Co	mg/kg	1.4	4.8	20.5	0.9	5.9	55.4	0.8	5.3	38.7	1.9	8.8	34.4
Cr	mg/kg	9	18	37	8	30	226	3	14	48	4	35.5	219
Cu	mg/kg	4.2	13.7	27.9	4.4	15.1	77.9	1.8	9.7	30.4	2.5	26.4	174
Fe	wt %	0.45	1.59	3	0.37	1.94	6.98	0.22	3.25	5.57	0.81	4.745	13.6
Ga	mg/kg	2.68	8.31	19.3	1.99	9.915	24.5	2.87	20.7	32.4	11.4	21.85	50.4
K	wt %	0.3	0.79	2.04	0.11	0.65	3.2	0.28	1.13	2.64	0.09	1.21	3.86
La	mg/kg	12.3	35.8	107	3.3	15.1	112	10.6	48	252	4.6	18.05	179
Li	mg/kg	8	22	39	3	10	41	10	24	57	4	21	93
Mg	wt %	0.04	0.13	0.34	0.04	0.175	2.37	0.04	0.22	0.35	0.04	0.33	1.78
Mn	mg/kg	133	278	1540	91	649.5	3170	30	63	950	27	263	1010
Mo	mg/kg	0.42	0.77	1.55	<0.05	0.495	2.21	0.07	1.05	2.67	0.07	0.76	10.1
Na	wt %	0.04	0.07	1.23	0.01	0.09	1.42	0.01	0.05	1.68	<0.01	0.09	3
Nb	mg/kg	10	23.1	42.1	<0.1	7.15	62.2	0.9	23.8	66.6	1.7	12.1	289
Ni	mg/kg	3.6	8.1	17.6	2	9.1	100	2.6	10.8	35.1	2.5	18.35	134
P	mg/kg	210	690	1660	110	515	1660	<50	350	1150	<50	220	1600
Pb	mg/kg	10.4	17.5	60.7	7.6	21	46.6	5.8	22.3	51.6	2.9	18.9	53.4
Rb	mg/kg	18.8	50.4	109	9.4	45	120	26	54.2	146	2.1	65.05	237
Sb	mg/kg	0.2	0.38	0.63	<0.05	0.28	2.26	<0.05	0.26	0.45	<0.05	0.275	3.33
Sc	mg/kg	1.5	4.7	9.5	1	5.8	22.5	1.6	9.5	16.1	2.8	17.65	70.8
Sn	mg/kg	1	2.1	4.5	0.4	1.05	3.9	0.6	2.1	7.3	0.3	1.85	5.2
Sr	mg/kg	20.7	45	81.2	9.7	49.1	630	20.9	36.9	197	2.2	33.15	1140
Th	mg/kg	3	7.9	13.1	0.8	4.65	56.2	3.3	13.9	28.3	0.8	9.85	34.1
Ti	wt %	0.2	0.51	0.72	0.06	0.34	1.21	0.18	0.37	0.65	0.06	0.5	1.56
Tl	mg/kg	0.2	0.3	0.5	<0.1	0.3	0.8	0.1	0.3	0.7	<0.1	0.4	1.2
U	mg/kg	1.4	7.1	47.4	0.4	1.55	6	2.5	17.4	194	0.2	2.4	19.5
V	mg/kg	15	45	135	11	49	293	9	91	440	8	114.5	603
W	mg/kg	0.6	0.9	4.1	<0.1	0.5	2.7	0.1	1	45.7	0.1	0.8	4.4
Y	mg/kg	4.7	11.2	44.2	2.2	9.35	74.9	4.2	10.6	53.9	2.7	10.45	59
Zn	mg/kg	19	40	88	11	51	164	6	33	104	16	58.5	184
C_{org}	wt %	1.12	3.59	16.48	0.46	5.755	41.2	0.02	0.07	0.13	<0.01	0.155	0.62

3.3.1 SOIL MINERALOGY AT COLES HILL

Phases detected by XRD in soils from Coles Hill include quartz, feldspar (albite and microcline), muscovite, kaolinite, chlorite, and hematite. Minerals were detected in the following percentages of samples in the A horizon: quartz (100%), feldspars (100%), kaolinite (36%), chlorite (80%), muscovite (56%), and hematite (36%); and in the following percentages of samples in the C horizon: quartz (100%), feldspars (84%), kaolinite (68%), chlorite (96%), muscovite (68%), and hematite (40%). The XRD analysis was qualitative and did not include amorphous content. Whereas the major minerals present do not show systematic variation in presence/absence within the Coles Hill soils, their relative amounts and the presence of amorphous and/or trace phases likely contributes to their different appearance and chemistry. Trace mineralogy from soils overlying Coles Hill was reported by Jerden (2001).

3.3.2 ENRICHMENT FACTORS

Enrichment factors of Coles Hill soils relative to reference soils of all elements normalized to Al are shown in Figure 3-5. Significant enrichment or depletion was defined as a $p < 0.0014$ (or 0.05/36) result of a Mann-Whitney-Wilcoxon test comparing the median of X/Al values between the Coles Hill and reference soils. Elements with significantly elevated EFs (>1) in the A horizon were, in decreasing order of median EF, U, Nb, Bi, La, Ce, Sn, W, Th, and Li. Elements with significantly elevated EFs in the C horizon were U, La, Ce, and Nb. Elements with significantly depleted EFs (<1) in the A horizon were Mg, Cr, and Mn, and in the C horizon were Fe, Sc, Zn, C_{org}, Cu, and Mn. In both horizons, U has the greatest median EF. Though a number of trace elements (e.g., U, Nb, La, Ce) are enriched in the Coles Hill soils relative to the reference soils, other trace elements, including many of those that have been reported as

pathfinders in other types of U deposits (e.g., As, Cu, Co) are depleted or not significantly different between the Coles Hill soils and surrounding reference soils.

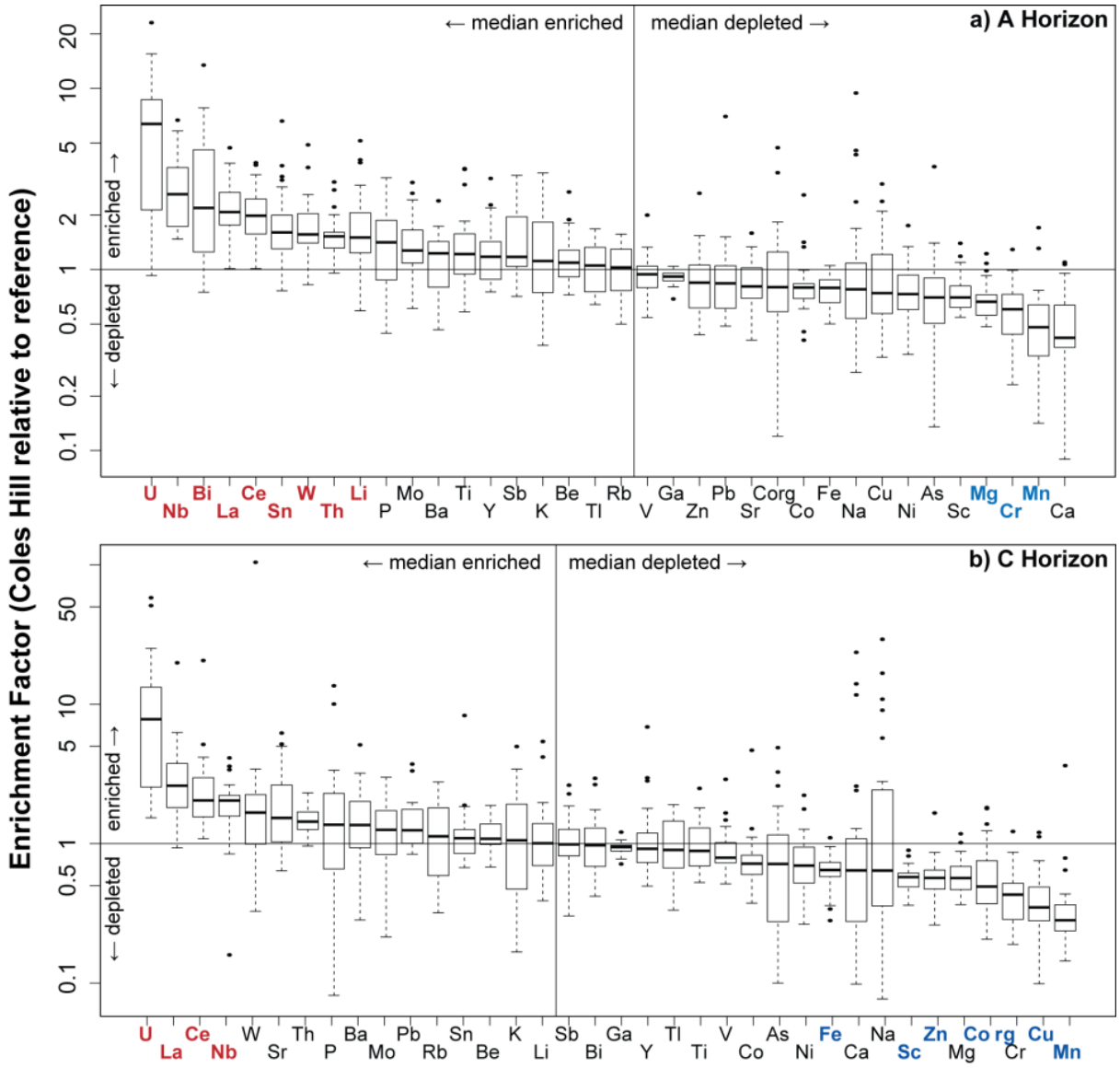

Figure 3-5: Tukey boxplots of median enrichment factors [$(X/Al)_{CH}$/median$(X/Al)_{ref}$] in the a) A horizon and b) C horizon. Elements are arranged horizontally in order of descending median. The horizontal line at EF=1 represents equivalence between the Coles Hill and reference ratios. The vertical line indicates the split between medians >1 (enriched) and <1 (depleted). Elements with significantly different *X*/Al ratios between Coles Hill and reference soils are labeled in bold, with significantly enriched elements in red and significantly depleted elements in blue. *p*-values can be seen in Table A-1.

3.3.3 CORRELATION AND VARIATION BETWEEN ELEMENTS

Table 3-2 presents correlation coefficients (Pearson's r) between alr values for U and the other elements, separated by dataset and horizon. In the reference soil data, for both horizons, ln U/Al is strongly correlated with ln Th/Al and moderately correlated with ln Rb/Al and ln Tl/Al, which are elements commonly associated with feldspar weathering (as K substitutes). In the Coles Hill soils, correlations show more differences between horizons. In the A horizon, ln U/Al is most strongly correlated with ln V/Al, ln Li/Al, and ln Mo/Al. In the C horizon, ln U/Al is most strongly correlated with ln La/Al and ln Ce/Al; the LREEs may coexist with U in phases such as rhabdophane, or, within the deposit, in U-rich apatites associated with the ore mineral coffinite (Jerden, 2001). For data from both horizons, a strong negative correlation with ln U/Al is seen with ln K/Al, ln Rb/Al, and ln Tl/Al, suggesting that U enrichment may be antipathetically related to feldspar/granite parent materials; Jerden (2001) noted depletion of K2O and Rb during alteration and mineralization of this deposit.

Table 3-2: Correlation coefficients (Pearson's r) between U alr values [ln(U/Al)] and other constituents' alr values [ln(*X*/Al)] in A and C horizons of Coles Hill (CH) and reference (ref) soils.

Analyte	A CH	A ref	C CH	C ref
As	0.416	0.305	-0.073	0.425
Ba	-0.461	0.325	0.050	0.345
Be	0.047	0.453	0.191	0.562
Bi	0.306	0.599	-0.494	0.667
Ca	0.255	-0.354	0.240	-0.395
Ce	-0.043	0.538	0.619	0.496
Co	-0.157	-0.329	-0.006	-0.329
Cr	0.145	-0.053	-0.584	0.128
Cu	0.155	-0.118	-0.477	0.065
Fe	0.196	-0.240	-0.169	-0.067
Ga	0.259	0.447	0.100	0.493
K	-0.634	0.418	-0.614	0.482
La	0.090	0.490	0.749	0.490
Li	0.539	0.550	0.037	0.688
Mg	0.183	-0.547	-0.176	-0.327
Mn	-0.284	-0.017	-0.256	-0.206
Mo	0.518	0.438	0.149	0.481
Na	0.137	-0.354	0.131	-0.214
Nb	0.252	0.421	0.071	0.591
Ni	0.152	-0.090	-0.439	0.016
P	0.462	0.092	-0.130	0.143
Pb	-0.188	0.532	0.354	0.721
Rb	-0.627	0.650	-0.744	0.661
Sb	0.210	0.370	-0.007	0.292
Sc	0.177	-0.290	-0.062	-0.108
Sn	-0.012	0.682	-0.370	0.776
Sr	0.196	-0.139	0.519	-0.107
Th	0.308	0.868	-0.517	0.860
Ti	0.232	0.188	-0.111	0.076
Tl	-0.312	0.689	-0.724	0.711
V	0.646	-0.180	0.311	-0.123
W	-0.082	0.319	-0.024	0.461
Y	-0.383	0.134	0.251	0.279
Zn	-0.138	-0.095	-0.187	0.042
C_{org}	-0.113	0.213	-0.290	0.446

To investigate the degree to which other elements may exhibit proportional relationships to U in the system, variances were calculated for each log-ratio involving U (Table 3-3). Variance tables consisting of all elements can be seen in Supplemental Information. Log-ratio variances give the variance of the natural log of the ratio between two elements. A powerful strength of this approach is that the results are independent of units, providing identical results for data in for instance, oxide percent versus ppm elemental concentrations. For the reference soil data, log-ratios involving U generally exhibit low to moderate variance, whereas for the Coles Hill data, they are high, suggesting higher variability in U in the Coles Hill soils; total log-ratio variance is greater for the Coles Hill samples than for the reference soil samples for each horizon despite the smaller Coles Hill sample size, and total log-ratio variance is greater for the C horizon samples than for the A horizon samples. In the reference soil sample data, Th exhibits the lowest low ratio variance with U, suggesting an association between these two elements, consistent with results from the correlation analysis (Table 3-2). For data from the Coles Hill soil samples, ln U/Th variance is much higher, consistent with the low correlation of ln U/Al vs. ln Th/Al. For data from the A horizon at Coles Hill, the lowest U log-ratio variances are with V, Li, Mo, and P, suggesting an association of U with phosphate minerals and the possibility of vanadate substitution for phosphate. Mineralogical analysis of Coles Hill ore and soil samples showed that U was associated with phosphate in crandallite-group minerals and was also sorbed on Fe oxides (Jerden and Sinha, 2006); the ratios of U and V to Fe and Al are relatively low among Coles Hill A horizon U and V ratios within the variation matrix. For data from the Coles Hill C horizon samples, the lowest U log-ratio variances are with La, Ce, and Sr. The presence of lanthanide-phosphate minerals in association with U and U-phosphates in Coles Hill soils, as suggested by the statistical relationship of these elements, was noted by Jerden (2001).

Table 3-3: U log-ratio variance values from variation matrices for Coles Hill (CH) and reference (ref) samples, A and C horizons. Values report the variance of ln(U/*X*). Values can be compared vertically but not horizontally due to differences in sample sizes. Results for all elements are reported in Table A-2.

Analyte	CH A	ref A	CH C	ref C
Al	0.74	0.27	1.06	0.76
As	0.72	0.48	2.22	1.10
Ba	1.22	0.40	1.39	1.04
Be	0.81	0.25	1.02	0.52
Bi	0.90	0.31	1.79	0.59
Ca	0.83	2.52	2.13	7.36
Ce	0.91	0.32	0.65	0.81
Co	0.96	1.18	1.28	1.61
Cr	0.79	0.76	1.86	1.37
Cu	0.92	0.60	1.94	1.26
Fe	0.71	0.62	1.28	1.07
Ga	0.71	0.22	1.05	0.63
K	1.74	0.41	3.06	0.99
La	0.82	0.35	0.49	0.82
Li	0.53	0.22	1.36	0.40
Mg	0.73	1.11	1.24	2.06
Mn	1.32	1.28	1.80	1.79
Mo	0.55	0.46	1.21	0.84
Na	1.31	1.65	3.14	5.15
Nb	0.74	0.74	1.35	0.62
Ni	0.78	0.70	1.72	1.44
P	0.60	0.76	2.62	1.42
Pb	1.19	0.26	0.93	0.36
Rb	1.23	0.22	2.43	0.63
Sb	0.77	0.62	1.31	1.30
Sc	0.73	0.67	1.13	1.24
Sn	0.97	0.23	1.67	0.31
Sr	0.74	1.27	0.79	3.36
Th	0.67	0.16	1.33	0.22
Ti	0.78	0.50	1.31	1.09
Tl	0.99	0.18	2.03	0.45
V	0.52	0.71	0.96	1.48
W	0.97	0.95	2.06	0.79
Y	1.13	0.56	1.10	0.94
Zn	1.01	0.52	1.31	0.99
C_{org}	1.40	2.60	1.70	0.87
Sum	**32.44**	**25.03**	**55.70**	**47.67**

3.3.4 MULTIVARIATE ANALYSIS

To evaluate similarities between elements, cluster dendrograms showing hierarchical variable clustering by average linkage for each soil horizon and dataset were created (Figure 3-6). Within the reference soil data, element clustering is similar in the A and C horizons. A major difference is the closer clustering in the A horizon of alkali and alkaline earth metals (e.g., Ca, Mg, Na, Sr), mobile elements associated with weathering, with organic carbon and P, elements associated with organic matter, a finding that likely reflects increased vegetation/agriculture-related influences in and weathering of the A horizon relative to the C horizon. In the reference soils, U is most closely clustered with Th, as well as Bi, Li, and Sn.

In the Coles Hill soil data, the C horizon dendrogram suggests three large clusters of elements. The first (from left to right, those between Ti and La, shown in red) includes Na, Ba, Ca, Pb, and Ti, elements associated with the deposit through the presence of Na-metasomatism, Ba-rich feldspars, apatite, Pb from radioactive decay of U, and Ti oxides. The second (P to Bi, green) is largely made up of elements associated with the granitic complex in which the mineralization occurs (e.g. K, Rb, Tl, Sn) but also includes P which is closely associated with U mineralization in the form of apatite. The third group (Mn to Ni, blue) consists of mainly redox-sensitive metals common in aluminosilicates and ferromagnesian minerals, either as major constituents or as trace elements. Notably, Th and U are not clustered as in the reference dendrograms.

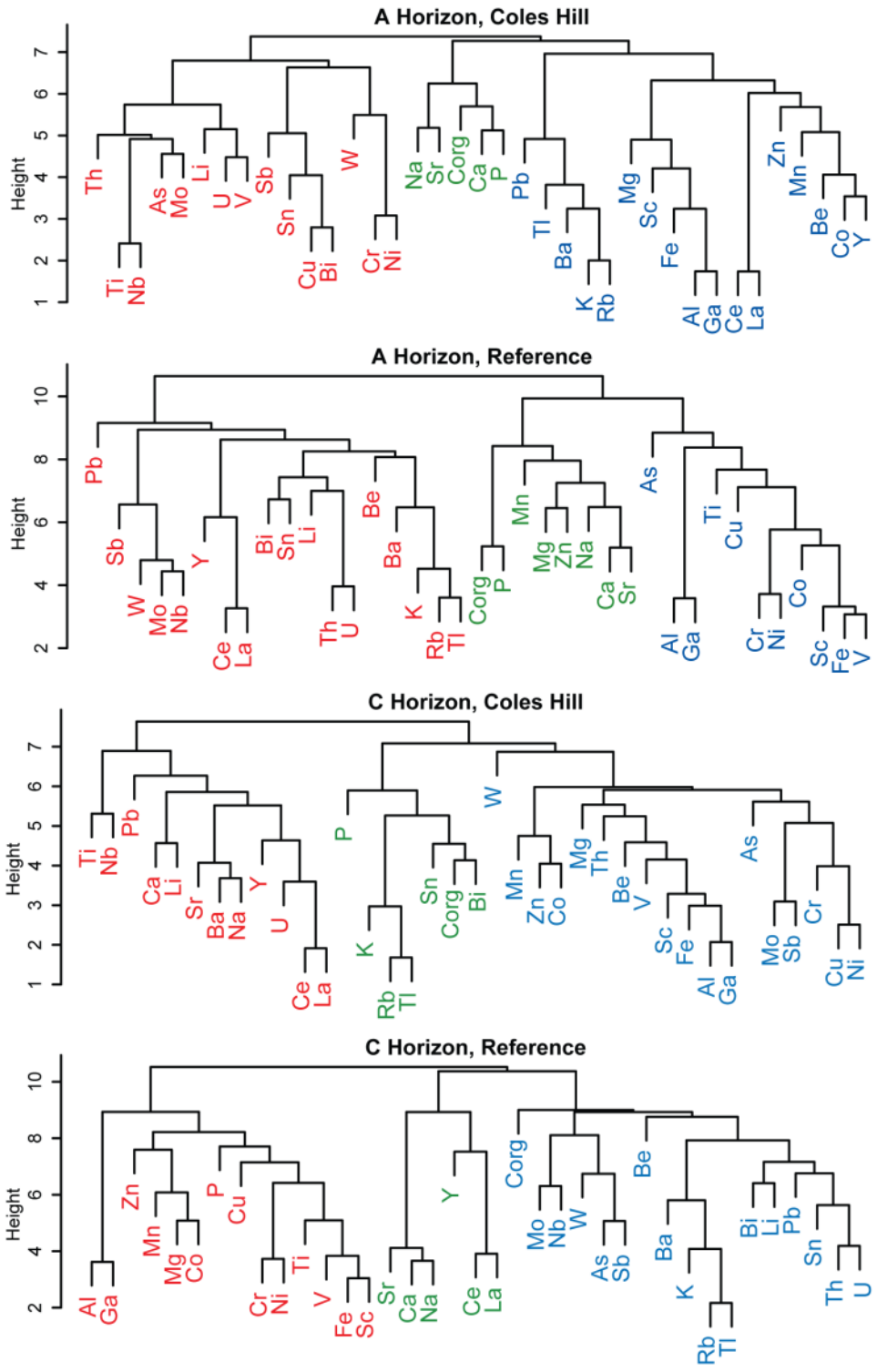

Figure 3-6: Dendrograms from hierarchical variable clustering for A and C horizons at Coles Hill and in reference soils. The highest three clusters for each dendrogram are colored from left to right in red, green, and blue.

Results from cluster analysis of data from A horizon samples preserve similar groups (U mineralization, granite complex, and common metals), with prominent exceptions: 1) the increased clustering of highly weatherable alkali and alkaline earth metals with organic carbon and P, indicating the increased association of weathering and the influence of agricultural practices and vegetation, and 2) the separation of U and rare earths Ce, La, and Y in favor of a closer clustering with V, indicating the possibility of U separating from ore minerals and/or LREE-bearing secondary assemblages that were preserved in the C horizon into phosphate/vanadate phases or co-adsorbed with these ions.

To further evaluate relationships between elements measured from the A horizon soils, principal component analysis was done for all A horizon samples (Figure 3-7a, Table A-3, Table A-4). Long rays for Ca and organic carbon suggest these constituents exhibit high clr-variance, mainly in negative PC2, indicating that PC2 may relate to vegetation and agricultural influence. Coles Hill data are concentrated in the area with low-magnitude PC2 and negative PC1. Loading vectors in this space include all elements that were significantly enriched in the Coles Hill A horizon relative to the reference (Figure 3-5), as well as K, Tl, and Rb. Elements with the shortest links to U are La, Th, and Ce, confirming results from the log-ratio variance calculations (cf. Figure 3-7a,

Table 3-3). Elements with positive loadings along PC1 and PC2 include those commonly found in clays and oxyhydroxides that are common weathering products (e.g. goethite, gibbsite), including Al and Fe (McBride, 1994), as well as those with similar ionic radii to Al and Fe that may substitute in such minerals (e.g. Co, Cu, Ni; Marques et al., 2004). Elemental loadings along PC2 suggest that it represents the balance of refractory minerals (i.e. resistant to weathering) versus weatherable constituents and is an important factor in distinguishing the variation between data for Coles Hill A horizon soil samples and data for reference A horizon soil samples. In addition, this analysis shows that the association of U and LREEs may persist in the A horizon, though with a weaker signal than in the C horizon.

