

Assessment of Statistical Methods Applied to Geochemical Data for Paleoenvironmental

Interpretation: *A Case Study from the Al-Azraq Basin, Jordan*

A Thesis  
in  
Environmental and Urban Geosciences

Presented to the Faculty of the University  
of Missouri-Kansas City in partial fulfillment of  
the requirements for the degree

Master of Science

By  
Jennifer Knipper

Geology, BS University of Missouri Kansas City, 2012

Kansas City, Missouri  
2021



ASSESSMENT OF STATISTICAL METHODS APPLIED TO GEOCHEMICAL DATA FOR  
PALEOENVIRONMENTAL INTERPRETATION: A CASE STUDY FROM THE AL-AZRAQ  
BASIN, JORDAN

Jennifer Knipper, Candidate for the Master of Science Degree

University of Missouri-Kansas City, 2021

ABSTRACT

This study investigates the appropriateness of statistical methods frequently applied to inorganic geochemistry data for indicating past environments. The data examined are two large geochemical data sets of cored lacustrine sediments from the Al-Azraq basin, Jordan. Lacustrine sediments contain valuable records of paleoenvironmental conditions, recording local and regional responses to environmental change. The Al-Azraq basin is a low lying an endorheic basin draining an area across eastern Jordan, southern Syria and northern western Saudi Arabia of approximately 12,700 km<sup>2</sup>. The basin is bounded to the west and east by silt dunes, and to the northwest by basaltic hills. Two cores, AZ1 and AZ2, are located on the eastern and western margins of the Al-Azraq Basin, respectively. AZ1 is located to the west of basalt fed springs and on the eastern edge of the centrally located oasis. AZ2 is located near the western margin where six major wadis enter the basin. Inorganic geochemical data generated by inductively coupled atomic emission spectroscopy (ICP-AES) show significant changes in elemental concentrations with depth throughout each sediment core. AZ1 cored sediments are described in previous work (Ahmad, 2013). ICP-AES data from the AZ2 sediments indicated periods of humidity and aridity, increased and decreased precipitation, and low and high

chemical weathering. This study contributes to understanding the paleoenvironments of the Al-Azraq basin during the mid-Pleistocene period.

Methods commonly used with geochemical data included data transformation, data standardization, outlier detection, and replacement. Both graphical and statistical methods examined data normality to assess the assumptions of the parametric requirement of normality, and determine the use of parametric or non-parametric statistical analyses. Both non-robust and robust estimates of location and scale, of which parametric and non-parametric statistical methods are based, respectively, are studied with geochemical data based on Rock (1988). Multivariate statistical methods included correlation analysis, principal components analysis, and hierarchical cluster analysis. Ahrens (1953), Rock (1988) and Reimann and Filzmoser (2000) stated geochemical data is commonly not normally or lognormally distributed. Overall the majority of geochemical elements in Al-Azraq sediment cores AZ1 and AZ2 are neither normal nor lognormal. Different graphical and statistical methods to determine data distribution largely yielded slightly different results. Neglecting that geochemical data are neither normal nor lognormally distributed may lead to biased or erroneous results with the use of parametric statistical methods. Non-robust estimates of location and scale used with AZ1 and AZ2 geochemical data corresponded to higher values and less consistent values compared to their robust counterparts. The results of parametric statistical methods with non-normally distributed data may yield geologically interpretable results; however, these methods are erroneous. The use of non-parametric statistical methods or the proper data standardization and transformation methods prior to the use of parametric statistical methods should be taken into consideration, as this may yield a better interpretation of the data.



The undersigned, appointed by the Dean of the College of Arts and Sciences, have examined a dissertation titled "*Assessment of Statistical Methods Applied to Geochemical Data for Paleoenvironmental Interpretation: A Case Study from the Al-Azraq Basin, Jordan*" presented by Jennifer Knipper, Candidate for the Master of Science Degree and certify that it is worthy of acceptance.

Caroline Davies, Ph.D., Committee Chair  
Department of Earth and Environmental Sciences

Wei Ji, Ph.D., Committee Member  
Department of Earth and Environmental Sciences

Dr. Fengpeng Sun, Ph.D., Committee Member  
Department of Earth and Environmental Sciences

## TABLE OF CONTENTS

ABSTRACT .....	iii
LIST OF ILLUSTRATIONS .....	xi
LIST OF TABLES .....	xiii
ACKNOWLEDGEMENTS .....	xv
CHAPTER	
1. INTRODUCTION	
Project Background .....	1
Research Objectives .....	2
Research Questions .....	3
Research Hypothesis .....	3
Research Significance .....	4
Thesis Organization .....	4
2. STUDY AREA	
Introduction .....	6
Geologic Setting .....	7
Geological History .....	11
Hydrology and Hydrogeology .....	18
Climate .....	20
Al-Azraq Basin Sediments .....	20
Al-Azraq Basin Chronology .....	22
3. STATISTICAL ANALYSES IN GEOCHEMISTRY	
Introduction .....	25
Geochemical Data .....	25
Normal Distribution .....	26
Censored Values .....	28
Outliers .....	29
Data Transformation .....	32
Data Standardization .....	33
Graphical Methods of Detecting Non-Normality .....	33
Skewness and Kurtosis .....	35
Shapiro-Wilk Test .....	36
Location and Scale Estimates .....	37
Correlation .....	43
Hierarchical Cluster Analysis .....	45
Principal Component Analysis .....	48
4. METHODOLOGY AND SAMPLING	
Introduction .....	53
Core Description .....	53
Inductively Coupled Plasma-Atomic Emission Spectroscopy (ICP-AES) .....	54
Statistical Methodology .....	58

5. AZ2 RESULTS	
AZ2 Sediment Core Description.....	65
Geochemical Concentration by Depth.....	66
Chemical Index of Alteration.....	77
Elemental Ratios.....	77
6. AZ2 DISCUSSION	
Introduction.....	80
Geochemical Elements.....	81
Chemical Index of Alteration.....	86
Weathering of AZ2 Sediments.....	88
7. AZ1 & AZ2 STATISTICAL RESULTS	
Introduction.....	92
Censored Values.....	92
Outliers.....	93
Geochemical Data Distribution.....	95
<i>Histograms</i> .....	95
<i>Quantile-Quantile Plots</i> .....	97
<i>Box Plots</i> .....	98
Statistical Methods of Determining Data Normality.....	98
<i>Skewness and Kurtosis</i> .....	98
<i>Shapiro-Wilk Test</i> .....	101
Estimates of Location.....	103
Estimates of Scale.....	109
Correlation.....	114
Hierarchical Cluster Analysis.....	120
Principal Component Analysis.....	124
8. AZ1 & AZ2 STATISTICAL DISCUSSION	
Geochemical Data Distribution.....	129
Performance of Location and Scale Estimates	
<i>Non-Robust Location Estimates</i> .....	130
<i>Robust Location Estimates</i> .....	131
<i>Scale Estimates</i> .....	133
Mutually Consistent Estimates of Location and Scale.....	134
<i>Non-Robust Location Estimates</i> .....	134
<i>Robust Location Estimates</i> .....	135
<i>Scale Estimates</i> .....	135
Multivariate Statistical Methods.....	136
<i>Correlation</i> .....	136
<i>HCA and PCA</i> .....	137
9. CONCLUSIONS	



Conclusions.....	138
APPENDIX	
A. AZ1 Inductively Coupled Plasma-Atomic Emission Spectroscopy (ICP-AES) Data .....	140
B. AZ2 Inductively Coupled Plasma-Atomic Emission Spectroscopy (ICP-AES) Data .....	145
C. AZ1 Box Plot Inner Fence Rule Outlier Charts .....	148
D. AZ2 Box Plot Inner Fence Rule Outlier Charts .....	151
E. AZ1 Median +/- 2MAD Rule Outlier Charts .....	154
F. AZ1 Histograms of Raw ICP-AES Data .....	156
G. AZ2 Histograms of Raw ICP-AES Data.....	160
H. AZ1 Histograms of Ln Transformed ICP-AES Data.....	164
I. AZ2 Histograms of Ln Transformed ICP-AES Data .....	168
J. AZ1 Q-Q and Detrended Q-Q Plots of ICP-AES Data.....	172
K. AZ2 Q-Q and Detrended Q-Q Plots of ICP-AES Data.....	177
L. Box Plots of Raw ICP-AES Data from AZ1 and AZ2.....	181
M. Box Plots of Ln Transformed ICP-AES Data from AZ1 and AZ2.....	183
REFERENCES .....	187
VITA.....	192



## List of Illustrations

Figure	Page
Figure 2.1. Map showing the approximate location of the Al-Azraq basin, Azraq oasis and Qa'Al-Azraq, adapted from Bender (1975). .....	7
Figure 2.2. Map of the Physiogeographic-Geologic Provinces of Jordan adapted from Bender (1975). .....	8
Figure 2.3. The geologic map of Jordan illustrating the Cenozoic sediment of dominating the eastern desert and the Neogene age basalt forming the northern margin around the Al-Azraq basin (modified from Naylor et al., 2013). .....	12
Figure 2.4. Faults of the Al-Azraq basin modified from Kaudse (2014). .....	18
Figure 3.1. Right Skewed Data Distribution of Calcium in AZ1. ....	28
Figure 4.1 Flow chart of data preprocessing methods. ....	58
Figure 4.2. Flow chart of graphical and statistical methods used to detect normality.....	60
Figure 4.3. Flow chart of multivariate statistical methods used. ....	63
Figure 5.1. Graphs of Major Elements with Depth of AZ2. ....	67
Figure 5.2. Graphs of Trace Elements with Depth of AZ2. ....	72
Figure 5.2. (Continued) Graphs of Trace Elements with Depth of AZ2. ....	73
Figure 5.3. Graphs of the CIA and Elemental Ratios with Depth of AZ2. ....	79
Figure 7.1. Pearson's Correlation Matrix of AZ1 Variables. ....	115
Figure 7.2. Spearman's Correlation Matrix of AZ1 Variables. ....	116
Figure 7.3. Pearson's Correlation Matrix of AZ2 Variables. ....	118
Figure 7.4. Spearman's Correlation Matrix of AZ2 Variables. ....	119
Figure 7.5. HCA Dendogram of Major Elements in AZ1. ....	121
Figure 7.6. HCA Dendogram of Trace Elements in AZ1. ....	122
Figure 7.7. HCA Dendogram of Major Elements in AZ2. ....	123
Figure 7.8. HCA Dendogram of Trace Elements in AZ2. ....	124