The same analysis was conducted for samples from the C horizon (Figure 3-7b, Table A-3, Table A-4). Again, long rays for Ca, as well as Na and Sr, indicate that these elements exhibit a relatively large clr-variance. Data for Coles Hill samples are concentrated in negative PC2 but vary in PC1. Loading vectors in this space include those for elements that were significantly enriched in data from the Coles Hill C horizon, including U, Ce, La, and Th, relative to reference soils, as well as Rb, Tl, and Pb. Uranium shows relatively high clr variance. Links to U from other elements are long, suggesting that variation in U does not vary proportionally with other elements. The association of these elements with the Coles Hill sample data suggests that feldspar-derived elements and refractory minerals are significant in Coles Hill C horizon soils, whereas the signatures of organic matter, other aluminosilicates, and ferromagnesian minerals are reflected more in a majority of the reference soils.

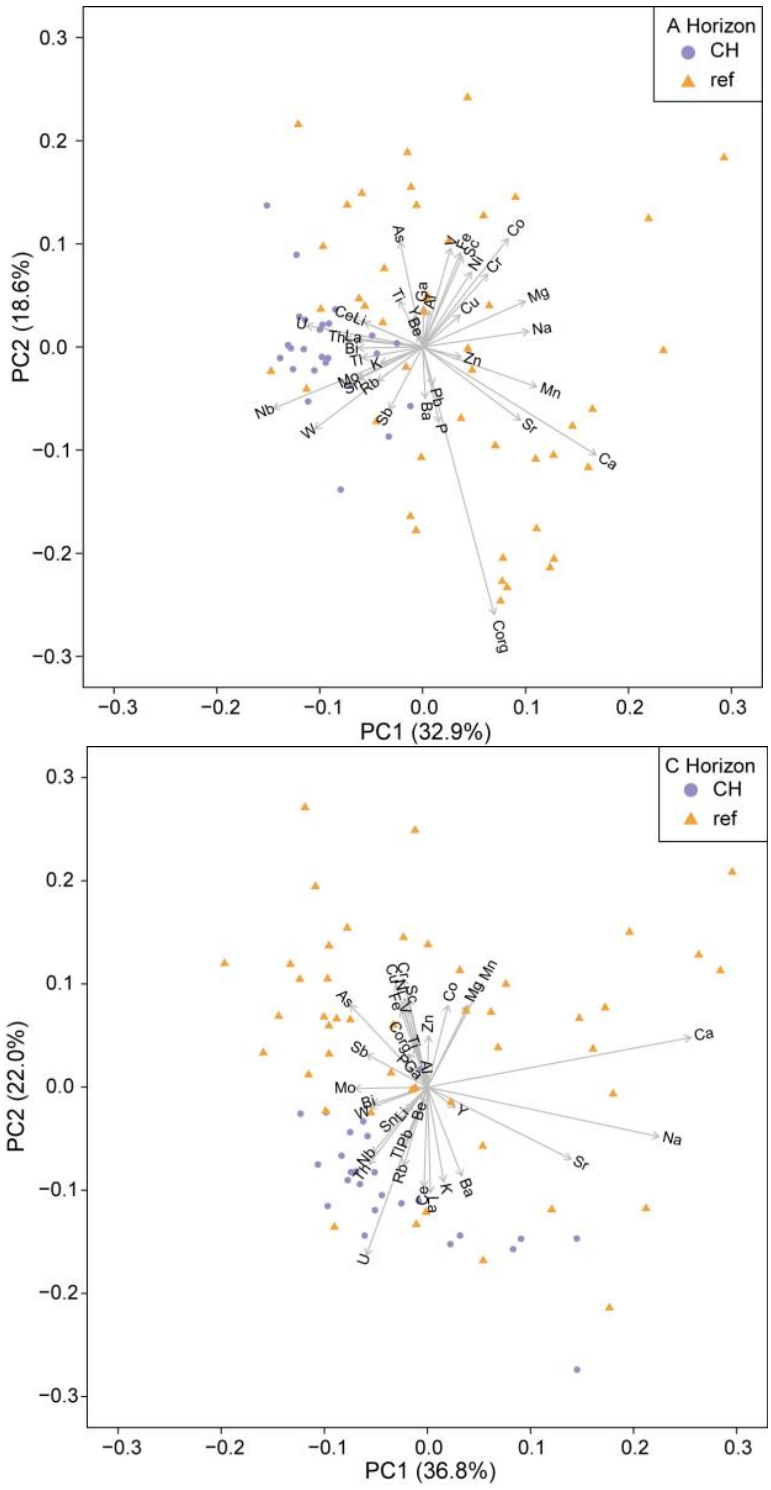

Figure 3-7: PCA biplots showing all a) A and b) C horizon samples. Blue circles indicate Coles Hill (CH) samples; green triangles indicate reference (ref) samples. Percent of variability explained by each component is given in parentheses. R code to create the biplot was modified from Vu (2011).

3.3.5 GEOCHEMICAL TRENDS

In the reference soil data, element groupings in soils are reflective of abundances in different soil/parent rock types. Elements largely associated with the weathering of feldspars in granites (K, Ba, Rb, Tl, Li) show association in the multivariate analysis. So too do redox-sensitive metals that may be released during the oxidative weathering of reduced phases in rocks and/or be present in mineral products of weathering such as clays and oxy-hydroxides (Fe, V, Cr, Ni, Cu, Mo, As); highly weatherable alkali and alkaline earth metals (Ca, Na, Sr, Mg) also exhibit association. Uranium is associated with Th as well as with feldspar/granite-related elements. In the A horizon data, the relationship of organic carbon and P, which can be organic or a remnant of fertilizers applied during cultivation, suggests a vegetative and/or agricultural influence.

In the Coles Hill data, the clustering of granitic/feldspathic, common metals, and highly weatherable elements remains, but there is a change in the elements with which U is associated. The change suggests that the fraction of U that is disseminated within granitic and meta-granitic rocks is not as influential on total U concentrations near the deposit as the reference soils, which is further substantiated by the negative correlation between ln U/Al and ln felsic elements/Al (Table 3-2). In the A horizon Coles Hill data, evidence for previously-observed association with phosphates is seen. In the C horizon data, a relationship between U and LREEs is apparent; this association likely reflects the presence of U-LREE phosphates derived from apatite enriched in these elements within the mineralization of the deposit (Jerden, 2001). In addition, the proportion of Nb is elevated in Coles Hill C horizon data, though its correlation with U is low to moderate and its PCA links with U are longer than those of other elements; in the Coles Hill mineralization

zone, Nb likely substitutes for Ti in anatase and other Ti phases associated with the ore (Jerden, 2001).

The Coles Hill site has been subjected to several centuries of agricultural activity, which includes the physical disturbance of the upper soils layers as well as the application of fertilizers and pesticides. Agricultural activity can impact the concentrations of organic carbon, the nutrients fertilizers and liming agents are applied to supply (e.g., P, K, Ca, Mg), trace elements found in fertilizers [e.g., U, V, Cd; Mortvedt (1995)], and the constituents of pesticides (e.g., Pb, As). As a result, C horizon soils in agricultural areas may provide more reliable geochemical prospecting information than the A horizon.

In addition, elements with a documented relationship to U, such as P (Jerden and Sinha, 2006), may not evidence useful geochemical relationships due to the varied sources and processes affecting these elements. Thus, statistical analysis of trace elements can elucidate relationships with less interference than more concentrated multi-source elements. For example, this study shows a relationship in Coles Hill soils between U and LREEs. Due to their location over the weathered deposit and the elevated concentration of U in these soils, it is assumed that their origin is linked to the deposit. Within the deposit mineralization, elevated LREEs are associated with high-U, high-P samples from the apatite-coffinite ore assemblage (Jerden 2001). Through the genesis of soils and their weathering, the relationship of U and LREEs appears to be preserved, while the relationship of U and P does not, as there are sources of P not associated with U mineralization, such as phosphate amendments (e.g., fertilizers) and non-U-bearing minerals.

3.3.6 PATHFINDER RELATIONSHIPS

Because the analysis of compositional data is based on ratios between elements and/or groups of elements, we present element ratios that may be characteristic pathfinders for a deposit similar to Coles Hill. In some cases, ratios contain a denominator that is present in fairly constant concentrations across all soils, such as Al. In the case of these ratios, for example U/Al, an elevated ratio is likely reflective of elevated U. However, in other cases, such as U/Th, typical ratios exist among different parent rock types (Rogers and Ragland, 1961), and changes in these ratios may be an indicator of mineralization similar to Coles Hill.

Potential pathfinder ratios suggested by the results of EFs, PCA, and correlation and cluster analysis in this study include elevated U/Al, LREEs/Al, Nb/Al, U/V, Nb/Ti, and U/Th (Figure 3-8). An elevated U/Mn or Al/Mn ratio appears from the data to be a potential indicator, but the presence of Mn deposits documented within the reference area (Espenshade, 1954) could provide a regional bias in this ratio.

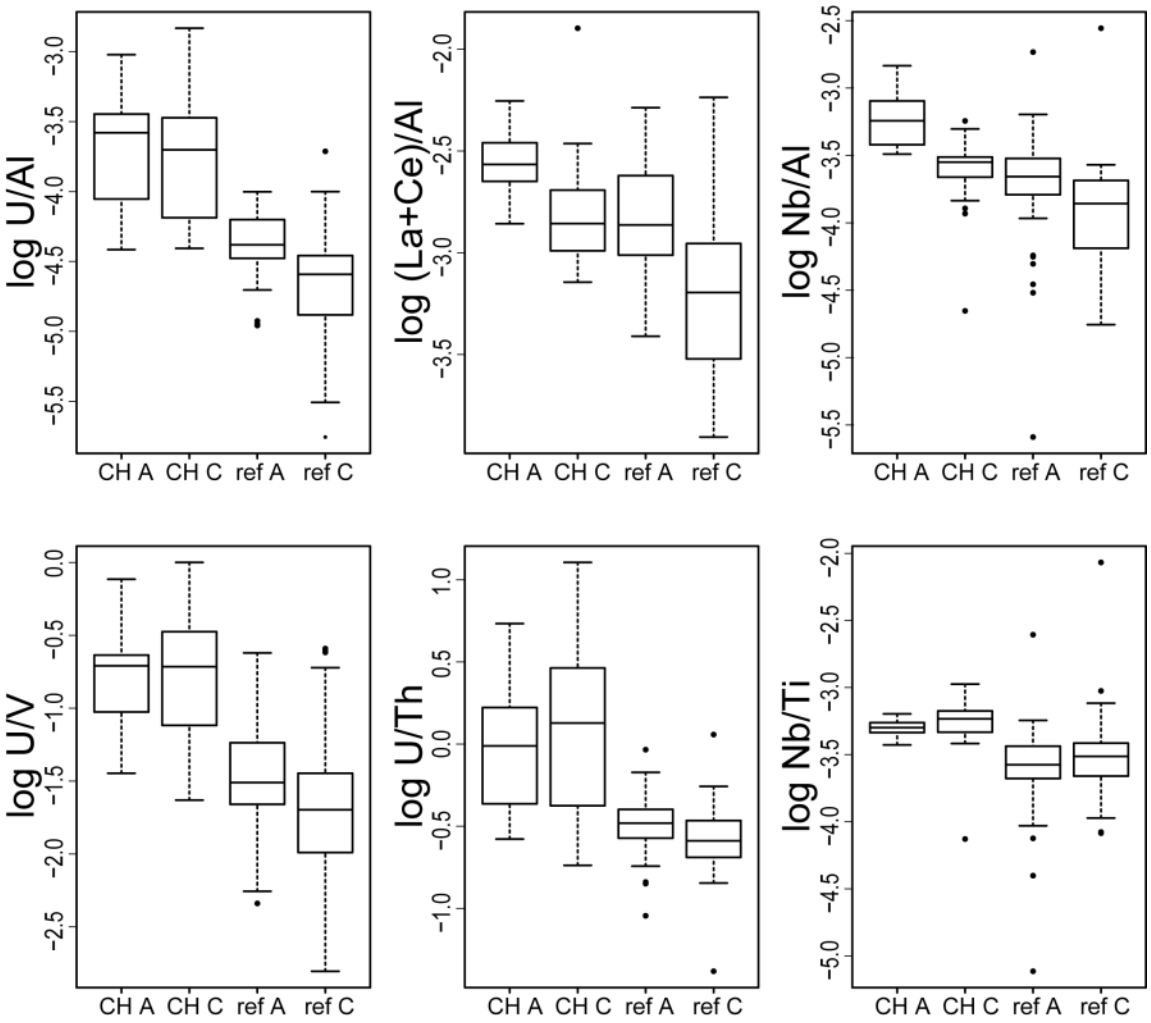

Figure 3-8: Tukey boxplots showing suggested pathfinder ratios in soils for deposits similar to Coles Hill. Boxes are shown for, from left to right, the Coles Hill (CH) A horizon, the Coles Hill C horizon, the reference (ref) A horizon, and the reference C horizon.

In addition to elevated values, the spread of Nb/Ti and U/Th are different between data for the Coles Hill and reference soil samples; in the Coles Hill soils data, Nb/Ti ratios are very uniform, whereas U/Th ratios range much more widely than in reference soils (Figure 9a). The typical U/Th ratio in the upper continental crust is 0.26 (Cuney, 2010). The reference soil U/Th values range from 0.041 to 1.1 with an average of 0.31 and standard deviation of 0.15, while the Coles Hill soil U/Th values range from 0.18 to 13 with an average of 1.8 and standard deviation

of 2.5. In some rocks, the U/Th ratio has been linked to the proportion of K-feldspar (Cuney, 2010; Rogers and Ragland, 1961). Figure 9b shows a plot of U, Th, and K relationships as isometric log-ratio balances (Egozcue and Pawlowsky-Glahn, 2005), in which the U/Th ratio in Coles Hill soils shows a relationship with K, while that of the reference soils does not. This relationship may be indicative of the K depletion associated with Na-metasomatism during mineralization.

Previous studies of pathfinders at U deposits (Snelling, 1984) suggested that U is not a good pathfinder due to its high mobility in near-surface conditions. However, results from the Coles Hill soil samples suggest otherwise, as significantly elevated U/Al and U/Th were seen in these data. Jerden et al. (2003) and Jerden and Sinha (2006) discussed the immobilization of U in the near-surface environment at Coles Hill by phosphate mineralization, and it is likely that this process is reflected by the elevated U ratios observed in this study.

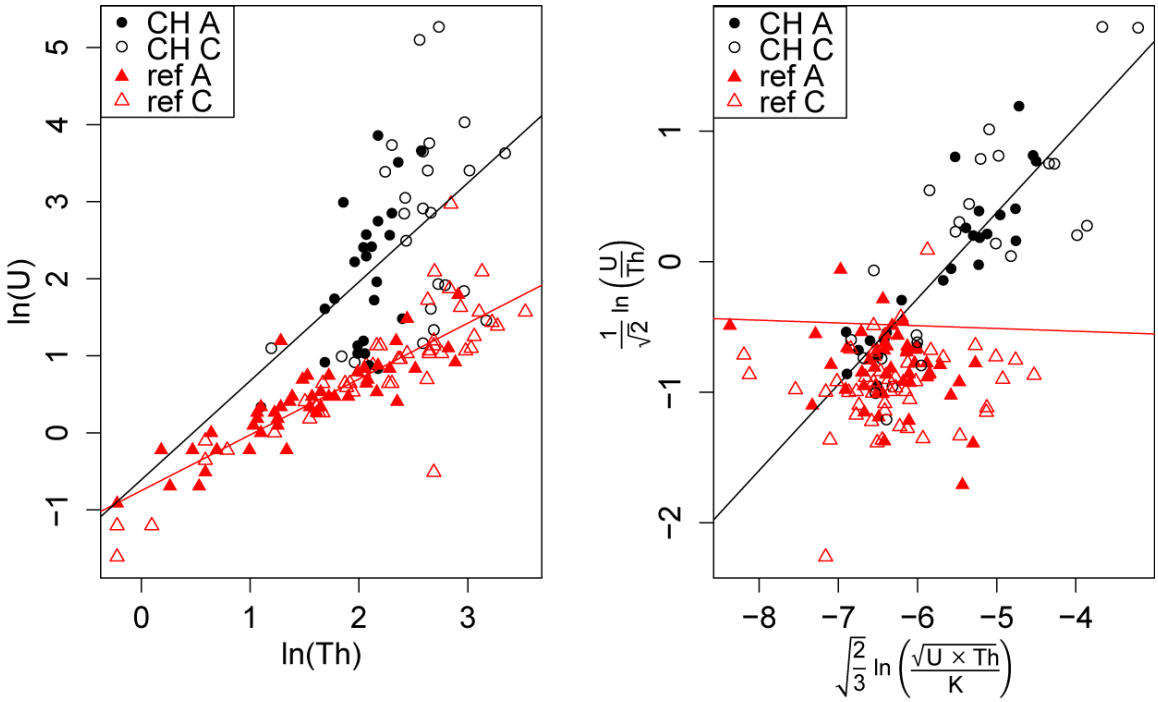

Figure 3-9: a) Log-log plot of Th vs. U concentrations (mg/kg). A horizon values are shown as closed symbols, and C horizon values are shown as open symbols. Black circles indicate Coles Hill (CH) samples, and red triangles indicate reference (ref) samples. Log-log linear best-fit lines are shown for Coles Hill and reference samples in their respective colors. b) Plot of isometric log-ratio balances for K, U, and Th.

3.3.7 RADIOISOTOPES

Radioisotopes related to the U and Th decay chains were not analyzed in the USGS soil survey. However, concentration-distance plots along the sampling transects show a very localized increase in concentrations of radionuclides of the U decay series over the Coles Hill deposit (Figure 3-10). Radioisotopes of the non-uranogenic Th-232 decay series (Th-232, Ra-228, and Th-228) do not show the same pattern of increase over the deposit. These observations suggest that concentrations of members of the U decay chain in soils are also indicative of underlying U mineralization, in addition to the trace elements discussed in the previous sections. The lack of this trend in Th progeny suggests that the elevated Th in Coles Hill soils may be a

radiogenic product of U decay, though the weak ln U/Th correlation in Coles Hill soils is not consistent with this hypothesis.

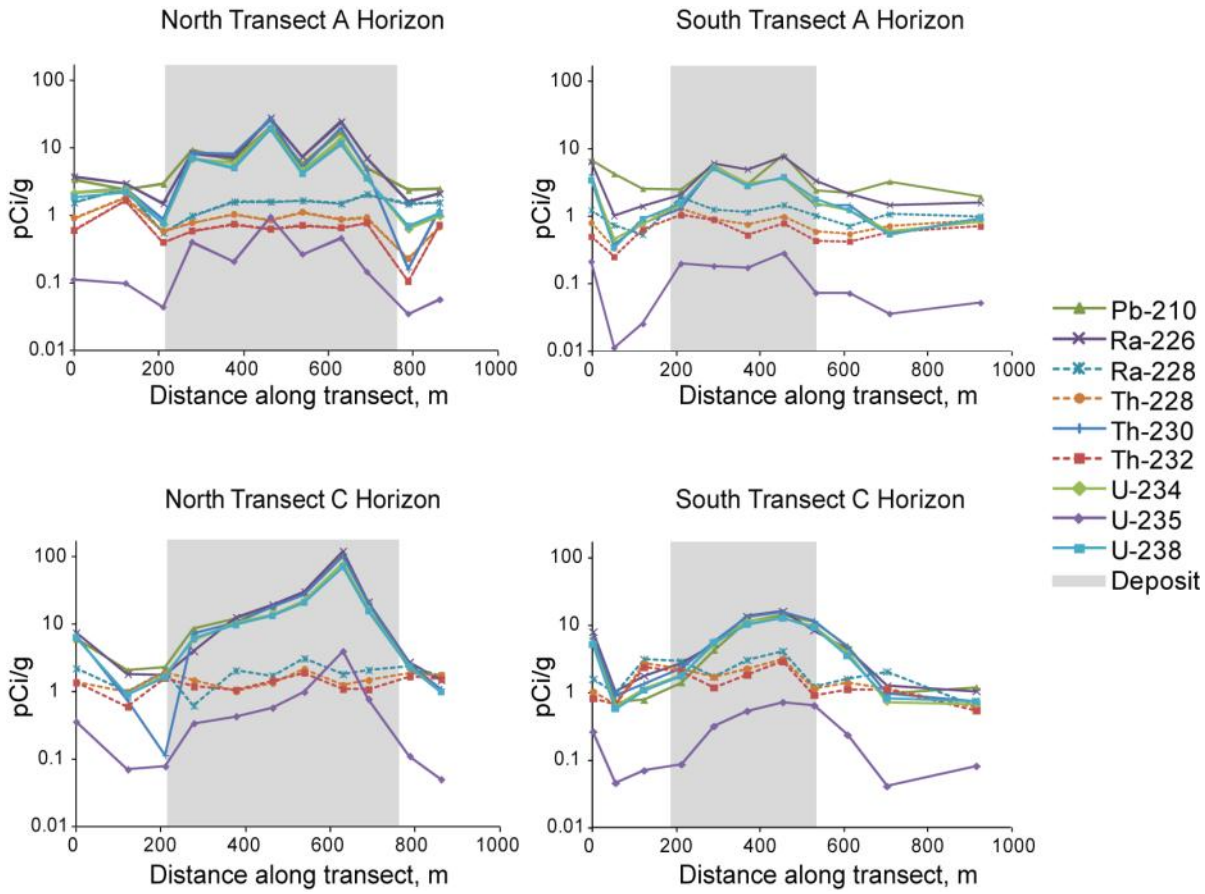

Figure 3-10: Radionuclide concentrations in the A and C horizons across transects over the North and South ore bodies. Grey shading reflects the extent of the surface projection of ore deposits.

3.4 CONCLUSIONS

Analysis of data from a baseline soil survey at the Coles Hill U deposit reveals geochemical anomalies relative to regional soils. Anomalies exist in both the A and C horizons; whereas there are significant differences in composition between samples from the A and C horizon, both horizons locally show the influence of the deposit, particularly in terms of U concentration. Soil concentrations and proportions (i.e., ratios) of certain elements associated with the deposit (e.g., U, Th, La, Ce, Nb) are higher in the samples of soils overlying the deposit

relative to regional reference soils. In some cases, elevated concentrations are likely due to proximity to the deposit and are the type of data that exploration geochemical surveys target. Based on the differences between Coles Hill and reference soils and on relationships between elements within those soils, elevated proportions (i.e., concentrations relative to Al or the geometric mean) of U, Th, La, Ce, and Nb in soils, as well as potentially those of U-progeny radioisotopes, may be indicative of underlying mineralization similar to Coles Hill and could be used as pathfinders for this type of deposit. Beyond simply elevated levels of certain elements, their relationship to U or other elements is indicative of a possible buried deposit. For example, our results show that this deposit is characterized by the positive relationship of U and LREEs in overlying soils and the negative or lack of relationship of the proportions to Al of U and Th, K, Rb, and Tl in overlying soils relative to other soils of the same type. These data patterns become evident through the use of multivariate compositional data analysis, which can be easily conducted using open-source software and allows for the rapid analysis of geochemical datasets with numerous interrelated parameters. This study also demonstrates the use of results of a national-scale geochemical survey as reference soils for comparison between a site of geologic interest, the Coles Hill U deposit, and regional soils.

ACKNOWLEDGMENTS

Randy Cosby was involved in soil sampling, and Virginia Tech soil laboratory analyses were led by Julie Burger and Trey Rauh. XRD preparation was completed by Nadine Piatak and Natalia Ainsfield. Virginia Uranium, Inc. provided partial financial support to Virginia Tech. VUI and Walter Coles Sr. allowed property access, and Stewart East provided valuable knowledge about the area. The authors appreciate the helpful comments from two anonymous reviewers and Larry Drew of the USGS. The use of product or trade names is for descriptive purposes only and does not constitute endorsement by the U.S. government.

REFERENCES

Aaron, L.G., 2009. Pittsylvania County, Virginia: A Brief History. The History Press, Charleston, SC.

Aitchison, J., 1982. The Statistical Analysis of Compositional Data. Journal of the Royal Statistical Society. Series B (Methodological) 44, 139-177.

Aitchison, J., 1986. The statistical analysis of compositional data. Chapman and Hall, London; New York.

Aitchison, J., Greenacre, M., 2002. Biplots of compositional data. Journal of the Royal Statistical Society: Series C (Applied Statistics) 51, 375-392.

Aylor, J., Bodnar, R.J., Beard, J., 2014. Geology of the Coles Hill uranium deposit, Virginia Piedmont, GSA Field Guide. Geological Society of America, Boulder, CO, pp. 271-284.

Bradshaw, P.M.D., Lett, R.E.W., 1980. Geochemical exploration for uranium using soils. Journal of Geochemical Exploration 13, 305-319.

Buccianti, A., Mateu-Figueras, G., Pawlowsky-Glahn, V., 2006. Compositional Data Analysis in the Geosciences, Geological Society Special Publication 264. Geological Society of London, London.

Carranza, E.J.M., 2011. Analysis and mapping of geochemical anomalies using logratio-transformed stream sediment data with censored values. Journal of Geochemical Exploration 110, 167-185.

Cohen, D.R., Rutherford, N.F., Morisseau, E., Christoforou, I., Zissimos, A.M., 2012. Anthropogenic versus lithological influences on soil geochemical patterns in Cyprus. Geochemistry: Exploration, Environment, Analysis 12, 349-360.

Comas-Cufí, M., Thió-Henestrosa, S., 2011. CoDaPack 2.0: a stand-alone, multi-platform compositional software, in: Egozcue, J., Tolosana-Delgado, R., Ortego, M. (Eds.), CoDaWork'11: 4th International Workshop on Compositional Data Analysis, Software Program.

Cox, D.P., Singer, D.A., 1986. Mineral deposit models, Bulletin 1693.

Cuney, M., 2009. The extreme diversity of uranium deposits. Mineralium Deposita 44, 3-9.

Cuney, M., 2010. Evolution of Uranium Fractionation Processes through Time: Driving the Secular Variation of Uranium Deposit Types. Economic Geology 105, 553-569.

Dahlkamp, F.J., 1993. Uranium ore deposits. Springer-Verlag, Berlin.

Dahlkamp, F.J., 2010. Uranium Deposits of the World. Springer Berlin Heidelberg.

Drew, L.J., Grunsky, E.C., Sutphin, D.M., Woodruff, L.G., 2010. Multivariate analysis of the geochemistry and mineralogy of soils along two continental-scale transects in North America. Science of the Total Environment 409, 218-227.

Egozcue, J.J., Pawlowsky-Glahn, V., 2005. Groups of parts and their balances in compositional data analysis. Mathematical Geology 37, 795-828.

Egozcue, J.J., Pawlowsky-Glahn, V., Mateu-Figueras, G., Barceló-Vidal, C., 2003. Isometric logratio transformations for compositional data analysis. Mathematical Geology 35, 279-300.

Engle, M.A., Bern, C.R., Healy, R.W., Sams, J.I., Zupancic, J.W., Schroeder, K.T., 2011. Tracking solutes and water from subsurface drip irrigation application of coalbed methane–produced waters, Powder River Basin, Wyoming. Environmental Geosciences 18, 169-187.

Espenshade, G.H., 1954. Geology and mineral deposits of the James River-Roanoke River manganese district, Virginia, Survey, U.S.G., Bulletin 1008.

Fayek, M., 2013. Uranium ore deposits: A Review, in: Burns, P.C., Sigmon, G.E. (Eds.), Uranium: Cradle to Grave. Mineralogical Association of Canada, pp. 121-146.

Gannon, J.P., Burbey, T.J., Bodnar, R.J., Aylor, J.G., 2011. Geophysical and geochemical characterization of the groundwater system and the role of Chatham Fault in groundwater movement at the Coles Hill uranium deposit, Virginia, USA. Hydrogeology Journal 20, 1-16.

Grunsky, E.C., 2010. The interpretation of geochemical survey data. Geochemistry: Exploration, Environment, Analysis 10, 27-74.

Henika, W.S., 2002. Geologic map of the Danville 30 by 60 minute quadrangle, Virginia, Publication 166.

Henika, W.S., Thayer, P.A., 1983. Geologic map of the Spring Garden quadrangle, Virginia, Publication 48.

IAEA, 2009. World Distribution of Uranium Deposits (UDEPO) with Uranium Deposit Classification, IAEA-TECDOC-1629.

Jefferson, C., Thomas, D., Quirt, D., Mwenifumbo, C., Brisbin, D., 2007. Empirical models for Canadian unconformity-associated uranium deposits, in: Milkereit, B. (Ed.), Exploration 07: Fifth Decennial International Conference on Mineral Exploration, Toronto, Canada.

Jerden, J.L., 2001. Origin of uranium mineralization at Coles Hill Virginia (USA) and its natural attenuation within an oxidizing rock-soil-ground water system,PhD Thesis.

Jerden, J.L., Sinha, A.K., 2003. Phosphate based immobilization of uranium in an oxidizing bedrock aquifer. Applied Geochemistry 18, 823-843.

Jerden, J.L., Sinha, A.K., 2006. Geochemical coupling of uranium and phosphorous in soils overlying an unmined uranium deposit: Coles Hill, Virginia. Journal of Geochemical Exploration 91, 56-70.

Jerden, J.L., Sinha, A.K., Zelazny, L., 2003. Natural immobilization of uranium by phosphate mineralization in an oxidizing saprolite-soil profile: chemical weathering of the Coles Hill uranium deposit, Virginia. Chemical Geology 199, 129-157.

Levitan, D.M., Schreiber, M.E., Seal, R.R., Bodnar, R.J., Aylor, J.G., 2014. Developing protocols for geochemical baseline studies: An example from the Coles Hill uranium deposit, Virginia, USA. Applied Geochemistry.

Lyntek Inc., 2010. Preliminary economic assessment on the Coles Hill Uranium Property, Lakewood, CO.
Marline Uranium Corporation, 1983. An evaluation of uranium development in Pittsylvania County, Virginia. Marline Uranium Corporation, USA.
Marques, J.J., Schulze, D.G., Curi, N., Mertzman, S.A., 2004. Trace element geochemistry in Brazilian Cerrado soils. Geoderma 121, 31-43.
Martín-Fernández, J.A., Hron, K., Templ, M., Filzmoser, P., Palarea-Albaladejo, J., 2012. Model-based replacement of rounded zeros in compositional data: Classical and robust approaches. Computational Statistics & Data Analysis 56, 2688-2704.
McBride, M.B., 1994. Environmental chemistry of soils. Oxford University Press, New York.
Meyertons, C.T., 1963. Triassic formations of the Danville basin, Resources, V.D.o.M., Report of Investigations 6.
Mortvedt, J.J., 1995. Heavy metal contaminants in inorganic and organic fertilizers. Fertilizer research 43, 55-61.
NCDC, 2012. National Climatic Data Center. National Oceanic and Atmospheric Administration, Asheville, NC.
NRCS, 2011. Soil Survey Geographic Database (SSURGO) Database, in: Natural Resources Conservation Service, U.D.o.A. (Ed.).
Palarea-Albaladejo, J., Martín-Fernández, J.A., Buccianti, A., 2013. Compositional methods for estimating elemental concentrations below the limit of detection in practice using R. Journal of Geochemical Exploration 141, 71-77.
Pawlowsky-Glahn, V., Buccianti, A., 2011. Compositional Data Analysis: Theory and Applications. Wiley, West Sussex, UK.
Pawlowsky-Glahn, V., Egozcue, J.J., 2006. Compositional data and their analysis: an introduction. Geological Society, London, Special Publications 264, 1-10.
Pincock Allen & Holt, 1982. Geologic reserves: Coles Hill South Uranium Deposit, Pittsylvania County, Virginia. Pincock Allen & Holt, Tucson, Arizona.
Power, M., Hattori, K., Sorba, C., Potter, E., 2012. Geochemical anomalies in soils and uppermost siliciclastic units overlying the Phoenix uranium deposit, Athabasca Basin, Saskatchewan, Canada, G.S.o., Open File 7257.
R Development Core Team, 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, Software Program.
Reimann, C., Filzmoser, P., Garrett, R.G., Dutter, R., 2008. Statistical data analysis explained : Applied environmental statistics with R. John Wiley & Sons, Chichester, England; Hoboken, NJ.
Rogers, J.J.W., Ragland, P.C., 1961. Variation of thorium and uranium in selected granitic rocks. Geochimica Et Cosmochimica Acta 25, 99-109.
Rose, A.W., Wright, R.J., 1980. Geochemical exploration models for sedimentary uranium deposits. Journal of Geochemical Exploration 13, 153-179.
Smith, D., Cannon, W., Woodruff, L., Garrett, R., Klassen, R., Kilburn, J., Horton, J., 2005. Major- and trace-element concentrations in soils from two continental-scale transects of the United States and Canada, Open-File Report 2005-1253.
Smith, D.B., Cannon, W.F., Woodruff, L.G., Solano, F., Kilburn, J.E., Fey, D.L., 2013. Geochemical and mineralogical data for soils of the conterminous United States, Data Series 801.

Snelling, A.A., 1984. A soil geochemistry orientation survey for uranium at Koongarra, Northern Territory. Journal of Geochemical Exploration 22, 83-99.
Tappa, M.J., Ayuso, R.A., Bodnar, R.J., Aylor, J.G., Beard, J., Henika, W.S., Vazquez, J.A., Wooden, J.L., 2014. Age of host rocks at the Coles Hill uranium deposit, Pittsylvania County, Virginia, based on zircon U-Pb geochronology. Economic Geology 109, 513-530.
Templ, M., Hron, K., Filzmoser, P., 2013. robCompositions: Robust estimation for compositional data, R package.
US EPA, 2013. Level III ecoregions of the continental United States. U.S. Environmental Protection Agency, Corvallis, Oregon.
Vu, V.Q., 2011. ggbiplot: A ggplot2 based biplot, R package.
Zuo, R., Xia, Q., Wang, H., 2013. Compositional data analysis in the study of integrated geochemical anomalies associated with mineralization. Applied Geochemistry 28, 202-211.

CHAPTER 4. MULTIVARIATE EXPLORATORY ANALYSIS TO IDENTIFY PROCESSES CONTROLLING THE GEOCHEMISTRY OF A ROCK-SOIL-SEDIMENT SYSTEM AT AN UNDEVELOPED ORE DEPOSIT

ABSTRACT

Geochemical data from analyses of rock cores, soils, and stream sediments at the Coles Hill uranium property were examined in order to determine statistical relationships that can be used to infer related geochemical processes governing the differentiation of these reservoirs. Comparisons were made between reservoirs using exploratory data techniques, both univariate (one chemical element at a time) and multivariate (multiple chemical elements simultaneously), including linear discriminant analysis and principal component analysis. Following multivariate analysis, isometric log-ratio balances from five elements (Al, Ca, K, Na, U) were calculated, allowing the data from the solid media to be compared with aqueous samples from a small number (n=7) of groundwater samples. Results from the statistical analyses suggest that two geochemical processes, weathering and U-enriching mineralization, are responsible for the majority of the geochemical variation within the data from the site. Based on these conclusions, further avenues for more focused study of these materials are suggested.

4.1 INTRODUCTION

Ore deposits and their weathering have the potential to create distinct geochemical signatures on their surrounding environment. As geologic materials evolve, changes in mineralogy can be reflected in their elemental compositions. Relative differences between the elemental compositions can provide insights on transport of these elements within a geologic system. Many methods have been applied to the study of this evolution and the rates at which it occurs, including stable and radiogenic isotopic studies (Baskaran, 2012; Michener, 2007), modeling of mineral kinetics and thermodynamics (Oelkers and Schott, 2009; Rimstidt, 2013), mass-balance calculations (Anderson et al., 2002), and various weathering indices (Price and Velbel, 2003).

A recent trend in environmental and exploration geochemistry is the use of exploratory data analysis to infer geochemical processes represented within large datasets. Exploratory data analysis (EDA) is the use of statistical methods to transform data from lists of numbers into more easily interpretable forms, particularly graphics that enable the visualization of data patterns. This concept was developed by J.W. Tukey in the 1970s (Tukey, 1977) and has been growing ever more popular and possible with the increase in data volume generated with improved analytical tools and increased computational efficiency. When working with a large dataset, EDA methods are often a first step in data interpretation, as they provide a means of rapid and comprehensive overview.

Exploratory data analysis is more common in geochemistry than many geochemists realize. Examples of its use include boxplots, scatterplots, and ternary diagrams depicting different variables used to look for correlations between variables and differences between groups. Exploratory data analysis is often a first step in the analysis and interpretation of

geochemical data, enabling researchers to detect patterns within the data. As a result of EDA, researchers are able to focus more in-depth analyses. For instance, scatterplots of all variables in a specific dataset may suggest which ones are correlated, spurring further study of those correlations or confirming that an expected association is present. Differences in magnitude and/or distribution between groups seen through a series of boxplots can give a researcher an idea of why these differences occur. However, these common EDA methods are often limited in the number of variables that can be analyzed simultaneously and thus may cause more complex relationships to be overlooked.

Multivariate data analysis involves the simultaneous analysis of multiple variables; in geochemistry, this can mean a collection of chemical elements, such as the results of analytical runs on instruments including inductively-coupled plasma mass spectrometry, which frequently include 30-50 elements. Multivariate EDA methods exist that can be used to investigate many variables within one figure or set of figures. An advantage of multivariate data analysis is that it provides an analysis of geochemical data as a whole rather than the more traditional element-by-element basis and, often, can reduce the number of dimensions (variables) used in interpreting the system (Grunsky, 2010; Reimann et al., 2008). Many of these multivariate methods also include forms of variable reduction, allowing researchers to complete analysis and interpretation with fewer variables while retaining the ability to include information from all the original variables. The downside of such methods is that their increased complexity compared to more traditional methods can make interpretation less straightforward.

Compositional data analysis, a set of methods that are often coupled with multivariate statistics, is the analysis of data that are reported as compositions, or parts of a whole (Aitchison, 1986; Buccianti et al., 2006; Pawlowsky-Glahn and Buccianti, 2011). In geochemistry, this type

of data is common and includes elements or minerals reported in units such as percent, parts per million, micrograms per liter, etc. Because these data are relative to a whole and subject to a constant sum, the use of traditional statistical methods requires them to be transformed prior to analysis, as discussed in Chapter 3 and Appendix C as well as in a recent issue of *Journal of Geochemical Exploration* (Buccianti and Grunsky, 2014).

Exploratory data analysis techniques employing multivariate and compositional methods have been applied to the analysis of data from large continental, national, and regional geochemical surveys, often in conjunction with geospatial statistics. Examples include the Kola Ecogeochemistry Project (Reimann et al., 2008), the North American Soil Geochemical Landscapes Project (Drew et al., 2010), the National Geochemical Survey of Australia (de Caritat and Grunsky, 2013), and others. However, these techniques have also been used at smaller scales, particularly at sites with expected geochemical anomalies, such as ore deposits (Šajn and Gosar, 2014; Zuo, 2014; Zuo et al., 2013). At these sites, multivariate data analysis is often used instead of or in addition to traditional geochemical modeling (Engle and Blondes, 2014; Gallo and Buccianti, 2013).

This study applies exploratory multivariate compositional data analysis to investigate the geochemical processes occurring at an undeveloped uranium deposit at Coles Hill, Virginia, USA. The goals of this study are to use data collected from a variety of geologic media to evaluate whether and how these media are geochemically distinct, determine which elements most strongly distinguish the materials, and identify possible processes controlling how these differences are related to the mineralization and weathering of the deposit for further study.

4.2 METHODS

4.2.1 DATA SOURCES

Geochemical data from the Coles Hill uranium deposit were compiled from a number of sources, including the studies detailed in earlier chapters. Stream sediment data collection and analysis are detailed in Chapter 2 (Levitan et al., 2014), and soil data collection and analysis are described in Chapter 3. Well water samples were provided by a Virginia Uranium, Inc., and spring water samples were collected following the same protocols as stream waters described in Chapter 2. Well and spring waters were analyzed by the same procedures as aqueous samples described in Chapter 2. Data from analyses of core samples were provided by Dr. James Beard of the Virginia Museum of Natural History. These data were divided into "ore" and "[other] rock" using a threshold value of 212 mg/kg U, which corresponds to the deposit cutoff grade of 0.025 wt. % U_3O_8 (Gannon et al., 2011; Levitan et al., 2014; Tappa et al., 2014).

Solid media were divided into five groups based on medium. These groups were labeled ore (O; n=21), rock (R; n=36), C horizon soils (C; n=25), A horizon soils (A; n=25), and stream sediments (S; n=42). A schematic of the Coles Hill system showing the relative locations of these various media is shown in Figure 4-1.

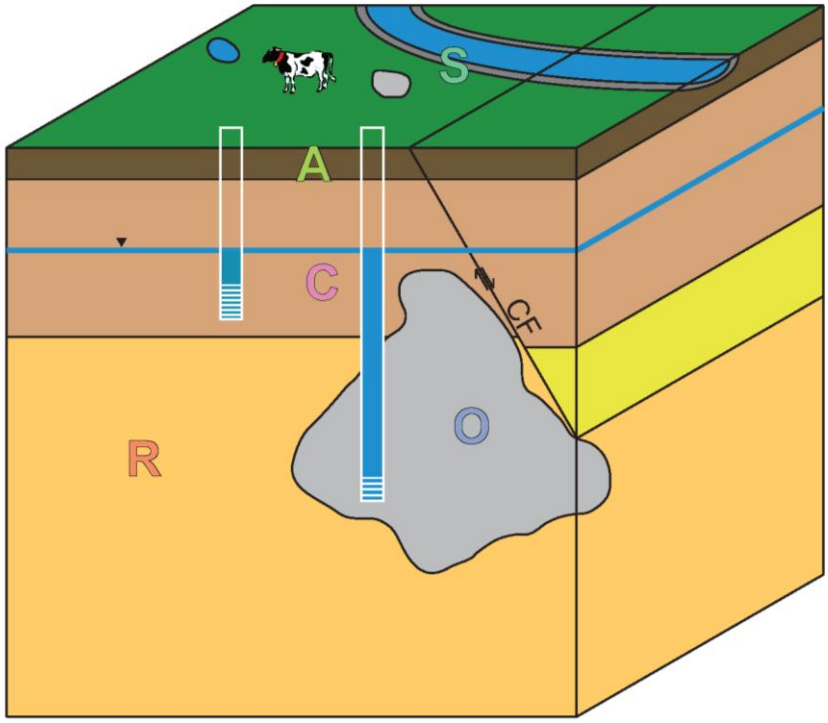

Figure 4-1: Schematic cross-section of the geoenvironmental media [ore (O), non-ore rock (R), C horizon soils (C), A horizon soils (A), and stream sediments (S)] at Coles Hill. Descriptions of the ore deposit, including its relationship to the Chatham Fault (CF), can be found in (Aylor et al., 2014; Dahlkamp, 2010; Gannon et al., 2011; Jerden and Sinha, 2003; Levitan et al., 2014; Tappa et al., 2014)

4.2.2 DATA PROCESSING

Parameters selected for data analysis were those elements associated with greater than 50% detected measurements in each solid medium, a total of 23 elements. Major elements that did not meet this criterion include Si, C, and S, which were not analyzed for in all datasets. Non-detect results for the 23 selected elements were simulated using the isometric log-ratio substitution method described in Chapter 3 (Martín-Fernández et al., 2012; Palarea-Albaladejo et al., 2013; Templ et al., 2013). If concentrations were reported as oxides, elemental concentrations were calculated from these data and used in the statistical analysis. All

concentrations were converted to parts per million (ppm; mg/kg or mg/L) for multivariate analysis.

Unless otherwise noted, statistical analyses were conducted on data opened by centered log-ratio (clr) transformation. The clr transformation is the natural log of the ratio of a particular element to the geometric mean of all elements, calculated independently for each sample (see Chapter 3 and Appendix C for a more in-depth discussion of log-ratio transformation and a list of related references).. To compensate for the range of precision and spread of the different elements, data were standardized by being centered (linearly transformed to set the overall mean to zero) and scaled (divided by standard deviation to set all variables to a comparable range).

4.2.3 STATISTICAL ANALYSIS

The first data analysis step in this study was to confirm that the groups' samples were geochemically distinct within the dataset acquired for the study, as this dataset did not include information on some common distinguishing geochemical factors between rocks, soils, and sediment, such as mineralogy, silica content, organic and carbonate carbon content, and redox species. To distinguish among the solid media groups, linear discriminant analysis (LDA) was conducted on the full suite of solid samples. Linear discriminant analysis creates linear functions that scale the variables within the dataset to maximize the separation between user-defined groups. The number of these functions (linear discriminants, LDs) is the number of groups minus one. The LD scaling functions can be used to classify additional samples into the most similar group. Ideally, if the original samples, or training data, are classified by the LDA results, the group membership predicted by the LDA will be the same as the user-defined group membership. The number of samples predicted to be in the wrong group divided by the total

number of samples is called the misclassification rate. Group membership was then predicted for all samples using the LDA results, and a misclassification rate was calculated. Mann-Whitney-Wilcoxon rank sum tests were used to determine whether group values were significantly different between each group. Scalings (the coefficient for each element for each LD function) for scaled clr-transformed data were compared by element.

Overall group differences were assessed on untransformed data by the Kruskal-Wallis rank sum test for each element. To evaluate between-group differences and to determine which elements drove the classification within the LDA, clr group centers were calculated and compared for each element. Group centers are standardized group means, meaning that they have been centered and scaled.

Principal component analysis (PCA) was conducted on solid media to gain a better understanding of how different elements within the Coles Hill system related to each other and to infer the dominant geochemical processes that may control these relationships. Links, or segments between loading vector tips, were visually determined. Within biplots of clr-transformed data, links can highlight patterns within the dataset, and these patterns can be inferred as representative of processes.

Based on the exploratory analyses, a subset of the elements that were determined to most strongly differentiate the various groups was used to develop isometric log-ratio (ilr) balances representing processes within the system (Egozcue and Pawlowsky-Glahn, 2005; Engle and Rowan, 2013b). Isometric log-ratio balances are log-ratios of elements calculated based on sequential divisions between elements. See Section C.1.1 in Appendix C for a mathematical explanation of ilr balances. The balances can be interpreted as processes, and correlations between them can be examined. Statistical differences between groups were again assessed using

the Kruskal-Wallis and Mann-Whitney-Wilcoxon tests, and correlations were evaluated using Pearson's *r*.

All data analysis was conducted in the R statistical environment (R Development Core Team, 2011). Some analyses and figures were done using code modified from (Reimann et al., 2008).

4.3 RESULTS AND DISCUSSION

4.3.1 ARE THE GROUPS GEOCHEMICALLY DIFFERENT?

Linear discriminant results show that the five groups can be clearly distinguished (Figure 4-2). The misclassification rate was less than 2%, with two of the 149 samples being assigned to new groups by LDA. The misclassified samples included one C horizon soil sample classified as a stream sediment and one rock sample classified as ore.

Because data were divided into five groups, four LD functions were calculated. The first LD accounts for 54.2% of the separation of the groups. However, LD1 does not significantly divide the ore and non-ore rock samples. The second LD accounts for 19.5% of the separation among groups, and all groups are significantly different for LD2. The third LD accounts for 16.8% of the separation among groups and does not significantly differentiate the C horizon soils and non-ore rocks. The fourth LD accounts for 9.50% of the separation and does not significantly distinguish the A horizon soils and the non-ore rock.