## List of Tables

Table	Page
Table 2.1. Stratigraphy of the East Jordan Limestone Plateau adapted from Ibrahim (1996). .....	14
Table 2.2. Stratigraphy of the Northern Plateau Basalt Province adapted from Ibrahim (1993). .....	16
Table 2.3. Dates from the Al-Azraq basin from Davies (2000; 2005). .....	24
Table 3.1. Chart of outlier detection rules, range of outliers addressed, and recommendations for application. ....	32
Table 4.1. ICP-AES major elements and corresponding units of concentration. ....	55
Table 4.2. ICP-AES trace elements and corresponding units of concentration. ....	56
Table 4.3. Number of samples and elements analyzed by ICP-AES in each sediment core. ....	58
Table 4.4. Graphical and statistical methods used to detect normality with univariate data. ....	59
Table 4.5. Estimates of Location and Scale used with univariate data of AZ1 and AZ2. ....	62
Table 7.1. Number and Percentage of Censored Values in the trace elements of AZ1 and AZ2. ....	93
Table 7.2. Number of outliers and extreme outliers for AZ1 and AZ2 data and Ln transformed data using outlier removal methods. ....	94
Table 7.3. Skewness and Kurtosis z-scores of AZ1. Bolded values are within +/- 1.96. ....	99
Table 7.4. Skewness and Kurtosis z-scores of AZ2. Bolded values are within +/- 1.96. ....	101
Table 7.5. Shapiro-Wilk Test Statistics of AZ1. ....	102
Table 7.6. Shapiro-Wilk Test Statistics of AZ2. ....	103
Table 7.7. Non-Robust Estimates of Location for Elements in AZ1. ....	104
Table 7.8. Robust Estimates of Location for Elements in AZ1. ....	106

## List of Tables

<b>Table</b>	<b>Page</b>
Table 7.9. Non-Robust Estimates of Location for Elements in AZ2.....	107
Table 7.10. Robust Estimates of Location for Elements in AZ2.....	109
Table 7.11. Estimates of Scale for Elements in AZ1.....	111
Table 7.12. Estimates of Scale for Elements in AZ2.....	113
Table 7.13. PCA Factors and Eigenvalues of AZ1 Variables.....	114
Table 7.14. PCA Factor Loadings of AZ1 Variables.....	114
Table 7.15. PCA Factors and Eigenvalues of AZ2 Variables.....	115
Table 7.16. PCA Factor Loadings of AZ2 Variables.....	116

## ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Caroline Davies for her encouragement, patience, and support throughout my undergraduate and graduate career. Her expertise was invaluable in the formulation of research questions, methodology, and the overall importance of my work.

I would like to thank the members of my committee, namely, Dr. Wei Ji and Dr. Fengpeng Sun for their support and availability.





## 1. INTRODUCTION

### **Project Background**

Sediments cored from the Al-Azraq basin (35° 30' – 37° 30' E, 29° 30' – 32° 00' N), but more specifically from the Qa', a mudflat ecosystem are the source of the geochemical data for this study. The Qa' Al-Azraq was once unique in that it was an oasis in the middle of the desert and a host to a very biodiverse group of flora and fauna, including migratory birds (Rollefson, et al., 1996). Over time, the extraction of water from the Al-Azraq groundwater basin has been detrimental to the oasis, depleting the springs that once discharged groundwater and degrading the surrounding environment. This has affected the people living in the nearby cities of Amman and Zarqa, as they rely on the Al-Azraq groundwater basin for municipal supply and irrigation. The Qa' Al-Azraq can now be characterized by saline groundwater and saline surficial soil. The AZ1 and AZ2 cored sediments extend from the surface to approximately 48.5 meters. Sediment samples analyzed by inductively coupled plasma-atomic emission spectroscopy (ICP-AES) had the objective of using these results to help explain past environments and climates of the Al-Azraq basin and broader Jordan Plateau.

Oftentimes with large geochemical data sets, statistical methods are used to aid in data interpretation. Many researchers may not understand the assumptions of the statistical methods they are using or know the reasons they chose the method in the first place. It is important to have an understanding of characteristics of the particular geochemical data set prior to choosing what statistical methods to use with the data. This can ultimately aid in a better interpretation of the results.

## Research Objectives

Exploring assessment methods commonly used with geochemical data is the objective of this thesis. This objective evaluates the data distribution using several statistical and graphical methods of determining normality of geochemical data. Also examined were outlier detection and replacement methods, including the box plot inner fence and the median  $\pm 2$  MAD methods.

This thesis explores estimates of location and scale. Parametric statistical methods rely on non-robust estimates of location and scale, the mean and standard deviation, respectively. If the geochemical data under study does not follow a normal distribution, the results of parametric statistical methods relying on the mean and standard deviation may be erroneous or contain little to no meaning. Nonparametric statistical methods rely on more robust estimates of location and scale. The results of nonparametric statistical methods used with geochemical data that does not follow a normal distribution may lead to a better interpretation of the data.

To determine the most mutually consistent estimates of location and scale, both non-robust and robust estimates of location and scale are studied with the geochemical data, based on Rock (1988). Non-robust estimates of scale included: (1) arithmetic mean, (2) geometric mean, (3) harmonic mean, and (4) midrange. Robust estimates of location included: (1) the 10, 20, and 25 % trimmed means, (2) median, (3) Gastwirth median, and (4) Trimean. Robust and non-robust estimates of scale included: (1) range, (2) standard deviation, (3) mean absolute deviation from the mean, (4) mean absolute deviation from the median, (5) median absolute deviation from the median, and (6) semi-interquartile range.

An additional objective of this thesis, using the geochemical data, is examining multivariate statistical methods including correlation analysis, principal components analysis and

hierarchical cluster analysis. Methods of data transformation and standardization, in addition to test parameters are chosen based on previous statistical studies using geochemical data.

Lastly, the ICP-AES statistical results and stratigraphy of the AZ2 sediment core provided information regarding past environments. Using core stratigraphy and looking at concentration of elements by depth provided insights including, but not limited to, past periods of aridity, humidity, evaporation, carbonate precipitation, aeolian and alluvial inputs, and chemical weathering.

### **Research Questions**

In this thesis, the research questions addressed, in regard to the geochemical data, are: (1) Does the data follow a normal or lognormal distribution? (2) How do non-robust and robust statistical estimates of location and scale perform and what are the most mutually consistent? (3) Are parametric statistical tests less valid compared to the non-parametric counterparts? And (4) What information does the statistical analyses of AZ2 sediments provide in terms of interpreting past environments?

### **Research Hypothesis**

Similar to previous studies, the majority of the two geochemical data sets will not follow a normal nor lognormal data distribution. Therefore non-robust estimates of location and scale used with these two data sets will correspond to higher and less consistent values compared to more robust estimates of location and scale. Smaller differences between robust estimates of location and scale will be observed in the data. Periods of humidity and aridity, increased and decreased precipitation, and low and high chemical weathering will be evident and identifiable based on the depletion and enrichment of relative concentrations of elements.

## **Research Significance**

Currently, there is a lack of studies of geochemical data distributions, the application of summary statistics with geochemical data, and even fewer examination of the statistical methods. This assessment of whether or not the ICP-AES data from two data sets, AZ1 and AZ2, follow a normal or lognormal distribution will contribute critical information to this lack of studies. In addition, this assessment of non-robust and robust estimates of location and scale with the two data sets will provide insight as to not only the understanding of the most mutually consistent estimates of location and scale, but to the selection of either parametric or non-parametric statistical tests. Furthermore, selecting the correct statistical test with geochemical data may lead to a better interpretation of the meaning of the data. Lastly, there are a lack of inorganic geochemical studies in the area of the Jordan Plateau. This research adds to the previous work in the Al-Azraq basin by using ICP-AES data as a main indicator of past environmental conditions represented by the AZ2 sediments and their correlation to AZ1 sediment record.

## **Thesis Organization**

Chapter 1 serves as an introduction to the research study. Chapter 2 details the geology, hydrology, hydrogeology, climate of the site area, chronology and previous research. Chapter 3 focuses on statistical analyses used in geochemistry and provides an overview of commonly used methods, as well as information on the nature of geochemical data, data transformation and standardization, outlier detection rules, and methods to account for censored data. Chapter 4 describes the geochemical and statistical methods used in this study and why these methods are chosen. Chapter 5 provides the results of ICP-AES data of AZ2. Chapter 6 discusses paleoclimate implications based on the results of the ICP-AES analysis of AZ2. Chapter 7 provides the results of statistical analyses results with the geochemical data of both AZ1 and

AZ2. Chapter 8 discusses the application of statistical methods used with the ICP-AES data of AZ1 and AZ2. Chapter 9 presents the conclusions of this study.