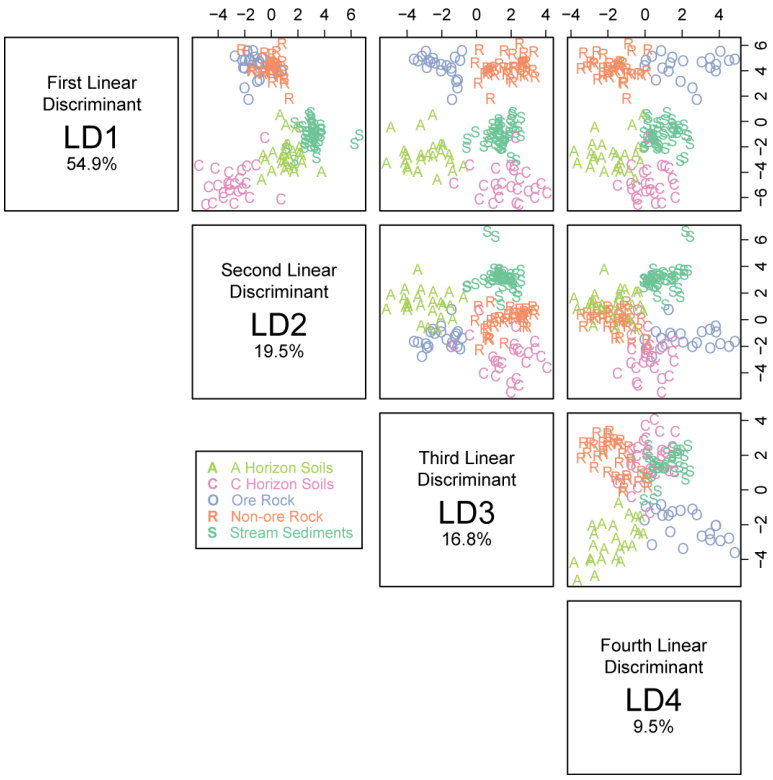

Figure 4-2: Scatterplot matrix of linear discriminant scores. Plotting color and symbol indicates original group membership. Percentage separation accounted for is given for each linear discriminant.

4.3.2 WHAT ELEMENTS MOST STRONGLY DIFFERENTIATE AMONG GROUPS?

Significant differences ($p << 0.05$ by Kruskal-Wallis test) among groups were found for untransformed concentrations of all elements analyzed. Figure 4-3 shows ordered test statistic results from these tests. Higher test statistic values indicate greater statistical difference among groups. Calcium, Na, P, and U yielded the highest values, whereas K, Tl, Rb, and Co yielded the lowest.

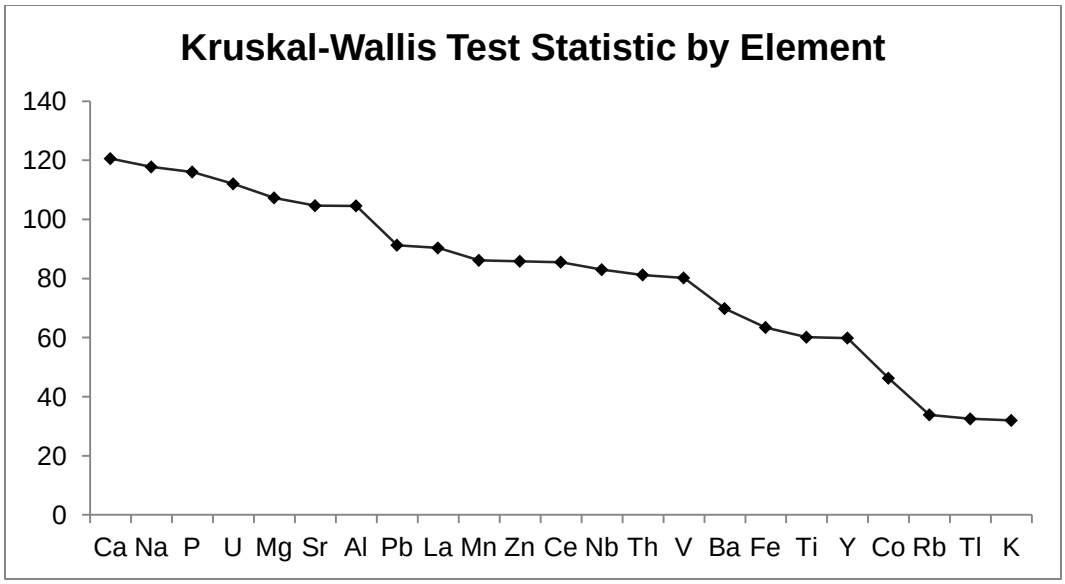

Figure 4-3: Ranked plot of Kruskal-Wallis test statistics by element. Analyses were conducted on untransformed data. All test scores were associated with significant group differences.

The scaled magnitude of linear discriminant scalings (coefficients of the LD functions) were compared by element (Figure 4-4). For LD1, which separates rock from soil/sediment samples, Co, Mg, and Ca have the highest magnitude scaling, suggesting that differences in the relative amounts of these elements may be of particular importance in differentiating the rocks from the soil/sediment samples. For LD2, in which all groups are statistically significantly different, the coefficients with the highest magnitudes are Ca, U, Mn, and Al. For LD3, Fe, Rb, and U have the highest magnitude scaling, and for LD4, Zn, U, and Ca have the highest. The repetition of Ca and U among the highest magnitude coefficients in addition to their high Kruskal-Wallis test scores suggests that these elements are of particular importance in differentiating groups within this system. The high scaling of elements with lower Kruskal-Wallis scores, such as Co and Rb, may be related to the difference in their importance as untransformed concentrations as opposed to compositional variables.

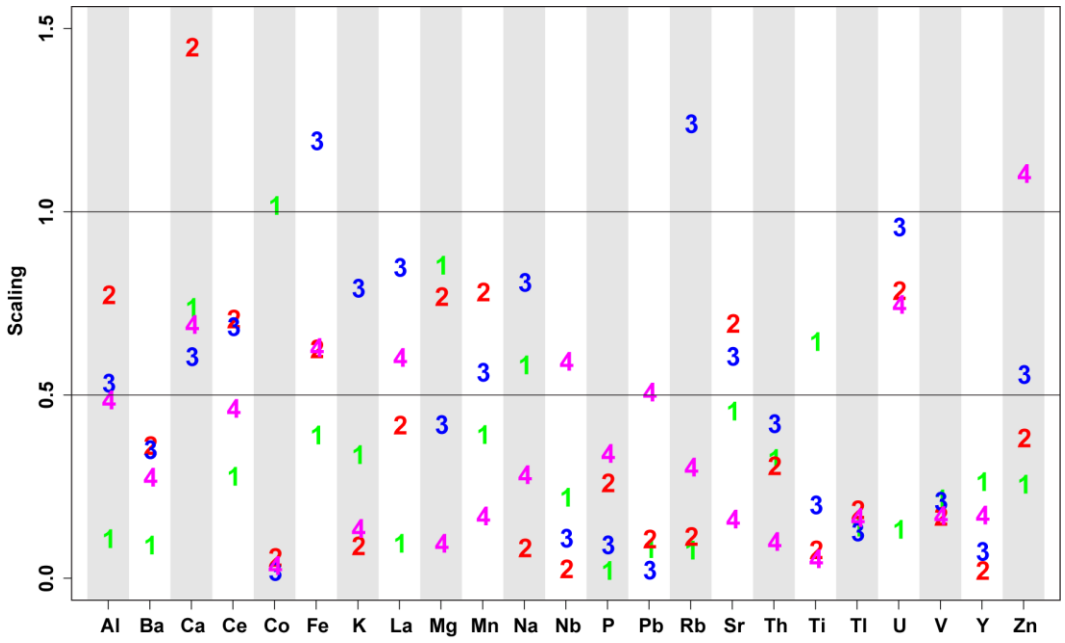

Figure 4-4: Plot of absolute values of scaled linear discriminant scalings by element. Number indicates LD associated with each value.

Group centers (means of clr-transformed data) by element were compared (Figure 4-5) to assess the relative importance of each element in distinguishing between groups. Higher values indicate higher relative concentrations of a given element in a particular medium and greater influence in the discriminant analysis. However, because these values are transformed and scaled, greater center values do not necessarily correspond to greater absolute values, and numerical values between elements cannot be compared. Notable results include 1) positive P, Pb, and U, likely related due to the presence of high-U phosphates in the ore and Pb being a U decay product; Jerden and Sinha (2003) and negative K, Rb, and Tl in the ore rock, with large gaps between ore and other media for the P and U; 2) positive Mg and negative Pb in the non-ore rock, with centers for this group generally being intermediate; 3) positive Ce, La, and Nb and negative Na in the soils and relatively large separation between the two horizons in Al, Ca, Fe, Mn, and V (consistent with findings in Chapter 3); and 4) positive K in the stream sediments,

with stream sediment centers being generally intermediate. In many cases, the ore rock and the C horizon soils are opposite extremes, suggesting that overall, these two reservoirs might be the most different geochemically, whereas the intermediate values for the non-ore rock and the stream sediments suggest that these groups might be overall geochemically "average." In addition, the ordering of the centers is consistent for some elements, which may suggest that these elements behave similarly in their geochemistry. For example, Ce and La have similar C>A positive values with much lower and similar R, S, and O values. Lanthanum and Ce are adjacent on the periodic table and were shown in Chapter 3 to behave similarly in the soils at Coles Hill. Other examples of similar ordering include K and Rb, which have much lower O centers than other media; Ca and Na, which have centers R≈O>S>A>C.

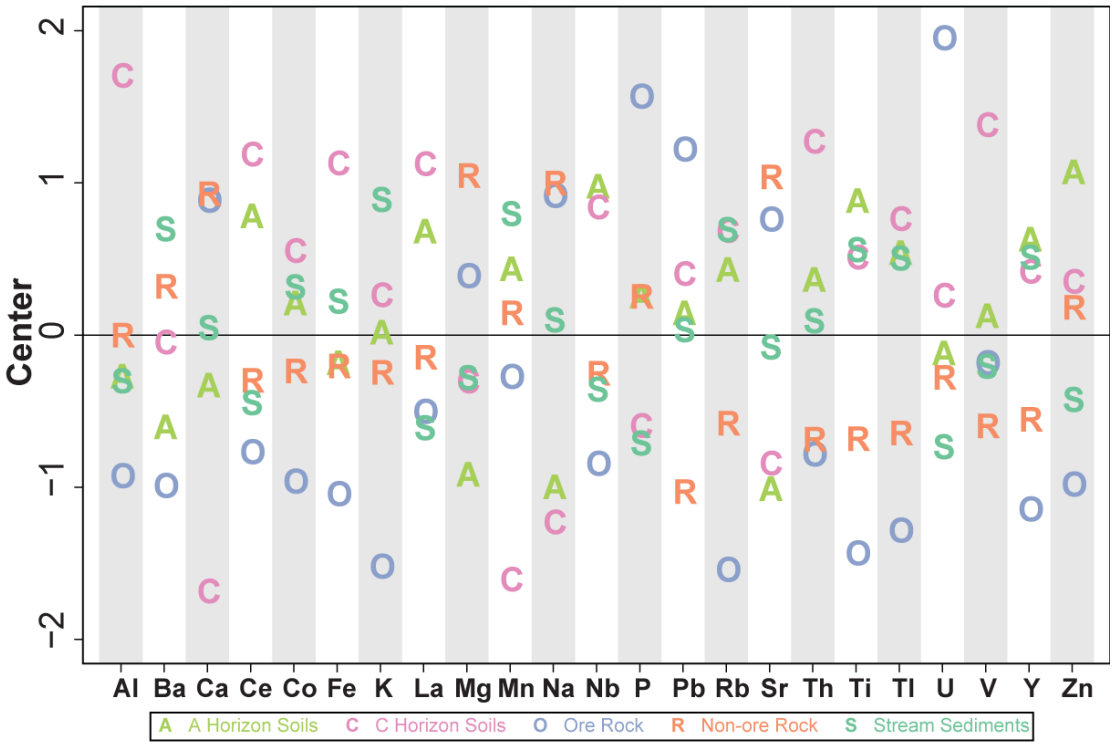

Figure 4-5: Element plot of scaled group centers. Plotting symbol and color indicate group membership: ore (O), rock (R), C horizon soils (C), A horizon soils (A), and stream sediments (S).

4.3.3 HOW DO THESE GROUP DIFFERENCES RELATE TO GEOCHEMICAL PROCESSES OCCURRING AT THE SITE?

Principal component analysis showed distinctive separation of the media groups by the first two PCs, which explained 74.3% of the variation in the data (Figure 4-6). The first principal component (PC) yielded negative values for the ore and non-ore rock and positive values for the stream sediments and soils. The second PC was mostly positive for the ore rock and the C horizon soils, intermediate for the A horizon soils, intermediate to negative for the unmineralized rock, and negative for the stream sediments. A screeplot with the percent of data variation explained by each of the 23 PCs is shown in Figure 4-7. The flattening of this plot after the first two PCs substantiates the suggestion that this system can be largely interpreted using PC1 and PC2.

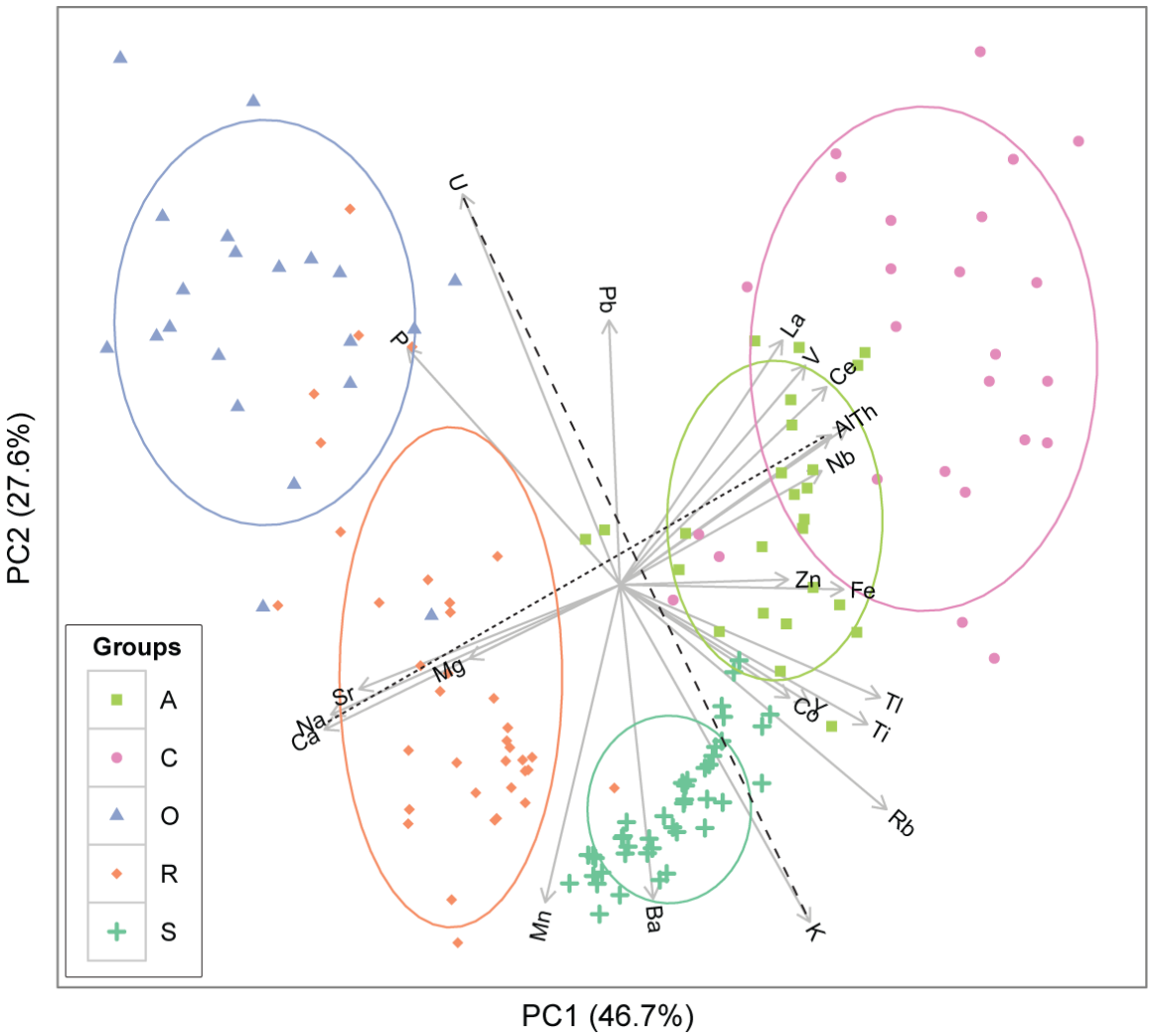

Figure 4-6: Principal component analysis biplot showing scaled results for the first two PCs. Values in parentheses represent the amount of variation explained by each PC. Plotting symbol and color represent group membership. 68% confidence concentration ellipses are shown for each group. Loading vectors originate from (0,0) and are labeled by element. Two links are shown as black dashed lines.

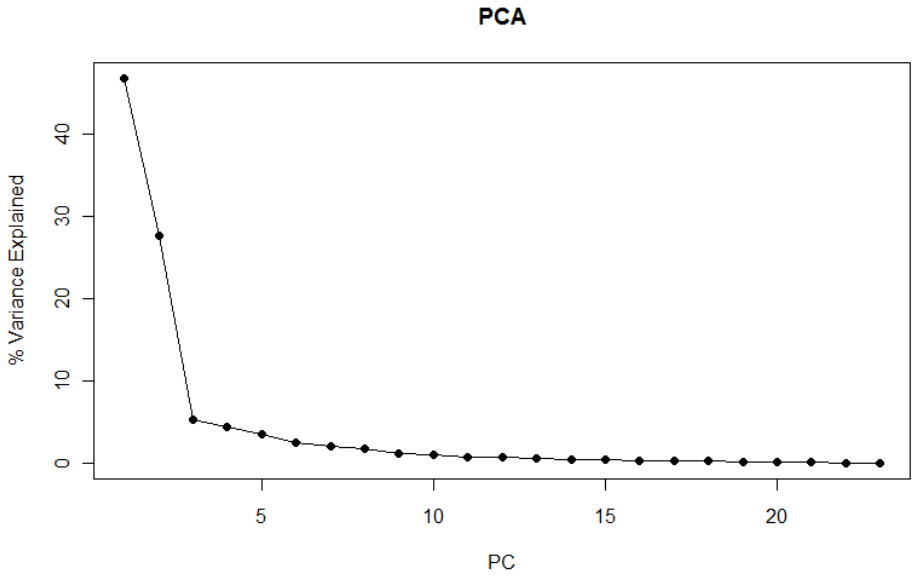

Figure 4-7: Principal component analysis screeplot. The x axis shows PCs in order, and the y axis shows the percent of variance explained by each PC.

Two perpendicular links were visually assessed and interpreted as representing two major geochemical processes occurring at the site (Figure 4-6). The first link is between Na/Ca and Al. Variation in this link is largely in PC1 and is interpreted to be weathering-related, as it trends between highly mobile alkali and alkali earth metals (Na, Ca, Sr, and Mg) and refractory Al and Th (Langmuir, 1997). A common mineral weathering reaction in keeping with this link is the weathering of albite to kaolinite, $2NaAlSi_3O_8+9H_2O+2H^+ \rightarrow Al_2Si_2O_5(OH)_4+2Na^++4H_4SiO_4^0$, in which acidic water leaches Na from the albite into the dissolved fraction, while Al remains in the solid phase. Mobilization of alkali earth metals from carbonate minerals is also a potential reaction represented by this link; calcite has been reported in the rock at Coles Hill (Aylor et al., 2014; Jerden, 2001), though carbonate concentration data are not available, but was not detected in the mineralogical analysis of soils described in Chapter 3, and unpublished soils data results show minimal (≤0.05 wt.%) carbonate carbon. Notably, the alkali (earth) elements had similar patterns for the ordering of the group centers, with the patterns being reversed for the alkali

(earth) elements versus Al and Th. Potassium, Rb, and Ba are alkali (earth) metals that behave similarly to each other but not to the previously discussed elements of the same groups; this difference is discussed further below.

The second link is between U and K and is interpreted to be related to the enrichment of uranium as part of the mineralization of the deposit. Mineralized (ore) rocks are defined as those with higher U concentrations, whereas the unmineralized rocks are generally felsic granites and metagranites (Aylor et al., 2014; Tappa et al., 2014). This differentiation results in variation in U/K values, which is represented by this link. While there is some variation in the angle of their loading vectors, several other elements show similar patterns to U and K in the PCA analysis. Lead and P are at low angles to the link in the U direction, while Ba and Rb are at low angles to the link in the K direction. These relationships suggest that the Ba/U and Rb/U ratios are correlated with the K/U ratio, and that the Pb/K and P/K ratios are correlated with U/K, which might indicate similar behavior. Within the Coles Hill system, P is associated with U in several of the media groups, such as in high-U apatites in the ore and in U-(LREE)-phosphates or associated with sorbed phosphates in the soils. Three Pb isotopes occur as part of the U-238 decay chain, and so an association of U and Pb might be expected. However, Th is also a U decay product but is not correlated with U in these materials, and no association has been found between U and Pb in the more weathered media. Chapter 3 discussed potential reasons for a lack of U-Th correlation, suggesting that non-uranogenic Th may play a role, though Pb is also a decay product of Th. Thus, the relationship of U and Pb appears somewhat enigmatic. As noted above, K, Rb, and Ba are alkali (earth) metals that show different behavior within the PCA and other statistical analyses than other elements of those groups. Possible geochemical reasons for differences in mobility of these elements include the rarity of carbonates of K, Ba, and Rb, as

well as an increased likelihood of their higher chemical reactivity and retention in weathering phases such as illite.

Because the two links are orthogonal, they are likely to represent independent processes. It is important to note, however, that the relationships between the elements defining these links are not necessarily independent or unrelated; as stated above, the first two PCs explain 74.3% of the data variation (Figure 4-7), and sample data vary along both links. This is notable because of suggestions of Na- metasomatism in conjunction with the U mineralization (Jerden, 2001); if this were the case, a more complex or dependent K-Na-U behavior would be expected in material related to the deposit. However, more recent descriptions of the deposit describe the Na-metasomatism as more uniform over a larger region (Aylor et al., 2014), which is consistent with the independence suggested by the PCA links in this study.

When the media groups are assessed with respect to the first link, the rock samples (both ore and non-ore rock) fall toward the mobile element end of the link, suggesting that these media may be more enriched in these elements and thus less weathered. The stream sediments fall towards the middle of this link, suggesting either moderate weathering or low importance of the degree of weathering in distinguishing these samples from other groups. The soil samples fall toward the Al end of the link, indicating that they are more weathered than the other samples. Unexpectedly, the C horizon soils plot further along this link than the A horizon, suggesting that they are more weathered. However, as weathering is interpreted in these data as the balance between "mobile" and "refractory" elements, alternative sources of these elements may confound this distinction. In the case of the soils, samples were collected from an area that has been used for several centuries for crop and grazing agriculture, including the application of various amendments, such as fertilizers and pesticides, which may contribute mobile elements to the

upper layers of the soils. It is also possible that the end of this link is representative of winnowing, or the sorting of materials based on density during weathering (Grunsky et al., 2009), and so the weathering link represents not only element mobility during mineral dissolution, but also sorting during physical transport.

The second link, between U and K, is interpreted to represent U enrichment at the expense of felsic host rock. At the U end of this link are the ore rock samples. Intermediate are the soils and some of the non-ore rock, which displays a large spread with respect to this link. At the felsic end of this link are the stream sediment samples. Because the soil samples were collected directly over the deposit with some extension into areas overlying unmineralized rock (see Chapter 3 for site map), their relationship to the second link is likely reflective of a mixing of weathered materials from both U-rich and U-poor parent rock. Stream sediments, however, were collected from a much wider area (see Chapter 2). Consequently, these samples are likely to be much more geologically diluted by other, particularly unmineralized, parent rock types.

Based on these links and on the significance and variability among groups, a subset of elements was selected for the creation of isometric log-ratio (ilr) balances (Egozcue and Pawlowsky-Glahn, 2005), which have been used in other studies to represent geochemical processes (Engle and Rowan, 2013a; Engle and Blondes, 2014). The elements selected were those used to delineate the PCA links: Al, Ca, K, Na, and U. Table 4-1shows the sequential binary partition scheme used to define the balances (see Appendix C). Balance ilr1 represents the division between weathering and mineralization processes. Balance ilr2 represents the mineralization interpreted from Link 2. Balance ilr3 represents the weathering interpreted from Link 1. Balance ilr4 represents differences in the behavior of Na and Ca and is not used further in this study as values are not significantly different between groups.

Table 4-1: Sequential binary partition matrix for ilr balances. A "+" indicates that an element is included in the numerator of a balance, a "–" that it is included in the denominator, and a 0 that it is not included.

Balance	Al	Ca	K	Na	U
ilr1	-	-	+	-	+
ilr2	0	0	+	0	-
ilr3	+	-	0	-	0
ilr4	0	+	0	-	0

Overall, the first three ilr balances are significantly different (p<0.05 for Kruskal-Wallis test) among the media groups (Figure 4-8). When compared group-by-group, ilr1 was significantly different (p<0.05 for Mann-Whitney-Wilcoxon test) between all pairings, ilr2 was significantly different for all pairings except the soil horizons with each other and with the non-ore rock, and ilr3 was significantly different for all pairings. Therefore, the ilr balances provide a way to separate the media groups by process (e.g., weathering, mineralization).

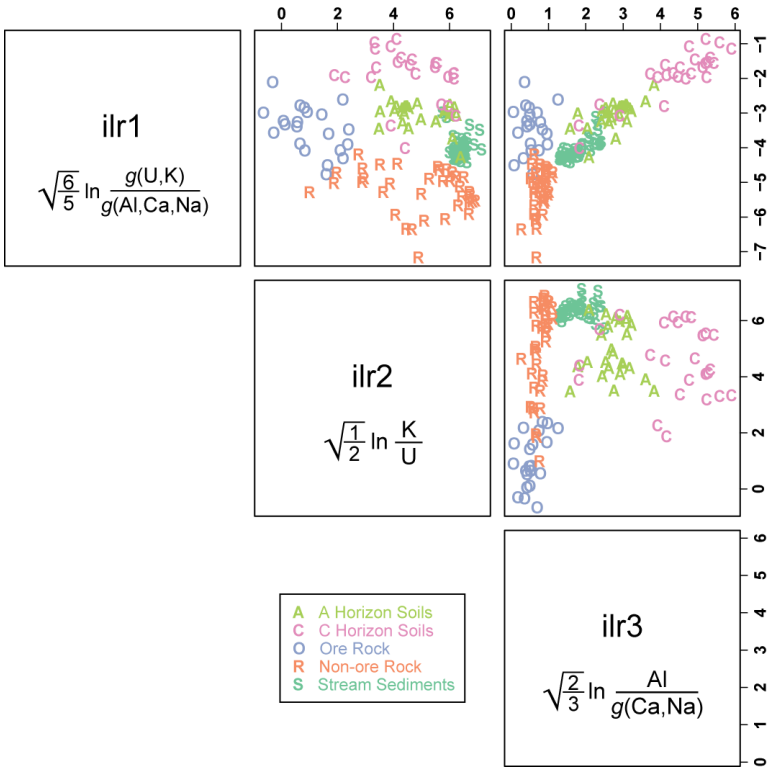

Figure 4-8: Scatterplot matrix of ilr balances for solid samples. Plotting symbol and color indicate group membership.

In addition to the ilr balance values differing among groups, a strong correlation (r^2=0.90) exists between ilr1 and ilr3 in the soils and sediment. Weaker correlations exist among these groups between ilr1 and ilr2 and between ilr2 and ilr3. Weak to moderate correlations also exist between balances for the rock samples, but only between ilr1 and ilr2 is the slope of this relationship significantly different from 0. The relationships between these variables within the sample data suggest that the inferred processes of weathering and mineralization may not be geochemically independent; the relationship of K to U may not be independent of the relationship of Al to Ca and Na. For example, the C horizon samples have the greatest ilr3 values (Al relative to Ca and Na), but they also have generally higher ilr1 (U and K relative to Na, Ca, and Al) and lower ilr2 (K relative to U) values than A horizon soils and sediment. Because the C horizon soils tend to be spatially closer to the mineralized zone but are also depleted in certain readily mobile elements, these two processes are somewhat conflated, thus resulting in a correlation between the balances.

4.3.4 THE ROLE OF WATER

The ilr balances were calculated for eight groundwater samples, two from wells drilled into and just above the deposit, one from another nearby well away from the deposit, and five from surface springs near the deposit (Figure 4-9). These water samples will be discussed in two groups: the deposit wells (drawn from the ore rock) and the shallow groundwater (drawn from the soils and unmineralized rock), which includes the springs and the well most distant from the deposit. The deposit well samples have higher ilr1, lower ilr2, and similar ilr3 values relative to the shallow groundwater. This suggests U mineralization creates a stronger signature on the deposit waters than does the relative mobility of elements (ilr1), that these waters are

geochemically more closely associated with mineralization than the shallow groundwaters (ilr2), and that the waters are fairly similar in their transport of labile elements relative to elements from less soluble sources (ilr3).

When compared with the ilr balances of the solid media, some similarities are observed. For ilr1 and ilr2, the values for the groundwater samples are fairly similar to the solid media from which they were drawn. For ilr2, this means that the K/U ratios are a similar order of magnitude, which may imply similar dissolution rates for minerals containing these two elements in the aquifer material. However, the ilr3 balances are much lower for the aqueous samples relative to the solids, likely as a result of mobility differences between Al and Ca/Na in the aquifer material. The implications of this pattern are that Na and Ca are indeed more mobile and/or easily dissolved than Al and that groundwater is a likely pathway for the transport of these elements after weathering from the solid phase.

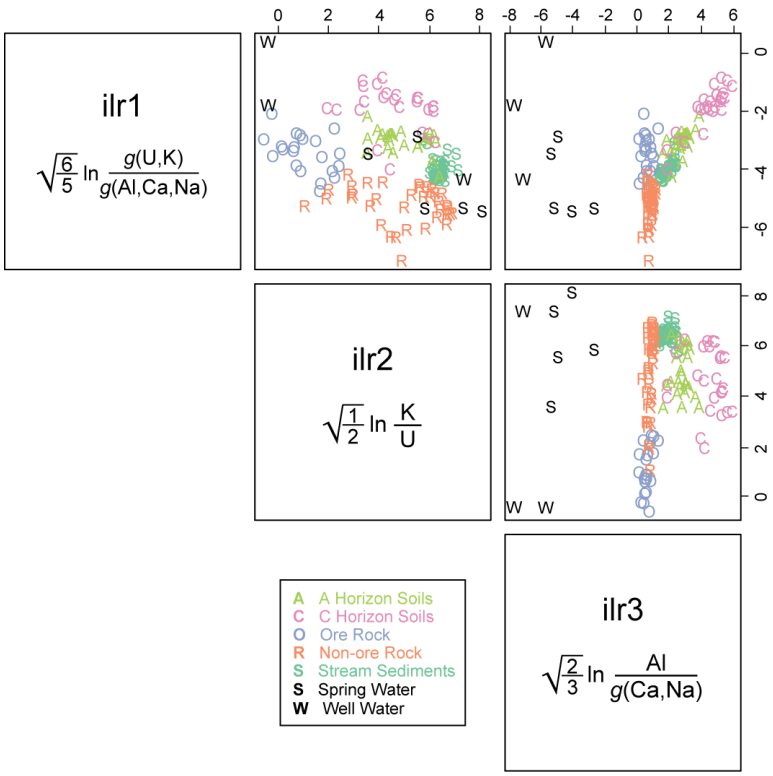

Figure 4-9: Scatterplot matrix of isometric log-ratio (ilr) balances for solid and aqueous samples. Plotting symbol and color indicate group membership; W represents well water samples, and S represents spring water samples.

4.4 IMPLICATIONS FOR FUTURE RESEARCH

The exploratory data analysis undertaken in this study suggested that two sets of element relationships were responsible for much of the geochemical differences among the various geologic reservoirs in the area around Coles Hill. These relationships were summarized as Ca and Na vs. Al and U vs. K and broadly interpreted to represent weathering and mineralization, respectively. The same relationships were investigated in groundwater samples, and results showed similar values for U/K but lower values for Al/√(CaNa). This difference is likely attributable to incongruent mineral dissolution reactions such as plagioclase dissolution to kaolinite. Geochemical thermodynamic modeling can be used to further investigate this hypothesis by calculating expected groundwater concentrations and mineral weathering products

based on the dissolution of the minerals present (described in Chapter 3; Aylor et al., 2014; Jerden, 2001) in the various geologic media. In addition, this modeling can provide insight into elements such as Fe and Ti which showed variation with respect to both links by identifying possible primary and secondary mineral and aqueous phases.

4.5 CONCLUSIONS

Through the use of multivariate compositional data analysis, large-scale patterns in geochemical data can be analyzed. By transforming data and using various methods of dimension reduction, such and LDA and PCA, these patterns can be observed and compared across different media. The most significant variables can be identified and used to further interpret processes.

In the case of the solid media collected from the Coles Hill area for this study, weathering of certain labile elements (e.g., Ca, Na) relative to more residual elements (e.g., Al, Th) and the changes in composition associated with the mineralization of the ore deposit were identified as the major processes driving geochemical differentiation among the different media sampled. The differentiation was simplified by identifying five elements (Al, Ca, K, Na, U) most characteristic of these processes and developing log-ratio transformed variables based on these elements in order to represent geochemical processes.

By using scale-invariant techniques such as compositional data analysis, data from different media, not only solid but also aqueous, could be combined. Using data patterns and knowledge of background geochemistry, processes occurring at the site within these media could

be inferred. Geochemical evidence of water serving as a pathway for mobilization and weathering of certain elements was observed using this type of data analysis.

Future work building on this initial exploratory data analysis will include the integration of geochemical modeling to enable comparison of actual and predicted geochemical signatures based on the inferences made in this initial study. Reactions and correlations will be calculated to determine whether the hypotheses about the importance of weathering and uranium mineralization processes in differentiating the geochemical signatures of the various media can be confirmed through more traditional geochemical methods.

REFERENCES

Aitchison, J., 1986. The statistical analysis of compositional data. Chapman and Hall, London; New York.

Anderson, S.P., Dietrich, W.E., Brimhall, G.H., 2002. Weathering profiles, mass-balance analysis, and rates of solute loss: Linkages between weathering and erosion in a small, steep catchment. Geological Society of America Bulletin 114, 1143-1158.

Aylor, J., Jr., Bodnar, R.J., Beard, J., 2014. Geology of the Coles Hill uranium deposit, Virginia Piedmont. GSA Field Guide 35, 271-284.

Baskaran, M., 2012. Handbook of Environmental Isotope Geochemistry: Vol I. Springer Berlin Heidelberg, Berlin, Heidelberg.

Buccianti, A., Grunsky, E., 2014. Compositional data analysis in geochemistry: Are we sure to see what really occurs during natural processes? Journal of Geochemical Exploration 141, 1-5.

Buccianti, A., Mateu-Figueras, G., Pawlowsky-Glahn, V., 2006. Compositional Data Analysis in the Geosciences, Geological Society Special Publication 264. Geological Society of London, London.

Dahlkamp, F.J., 2010. Uranium Deposits of the World. Springer Berlin Heidelberg.

de Caritat, P., Grunsky, E.C., 2013. Defining element associations and inferring geological processes from total element concentrations in Australian catchment outlet sediments: Multivariate analysis of continental-scale geochemical data. Applied Geochemistry 33, 104-126.

Drew, L.J., Grunsky, E.C., Sutphin, D.M., Woodruff, L.G., 2010. Multivariate analysis of the geochemistry and mineralogy of soils along two continental-scale transects in North America. Science of the Total Environment 409, 218-227.

Egozcue, J.J., Pawlowsky-Glahn, V., 2005. Groups of parts and their balances in compositional data analysis. Mathematical Geology 37, 795-828.

Engle, M., Rowan, E., 2013a. Interpretation of Na–Cl–Br Systematics in Sedimentary Basin Brines: Comparison of Concentration, Element Ratio, and Isometric Log-ratio Approaches. Mathematical Geosciences 45, 87-101.

Engle, M.A., Blondes, M.S., 2014. Linking compositional data analysis with thermodynamic geochemical modeling: Oilfield brines from the Permian Basin, USA. Journal of Geochemical Exploration 141, 61-70.

Engle, M.A., Rowan, E.L., 2013b. Interpretation of Na–Cl–Br Systematics in Sedimentary Basin Brines: Comparison of Concentration, Element Ratio, and Isometric Log-ratio Approaches. Mathematical Geosciences 45, 87-101.

Gallo, M., Buccianti, A., 2013. Weighted principal component analysis for compositional data: application example for the water chemistry of the Arno river (Tuscany, central Italy). Environmetrics 24, 269-277.

Gannon, J.P., Burbey, T.J., Bodnar, R.J., Aylor, J.G., 2011. Geophysical and geochemical characterization of the groundwater system and the role of Chatham Fault in groundwater movement at the Coles Hill uranium deposit, Virginia, USA. Hydrogeology Journal 20, 1-16.

Grunsky, E.C., 2010. The interpretation of geochemical survey data. Geochemistry: Exploration, Environment, Analysis 10, 27-74.

Grunsky, E.C., Drew, L.J., Sutphin, D.M., 2009. Process recognition in multi-element soil and stream-sediment geochemical data. Applied Geochemistry 24, 1602-1616.
Jerden, J.L., 2001. Origin of uranium mineralization at Coles Hill Virginia (USA) and its natural attenuation within an oxidizing rock-soil-ground water system,PhD Thesis.
Jerden, J.L., Sinha, A.K., 2003. Phosphate based immobilization of uranium in an oxidizing bedrock aquifer. Applied Geochemistry 18, 823-843.
Langmuir, D., 1997. Aqueous environmental geochemistry. Prentice Hall, Upper Saddle River, N.J.
Levitan, D.M., Schreiber, M.E., Seal II, R.R., Bodnar, R.J., Aylor Jr, J.G., 2014. Developing protocols for geochemical baseline studies: An example from the Coles Hill uranium deposit, Virginia, USA. Applied Geochemistry 43, 88-100.
Martín-Fernández, J.A., Hron, K., Templ, M., Filzmoser, P., Palarea-Albaladejo, J., 2012. Model-based replacement of rounded zeros in compositional data: Classical and robust approaches. Computational Statistics & Data Analysis 56, 2688-2704.
Michener, R.H.L.K., 2007. Stable isotopes in ecology and environmental science, 2 ed. Blackwell Publishing, Malden, MA.
Oelkers, E.H., Schott, J., 2009. Thermodynamics and Kinetics of Water-Rock Interaction, Reviews in Mineralogy and Geochemistry.
Palarea-Albaladejo, J., Martín-Fernández, J.A., Buccianti, A., 2013. Compositional methods for estimating elemental concentrations below the limit of detection in practice using R. Journal of Geochemical Exploration 141, 71-77.
Pawlowsky-Glahn, V., Buccianti, A., 2011. Compositional Data Analysis: Theory and Applications. Wiley, West Sussex, UK.
Price, J.R., Velbel, M.A., 2003. Chemical weathering indices applied to weathering profiles developed on heterogeneous felsic metamorphic parent rocks. Chemical Geology 202, 397-416.
R Development Core Team, 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, Software Program.
Reimann, C., Filzmoser, P., Garrett, R.G., Dutter, R., 2008. Statistical data analysis explained : Applied environmental statistics with R. John Wiley & Sons, Chichester, England; Hoboken, NJ.
Rimstidt, J.D., 2013. Geochemical rate models : an introduction to geochemical kinetics. Cambridge University Press, New York.
Šajn, R., Gosar, M., 2014. Multivariate statistical approach to identify metal sources in Litija area (Slovenia). Journal of Geochemical Exploration 138, 8-21.
Tappa, M.J., Ayuso, R.A., Bodnar, R.J., Aylor, J.G., Beard, J., Henika, W.S., Vazquez, J.A., Wooden, J.L., 2014. Age of host rocks at the Coles Hill uranium deposit, Pittsylvania County, Virginia, based on zircon U-Pb geochronology. Economic Geology 109, 513-530.
Templ, M., Hron, K., Filzmoser, P., 2013. robCompositions: Robust estimation for compositional data, R package.
Tukey, J., 1977. Exploratory Data Analysis. Addison-Wesley, Reading, MA.
Zuo, R., 2014. Identification of geochemical anomalies associated with mineralization in the Fanshan district, Fujian, China. Journal of Geochemical Exploration 139, 170-176.

Zuo, R., Xia, Q., Wang, H., 2013. Compositional data analysis in the study of integrated geochemical anomalies associated with mineralization. Applied Geochemistry 28, 202-211.

APPENDIX A. SUPPLEMENTARY INFORMATION FOR CHAPTER 3

Table A-1: *p*-values for enrichment factor (EF) median comparisons between numerator (*X*/Al at Coles Hill) and denominator (*X*/Al in reference soils) by the Mann-Whitney-Wilcoxon test (see Figure 3-5). Statistically significant ($p<0.05/36$, or 0.0014) values are highlighted.

	A Horizon EF *p*	C Horizon EF *p*
As	0.205	0.026
Ba	0.575	0.031
Be	0.422	0.240
Bi	0.000	0.442
Ca	0.014	0.200
Ce	0.000	0.000
Co	0.004	0.022
Cr	0.000	0.001
Cu	0.031	0.000
Fe	0.004	0.000
Ga	0.007	0.091
K	0.274	0.690
La	0.000	0.000
Li	0.000	0.607
Mg	0.000	0.032
Mn	0.001	0.000
Mo	0.034	0.235
Na	0.166	0.295
Nb	0.000	0.000
Ni	0.031	0.025
P	0.065	0.192
Pb	0.076	0.019
Rb	0.733	0.567
Sb	0.133	0.940
Sc	0.004	0.000
Sn	0.000	0.169
Sr	0.254	0.124
Th	0.000	0.002
Ti	0.059	0.786
Tl	0.977	0.931
U	0.000	0.000
V	0.435	0.269
W	0.001	0.095
Y	0.259	0.995
Zn	0.054	0.000
C_{org}	0.284	0.000

Table A-2: log-ratio variance matrices for all elements. Values report the variance of the natural log of the ratio of elements. Values for U and other elements from

Table 3-3 are repeated in this table. Top right half shows Coles Hill values, and bottom left half shows reference values for a) A horizon, and b) C horizon. Cells are color-coded such that maximum values are red and minimum values are green. (Next two pages.)