## 2. STUDY AREA

### **Introduction**

The Al-Azraq Basin is located in the northeastern portion of Jordan and extends to the northwest towards Syria and to the southeast towards Saudi Arabia. The Al-Azraq basin is located in a tectonically active depression, as the basin contains many faults and fault lineaments (Ibrahim, 1996). Igneous and sedimentary rocks crop out north of the Al-Azraq basin. Sedimentary rocks of marine origin crop out along basin margins, including limestone, chert, marl, clay and evaporates. In the north and northeastern portion of the basin, basalt covers a large area. The Al-Azraq oasis or central mudflat, also known as the Qa' Al-Azraq is unique in that it was a spring fed wetland located in a desert. The Qa' Al-Azraq was once very important for migratory birds, which utilized the oasis during periods of migration. The oasis was an important ecosystem in the middle of a desert, supporting a biodiversity of flora and fauna as far back as the Pleistocene (Rollefson, et al., 1996). Two sediment cores drilled in the Al-Azraq basin ( $35^{\circ} 30' - 37^{\circ} 30' \text{ E}$ ,  $29^{\circ} 30' - 32^{\circ} 00' \text{ N}$ ), specifically from the Qa' Al-Azraq, represent sediments extending from the surface to 51 m (AZ1) and from the surface to 48.5 m (AZ2). Figure 2.1. is a map of Jordan, including the boundary of the Al-Azraq basin, Qa Al-Azraq and the Al-Azraq oasis.

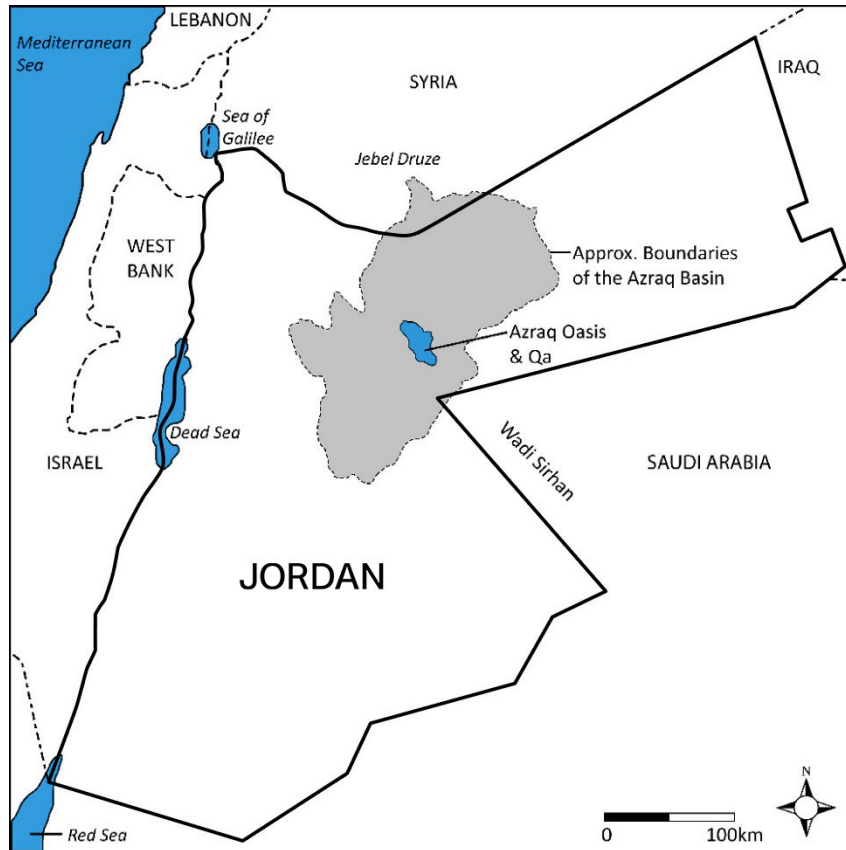


Figure 2.1. Map showing the approximate location of the Al-Azraq basin, Al-Azraq oasis and Qa' Al-Azraq, adapted from Bender (1975).

### Geologic Setting

Bender (1975) divides Jordan into the following seven physiographic provinces: Southern Mountainous Desert, Mountain Ridge and Northern Highlands East of the Rift, Central Plateau, Northern Plateau Basalt, Northeastern Plateau, Wadi al 'Arabah Jordan Rift, and the Highlands West of the Rift. The physiogeographic and geologic provinces overlap. The Northern Plateau Basalt and the Al-Azraq Wadi as Sirhan Basin are of particular importance, since these areas surround and form the watershed of the Qa' Al-Azraq. Figure 2.2. is a map of the physiographic-geologic provinces of Jordan.

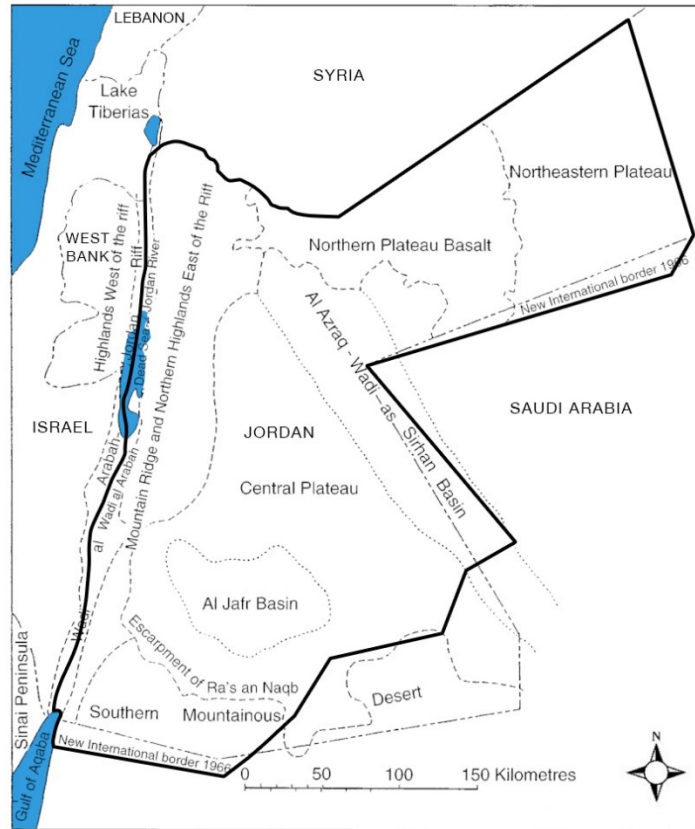


Figure 2.2. Map of the Physiogeographic-Geologic Provinces of Jordan adapted from Bender (1975).

The Central Plateau includes the Al-Jafr basin and the Al-Azraq-Wadi-as-Sirhan basin. In the Central Plateau, the Al-Jafr basin is located in the southwest, while the Al-Azraq-Wadi-as-Sirhan basin is located in the north and east. Horizontal and slightly dipping Upper Cretaceous and Tertiary limestone, marl and chert beds characterize the Central Plateau (Bender, 1975). In the west Central Plateau, drainage is to the west into the Rift Valley, while near the Al-Jafr basin, drainage is from all sides (Bender, 1975). In the east and north Central Plateau, drainage is northeast into the Al-Azraq-Wadi-as-Sirhan basin (Bender, 1975).

The Al-Azraq-Wadi-as-Sirhan basin occupies the area from the southern rim of the Jebel ed Drouz basalt shield southeast across the Jordan-Saudi Arabia border, and extends 300 km east into the An Nafud Desert in Saudi Arabia (Bender, 1974). In this southeast striking basin, the



lowest elevation is 500 meters above sea level (masl) in the northeast and the highest elevation is 700 masl in the southeast (Bender, 1974). The Al-Azraq-Wadi-as-Sirhan basin is a shallow, elongated structural depression characterized by several faults in addition to silt dunes, alluvium, the Azraq Formation and Neogene-Quaternary sediments (Ibrahim, 1996). The Al-Azraq basin contains 5,550 m of Paleozoic to Pleistocene carbonates, sandstone and shale (Core Lab, 1987).

The Al-Azraq-Wadi-as-Sirhan basin is a major internal drainage basin of 35,500 km<sup>2</sup> (Edgell, 2006). The catchment area of the northeast oriented wadi systems drain into the depression and extend 100 km west into the Central Plateau (Bender, 1974). The Al-Azraq depression is located within the Al-Azraq-Wadi-as-Sirhan basin and is smaller than the Al-Jafr basin, measuring 50 km in length and 30 km in width (Ibrahim, 1996). The Qa' Al-Azraq is a 100 km<sup>2</sup> playa located at an elevation of 505 masl in the center of the Al-Azraq depression (Ibrahim, 1996). The Qa' Al-Azraq is the second largest playa located on the Central Plateau (Ibrahim, 1996). Al-Jafr basin is 150 km in length and 100 km in width and contains the largest playa on the Central Plateau, with a total area of 240 km<sup>2</sup> (Bender, 1975).

Cuesta topography characterizes the area in the Central Plateau, located to the west and southwest of the Al-Azraq-Wadi-as-Sirhan basin (Bender, 1974). The cuestas are composed of Cretaceous and Tertiary limestone, marl and chert beds that exhibit a dendritic drainage pattern (Bender, 1974; Ibrahim, 1996). The cuestas are mostly covered by chert detritus, a result of arid weathering (Bender, 1974). The Fuluq fault scarp and pediment is located east of the Al-Azraq-Wadi-as-Sirhan basin, and hills of Umm Rijam Chert Limestone and Mawaqqar Chalk Marl Formations are covered with flint debris (Ibrahim, 1996).

The Northern Plateau Basalt province borders the Central Plateau to the north and northeast. This province occupies 11,000 km<sup>2</sup> in Jordan and 34,000 km<sup>2</sup> in Saudi Arabia

(Bender, 1974). In total, the basalt shield, Harrat esh Shaba, is 45,000 km<sup>2</sup> in size (Bender, 1974). Neogene-Quaternary flood lava sloping mainly to the south and southeast characterizes the basalt shield (Ibrahim, 1996).

The Harrat esh Shaba basalt shield is the southern extension of the Jebel ed Drouz-Basalt-Highlands (Bender, 1974). The area of Harrat esh Shaba is located at the southern boundary of the Damascus basin in Syria and extends east of the Al-Azraq-Wadi-as-Sirhan basin (Bender, 1974). In the north and northwest, the elevation is 1,100 masl near the northern Jordan border and the elevation is 500 masl at the eastern edge of the Al-Azraq-Wadi-as-Sirhan basin (Bender, 1974). The western part of the Northern Plateau Basalt province drains to the south and southwest into the Al-Azraq-Wadi-as-Sirhan basin (Bender, 1975). The drainage pattern on the basalt shield outlines joints and faults (Ibrahim, 1996).

The Northeastern Plateau province is a limestone plateau that sits adjacent to the Northern Plateau Basalt province (Bender, 1974). The Northeastern Plateau province extends east across the Jordan-Iraq border, north across the Jordan-Syria border and south across the Jordan-Saudi Arabia border (Bender, 1975). Drainage in this area is northwest, north-northwest and east-northeast and follows zones of structural weakness (Bender, 1975). Calcareous and siliceous Tertiary rocks overlain by Tertiary calcareous rocks and sandstone along with Quaternary sediments characterize the Northern Plateau province (Bender, 1975).

The Southern Mountainous Desert province borders the Central Plateau to the south. The Southern Mountainous Desert province is located south of the Ra's en Naqb escarpment and extends south into Saudi Arabia (Bender, 1975). The Southern Mountainous Desert province drains south into the Gulf of Aqaba and east into the Al-Azraq-Wadi-as-Sirhan basin (Bender, 1975). The cuestas located in the Central Plateau are bound in the southwest by the west-

northwest striking, 1,700 masl escarpment at Ra's en Naqb (Bender, 1974). This escarpment divides the Central Plateau and the Southern Mountainous provinces.

The Mountain Ridge and Northern Highlands East of the Rift province borders the Central Plateau to the west. The Mountain Ridge and Northern Highlands East of the Rift province extends north and northeast for 370 km from the Gulf of Aqaba to Lake Tiberias (Bender, 1975). The highest altitude in Jordan is located in this province at 1,850 m in Jibal ash Sharah (Bender, 1975). This province slopes towards the Central Plateau in the east and towards the Wadi Al-'Arabah Jordan Rift province in the west (Bender, 1974). The Wadi Al-'Arabah Jordan Rift province, located west of the Mountain Ridge and Northern Highlands East of the Rift province drains west into the rift valley (Burdon, 1959; Bender, 1974).

### **Geological History**

Since the Cambrian, the following three major events are responsible for the paleogeographic evolution of Jordan: (1) multiple transgressions of the Tethys Sea located in the north and northwest, over parts of Jordan, (2) the formation of the Wadi Al-'Arabah-Jordan geosuture, which divides the Palestine block and the Transjordan block, and (3) the dipping of the Nubo-Arabian shield located in the south and southeast, transporting chemical and mechanical weathering products and the deposition of continental clastic sediments to the shelf of the Tethys Sea (Bender, 1974).

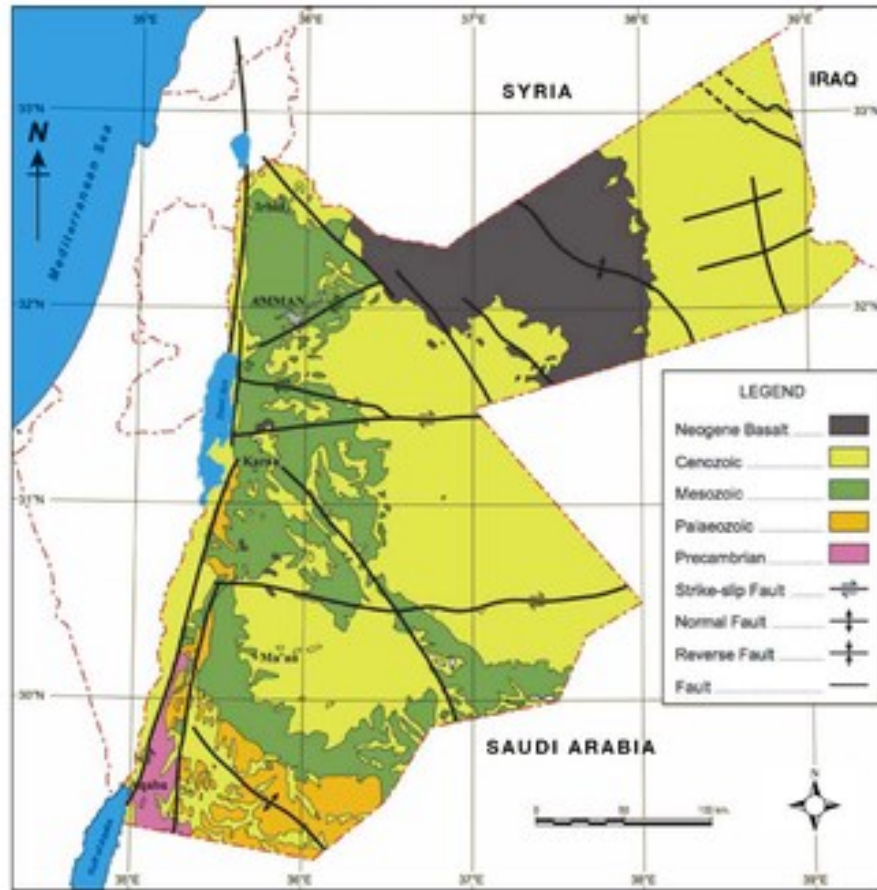


Figure 2.3 The geologic map of Jordan illustrating the Cenozoic sediment of dominating the eastern desert and the Neogene age basalt forming the northern margin around the Al-Azraq Basin (modified from Naylor et al., 2013).

During the Oligocene, tectonic movements caused land area and shallow basin separation in the East Jordan Limestone Plateau, as well as an accumulation of alternating marine and continental sediments (Bender, 1974). During the Neogene, calcareous and sandy marine deposits developed in the Al-Azraq-Wadi-as-Sirhan basin (Bender, 1974). Pleistocene lacustrine and fluvial deposits lie on top of the uplifted areas of the East Jordan Limestone Plateau (Bender, 1974). The Quaternary deposits on the limestone plateau are of continental origin (Bender, 1974).

From the Cambrian to the Mid-Cretaceous, erosional phases dominated in the Al-Azraq basin, resulting in the missing Cambrian to Mid-Cretaceous sedimentary sequence (Ibrahim, 1996). From the late Cretaceous to Oligocene, epirogenic events influenced the sedimentation and erosion of the Al-Azraq basin (Ibrahim, 1996). Regional uplift occurred from the Eocene to Oligocene after the marine sedimentation of the Umm Rijam and Wadi Shallala Formations of the Belqa Group (Ibrahim, 1996). Erosion, transport and deposition of sediments from the Qirma Calcareous Sandstone and Azraq Formations to the Al-Azraq basin occurred during the Oligocene to Miocene (Ibrahim, 1996). During the late Miocene, a shallow marine and lacustrine environment characterized the Al-Azraq basin (Ibrahim, 1996). From the late Eocene to Pleistocene, uplift and erosion removed most of the Belqa series (Powell, 1989). From the Pliocene to the Pleistocene, the Al-Azraq basin was a closed, saline alkaline lake (Ibrahim, 1996). Epirogenic swells and basins, including the Al-Azraq-Wadi-as-Sirhan and Al-Jafr that strike northwest and west-northwest, characterize block faulting in central and southeast Jordan (Bender, 1974). The Neogene-Quaternary sediments found in the Al-Azraq basin formed from subsurface block faulting (Ibrahim, 1996). Bender (1974) describes the East Jordan Limestone Plateau as a 45,000 km<sup>2</sup> area extending from the arched rim of the Wadi Araba-Jordan graben to beyond the eastern border of Jordan. The plateau dips from the arched rim of the Wadi Al-Araba-Jordan graben to the east and from the Ra's en Naqb escarpment to the northeast, while the plateau dips in the east toward the Al-Azraq-Wadi-as-Sirhan basin (Bender, 1974). The plateau consists of mostly Upper Cretaceous and Tertiary calcareous sediments in which the thickness decreases from 800 m in the west to 300 m and 150 m in the south and east, respectively (Bender, 1974). The upper part of the Belqa Group consists of the Umm Rijam Chert-Limestone and the overlying Wadi Shallala Formations and is the bedrock of the Al-Azraq

basin (Ibrahim, 1996). The Wadi Shallala Formation consists of chalk with barite concretions containing chert layers and is overlain by the Qirma Calcareous Sandstone Formation which consists of sandy and calcareous limestone containing pebbles of chert and quartzite (Ibrahim, 1996). The Azraq Formation lies above the Qirma Calcareous Sandstone Formation and contains mainly evaporates, gypsiferous clays and micrites (Ibrahim, 1996). The Qirma Calcaerous Sandstone Formation is covered by Pleistocene and Holocene alluvial sediments (Ibrahim, 1996). Table 2.1. presents the stratigraphy of the East Jordan Limestone Plateau.