(a)	C_{org}	Al	As	Ba	Ca	Cu	Fe	K	Mg	Na	Mn	Pb	Th	Ti	U	V	Zn	P	Be	Bi	Ce	Co	Cr	Ga	La	Li	Mo	Nb	Ni	Rb	Sb	Sc	Sn	Sr	Tl	W	Y
C_{org}		0.51	0.76	0.40	0.32	0.58	0.57	0.50	0.46	1.05	0.36	0.30	0.43	0.43	1.40	0.63	0.43	0.41	0.48	0.87	0.54	0.61	0.42	0.55	0.51	0.54	0.40	0.42	0.48	0.42	0.41	0.50	0.44	0.31	0.37	0.43	0.51
Al	2.70		0.33	0.16	0.34	0.33	0.04	0.35	0.06	0.78	0.31	0.28	0.07	0.23	0.74	0.07	0.17	0.26	0.09	0.55	0.14	0.12	0.15	0.01	0.13	0.23	0.13	0.18	0.13	0.12	0.18	0.06	0.22	0.12	0.09	0.17	0.14
As	3.02	0.53		0.47	0.63	0.59	0.36	0.70	0.43	1.17	0.68	0.46	0.22	0.21	0.72	0.28	0.43	0.40	0.37	0.64	0.40	0.53	0.40	0.34	0.39	0.36	0.19	0.19	0.42	0.44	0.26	0.33	0.38	0.34	0.33	0.45	0.51
Ba	1.95	0.32	0.96		0.38	0.41	0.22	0.10	0.18	0.57	0.20	0.16	0.19	0.27	1.22	0.28	0.17	0.35	0.17	0.70	0.17	0.23	0.20	0.19	0.19	0.35	0.30	0.24	0.25	0.06	0.24	0.21	0.23	0.09	0.08	0.20	0.16
Ca	2.01	1.76	2.58	1.75		0.28	0.32	0.61	0.20	0.69	0.33	0.50	0.26	0.36	0.83	0.38	0.28	0.19	0.27	0.52	0.31	0.39	0.28	0.33	0.27	0.31	0.30	0.32	0.33	0.38	0.32	0.30	0.26	0.19	0.32	0.35	0.36
Cu	2.59	0.26	0.55	0.65	1.62		0.27	0.56	0.27	1.10	0.54	0.46	0.24	0.34	0.92	0.36	0.18	0.21	0.24	0.12	0.29	0.37	0.17	0.30	0.31	0.24	0.25	0.31	0.16	0.33	0.17	0.24	0.12	0.22	0.28	0.28	0.37
Fe	3.31	0.23	0.55	0.79	2.02	0.28		0.45	0.05	0.83	0.33	0.36	0.09	0.22	0.71	0.07	0.16	0.23	0.06	0.50	0.16	0.08	0.11	0.02	0.14	0.23	0.13	0.21	0.10	0.17	0.21	0.03	0.19	0.16	0.12	0.18	0.13
K	3.07	0.42	1.11	0.28	2.77	0.95	0.99		0.40	0.75	0.38	0.23	0.36	0.51	1.74	0.55	0.34	0.60	0.38	0.89	0.38	0.49	0.34	0.40	0.44	0.54	0.53	0.46	0.38	0.08	0.39	0.41	0.39	0.30	0.17	0.34	0.33
Mg	2.91	0.45	1.11	0.94	1.00	0.47	0.43	1.33		0.61	0.24	0.36	0.12	0.29	0.73	0.13	0.15	0.18	0.06	0.55	0.14	0.09	0.15	0.05	0.12	0.24	0.17	0.24	0.14	0.15	0.24	0.08	0.21	0.10	0.11	0.21	0.12
Na	3.64	1.01	2.04	1.30	1.50	1.39	1.35	1.53	0.84		0.72	0.87	0.85	1.02	1.31	0.91	0.88	0.77	0.72	1.46	0.80	0.79	0.86	0.81	0.76	0.91	1.00	0.97	0.93	0.74	1.00	0.86	0.96	0.60	0.75	0.91	0.79
Mn	2.06	0.99	1.19	0.93	1.13	0.79	0.96	1.77	0.98	1.54		0.36	0.32	0.39	1.32	0.43	0.23	0.43	0.23	0.97	0.29	0.21	0.32	0.32	0.28	0.51	0.37	0.40	0.35	0.25	0.44	0.33	0.39	0.22	0.23	0.26	0.13
Pb	1.79	0.28	0.67	0.24	1.71	0.49	0.64	0.53	0.83	1.33	0.87		0.23	0.25	1.19	0.34	0.26	0.43	0.33	0.70	0.33	0.42	0.22	0.33	0.33	0.32	0.24	0.25	0.27	0.19	0.15	0.30	0.29	0.17	0.13	0.22	0.31
Th	2.84	0.58	0.75	0.69	3.18	0.93	0.95	0.59	1.60	2.14	1.73	0.49		0.09	0.67	0.06	0.15	0.21	0.09	0.41	0.10	0.21	0.09	0.06	0.10	0.10	0.05	0.06	0.11	0.14	0.09	0.05	0.12	0.09	0.07	0.11	0.17
Ti	2.71	0.34	0.47	0.72	2.46	0.46	0.25	0.95	0.90	1.71	1.08	0.58	0.81		0.78	0.15	0.20	0.29	0.18	0.46	0.14	0.31	0.18	0.22	0.12	0.17	0.09	0.02	0.24	0.28	0.12	0.14	0.14	0.15	0.18	0.16	0.25
U	2.60	0.27	0.48	0.40	2.52	0.60	0.62	0.41	1.11	1.65	1.28	0.26	0.16	0.50		0.52	1.01	0.60	0.81	0.90	0.91	0.96	0.79	0.71	0.82	0.53	0.55	0.74	0.78	1.23	0.77	0.73	0.97	0.74	0.99	0.97	1.13
V	3.33	0.33	0.53	0.97	2.05	0.31	0.07	1.26	0.50	1.45	0.95	0.73	1.07	0.28	0.71		0.22	0.29	0.12	0.52	0.14	0.18	0.15	0.05	0.12	0.13	0.08	0.13	0.15	0.25	0.18	0.06	0.22	0.16	0.15	0.17	0.21
Zn	2.01	0.20	0.63	0.41	1.21	0.22	0.38	0.73	0.37	1.11	0.65	0.31	0.83	0.52	0.52	0.47		0.28	0.09	0.40	0.13	0.15	0.17	0.17	0.14	0.29	0.18	0.19	0.16	0.14	0.15	0.14	0.10	0.11	0.12	0.17	0.11
P	1.23	0.57	0.90	0.48	1.02	0.53	0.87	1.04	0.71	1.60	0.81	0.46	1.07	0.91	0.76	0.96	0.30		0.20	0.36	0.29	0.34	0.21	0.25	0.27	0.22	0.19	0.26	0.22	0.34	0.22	0.22	0.23	0.16	0.26	0.30	0.35
Be	2.50	0.19	0.53	0.30	1.87	0.40	0.44	0.43	0.71	1.15	0.75	0.28	0.51	0.52	0.25	0.54	0.25	0.55		0.49	0.09	0.07	0.13	0.07	0.09	0.23	0.15	0.17	0.13	0.13	0.21	0.04	0.14	0.10	0.09	0.12	0.07
Bi	2.23	0.46	0.43	0.64	2.21	0.60	0.64	0.79	1.04	1.78	0.99	0.38	0.51	0.56	0.31	0.66	0.45	0.72	0.35		0.49	0.70	0.42	0.52	0.52	0.40	0.40	0.40	0.42	0.62	0.26	0.44	0.22	0.43	0.53	0.56	0.69
Ce	3.06	0.40	0.70	0.56	2.86	0.71	0.66	0.53	1.36	1.83	1.39	0.46	0.36	0.60	0.32	0.87	0.67	0.95	0.38	0.65		0.16	0.22	0.13	0.01	0.23	0.20	0.11	0.25	0.17	0.20	0.12	0.12	0.12	0.13	0.18	0.11
Co	3.83	0.64	0.90	1.20	1.83	0.53	0.35	1.64	0.55	1.48	0.86	1.03	1.68	0.76	1.18	0.31	0.67	1.19	0.78	1.17	1.14		0.22	0.10	0.15	0.38	0.25	0.31	0.19	0.21	0.34	0.11	0.26	0.20	0.17	0.19	0.06
Cr	2.92	0.45	0.60	0.97	1.96	0.32	0.23	1.37	0.47	1.55	0.98	0.65	1.13	0.42	0.76	0.21	0.47	0.81	0.63	0.70	0.94	0.43		0.14	0.21	0.11	0.11	0.19	0.03	0.16	0.13	0.08	0.15	0.12	0.09	0.09	0.21
Ga	2.76	0.02	0.51	0.34	1.97	0.29	0.23	0.37	0.57	1.14	1.07	0.28	0.48	0.33	0.22	0.37	0.24	0.61	0.17	0.44	0.31	0.72	0.48		0.12	0.22	0.13	0.18	0.12	0.14	0.19	0.04	0.20	0.13	0.10	0.17	0.13
La	2.74	0.40	0.82	0.46	2.65	0.75	0.70	0.49	1.26	1.69	1.30	0.42	0.34	0.68	0.35	0.92	0.62	0.79	0.33	0.62	0.08	1.23	0.93	0.34		0.21	0.17	0.11	0.24	0.20	0.20	0.11	0.14	0.11	0.14	0.18	0.11
Li	2.75	0.22	0.43	0.51	2.39	0.42	0.46	0.52	0.88	1.72	1.23	0.40	0.36	0.41	0.22	0.53	0.38	0.69	0.25	0.30	0.45	1.00	0.64	0.20	0.41		0.11	0.17	0.14	0.31	0.14	0.16	0.22	0.17	0.20	0.19	0.36
Mo	2.19	0.51	0.60	0.64	2.55	0.66	0.91	0.74	1.45	2.14	1.43	0.48	0.64	0.62	0.46	0.96	0.55	0.70	0.52	0.58	0.52	1.50	1.04	0.47	0.55	0.44		0.09	0.10	0.25	0.08	0.10	0.17	0.12	0.14	0.16	0.24
Nb	2.72	0.92	1.19	0.88	3.50	1.28	1.52	0.79	2.25	2.79	2.21	0.97	0.90	0.98	0.74	1.67	1.15	1.32	0.97	1.08	0.81	2.38	1.74	0.84	0.86	0.78	0.33		0.23	0.24	0.09	0.14	0.12	0.11	0.16	0.16	0.26
Ni	3.11	0.37	0.69	0.90	1.88	0.38	0.31	1.20	0.39	1.46	1.18	0.64	1.09	0.60	0.70	0.32	0.43	0.78	0.59	0.75	0.90	0.44	0.12	0.42	0.89	0.63	1.03	1.73		0.16	0.14	0.09	0.18	0.15	0.11	0.11	0.21
Rb	2.38	0.36	0.79	0.24	2.54	0.76	0.89	0.15	1.28	1.78	1.43	0.36	0.37	0.72	0.22	1.09	0.55	0.73	0.30	0.46	0.45	1.58	1.17	0.32	0.39	0.28	0.44	0.53	1.08		0.21	0.15	0.20	0.14	0.04	0.16	0.13
Sb	1.65	0.66	0.56	0.68	1.96	0.62	1.03	1.02	1.26	2.23	1.08	0.47	0.87	0.79	0.62	0.98	0.47	0.48	0.59	0.50	0.93	1.41	0.99	0.69	0.90	0.52	0.31	0.78	1.03	0.59		0.17	0.10	0.10	0.12	0.18	0.29
Sc	3.24	0.25	0.54	0.87	1.97	0.29	0.07	1.06	0.42	1.33	1.02	0.67	1.01	0.34	0.67	0.09	0.38	0.85	0.45	0.59	0.73	0.35	0.22	0.26	0.77	0.49	0.90	1.58	0.24	0.97	0.98		0.13	0.13	0.09	0.11	0.13
Sn	1.98	0.42	0.57	0.48	2.35	0.62	0.74	0.63	1.21	1.92	1.09	0.33	0.36	0.51	0.23	0.82	0.48	0.62	0.35	0.25	0.51	1.42	0.84	0.37	0.49	0.27	0.36	0.60	0.88	0.29	0.41	0.76		0.13	0.14	0.16	0.22
Sr	1.91	0.86	1.84	0.61	0.78	1.07	1.31	1.33	0.81	0.73	0.91	0.74	1.77	1.45	1.27	1.38	0.76	0.73	0.99	1.47	1.61	1.36	1.28	1.00	1.43	1.44	1.65	2.23	1.20	1.34	1.43	1.34	1.37		0.08	0.15	0.17
Tl	2.52	0.40	0.68	0.30	2.73	0.77	0.83	0.23	1.41	1.94	1.33	0.32	0.26	0.68	0.18	1.00	0.59	0.83	0.27	0.39	0.31	1.44	1.12	0.33	0.31	0.31	0.40	0.61	1.06	0.08	0.59	0.89	0.29	1.48		0.11	0.10
W	2.58	1.07	1.13	1.19	3.04	1.19	1.64	1.26	2.10	2.93	1.92	1.13	1.19	1.23	0.95	1.66	1.07	1.19	1.01	1.06	1.12	2.22	1.73	1.07	1.11	0.81	0.38	0.36	1.67	0.82	0.53	1.61	0.71	2.34	0.89		0.15
Y	3.01	0.37	0.67	0.62	2.18	0.55	0.50	0.74	0.99	1.48	0.89	0.57	0.78	0.62	0.56	0.66	0.51	0.82	0.33	0.61	0.25	0.79	0.78	0.34	0.27	0.53	0.63	1.03	0.77	0.61	0.88	0.54	0.61	1.44	0.54	1.09	

(b)	C_{org}	Al	As	Ba	Ca	Cu	Fe	K	Mg	Na	Mn	Pb	Th	Ti	U	V	Zn	P	Be	Bi	Ce	Co	Cr	Ga	La	Li	Mo	Nb	Ni	Rb	Sb	Sc	Sn	Sr	Tl	W	Y
Corg		0.31	1.07	0.48	1.71	0.34	0.53	0.68	0.34	2.66	0.51	0.28	0.27	0.31	1.70	0.53	0.33	1.07	0.35	0.15	0.81	0.53	0.24	0.32	0.82	0.30	0.68	0.91	0.26	0.27	0.56	0.31	0.22	0.66	0.22	1.00	0.61
Al	0.81		1.01	0.40	1.72	0.32	0.11	0.84	0.08	2.51	0.40	0.15	0.04	0.16	1.06	0.14	0.12	1.26	0.06	0.24	0.36	0.22	0.23	0.01	0.37	0.34	0.33	0.38	0.23	0.40	0.25	0.04	0.24	0.43	0.24	0.96	0.34
As	1.25	1.13		2.11	3.65	0.88	0.90	2.31	1.13	5.30	1.62	1.29	0.97	1.31	2.22	1.07	1.09	1.71	1.10	0.86	1.99	1.21	0.87	0.95	1.91	1.39	0.72	1.46	0.78	1.34	0.76	0.96	1.30	2.04	1.15	1.88	1.99
Ba	1.46	0.83	2.18		1.19	0.80	0.74	0.47	0.34	1.33	0.49	0.34	0.42	0.26	1.39	0.60	0.49	1.33	0.41	0.57	0.42	0.65	0.62	0.45	0.45	0.34	1.12	0.76	0.67	0.46	0.91	0.47	0.40	0.23	0.42	1.07	0.24
Ca	6.40	5.06	8.30	5.12		2.44	2.44	1.86	1.58	1.22	1.77	1.49	1.72	1.34	2.13	2.12	2.00	2.45	1.54	1.71	1.20	1.91	1.94	1.82	1.19	1.00	2.57	1.97	2.03	1.97	2.33	1.86	1.70	0.91	1.80	2.24	1.09
Cu	0.94	0.60	1.27	1.78	6.70		0.26	0.92	0.23	3.49	0.39	0.52	0.28	0.50	1.94	0.42	0.21	1.19	0.29	0.36	0.92	0.32	0.20	0.26	0.97	0.44	0.34	0.69	0.07	0.33	0.34	0.24	0.55	1.14	0.27	0.68	0.72
Fe	0.93	0.25	0.93	1.46	6.00	0.47		1.16	0.16	3.19	0.43	0.40	0.19	0.38	1.28	0.16	0.10	1.47	0.14	0.45	0.60	0.17	0.25	0.08	0.62	0.62	0.22	0.48	0.23	0.56	0.22	0.09	0.46	0.83	0.37	1.08	0.62
K	1.58	1.12	2.31	0.28	6.09	1.98	1.92		0.70	1.94	0.82	0.93	0.71	0.57	3.06	1.27	0.94	1.40	0.88	0.67	1.28	1.18	0.70	0.90	1.41	0.72	1.60	1.10	0.85	0.20	1.32	0.91	0.57	1.12	0.32	1.36	0.85
Mg	1.89	0.80	2.09	1.46	4.23	1.33	0.78	1.99		2.30	0.31	0.25	0.11	0.20	1.24	0.18	0.11	1.29	0.06	0.32	0.40	0.19	0.21	0.07	0.43	0.24	0.36	0.39	0.17	0.31	0.26	0.10	0.37	0.51	0.20	0.72	0.30
Na	4.90	3.68	6.45	2.93	1.25	5.58	4.84	3.55	3.61		2.10	2.24	2.60	1.95	3.14	2.81	2.60	3.59	2.36	2.70	1.89	2.70	2.80	2.67	2.00	2.06	3.84	2.95	3.13	2.54	3.56	2.66	2.20	1.31	2.48	3.58	1.59
Mn	1.55	0.72	2.02	1.35	3.68	1.19	0.68	1.82	0.53	3.23		0.53	0.43	0.32	1.80	0.50	0.22	1.51	0.31	0.57	0.64	0.21	0.29	0.39	0.80	0.44	0.67	0.72	0.30	0.48	0.73	0.30	0.50	0.71	0.39	1.08	0.48
Pb	0.75	0.37	1.33	0.81	6.22	0.83	0.79	0.76	1.59	4.32	1.27		0.17	0.28	0.93	0.23	0.30	1.20	0.17	0.30	0.33	0.42	0.40	0.17	0.32	0.30	0.59	0.69	0.40	0.55	0.47	0.18	0.22	0.34	0.38	0.88	0.28
Th	0.96	0.80	1.23	1.05	7.87	1.16	1.16	0.85	2.30	5.53	1.99	0.38		0.18	1.33	0.21	0.19	1.16	0.11	0.16	0.46	0.29	0.19	0.06	0.50	0.30	0.41	0.39	0.19	0.30	0.27	0.08	0.24	0.53	0.18	0.85	0.37
Ti	0.95	0.42	1.12	1.42	5.89	0.69	0.19	1.97	0.91	4.70	0.72	0.83	1.27		1.31	0.42	0.30	1.33	0.23	0.29	0.38	0.34	0.22	0.19	0.43	0.21	0.53	0.33	0.32	0.32	0.45	0.21	0.24	0.34	0.20	1.02	0.31
U	0.87	0.76	1.10	1.04	7.36	1.26	1.07	0.99	2.06	5.15	1.79	0.36	0.22	1.09		0.96	1.31	2.62	1.02	1.79	0.65	1.28	1.86	1.05	0.49	1.36	1.21	1.35	1.72	2.43	1.31	1.13	1.67	0.79	2.03	2.06	1.10
V	1.17	0.56	1.04	2.05	5.95	0.62	0.20	2.66	0.78	5.20	0.83	1.23	1.68	0.26	1.48		0.18	1.64	0.16	0.47	0.51	0.27	0.39	0.11	0.51	0.54	0.34	0.57	0.33	0.68	0.25	0.12	0.50	0.61	0.50	0.93	0.48
Zn	1.01	0.27	1.25	0.96	5.24	0.85	0.33	1.19	0.69	3.98	0.49	0.61	1.10	0.51	0.99	0.70		1.26	0.11	0.35	0.52	0.12	0.21	0.10	0.57	0.42	0.38	0.62	0.18	0.44	0.38	0.09	0.33	0.62	0.30	1.03	0.47
P	1.45	0.90	1.47	1.68	6.39	1.11	0.60	2.25	1.33	5.04	0.99	1.11	1.60	0.56	1.42	0.80	0.83		1.22	0.86	1.49	1.60	1.24	1.31	1.50	1.14	1.86	2.08	1.24	1.15	1.76	1.32	0.97	1.68	1.06	1.90	1.35
Be	0.78	0.33	1.24	0.65	5.42	0.83	0.58	0.85	0.98	3.70	0.82	0.38	0.74	0.61	0.52	0.91	0.39	0.90		0.32	0.30	0.18	0.25	0.05	0.34	0.29	0.34	0.43	0.22	0.44	0.29	0.05	0.31	0.46	0.28	0.80	0.25
Bi	0.96	0.99	0.76	1.75	7.48	1.40	1.01	1.76	2.13	5.68	1.81	0.78	0.77	1.12	0.59	1.24	1.16	1.65	0.95		0.82	0.59	0.18	0.27	0.85	0.37	0.68	0.80	0.27	0.25	0.51	0.26	0.15	0.72	0.16	1.05	0.60
Ce	1.28	0.84	1.90	0.87	6.08	1.26	1.16	0.94	1.87	4.00	1.37	0.51	0.83	1.12	0.81	1.76	0.85	0.96	0.63	1.58		0.48	0.80	0.39	0.05	0.40	0.78	0.57	0.78	1.01	0.74	0.44	0.68	0.33	0.78	1.11	0.14
Co	1.27	0.47	1.46	1.45	4.47	0.73	0.39	2.00	0.37	3.79	0.41	1.02	1.79	0.45	1.61	0.38	0.46	0.75	0.74	1.64	1.24		0.28	0.19	0.54	0.42	0.29	0.53	0.21	0.64	0.35	0.19	0.58	0.69	0.45	1.05	0.53
Cr	1.06	0.81	1.14	1.92	6.53	0.60	0.40	2.48	1.03	5.62	0.98	1.17	1.58	0.37	1.37	0.37	0.84	0.77	0.82	1.26	1.36	0.51		0.20	0.85	0.35	0.40	0.59	0.08	0.23	0.30	0.18	0.27	0.80	0.11	0.88	0.63
Ga	0.78	0.04	1.06	0.81	5.50	0.60	0.24	1.03	0.91	3.93	0.79	0.32	0.67	0.42	0.63	0.61	0.25	0.83	0.29	0.94	0.69	0.57	0.80		0.39	0.35	0.26	0.36	0.18	0.41	0.19	0.02	0.29	0.50	0.24	0.84	0.37
La	1.19	0.85	1.91	0.79	5.66	1.33	1.16	0.97	1.63	3.67	1.22	0.55	0.90	1.05	0.82	1.63	0.91	0.89	0.62	1.43	0.17	1.09	1.25	0.78		0.42	0.74	0.64	0.82	1.09	0.70	0.48	0.70	0.29	0.84	1.14	0.22
Li	0.82	0.46	0.92	0.86	6.37	0.91	0.70	0.86	1.24	4.53	1.07	0.34	0.50	0.76	0.40	0.99	0.57	1.15	0.37	0.51	0.87	0.87	0.98	0.44	0.69		0.63	0.57	0.30	0.42	0.51	0.38	0.47	0.46	0.32	0.63	0.27
Mo	0.89	0.87	0.94	1.71	8.08	0.95	0.79	1.71	2.28	6.01	1.78	0.84	0.83	0.99	0.84	1.30	0.91	1.27	0.89	1.00	0.96	1.53	1.13	0.69	1.21	0.84		0.46	0.28	0.82	0.14	0.28	0.83	1.08	0.62	0.99	0.88
Nb	1.10	0.76	1.35	1.20	7.50	1.02	0.83	1.07	2.01	5.34	1.48	0.58	0.58	0.82	0.62	1.45	0.68	1.18	0.60	1.14	0.54	1.47	1.19	0.55	0.83	0.70	0.53		0.55	0.72	0.43	0.42	0.90	0.84	0.59	0.91	0.60
Ni	1.18	0.71	1.23	1.82	6.40	0.57	0.49	2.31	1.00	5.39	1.08	1.02	1.54	0.54	1.44	0.54	0.80	0.72	0.84	1.52	1.13	0.39	0.18	0.72	1.07	0.98	1.15	1.27		0.29	0.22	0.17	0.44	0.89	0.19	0.64	0.63
Rb	1.13	1.05	1.84	0.61	6.94	1.65	1.67	0.22	2.17	4.57	1.84	0.53	0.53	1.72	0.63	2.35	1.12	2.13	0.71	1.20	1.01	1.96	2.13	0.95	1.01	0.54	1.32	0.82	2.09		0.62	0.41	0.31	0.96	0.04	0.89	0.65
Sb	1.47	1.07	0.53	2.07	7.68	1.38	0.93	2.15	2.05	5.87	1.96	1.34	1.37	1.11	1.30	1.08	1.22	1.92	1.31	0.84	2.20	1.58	1.48	1.02	2.19	0.97	1.19	1.34	1.81	1.75		0.24	0.69	0.94	0.43	0.69	0.72
Sc	1.12	0.37	1.02	1.63	5.95	0.57	0.08	2.17	0.80	4.97	0.73	0.91	1.39	0.23	1.24	0.16	0.41	0.75	0.65	1.02	1.44	0.41	0.43	0.39	1.40	0.80	1.05	1.07	0.60	1.86	0.91		0.27	0.52	0.25	0.86	0.39
Sn	0.73	0.37	0.84	0.87	6.56	0.84	0.53	0.96	1.46	4.53	1.23	0.36	0.34	0.64	0.31	0.91	0.54	1.11	0.31	0.46	0.85	1.10	0.96	0.29	0.86	0.31	0.67	0.47	1.08	0.68	0.78	0.61		0.53	0.19	1.31	0.51
Sr	3.46	2.32	4.65	1.74	1.99	3.68	3.27	2.53	2.52	0.98	2.30	2.69	3.73	3.11	3.36	3.51	2.81	3.04	2.41	4.04	2.29	2.38	3.59	2.50	2.03	3.09	4.29	3.79	3.28	3.35	4.52	3.37	3.13		0.77	1.51	0.33
Tl	0.89	0.81	1.53	0.59	6.76	1.41	1.33	0.34	1.97	4.53	1.62	0.33	0.41	1.32	0.45	1.91	0.91	1.75	0.51	0.94	0.82	1.63	1.72	0.70	0.84	0.38	1.10	0.63	1.71	0.08	1.47	1.48	0.48	3.19		0.81	0.52
W	1.12	0.72	0.78	1.51	7.37	1.04	0.72	1.46	1.70	5.48	1.53	0.79	0.94	0.88	0.79	1.15	0.76	1.43	0.70	0.61	1.26	1.37	1.09	0.65	1.33	0.59	0.74	0.63	1.25	1.09	0.77	0.84	0.45	4.03	0.95		0.82
Y	1.07	0.54	1.60	0.81	4.68	1.12	0.70	1.02	0.95	3.10	0.68	0.67	1.09	0.75	0.94	1.16	0.46	0.91	0.43	1.32	0.49	0.72	1.07	0.49	0.50	0.72	1.09	0.73	1.05	1.06	1.61	0.84	0.68	2.08	0.89	1.01	

Table A-3: Statistics for principal component analysis (see Figure 3-7), including standard deviations (SD), eigenvalues (EC), percentage of variation explained (VE) by each PC, and cumulative percentage of variation explained by successive PCs.

	A Horizon				C Horizon			
	SD	**EV**	**% VE**	**Cum. % VE**	**SD**	**EV**	**% VE**	**Cum. % VE**
PC1	2.345	5.501	32.86%	32.86%	3.146	9.895	36.76%	36.76%
PC2	1.764	3.112	18.59%	51.45%	2.432	5.913	21.97%	58.73%
PC3	1.333	1.777	10.61%	62.07%	1.547	2.393	8.89%	67.62%
PC4	1.046	1.095	6.54%	68.61%	1.243	1.545	5.74%	73.36%
PC5	0.980	0.960	5.73%	74.34%	1.008	1.015	3.77%	77.13%
PC6	0.893	0.797	4.76%	79.11%	0.896	0.803	2.98%	80.12%
PC7	0.773	0.597	3.57%	82.67%	0.853	0.728	2.70%	82.82%
PC8	0.682	0.466	2.78%	85.45%	0.813	0.662	2.46%	85.28%
PC9	0.608	0.370	2.21%	87.66%	0.758	0.575	2.14%	87.41%
PC10	0.553	0.306	1.83%	89.49%	0.681	0.463	1.72%	89.13%
PC11	0.513	0.263	1.57%	91.06%	0.649	0.421	1.56%	90.70%
PC12	0.483	0.233	1.39%	92.46%	0.587	0.344	1.28%	91.98%
PC13	0.406	0.165	0.99%	93.44%	0.549	0.302	1.12%	93.10%
PC14	0.384	0.148	0.88%	94.33%	0.500	0.250	0.93%	94.03%
PC15	0.370	0.137	0.82%	95.14%	0.469	0.220	0.82%	94.85%
PC16	0.349	0.122	0.73%	95.87%	0.436	0.190	0.71%	95.55%
PC17	0.310	0.096	0.57%	96.45%	0.431	0.186	0.69%	96.24%
PC18	0.284	0.081	0.48%	96.93%	0.404	0.163	0.61%	96.85%
PC19	0.277	0.077	0.46%	97.39%	0.367	0.135	0.50%	97.35%
PC20	0.252	0.064	0.38%	97.77%	0.331	0.109	0.41%	97.76%
PC21	0.236	0.056	0.33%	98.10%	0.305	0.093	0.35%	98.10%
PC22	0.221	0.049	0.29%	98.39%	0.285	0.081	0.30%	98.40%
PC23	0.204	0.042	0.25%	98.64%	0.270	0.073	0.27%	98.67%
PC24	0.197	0.039	0.23%	98.87%	0.243	0.059	0.22%	98.89%
PC25	0.181	0.033	0.20%	99.07%	0.232	0.054	0.20%	99.09%
PC26	0.172	0.030	0.18%	99.25%	0.223	0.050	0.19%	99.28%
PC27	0.164	0.027	0.16%	99.41%	0.212	0.045	0.17%	99.45%
PC28	0.142	0.020	0.12%	99.53%	0.189	0.036	0.13%	99.58%
PC29	0.138	0.019	0.12%	99.64%	0.174	0.030	0.11%	99.69%
PC30	0.121	0.015	0.09%	99.73%	0.152	0.023	0.09%	99.78%
PC31	0.120	0.014	0.09%	99.82%	0.136	0.018	0.07%	99.84%
PC32	0.106	0.011	0.07%	99.88%	0.130	0.017	0.06%	99.91%
PC33	0.088	0.008	0.05%	99.93%	0.103	0.011	0.04%	99.95%
PC34	0.074	0.005	0.03%	99.96%	0.086	0.007	0.03%	99.97%
PC35	0.069	0.005	0.03%	99.99%	0.063	0.004	0.02%	99.99%
PC36	0.042	0.002	0.01%	100.00%	0.055	0.003	0.01%	100.00%
PC37	0.000	0.000	0.00%	100.00%	0.000	0.000	0.00%	100.00%

Table A-4: Eigenvectors for the first two principal components (see Figure 3-7) for the A and C horizons.