Table 2.1. Stratigraphy of the East Jordan Limestone Plateau adapted from Ibrahim (1996).

Era	Period	Epoch	Group	Formation	
Cenozoic		Holocene		Alluvium and Wadi Sediments Alluvial Mudflat Alluvial Silt and Siltflat Alluvial Fan Dashsha Silt	
		Pleistocene		Alluvium or Colluvium	
		Pleistocene		Azraq	
	Neogene	Pliocene Miocene			
			Miocene		Qirma Calcareous Sandstone
		Paleogene	Eocene	Belqa	Wadi Shallala
			Eocene		Umm Rijam Chert-Limestone
Mesozoic	Cretaceous				Muwaqqar Chalk-Marl
				Usaykhim Sandy Limestone	

The Northern Plateau Basalt Province extends from the southern rim of the Damascus basin in Syria south to the northern rim of the Al-Azraq basin (Bender, 1974). The eastern Badia basalt plateau is part of the North Arabian Volcanic Province, extending from the southern rim of the Damascus Basin in Syria, across Jordan and into Saudi Arabia (Barberi et al., 1979). The

basalt plateau consists of late Tertiary and Quaternary volcanic basalt lava flows (Burdon, 1959; Bender, 1974; 1975). The basalt plateau is between 50 to 170 km wide and extends over 180 km in the north-south direction. The exposed Neogene-Quaternary basalts of the Northern Plateau Basalt Province include the Harrat ash Shaam Basaltic Super-Group, which is classified into five groups consisting of flood lavas, dyke systems, pyroclastic sediments and volcanic centers (Ibrahim, 1993). K-Ar dating techniques concluded that the absolute ages of the basalts range from 13.7 Ma to less than 0.5 Ma (Barberi et. al., 1979, Moffat, 1988, Ibrahim, 1993). The Wisad Group is the oldest and consists of flood flows, shield volcanoes and dyke systems (Ibrahim, 1993). The Wisad Group lies underneath the Safawi Group or the Qirma Calcareous Sandstone Formation (Ibrahim, 1993). Sandy limestone and calcareous sandstone with pebbles of chert and quartzite composes the Qirma Calcaerous Sandstone Formation (Ibrahim, 1996). The Safawi Group is comprised of flood flows and feeder dyke systems and outcrops near the Al-Azraq basin (Ibrahim, 1993). The Asfar Group lies above the Safawi Group and includes dyke systems and flood lavas that once flowed from volcanic centers such as shield volcanoes, stratovolcanoes and scoria cones (Ibrahim, 1993). The Rimah Group lies above the Asfar Group and is an artificial group and includes volcanosedimentary and scoriaceous deposits associated with the volcanic center eruptions of the other groups (Ibrahim, 1993). The Bishriyya Group is the youngest and is composed of 'a 'a and pahoehoe flows (Ibrahim, 1993). Table 2.2. presents the stratigraphy of the Northern Plateau Basalt Province.

Table 2.2. Stratigraphy of the Northern Plateau Basalt Province adapted from Ibrahim (1993).

Formation	Group	Age
Fahda Vesicular Basalt	Bishriyya	1.45-0.1
Wadi Manasif Basalt		
Aritayn Volcaniclastic	Rimah	2.94-
Hassan Scoriaceous		2.01
Mahada Basalt	Asfar	3.41- 1.96
Madhal Olivine Phyric Basalt		
Hashimyya Aphanetic		
Ushayhib Pyroxene Phyric		
Ufayhim Xenolithic Basalt		
Salaman Basalt	Safawi	9.30- 8.45
Abed Olivine Phyric Basalt		
Ali Doleritic Basalt		

The East Jordan Limestone Plateau contains many faults with minor displacements (Bender, 1974). The western part of the plateau near the arched rim of the Wadi Al-‘Araba-Jordan graben contains faults that trend north-northwest to north, northwest and eastwest (Bender, 1974). In the central part of the plateau, epeirogenic tilting and normal faults occur striking northwest (Bender, 1974). Epeirogenic large scale tilting is responsible for the dip of the arched rim of the Wadi Al-‘Araba-Jordan graben to the east and the dip of the East Jordan Limestone Plateau to the east/northeast towards the Al-Azraq-Wadi-as-Sirhan basin (Bender, 1974).

The Al-Azraq Wadi as Sirhan and Al-Azraq basins are both fault bounded and trend northwest-southeast. Several periods of tectonic faulting formed the Al-Azraq basin (Ibrahim, 1996). The Al-Azraq basin is the northwestern extension of the Al-Azraq-Wadi-as-Sirhan basin (Ibrahim, 1996). A gentle arch separates the two basins (Ibrahim, 1996).

Three major tectonic periods occurred in the Al-Azraq basin with the possibility of more tectonic periods due to the missing sedimentary sequence (Ibrahim, 1996). These periods



occurred in the late Cretaceous, Paleocene and Eocene (Core Lab, 1987). During the Cenozoic, the Al-Azraq-Wadi-as-Sirhan and Al-Azraq basins experienced epeirogenic subsidence with no large scale faulting (Bender, 1974). Synsedimentary faulting occurred in the Cretaceous; however, these faults show little to no surface expression (Bender, 1974).

Corelab (1987) describes the Al-Azraq basin as being bounded to the north and northeast by the Fuluq fault, to the south by the Siwaqa fault, to the east by the Fuluq fault extending towards the Siwaqa fault and to the west by a sedimentary outcrop near Amman. Both the Fuluq and Siwaqa faults are normal faults; however, strike slip movement occurred on the Fuluq fault in the past (Core Lab, 1987; Ibrahim, 1996).

In the Al-Azraq basin, the sedimentary column consists of 5,500 m of Paleozoic to Pleistocene carbonates, sandstone and shale (Corelab, 1987). Other structural features identified in the Al-Azraq basin include: the Hamza Graben, Wadi Rajil Homocline, Umary Road Fault Zone and the Dahikiyah-Wadi Ghandaf Platform (Corelab, 1987). The Al Bayda and Baqaqiyya faults bound the playa in the northeast and the Qaisiyeh fault bounds separates the playa from the Northern Plateau Basalt Province and the outcrop of the Umm Rijam Chert-Limestone Formation (Ibrahim, 1996).

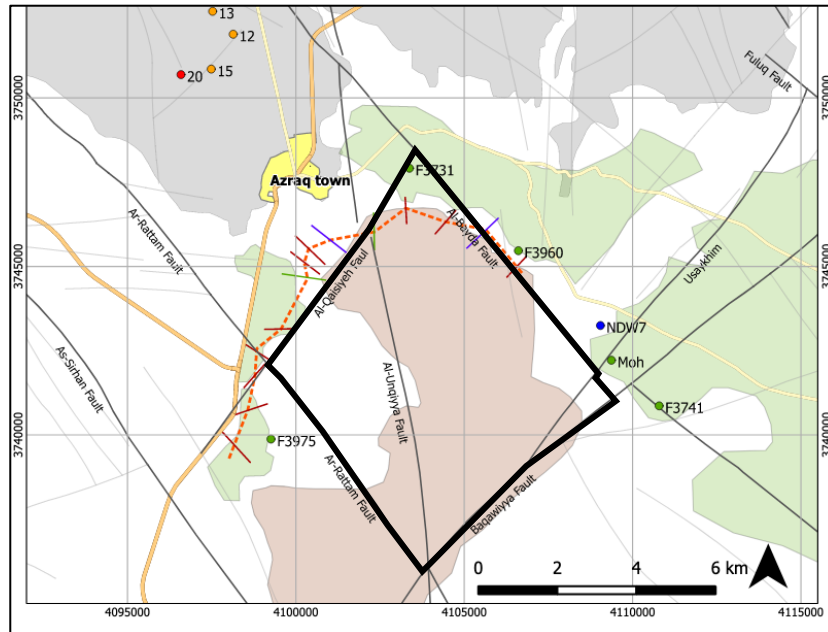


Figure 2.4. Faults bounding the central Al-Azraq basin containing the central mudflat in brown, northern basalt field in grey modified from Kaudse (2014). The Fuluq and Ar-Rattam and Sirhan faults are parallel and run NW to SE. These faults are intersected by the parallel NE-SW running Baqawiyya and Al-Qaysiyeh faults. These form the Hamza Grabben, the basin depression in which sediments accumulated. Modified from Kaudse (2014).

## Hydrology and Hydrogeology

The Al-Azraq basin is one of the largest and most important groundwater basins in Jordan, as this basin supplies those people living in the cities of Amman, Zarqa and a very large refugee population near the Syrian border (Frances 2015; Hegazi 2016; Deutsche Welle 2016). Groundwater in the Al-Azraq basin is an important source of freshwater for domestic supply and surrounding agriculture. Groundwater in the Al-Azraq basin has been utilized since the 1970s and over time the basin is increasingly overexploited. Proper management of the basin is critical to groundwater supply. The Al-Azraq basin is poorly managed and has been exploited since the 1990s (Ibrahim, 1996). Over pumping has resulted in the degradation of groundwater quality and

lowering of the water table. Access to water is a critical issue in the Middle East, with Jordan the fifth most water stressed country in the world (Hofste et al., 2019); thus, understanding groundwater recharge is essential for future management of the basin (Davies, 2000).

Surface water in the Al-Azraq basin is dependent upon runoff and wadi inputs, but is limited to rain mainly during the winter. The drainage network consists of six main wadis that are oriented towards the basin (Sahawneh, 1996). The Al-Azraq basin was once a permanent oasis; however, groundwater extraction has affected the springs, which used to be the source of water supply during times of no precipitation. In 1986, the Drouze springs dried completely, while the Shishan springs discharge was reduced by 80 percent (El-Naqa et al., 2007). Four springs once existed, of which their discharge was from the upper aquifer; however, discharge of the springs stopped in the 1990s due to over pumping (Bajjali and Al-Hadidi, 2006).

Three aquifer systems are identified in the Al-Azraq groundwater basin by depth below the surface, the deep, middle, and shallow. The deep and middle aquifers are both confined. The middle aquifer system contains fresh to brackish water and the deep aquifer system contains brackish water (Bajjali and Al-Hadidi, 2005). The deep aquifer consists of sandstone, while the middle aquifer consists of limestone and dolomitic limestone with layers of chalk and evaporates (Bajjali and Al-Hadidi, 2005). The middle aquifer is recharged from the Jebel Druze area located to the north (Bajjali and Al-Hadidi, 2005). The shallow aquifer consist of four hydraulically connected formations and include the following: Rijam, Shallala, Basalt, and Quaternary deposits (Bajjali and Al-Hadidi, 2005). The shallow aquifer system is recharged from the north as well (Ibrahim, 1996). The springs discharged from the Rijam formation, which consists of phosphatic and chalky marl; however, overexploitation has resulting in the springs completely drying up (Bajjali and Al-Hadidi, 2005). The Shallala formation consists of chalk, chalky marl,

marly limestone, and sandy layers (Bajjali and Al-Hadidi, 2005). The basalt aquifer is exploited, as is the shallow aquifer which is used for domestic supply, agriculture and the local salt industry.

### **Climate**

Northern and western Jordan highlands fall within a Mediterranean climate, with an average annual rainfall of more than 300 mm (Bender, 1974). Summer extends from April to October and winter is from November to March, with an average annual temperature of 38.8 ° C and 0.5 ° C, respectively (Bender, 1974). The deserts in eastern and southern Jordan lie within a semi-arid transition zone in which the average annual rainfall is between 50 mm to 300 mm (Bender, 1974). Here, the average annual temperatures are between -1.6 ° C and 40 ° C (Bender, 1974). To the east and south of the semi-arid transition zone is the arid zone in which the average annual rainfall is less than 50 mm (Bender, 1974).

The climate of the Al-Azraq basin is arid to semi-arid, with minimum and maximum temperatures of 10 ° C to 24.5 ° C, respectively (Ibrahim, 1996). The Al-Azraq basin experiences winters, from October to April, and summers, from May to October. The Al-Azraq Basin climate is characterized by hot and dry summers, and wet and cold winters. The annual maximum rainfall is 100 mm to the north and 50 mm to the east and south, and some years with no precipitation at all (Ibrahim, 1996).

### **Al-Azraq Basin Sediments**

Cane (1992) performed a paleoclimate reconstruction based on a 73 meter sediment core collected from an unspecified location in the Al-Azraq basin for a Master's thesis. X-ray diffraction, ICP-AES,  $\delta^{18}\text{O}$  and  $\delta^{13}\text{C}$  isotopes and radiocarbon dating were the analytical methods. This study concludes a change in climate occurred between the Mid-Pleistocene to Late

Holocene in north central Jordan (Cane, 1992). During the Mid-Pleistocene, humid, pluvial conditions dominated and were punctuated by three semi-arid periods (Cane, 1992). During the Late Pleistocene, an extensive arid period dominated (Cane, 1992). During the Early Holocene, pluvial conditions dominated (Cane, 1992). During the Late Holocene, arid conditions dominate, as seen in today's climate in north central Jordan (Cane, 1992).

Ahmad (2013) performed a high resolution study of Al-Azraq cored sediments, focusing on lithology and the stable isotopic composition of organic matter and carbonate to identify environmental zones in AZ1, a sediment core extending from the surface to 51 m, located in Qa' Al-Azraq. Methods used include oxygen and carbon isotopes of carbonate, carbon and nitrogen isotopes of organic matter, ICP-AES, magnetic susceptibility, smear slides, grain size analysis, x-ray diffraction and scanning electron microscopy (Ahmad, 2013). Three paleoenvironmental zones are present in the sediments (Ahmad, 2013).

Ahmad and Davies (2017) modeled basin evolution of the Qa' Al-Azraq, focused on sulfur isotope analysis to distinguish sources of sulfur and gypsum. This study identified processes including algae production, evaporation, pyrite oxidation, dissolution and recrystallization processes, by looking at the isotope chemistry of gypsum and sulfur, among other methods such as smear slides, XRD, and SEM. This study identified distinct environments from sulfate mineral processes and provides a model of basin evolution in the Al-Azraq basin.

Marsh, lake, and playa environments along with cycles of seasonal deposition indicate climate shifts occurring during the Middle to Late Pleistocene (Ahmad, 2013). Using oxygen isotope data and regional eastern Mediterranean records, the upper sediment corresponds to Marine Isotope Stage (MIS) 5 and the end of the Pleistocene, while basinal sediments are potentially between MIS 15 and 9 (Ahmad, 2013). Low lake levels with a short high

precipitation phase characterizes zone 1 (Ahmad, 2013). High lake levels, presence of diatoms and abundance of organic matter dominate in Zone 2 (Ahmad, 2013). High organic matter indicates a marsh deposit in Zone 3, as well as a dry period at the base of the core (Ahmad, 2013).

Ahmad (2010) performed an analysis of the past environments and climates of the Al-Azraq basin from cored sediments, AZ3, extending from 33 to 56 m located in the Qa'Al-Azraq. Methods in this study included percent organic carbon,  $\delta^{15}\text{N}$  and  $\delta^{13}\text{C}$  isotopes of organic matter, smear slides, grain size analysis and x-ray diffraction (Ahmad, 2010). This study identifies an extended period of humid climate and a much shorter period of arid climate, (Davies, 2000; Ahmad, 2010). Two species of diatom are identified, *Aulacoseira* sp. and *Stephanodiscus* sp. (Ahmad, 2010). The sources of organic matter include marine algae and terrestrial plants (Ahmad, 2010). Carbon and nitrogen stable isotope data and the distribution of organic matter characterize five environmental zones. In zone 1, aquatic algae, low organic matter and a moist climate dominates. Zones 2 and 3 are characterized by a moist climate and an increased amount of organic matter relative to zone 1. Zones 4 and 5 represent low organic matter and a dry climate (Ahmad, 2010).

### **Al-Azraq Basin Chronology**

Dating methods including optically stimulated luminescence (OSL) and U/Th are used on samples surrounding the Al-Azraq basin. Root bound sandstone is exposed in the Azraq Formation in Dahikiya, located in the southern Badia Region, northeastern Jordan (Turner and Makhlof, 2005). One sample of sandstone was dated using OSL to a minimum age of 652 +/- 47 ka and deposition was during MIS 17 (Turner and Makhlof, 2005). The Cardium horizon in the uppermost sediment of the Azraq Formation is exclusively found in two locations

surrounding the Al-Azraq basin (Abed et al., 2008). The first Cardium horizon is located 10 km northeast of the Al-Azraq basin at 31 55' N, 36 45' E, while the second horizon is located on the potential southern edge of the Al-Azraq basin at 31 32' N, 37 06' E (Abed et al., 2008). U/Th dating on two samples from the first horizon produced an average age of 330 ka, concluding that marine deposition occurred during MIS 9 (Abed et. al, 2008). The mineralogy and types of fossils present as well as distribution of the two Cardium horizons found in the Azraq Formation are evidence of the existence of one large lake near the Saudi border or several smaller shallow lakes occupying the Al-Azraq basin during MIS 9 (Abed et al., 2008).

Radiocarbon dating on a charcoal sample taken at approximately 6.3 m below the surface from a 73 m sediment core located in an unspecified location in the Al-Azraq basin date to 40,000 ka (Cane, 1992). Cane (1992) extrapolated a sedimentation rate of 40,000 years per 6.3 m of sediment. Cane (1992) determined the length of time to fill the Al-Azraq basin was 500,000 years.

Dating methods including radiocarbon and IRSL used on samples within the Al-Azraq basin reveal slightly different results. Radiocarbon dating on a charcoal sample taken at approximately five meters below the surface from AZ1 cored sediments, located at 35 24613 N, 37 297218 E, date to  $11,460 \pm 40$  ka (Davies, 2000). IRSL on three sediment samples from AZ1 are internally coherent; at 0.20-0.28, 5.37-5.46 m and 11.62 m below the surface date to  $24.2 \pm 2.0$  ka,  $163.3 \pm 11.7$  ka and greater than 250 ka, respectively (Davies, 2005). A minimum basal age determined to be 250 ka at 11.62 m was the limit of IRSL. At present, there is no maximum date (Davies, 2005). Table 2.3. presents radiocarbon and IRSL dates from the Al-Azraq basin.

During the Pleistocene, the Al-Azraq basin was a lake (Bender, 1968). Davies (2000) concluded that the Al-Azraq basin contains lacustrine clays deposited during a period of increased moisture, and calcium carbonate nodules, calcrete bands, and aragonite deposited during a dry period. Clay layers from lacustrine deposits of the endorheic basins of the Central Plateau indicate that deep-water environments once existed (Davies 2000; 2005). Sediments from a 31 meter sediment core obtained from the Qa' Al-Jafr shows various types of deposition occurred in the Central Plateau, including alluvial, aeolian and lacustrine (Davies, 2005). Abed *et al.* (2008) also concludes the Al-Azraq area was once much wetter compared to the current climate conditions based on his work and the work of Davies (2000), Khoury (2003), and Turner and Makhlouf (2005). A lake occupied the central part of the Al-Azraq basin between 250-500 ka (Davies, 2000, Khoury, 2003). Turner and Makhlouf (2005) dated sandstone of the Azraq Formation to MIS 17, which was a warm, wet period.