	A Horizon		C Horizon	
	PC1	**PC2**	**PC1**	**PC2**
Al	0.015	0.082	-0.003	0.032
As	-0.052	0.241	-0.178	0.189
Ba	0.006	-0.118	0.079	-0.204
Be	-0.011	0.028	-0.011	-0.038
Bi	-0.149	-0.002	-0.125	-0.038
Ca	0.398	-0.247	0.604	0.113
Ce	-0.162	0.067	-0.008	-0.229
Co	0.196	0.248	0.047	0.188
Cr	0.150	0.167	-0.056	0.249
Cu	0.086	0.073	-0.072	0.235
Fe	0.087	0.216	-0.060	0.176
Ga	0.003	0.094	-0.026	0.029
K	-0.097	-0.037	0.038	-0.217
La	-0.151	0.019	0.007	-0.242
Li	-0.135	0.057	-0.049	-0.056
Mg	0.236	0.105	0.093	0.188
Mn	0.262	-0.091	0.126	0.238
Mo	-0.150	-0.072	-0.167	-0.003
Na	0.244	0.036	0.531	-0.114
Nb	-0.344	-0.141	-0.126	-0.152
Ni	0.112	0.172	-0.053	0.203
P	0.037	-0.172	-0.042	0.044
Pb	0.024	-0.087	-0.048	-0.122
Rb	-0.104	-0.078	-0.054	-0.180
Sb	-0.077	-0.144	-0.139	0.075
Sc	0.091	0.208	-0.042	0.201
Sn	-0.148	-0.075	-0.074	-0.072
Sr	0.225	-0.168	0.330	-0.165
Th	-0.180	0.019	-0.134	-0.177
Ti	-0.053	0.106	-0.029	0.091
Tl	-0.140	-0.025	-0.056	-0.135
U	-0.267	0.049	-0.140	-0.385
V	0.064	0.226	-0.054	0.181
W	-0.248	-0.188	-0.131	-0.047
Y	-0.019	0.071	0.064	-0.047
Zn	0.088	-0.023	0.003	0.117
C_{org}	0.165	-0.613	-0.043	0.076

APPENDIX B. KINETICS EXPERIMENTS: APATITE DISSOLUTION AND AUTUNITE PRECIPITATION

B.1 INTRODUCTION

Preliminary laboratory studies were done on the dissolution of apatite and its reaction with uranium-bearing solutions. This reaction occurs in the saprolite at Coles Hill (Jerden and Sinha, 2003) and has been suggested as a potential treatment option for uranium-contaminated groundwater.

B.1.1 URANIUM IMMOBILIZATION BY PHOSPHATE

Autunite is a stable uranium weathering product found within the Coles Hill deposit. Measurements in the area around the deposit have found low levels of dissolved uranium in groundwater, and the low concentrations of uranium in this system have been attributed to immobilization by autunite (Jerden and Sinha, 2003). The apparent effectiveness of this mineral in controlling uranium transport at Coles Hill may provide a natural analogue for a contamination treatment system.

One promising technology which inhibits the transport of uranium in groundwater is the use of subsurface permeable reactive barriers. In these systems, contaminated groundwater flows through a downgradient trench filled with a material known to react with the contaminant of concern. The contaminant reacts with the barrier material and removed from solution through sorption or precipitation into an immobile phase. Materials currently under investigation for uranium treatment include apatite ($Ca_5(PO_4)_3(F,Cl,OH)$). Phosphate released from apatite dissolution reacts with uranium to precipitate uranyl phosphate minerals such as those of the autunite and meta-autunite groups ($Ca(UO_2)_2(PO_4)_2 \cdot nH_2O$). Results of previous studies

examining the effectiveness of using apatite as a material in permeable reactive barriers show that uranium at low concentrations adsorbs to apatite surfaces, while at higher concentrations, uranyl phosphate minerals can precipitate (e.g. Fuller et al., 2003; Ohnuki et al., 2004). As uranyl phosphate minerals such as autunite have low solubility and exhibit long-term stability (Sandino and Bruno, 1992), they can persist for geologic timescales, as at Coles Hill (Jerden and Sinha, 2003); thus, promoting their formation offers the prospect of effective, permanent uranium sequestration.

Although there have been many studies on dissolution rates of autunite (Gudavalli et al., 2013; Gunderson and Wellman, 2008; Smeaton et al., 2008; Wellman et al., 2007; Wellman et al., 2006), to date there have been no studies focused on the precipitation rates of autunite. As effective immobilization of uranium will depend on the balance between the rates of apatite dissolution and autunite precipitation, determining the kinetics of autunite precipitation is critical in the design of these types of reactive barriers.

B.1.2 APATITE DISSOLUTION

Fluorapatite dissolution rates have been measured under a variety of pH and temperature conditions in a number of studies. A plot of pH versus log rate is shown in Figure B-1. In these studies and others, dissolution mechanisms have been proposed. Chin and Nancollas (1991) measured apatite dissolution rates in batch reactors at constant pH values of 4.5, 5.0, 5.5, and 6.0 at 37° C. They found the dissolution reaction to be surface-controlled, probably following a spiral dislocation mechanism. Rates were calculated using BET surface area. Valsami-Jones et al. (1998) studied the dissolution of apatite in batch reactors with and without Pb^{2+} and Cd^{2+} in the pH range of 2-7 at 25° C. Rates were calculated using BET surface area. Guidry and

Mackenzie (2003) measured dissolution rates of apatite at a variety of pH between 2 and 8.5. They used a fluidized bed reactor and a stirred tank reactor (MFR). Surface area was measured by BET. They found that initially, Ca was released preferentially relative to P and F and that Ca was exchanged for H+ at low pH, suggesting a surface-controlled reaction. They also found that the degree of undersaturation affected rate measurements. Chaïrat et al. (2007b) measured dissolution rates of apatite at 25 C and pH between 3 and 12. They used mixed-flow reactors. Their findings suggest that apatite dissolution rates are pH-dependent for pH < 7 and pH > 10 and pH-independent for 7 < pH < 10. They also suggest that dissolution initially consists of leaching of Ca and F for H+ and OH-, creating a leached layer, whose dissolution is controlled by different processes at different pH, with adsorption/penetration of protons into the surface occurring at low pH and surface hydration occurring at higher pH. Another study by the same group (Chaïrat et al., 2007a) suggested the same processes. Harouiya et al. (2007) measured apatite dissolution rates in acidic (pH 1-6) solutions at temperatures of 5, 25, and 50°C, using closed-system batch reactors. Surface area was measured by BET. Zhu et al. (2009) did not measure dissolution rates but looked at the dissolution mechanism of apatite. They found that there is ion exchange between F^- and OH^- and that in the earlier stages of dissolution, Ca, F, and P are released non-stoichiometrically. Dorozhkin (2012) presents a review of eight different models of apatite dissolution in acid and proposes a model of the general dissolution mechanism of apatite. In this model, ions are transported to the surface, where H^+ and the acidic anion are adsorbed, potentially forming surface complexes. Then the X-ions (e.g. F^-) detach, followed by Ca^{2+} detachment and diffusion along the surface. Ca^{2+} is replaced by H^+ and may subsequently adsorb back on the surface. Following the loss of Ca, phosphate detaches and diffuses from the surface.

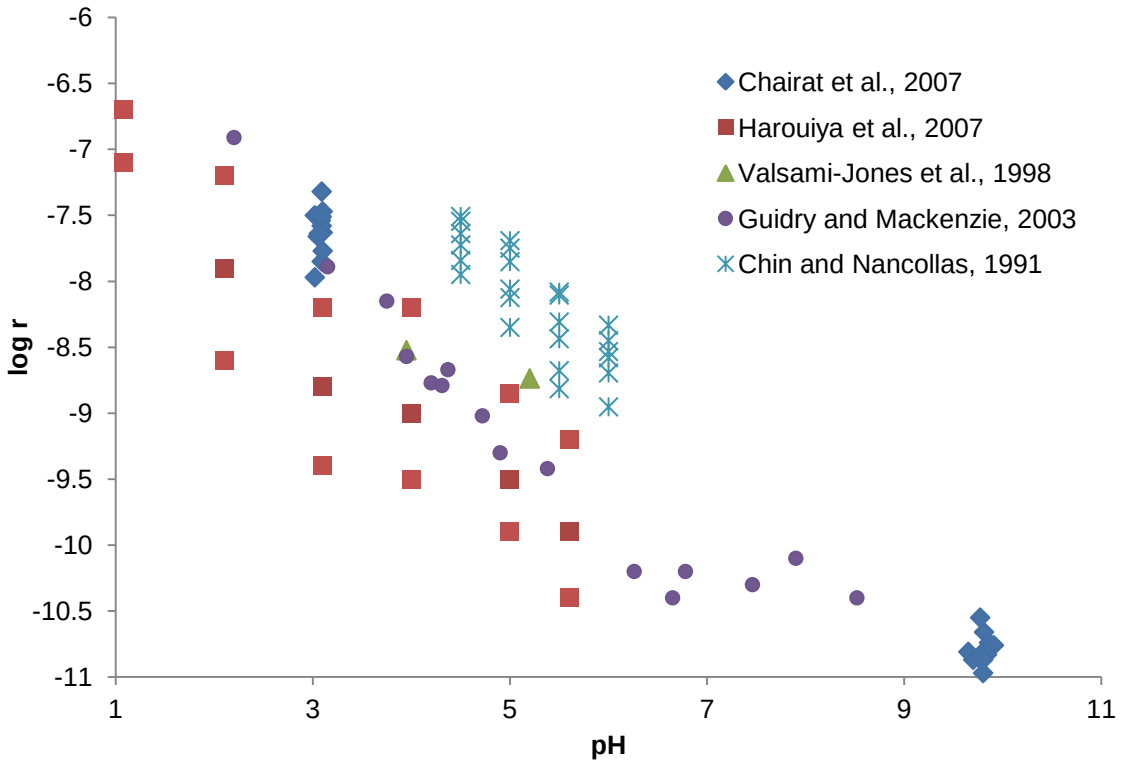

Figure B-1: Published apatite dissolution rates as a function of pH. Rate (r) units are mol/m^{2}s.

B.2 EXPERIMENTAL METHODS

B.2.1 APATITE PREPARATION

Natural igneous fluorapatite from Wilberforce, ON, was obtained from Ward's Scientific (Product #460628). The apatite was crushed in a steel percussion mortar to <2 mm and rolled with a steel rolling pin, the sieved to 250-425 μm. To remove attached smaller particles, the sieved fraction was rinsed in water, then sonicated in ethyl alcohol for several minutes before discarding the supernatant. This process was repeated until the liquid appeared clear, a minimum of ten times. The apatite was then dried in a drying oven at 60° C. The geometric surface area of the particles was calculated assuming spherical particles by

$$D_e = \frac{D_{max} - D_{min}}{ln\frac{D_{max}}{D_{min}}}, \qquad A_{geo} = \frac{6V_m}{D_e W_m} \qquad \text{Eq. B-1}$$

where D is diameter, A_{geo} is the geometric surface area, V_m is the molar volume of apatite, and W_m is the molecular weight of apatite.

B.2.2 SOLUTIONS

Sodium chloride was added to deionized water to a concentration of 0.001 M. The solution was then titrated with HCl and/or NaOH to reach a desired pH.

For uranium-bearing solutions, uranyl nitrate was added at the desired concentration.

B.2.3 REACTOR SETUP

Millipore Swinnex 25 mm filter holders were used as reaction vessels. Mesh screen was added to the outflow port and held in place by a fitting and PTFE tape. A silicone O-ring was placed on the inner lip. Approximately 3 g of apatite was weighed into the filter top, covered with a mesh screen and another O-ring. The screw threads of the filter holder bottom were covered in PTFE tape, and the holder was closed tightly.

Feed solution was contained in an HDPE bottle with an outlet port near the bottom. The port was connected by Tygon tubing through a mini peristaltic pump to the inflow port on the bottom of the filter holder. Tubing was also connected from the top of the filter holder and held in place above glass test tubes in a fraction collector. A photograph of the experimental setup is shown in Figure B-2.

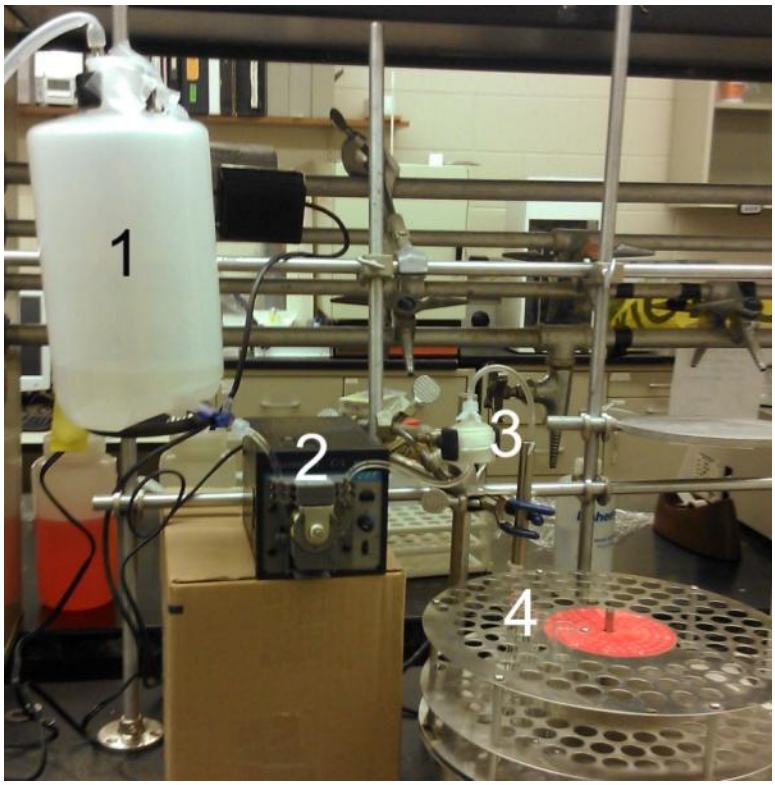

Figure B-2: Experimental setup. 1) feed solution, 2) pump, 3) reactor, 4) fraction collector.

Solution was run through the entire system at a relatively high flow rate to provide enough fluid pressure to saturate the system. Flow rate was then turned down, and the system ran continuously until the experiment was deemed complete. The pH of the feed and effluent was measured periodically using a pH meter.

B.2.4 SOLUTION ANALYSIS

Following collection, calcium concentrations in the effluent were measured using atomic absorption spectroscopy (AAS). For selected experiments, Ca, P, and U concentrations were measured by inductively-coupled plasma-optical emission spectroscopy (ICP-OES) in the VT Soils Testing Laboratory under the direction of Athena Tilley.

B.2.5 CALCULATIONS

The experimental determination of rate laws uses a concentration change over time to find a reaction rate and the change in reaction rate with concentration to find reaction coefficients and orders. For reactions in which the initial concentration of only one parameter is changed, the rate law can be expressed as:

$$\text{Rate} = k'[\text{C}]_o^{n_A} \qquad \text{Eq. B-2}$$

where *k'* is a constant, $[\text{C}]_o$ is the initial concentration of a particular constituent, and n_A is the reaction order (Lasaga, 1998). Taking the log of Eq. B-2 yields:

$$\log(\text{Rate}) = \log(k') + n_A \log([\text{C}]_o) \qquad \text{Eq. B-3}$$

(Lasaga, 1998). The rate of a reaction in a mixed flow reactor can be determined using the equation:

$$\text{Rate} = \frac{[\text{C}]_o r_f}{MA} \qquad \text{Eq. B-4}$$

(Walker et al., 2006), where r_f is the flow rate through the reactor, *M* is the mass of apatite, and *A* is the specific surface area of the apatite. Using a series of rates for different initial concentrations, log(*k'*) and n_A can be calculated by Eq. B-3 from the intercept and slope, respectively, of a log(Rate)-log($[\text{C}]_o$) plot. If multiple parameters are varied, multiple linear regression modeling can yield the same results.

Apatite dissolution was calculated based on Ca removal following the reaction:

$$Ca_5(PO_4)_3F = 5Ca^{2+} + 3\ PO_4^{3-} + F^- \qquad \text{Eq. B-5}$$

B.3 RESULTS

Apatite dissolution experimental results are summarized in Table B-1 and shown in Figure B-3.

Table B-1: Calculated rates (r) from apatite dissolution experiments. Rate units are mol/m^{2}s.

Sample	Starting pH	log r Ca loss	log r apatite
030113	4.4	-8.4	-9.10
030313	4.4	-8.2	-8.90
031813	2.7	-7.4	-8.10
032013	2.7	-8.0	-8.70
032213	7.98	-8.3	-9.00
032613	3.3	-6.7	-7.40
032713a	2.56	-6.4	-7.10
032713b	2.79	-6.2	-6.90
032913	4.99	-7.6	-8.30
040113	5.6	-6.5	-7.20
082813a	7.5	-7.6	-8.30
082813b	7.5	-7.9	-8.60
090213	7.6	-8.0	-8.70
091613a	4.3	-8.1	-8.80
091613b	4.16	-8.2	-8.90
091813	4.4	-8.3	-9.00
092513	4.4	-8.3	-9.00
101113	5.01	-8.7	-9.40
101813	3.05	-7.7	-8.40

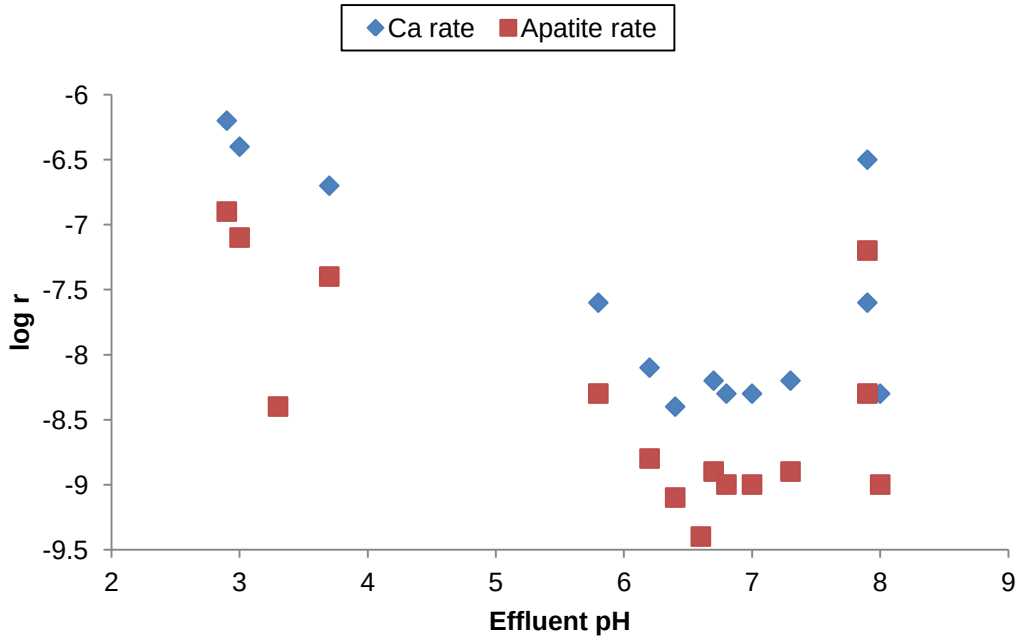

Figure B-3: Rates of Ca release and apatite dissolution as a function of pH. Rate units are mol/m^{2}s.

Two uranium-apatite experiments were conducted, one with higher (hi, 10^{-3} M U) U and one with lower (lo, 10^{-5} M U) U. Results of ICP-OES are shown in Table B-2 for the lower U experiment and in Table B-3 for the higher U experiment.

Table B-2: ICP-OES results for lower U experiment.

Sample	Time (hours)	Ca mg/L	P mg/L	U mg/L
110613-F1	feed	0.026	0.026	2.7017
110613-F2	feed	0.3142	0.1443	2.769
110613-01	6	3.233	0.0341	0.3047
110613-02	24	3.0652	0.0518	0.4811
110613-03	42	2.5954	0.0149	0.5229
110613-04	60	2.7758	0.259	0.5362
110613-05	78	2.1795	<0.026	0.5719
110613-06	96	2.0068	<0.026	0.6344
110613-07	114	1.8837	<0.026	0.5205
110613-08	132	1.8036	<0.026	0.5423
110613-9	150	1.5827	<0.026	1.6857
110613-10	168	1.6044	<0.026	1.2873
110613-11	186	1.5981	<0.026	1.1886
110613-12	204	1.489	<0.026	1.3581
110613-13	222	1.4952	<0.026	1.3795
110613-14	240	1.5162	<0.026	1.2088
110613-15	258	1.5469	<0.026	1.0441
110613-16	276	1.4806	<0.026	1.1223
110613-17	294	1.5307	<0.026	1.0409
110613-18	312	1.5223	<0.026	1.4213
110613-19	330	1.3977	<0.026	1.8291

Table B-3: ICP-OES results for higher U experiment

Sample	Time (hours)	Ca mg/L	P mg/L	U mg/L
011314-F1	Feed	0.1978	<0.026	24.709
011314-F2	Feed	0.0425	<0.026	25.629
011314-01	6	5.925	<0.026	20.041
011314-03	42	4.232	<0.026	23.379
011314-05	78	3.622	<0.026	24.074
011314-07	114	3.194	<0.026	23.622
011314-09	150	3.335	<0.026	24.962
011314-11	186	3.102	<0.026	24.067
011314-13	222	2.737	<0.026	23.908
011314-15	258	2.64	<0.026	22.782
011314-17	294	2.704	<0.026	22.783
011314-19	330	2.429	<0.026	23.82
011314-21	366	2.318	<0.026	23.243
011314-23	403.5	2.201	<0.026	22.51

Uranium depletion relative to the feed solution was calculated for each sample and plotted in Figure B-4.

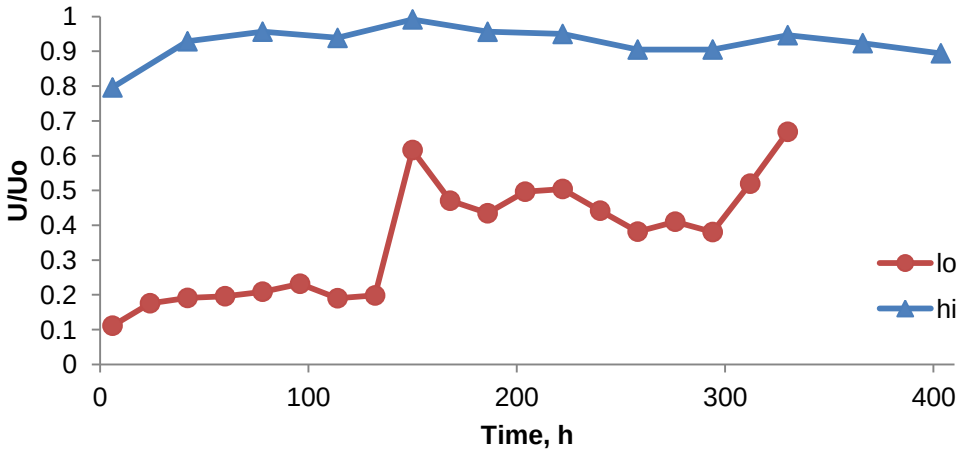

Figure B-4: Plot of uranium concentration (U) relative to average feed uranium concentration (U_O) over the course of experiments.