Table 2.3. Dates from the Al-Azraq basin from Davies (2000; 2005).

Sample Composition	Dating Method	Depth (m)	Time (ka)	Author, date
charcoal	Radiocarbon	5	11.46 ± 40	Davies, 2000
sediment	IRSL	0.20-0.28	24.2 ± 2	Davies, 2005
sediment	IRSL	5.37-5.46	163.3 ± 11.7	Davies, 2005
sediment	IRSL	11.62	> 250	Davies, 2005



### 3. STATISTICAL ANALYSES IN GEOCHEMISTRY

#### **Introduction**

This chapter provides information regarding the definitions and characteristics of statistical methods used in geochemical studies and examined in this research. The background information on how statistical methods have been previously applied to geochemical data and methods of accounting for censored values and outliers are discussed, as well as data standardization and transformation methods. Estimates of location and scale are explained, and lastly multivariate statistical methods are included in this chapter, including those methods examined in this study.

#### **Geochemical Data**

Geochemical data are closed data sets, expressed as a portion of a whole, such as weight percent or parts per million. Geochemical data are compositional, as parts of a whole being quantified. The variable that accounts for the largest proportion of the whole is not constrained; however, each consecutive variable is more constrained than the one preceding. Individual variables in the geochemical data set are not free to vary independently, therefore, the geochemical data set is closed (Filzmoser et al., 2009). As for data distribution, the constant sum constraint does not allow the compositional data to be truly univariate (Filzmoser et al., 2009). The compositional nature of geochemical data should be considered prior to data analysis (Reimann and Filzmoser, 2000).

## Normal Distribution

A normal distribution, also known as a bell-shaped or Gaussian distribution, is a theoretical distribution that is often assumed to be true in parametric statistics. A standard normal distribution is a class of probability curves that is defined based upon the population mean,  $\mu$ , and the population standard deviation,  $\sigma$ , for the continuous random variable. In the standard normal distribution,  $\mu = 0$  and  $\sigma = 1$ , and skewness and kurtosis values are equal to zero. A standard normal distribution is unimodal, single peaked and symmetrical around  $x = \mu$ . The range of a standard normal distribution is between negative to positive infinity.

Continuous random variables that are normally distributed use the empirical rule, which determines probabilities. The empirical rule for normal distributions states that: approximately 68% of the area under the curve is above  $x$  values in the range  $(\mu - 1\sigma)$  to  $(\mu + 1\sigma)$ , approximately 95% of the area under the curve is above  $x$  values in the range  $(\mu - 2\sigma)$  to  $(\mu + 2\sigma)$  and approximately 99% of the area under the curve is above  $x$  values in the range  $(\mu - 3\sigma)$  to  $(\mu + 3\sigma)$  (Pukelsheim, 1994). The total area under a normal distribution curve is equal to 1. The standard normal distribution curve includes two inflection points located at  $x = \mu - \sigma$  and  $x = \mu + \sigma$ . The horizontal distance along the  $x$ -axis from  $x = \mu$  to either one of the two inflection points, located at  $x = \mu - \sigma$  and  $x = \mu + \sigma$ , respectively, is equal to 1 standard deviation.

Calculating the curve and the proportional area under the normal distribution curve that is in between the selected range generates the range of data. There is high probability associated with values of the random variable near the mean or center of the normal distribution, and a lower probability associated with values of the random variable further away from the mean. The Central Limit Theorem states that for a large sample size, the random samples taken from the population will follow a normal distribution. Using a standard normal table determines the

probabilities associated with the values of a random variable; however, the values need to first be converted into z-scores.

Rock (1988) states geochemical data should never be assumed to be normal. If different geochemical processes influence the measurement, the samples will not have the same distribution (Reimann et al., 2005). It is common for a distribution to be a mixture of distributions caused by multiple sub-populations within the geochemical data set (Filzmoser et al., 2009). Violation of the assumption of normality occurs when the distribution consists of a mixture of populations or contains outliers (Grunsky, 2010). Generally, the distribution of geochemical data sets skew positively and contain large outlier values; however, this may be even more observable when looking at the distribution of trace element data (Rock, 1988).

Skewed or polymodal populations indicate the distribution consists of a mixture of populations (Grunsky, 2010). Right skewed distributions are the most commonly observed distributions with geochemical data (Grunsky, 2010). In a right skewed distribution, the mean is greater than the median. Figure 3.1. presents an example of a right skewed distribution using Calcium in AZ2. Most of geochemical distributions are not normal or lognormally distributed, but are binomial, Poisson, bimodal, polymodal or irregular (Krumbein and Graybill, 1965, Rock, 1988).

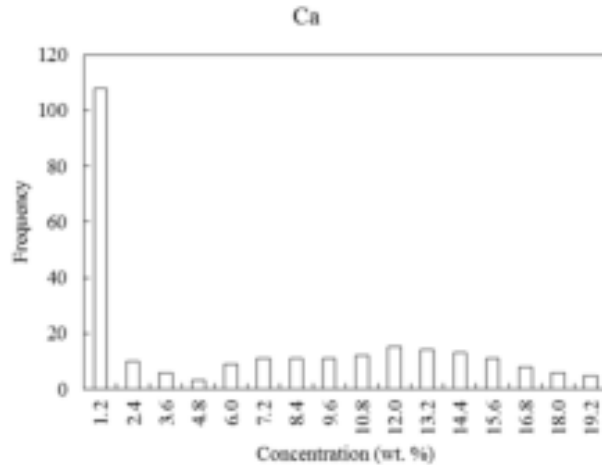


Figure 3.1. Right Skewed Data Distribution of Calcium in AZ1.

### Censored Values

Reports of censored geochemical data occur as either less than or greater than the lower and upper limits of detection and denoted by a '<' or '>', respectively. The limits of detection correspond to the limits of the analytical procedure, such as ICP-AES. Detection limit problems are often a result of trace element analysis resulting in a loss of precision at both very low and very high concentrations (Reimann and Filzmoser, 2000). According to Reimann and Filzmoser (2000), geochemical data are imprecise measurements. Censored values, including those below and above the limits of detection have no true value (Reimann and Filzmoser, 2000). Sampling and instrument bias can also influence geochemical data and are the results of error introduced during the sampling, sample preparation, and/or analytical process (Reimann and Filzmoser, 2000).

Truncated data sets are different from continuous data sets in that truncated data sets contain values that are above or below the limits of detection (Rock, 1988). It is common in geochemical data sets, for data to be truncated in the lower tail (Rock, 1988). It is rare for geochemical data to be truncated in the upper tail (Rock, 1988). A high proportion of

geochemical samples include values below the lower limit of detection (Reimann and Filzmoser, 2000). The concentrations of the elements in the lower tails may interfere when there are multiple elements containing values close to zero, such as trace elements (Grunsky, 2010). Rock (1988) suggests approaching censored values below the lower limit of detection by either removing them or replacing them with a value between zero and the lower limit of detection. When there is a significant number of censored values less than the lower limit of detection, the recommendation is to exclude that element from statistical analysis (Rawlins et al., 2002).

### **Outliers**

Outliers can originate from multiple populations and are common in geochemical data, as well as skewed distributions (Reimann and Filzmoser, 2000). Multiple processes influence the regional distribution of elements; therefore, the data distribution is often times skewed and multi-modal (Reimann and Filzmoser, 2000). Outliers in geochemistry commonly originate from secondary processes, such as mineralization or contamination (Filzmoser et al., 2005). Outliers in environmental geochemistry can be attributed to one or more contaminated distributions (Reimann et al., 2005). Two or more modes identified in the frequency histogram can be an indication of mineralization (Singer and Kouda, 2001). Trace elements exhibiting a non-lognormal distribution or a large standard deviation ( $> 1.0$ ) can be an indication of mineralization or more than one process, such as aeolian or alluvial processes (Singer and Kouda, 2001).

True outliers in geochemical data sets can be geologically explained and reproduced (Rock, 1988). Some outliers in geochemical data sets cannot be reproduced and are a result of machine error, contamination and/or sampling (Rock, 1988). Outliers in geochemical data could be an indication of unusual rock types or an ore deposit in the area surrounding the sampling site (Reimann and Filzmoser, 2000). Reimann et al. (2005) identified true outliers as those

observations attributed to one or more different distributions and extreme values such as those attributed to the same distribution. Due to the presence of extreme outliers in geochemical data, the univariate data distribution of most geochemical data is strongly right skewed (Filzmoser et al., 2009).

In geochemistry, currently there is not a universally accepted method of outlier detection (Reimann et al., 2005). In a geochemical data set, it is common to detect outliers univariately (Reimann et al., 2002). However, it is difficult to differentiate outliers belonging to a secondary process from those that are extreme when looking at the geochemical data univariately (Filzmoser et al., 2005). Detection of outliers is commonly based on estimates of location and scale, with outliers being attributed to a larger distance from the center of the distribution (Filzmoser et al., 2005). Outlier detection methods are necessary to identify outlier values in a way that minimizes data subjectivity, while accounting for multiple populations in the geochemical data (Reimann et al., 2005). The removal of outliers is subjective and often based on the background knowledge of the researcher (Grunfeld, 2005).