The rate of U removal and apatite dissolution was calculated over the course of both experiments (Figure B-5).

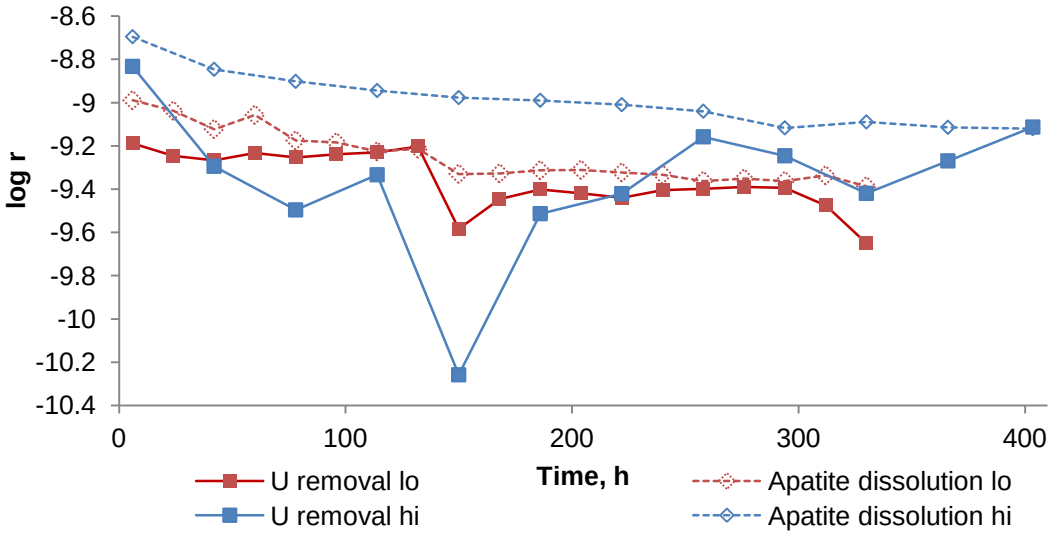

Figure B-5: Calculated rates (r) of uranium removal and apatite dissolution over the course of experiments. Apatite dissolution was calculated from the rate of Ca release assuming no Ca uptake. Rate (r) units are mol/m^2s.

REFERENCES

Chaïrat, C., Oelkers, E.H., Schott, J., Lartigue, J.-E., 2007a. Fluorapatite surface composition in aqueous solution deduced from potentiometric, electrokinetic, and solubility measurements, and spectroscopic observations. Geochimica Et Cosmochimica Acta 71, 5888-5900.

Chaïrat, C., Schott, J., Oelkers, E.H., Lartigue, J.-E., Harouiya, N., 2007b. Kinetics and mechanism of natural fluorapatite dissolution at 25°C and pH from 3 to 12. Geochimica Et Cosmochimica Acta 71, 5901-5912.

Chin, K.O.A., Nancollas, G.H., 1991. Dissolution of fluorapatite - A constant-composition kinetics study. Langmuir 7, 2175-2179.

Dorozhkin, S.V., 2012. Dissolution Mechanism of Calcium Apatites in Acids: A review of literature. World Journal of Methodology 2, 1-17.

Fuller, C.C., Bargar, J.R., Davis, J.A., 2003. Molecular-scale characterization of uranium sorption by bone apatite materials for a permeable reactive barrier demonstration. Environmental Science & Technology 37, 4642-4649.

Gudavalli, R.K.P., Katsenovich, Y.P., Wellman, D.M., Idarraga, M., Lagos, L.E., Tansel, B., 2013. Comparison of the kinetic rate law parameters for the dissolution of natural and synthetic autunite in the presence of aqueous bicarbonate ions. Chemical Geology 351, 299-309.

Guidry, M.W., Mackenzie, F.T., 2003. Experimental study of igneous and sedimentary apatite dissolution: Control of pH, distance from equilibrium, and temperature on dissolution rates. Geochimica Et Cosmochimica Acta 67, 2949-2963.

Gunderson, K.M., Wellman, D.M., 2008. Effect of co-contaminants on in situ phosphate-based immobilization of uranium: Dissolution kinetics of autunite in the presence of tributyl phosphate, 8th Annual V M Goldschmidt Conference, Vancouver, Canada, pp. A335-A335.

Harouiya, N., Chairat, C., Koehler, S.J., Gout, R., Oelkers, E.H., 2007. The dissolution kinetics and apparent solubility of natural apatite in closed reactors at temperatures from 5 to 50 degrees C and pH from 1 to 6. Chemical Geology 244, 554-568.

Jerden, J.L., Sinha, A.K., 2003. Phosphate based immobilization of uranium in an oxidizing bedrock aquifer. Applied Geochemistry 18, 823-843.

Lasaga, A.C., 1998. Kinetic theory in the earth sciences. Princeton University Press, Princeton, N.J.

Ohnuki, I., Kozai, N., Samadfam, M., Yasuda, R., Yamamoto, S., Narumi, K., Naramoto, H., Murakami, T., 2004. The formation of autunite ($Ca(UO_2)_2(PO_4)_2 \bullet nH_2O$)within the leached layer of dissolving apatite: incorporation mechanism of uranium by apatite. Chemical Geology 211, 1-14.

Sandino, A., Bruno, J., 1992. The solubility of $(UO_2)_3(PO_4)_2 \bullet 4H_2O_{(S)}$ and the formation of U(VI) phosphate complexes - their influence in uranium speciation in natural-waters. Geochimica Et Cosmochimica Acta 56, 4135-4145.

Smeaton, C.M., Weisener, C.G., Burns, P.C., Fryer, B.J., Fowle, D.A., 2008. Bacterially enhanced dissolution of meta-autunite. American Mineralogist 93, 1858-1864.

Valsami-Jones, E., Ragnarsdottir, K.V., Putnis, A., Bosbach, D., Kemp, A.J., Cressey, G., 1998. The dissolution of apatite in the presence of aqueous metal cations at pH 2–7. Chemical Geology 151, 215-233.

Walker, F.P., Schreiber, M.E., Rimstidt, J.D., 2006. Kinetics of arsenopyrite oxidative dissolution by oxygen. Geochimica Et Cosmochimica Acta 70, 1668-1676.

Wellman, D.M., Gunderson, K.M., Icenhower, J.P., Forrester, S.W., 2007. Dissolution kinetics of synthetic and natural meta-autunite minerals, $X_{3-n}^{(n)+}[(UO_2)(PO_4)]_2 \bullet xH_2O$, under acidic conditions. Geochemistry Geophysics Geosystems 8.

Wellman, D.M., Icenhower, J.P., Gamerdinger, A.P., Forrester, S.W., 2006. Effects of pH, temperature, and aqueous organic material on the dissolution kinetics of meta-autunite minerals, (Na, Ca)(2-1) (UO2)(PO4) (2)•3H(2)O. American Mineralogist 91, 143-158.

Zhu, Y., Zhang, X., Chen, Y., Xie, Q., Lan, J., Qian, M., He, N., 2009. A comparative study on the dissolution and solubility of hydroxylapatite and fluorapatite at 25 degrees C and 45 degrees C. Chemical Geology 268, 89-96.

APPENDIX C. DATA ANALYSIS

C.1 COMPOSITIONAL DATA

Compositional data are data that represent relative parts of a whole and often sum (or potentially sum) to a constant value (Aitchison, 1986; Pawlowsky-Glahn and Egozcue, 2006). These data are also called "closed" data. Common examples in geochemistry include weight percent or mg/kg; the constant sum values in those cases are 100% and 1 kg. In the analysis of compositional data, the issue of closure can lead to the misinterpretation of statistical results, such as spurious correlations. A daily example of compositional data might be the number of hours per day spent doing different tasks. Hours per day are parts of a whole (the day) and subject to a constant sum constraint (24). An increase in the amount of time spent on one task necessarily impacts the amount of time spent on one of more other tasks, as shown in Figure C-1.

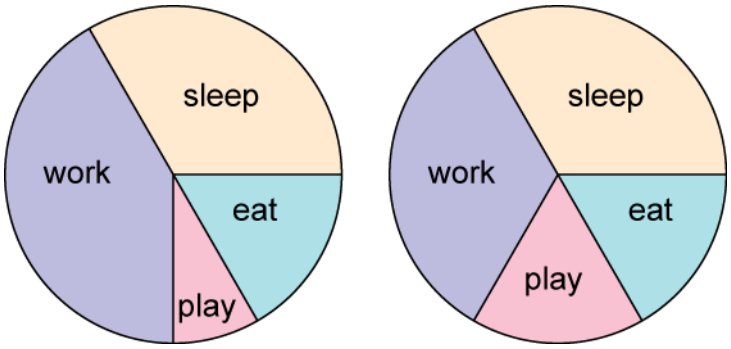

Figure C-1: Pie charts showing potential hours per day spent on different tasks.

C.1.1 LOG-RATIO TRANSFORMATIONS

As a result, methods to "open" such data have been proposed. The most common methods are log-ratio transformations (Aitchison, 1986; Egozcue et al., 2003), including additive log-ratio (alr), centered log-ratio (clr), and isometric log-ratio (ilr) transformations. The formulas given below are based on a group x consisting of D parts. In a standard geochemical study, x

may be concentrations of Ca, K, Mg, and Na. In this example, D is 4, and Ca is represented as x_1, K as x_2, and so on until Na as x_4 or x_D.

The alr transformation uses the natural log of an element concentration relative to a selected divisor (Aitchison, 1986). The use of an alr transformation requires the "sacrifice" of one variable (denoted as x_D) to serve as the divisor in the transformation, and so the alr-transformed data have one fewer variable than the original data. An ideal divisor will have values well above detection and a degree of precision that provides a range of values.

$$\mathrm{alr}(x) = \ln\frac{x_1}{x_D}, \ln\frac{x_2}{x_D}, \ldots, \ln\frac{x_{D-1}}{x_D} \qquad \text{Eq. C-1}$$

The clr transformation takes the natural log of an element concentration divided by the geometric mean [g; $(x_1x_2\ldots x_D)^{1/D}$] of the other element concentrations for that sample (Aitchison, 1986).

$$\mathrm{clr}(x) = \ln\frac{x_1}{g(x)}, \ln\frac{x_2}{g(x)}, \ldots, \ln\frac{x_D}{g(x)} \qquad \text{Eq. C-2}$$

The ilr transformation is based on parts and is here employed as balances (Egozcue and Pawlowsky-Glahn, 2005; Egozcue et al., 2003). Variables are divided sequentially between groups r and s (Figure C-2), and ilr balance values are calculated for each sample point by Eq. C-3.

$$\mathrm{ilr} = \sqrt{\frac{D_r D_s}{D_r + D_r}} \ln\left[\frac{g(x_r)}{g(x_s)}\right] \qquad \text{Eq. C-3}$$

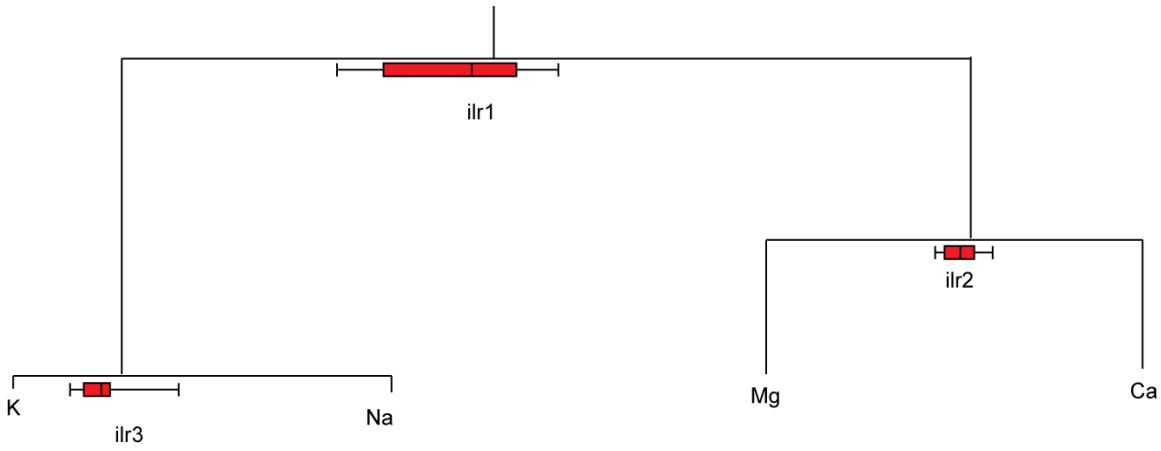

Figure C-2: Sample balance dendrogram showing sequential binary partitioning of four elements. Using four elements, three ilr balances are formed, each represented by one of the dendrogram junctions. The red boxes and lines represent boxplots of ilr values calculated for a particular dataset.

Table C-1: Table showing sequential binary partitioning of four elements as illustrated in Figure C-2.

Balance	Mg	Ca	K	Na
ilr1	+	+	-	-
ilr2	-	+	0	0
ilr3	0	0	-	+

In the example shown in Figure C-2, $\text{ilr1} = \sqrt{\frac{2\times2}{2+2}}\ln\left[\frac{(\text{Mg}\times\text{Ca})^{1/2}}{(\text{K}\times\text{Na})^{1/2}}\right]$, $\text{ilr2} = \sqrt{\frac{1\times1}{1+1}}\ln\left[\frac{\text{Ca}}{\text{Mg}}\right]$, and $\text{ilr3} = \sqrt{\frac{1\times1}{1+1}}\ln\left[\frac{\text{Na}}{\text{K}}\right]$.

C.2 COMMON MULTIVARIATE METHODS AND FIGURES

This section includes examples and explanations of some of the multivariate analyses used. Analyses and figure plotting were done in R (R Development Core Team, 2011), often using code and concepts modified from Reimann et al. (2008).

C.2.1 HIERARCHICAL CLUSTER ANALYSIS (HCA)

Hierarchical cluster analysis takes all the data points and calculates the distance between them (mathematically, this is sort of like a multidimensional Pythagorean theorem) and then groups the samples or variables by which have the most similar average distances. The results are often shown in a dendrogram like Figure C-3, an example HCA dendrogram showing hexadecimal color codes. In some ways, these are similar to phylogenetic trees or cladograms, though they are not the same. Each color is defined by the values of the three colors of light, red, green, and blue, which can be reported as a string of 3 2-digit numbers in base 16 (00 is a value of 0, while FF is a value of 255; data were not log-ratio transformed here). The colors that are more similar are separated by fewer junctions and are joined at lower heights. For example, #82CA FA and FF are very similar—a primary junction at low height. The highest junction here represents the split between the purples and the blues. Similar to how we can use what we know about color to conclude that difference, we can infer geochemical differences or processes using this method.

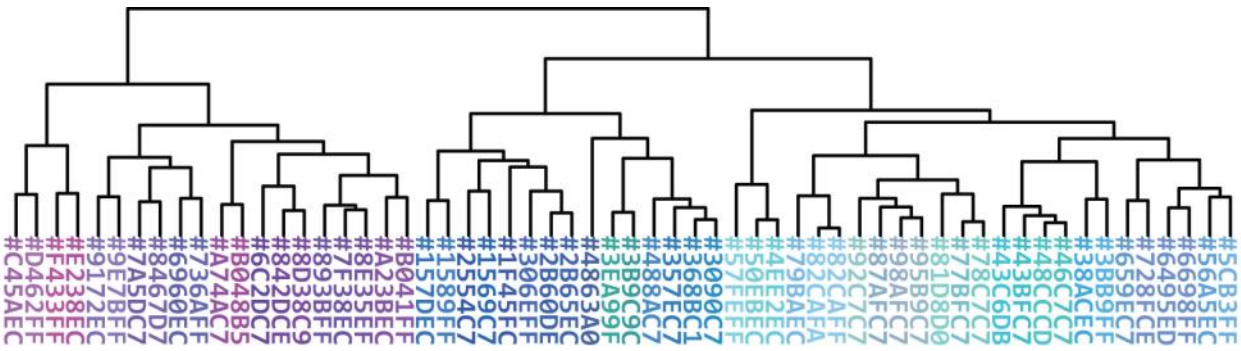

Figure C-3: HCA dendrogram showing hexadecimal color codes displayed in their corresponding colors.

C.2.2 PRINCIPAL COMPONENT ANALYSIS (PCA)

Principal component analysis is another common multivariate method. When using compositional data, a clr transformation is appropriate. Principal compenent analysis recasts varaibles into an identical number of new variables, called principal components (PCs), that maximize the amount of data variation they account for. The two that account for the most variation comprise the axes of the diagram in Figure C-4, called a biplot. Biplots can also be created using subsequent PCs. To determine which PCs may be valuable, the eigenvalues, standard deviations, or percent variation explained is plotted for each PC. The PC number is based on the sequential ordering of these parameters from highest to lowest. The plot of these values is called a screeplot (Figure C-5). A screeplot generally has several values at a steep slope and then levels off.

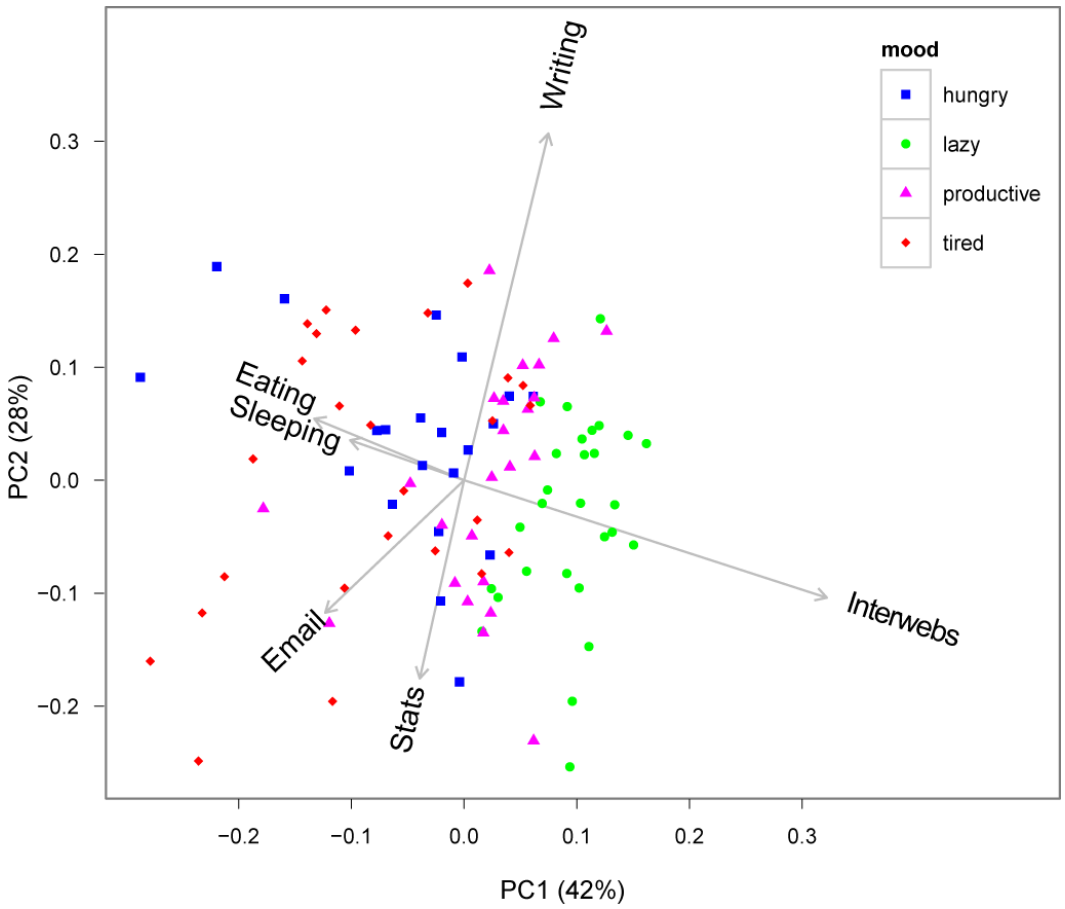

Figure C-4: PCA biplot of time spent per day on various activities.

In this PCA biplot (Figure C-4), the grey vectors in represent the original variables within each PC, with their length relating to the variance. The colored points represent the PC scores for 100 samples, which in this example represent the number of hours per day spent on various tasks. In general, the vectors that are dominant along the x-axis, or PC1, are how I spend my time when not working, whereas those that dominate PC2, the y-axis, are how I spend my time when working. Links between vector tips, as discussed in Chapter 3, can also be used for interpretation when working with clr-transformed data.

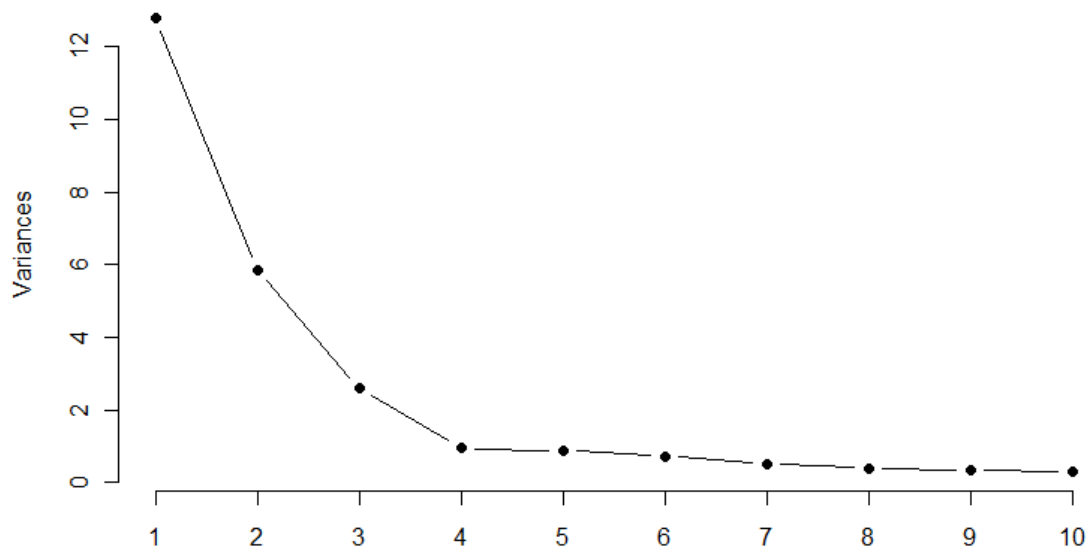

Figure C-5: Example screeplot of ten PCs by variance (eigenvalue). In this example, the first three PCs are likely to yield useful information.

C.2.3 LINEAR DISCRIMINANT ANALYSIS (LDA)

Linear discriminant analysis creates linear functions that scale the variables within the dataset to maximize the separation between user-defined groups. The number of these functions (linear discriminants, LDs) is the number of groups minus one. The LD scaling functions can be used to classify additional samples into the most similar group. Ideally, if the original samples, or training data, are classified by the LDA results, the group membership predicted by the LDA will be the same as the user-defined group membership. The number of samples predicted to be in the wrong group divided by the total number of samples is called the misclassification rate. Like PCs, LDs can be used as variables in further analyses and thus present an opportunity for dimension reduction. The results of LDA can be shown as scatterplots of the LD scores (values for each sample calculated by the scaling functions) for each sample for two LDs, analogous to the scores plotted on a PCA biplot.

REFERENCES

Aitchison, J., 1986. The statistical analysis of compositional data. Chapman and Hall, London; New York.

Egozcue, J.J., Pawlowsky-Glahn, V., 2005. Groups of parts and their balances in compositional data analysis. Mathematical Geology 37, 795-828.

Egozcue, J.J., Pawlowsky-Glahn, V., Mateu-Figueras, G., Barceló-Vidal, C., 2003. Isometric logratio transformations for compositional data analysis. Mathematical Geology 35, 279-300.

Pawlowsky-Glahn, V., Egozcue, J.J., 2006. Compositional data and their analysis: an introduction. Geological Society, London, Special Publications 264, 1-10.

R Development Core Team, 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, Software Program.

Reimann, C., Filzmoser, P., Garrett, R.G., Dutter, R., 2008. Statistical data analysis explained : Applied environmental statistics with R. John Wiley & Sons, Chichester, England; Hoboken, NJ.