Visually inspecting the Q-Q plot of the univariate geochemical data is a necessary step prior to selecting a method of outlier detection (Reimann et al., 2005). Reimann et al. (2005) recommend first inspecting the geochemical data graphically and using these rules with data that are as close to a normal distribution as possible. Reimann et al. (2005) tested the following three commonly used outlier detection rules in geochemistry to determine the procedure best suited for use with a particular geochemical data set: (1) mean  $\pm$  2 absolute deviation, (2) median  $\pm$  2 standard deviation and (3) the boxplot inner fence. The box plot inner fence rule considers outliers as those values that are outside the upper and lower inner fences (Reimann et al., 2005).

Reimann et al. (2005) applies these rules to normal and lognormal data, in addition to data of unknown distribution.

The median  $\pm 2$  median absolute deviation resulted in the highest amount of detected extreme values and outliers, followed by the boxplot inner fence, and then the mean  $\pm 2$  standard deviation when used with normal and lognormal data (Reimann et al., 2005). When used with data from an unknown distribution the median  $\pm 2$  standard deviation identified the most outliers, followed by the boxplot inner fence (Reimann et al., 2005). The boxplot inner fence rule identified more outliers with data from the unknown distribution compared to the mean  $\pm 2$  standard deviation rule with log-transformed data from the unknown distribution (Reimann et al., 2005). The proportion of detected extreme values using the median  $\pm 2$  median absolute deviation or the mean  $\pm 2$  standard deviation rules are high when working with non-symmetrically distributed data (Reimann et al., 2005).

If the number of outliers in a data set is below 10%, the boxplot inner fence rule is most informative (Reimann et al., 2005). If the number of outliers is between 15-50% in a data set, then the median  $\pm 2$  median absolute deviation is the most informative (Reimann et al., 2005). When there are a high number of outliers in a data set, the mean  $\pm 2$  standard deviation was not a preferred method and should not be used (Reimann et al., 2005).

The mean  $\pm 2$  standard deviation rule identifies approximately 2.3% of extreme values at the upper and lower ends of ranked geochemical data, respectively (Reimann et al., 2005). The mean  $\pm 2$  standard deviation rule to detect outliers in a data set should not be used because it relies on a normal distribution (Reimann et al., 2005). Since the mean and standard deviation are influenced by extreme values, the mean  $\pm 2$  standard deviation rule is invalid (Reimann et al.,

2005). Table 3.1. presents general rules and recommendations for application of outlier detection methods.

Table 3.1. Outlier detection rules, range of outliers addressed, and recommendations for application.

Rule	Range of outliers	Recommendation
logarithmically transformed Q-Q plots		First visual observation
median +/- 2 median absolute deviation	15-50%	No preferred for >50%
boxplot inner fence	<10%	Informative
mean +/- 2 standard deviation	<50%	Do not use, relies on normal distribution

### Data Transformation

A common non-linear data transformation used in geochemical studies is a logarithmic transformation; however, a logarithmic transformation rarely results in a normal distribution (Reimann and Filzmoser, 2000). A logarithmic transformation of geochemical data does not account for the compositional nature of the geochemical data; however, this transformation often reduces skewness in the data (Filzmoser et al., 2009). Grunsky (2010) recommends visually inspecting the logarithmically transformed Q-Q plots if the data transformation minimizes the effect of skewness.

A logarithmic transformation changes the data to either the natural logarithm (ln) or the base ten logarithm (log10). Logarithmic transformations change the shape and the dispersion characteristics of a distribution (Grunsky, 2010). A logarithmic transformation asymmetrically



changes the distances of the observations from the center, thus, the variance and standard deviation cannot be computed from logarithmically transformed data and then back transformed to the original data scale (Filzmoser et al., 2009). If visual methods of detecting normality cannot be justified, it is common in geochemical studies to use a non-linear data transformation such as a logarithmic transformation to change the distribution of the raw data to approximately normal (Grunsky, 2010). Reimann and Filzmoser (2000) state that a histogram of logarithmically transformed geochemical data could result in normality.

### **Data Standardization**

A standard score or z-score linearly transforms the data values that are expressed in standard deviations above or below the mean. These values are either positive or negative. A z-score is calculated by subtracting the sample mean from the data value and dividing the result by the standard deviation of the sample using the following equation:

$$z = (x - \mu) / \sigma$$

where  $z$  = the z-score,  $x$  = ,  $\mu$  = , and  $\sigma$  = . The mean of the z-scores of a sample is equal to 0 and the unit variance is equal to 1. The z-score standardization however does not change the shape of the distribution. Z-score standardization removes the effects of differences in means and measurement scales (Singer and Kouda, 2001).

### **Graphical Methods of Detecting Non-Normality**

Reimann and Filzmoser (2000) suggest to first study the distribution of each element graphically. The following graphical methods are subjective and are used to visually detect non-normality in a univariate distribution: histogram, box plot, normal quantile-quantile plot and detrended normal quantile-quantile plot. A histogram is a graphical representation of a frequency distribution. The x-axis represents the data values of the variable and the y-axis represents the

frequency of occurrence of each data value of the variable. Equidistant intervals called bins are selected based on Euclidean distance. The number of data points that fall into each bin are shown in the histogram. The Euclidean distance between two points, a and b, is defined by the following equation:

$$d(a, b) = ((a_x - b_x)^2 + (a_y - b_y)^2)^{1/2}$$

A box plot, originally defined by Tukey (1977), is a graphical representation of the distribution of the variable that utilizes the median, 25<sup>th</sup> and 75<sup>th</sup> quantiles and the interquartile range (IQR). The upper and lower hinge points form the upper and lower bounds of the box and correspond to the 75<sup>th</sup> and 25<sup>th</sup> percentile scores, respectively. The median or 50<sup>th</sup> percentile score is the line located in the middle of the box. The box represents the middle 50% of the data values, otherwise known as the IQR. The lines extending from the top and bottom of the box plot are called whiskers and the ends of these lines correspond to the minimum and maximum values within 1.5 times above or below the IQR, respectively. Mild outliers occur in a box plot if the values are less than  $Q1 - (1.5 * IQR)$  or greater than  $Q3 + (1.5 * IQR)$ , but are not considered extreme outliers. Extreme outliers occur in a box plot if the values are less than  $Q1 - (3 * IQR)$  or greater than  $Q3 + (3 * IQR)$ . Extreme outliers are denoted by an asterisk on the boxplot. In a box plot, the position of the median and quartiles do not change with a logarithmic transformation, since the box plot is based on the median and quantiles (Filzmoser et al., 2009).

A normal Q-Q plot and corresponding de-trended Q-Q plot serve as graphical representations of the distribution of a variable. In a normal Q-Q plot the x-axis represents the quantiles of the observed data and the y-axis represents the quantiles of the expected data from a standard normal distribution; however, the units of the x- and y- axes are units of the data set, not quantiles. If the data values plot on a straight line with a slope equal to 1, then the data is

normally distributed; however, if the data points on the plot deviate from the straight line, then the data is non-normally distributed. A skewed or polymodal frequency distribution will result in a discontinuous or curved Q-Q plot (Grunsky, 2010). Slope changes or discontinuous lines on the Q-Q plot could be an indication of multiple populations (Grunsky, 2010). A Q-Q plot is advantageous in that each data observation is plotted, allowing for observation of groups as well as identification of extreme data values (Grunsky, 2010).

In a de-trended Q-Q plot the x-axis represents the quantiles of the observed data and the y-axis represents the deviation between the observed value and the expected value following a standard normal distribution. If the data values plot on a horizontal line, with an equal number of data points above and below the horizontal line, then the data follows a normal distribution. Grunsky (2010) suggests looking at the Q-Q plots of logarithmically transformed data for breaks in the line or changes in the slope of the line, indicating multiple populations.

### **Skewness and Kurtosis**

Skewness and kurtosis statistics are equal to 0 in normal distribution; therefore, these two statistics are used to test for a normal distribution. Skewness is a measure of symmetry, while kurtosis is a measure of peakedness of a distribution (Hill and Lewicki, 2006). A distribution is positively skewed when the tail on the right side is longer than that on the left. The opposite is true for a negatively skewed distribution. Thus, positive and negative skewness values indicate a skew to the right or left, respectively. A kurtosis value greater than 1 indicates a pointed distribution and a kurtosis value less than one indicates a flat or platykurtic distribution. A negative and positive kurtosis value indicates a platykurtic and leptokurtic distribution, respectively. A kurtosis value of 0 indicates a mesokurtic distribution.

A critical value or z-score for the skewness and kurtosis is calculated by dividing the skewness and kurtosis values by the corresponding standard error values. Approximately 95% of the data values in a normal distribution fall between -1.96 and +1.96 standard deviations from the mean. Therefore, if a critical value is outside the range of -1.96 to +1.96, then this is an indication of a non-normal distribution. This is a general method to numerically detect non-normality in a univariate data set; however, the Shapiro-Wilk test is perhaps a more widely used statistical method to test for non-normality in geochemistry (Reimann and Filzmoser, 2000).

### **Shapiro-Wilk Test**

The Shapiro-Wilk test is a goodness of fit test for univariate and continuous data introduced by Shapiro and Wilk (1965). Royston (1982) extended the Shapiro-Wilk test for sample sizes greater than fifty and up to two thousand. The null hypothesis states that the univariate data are normally distributed. The alternative hypothesis states that the univariate data are non-normally distributed.

The Shapiro-Wilk W-test statistic is defined by the equation:

$$W = (\sum a_i x_i)^2 / \sum (x_i - \bar{x})^2$$

where the constants associated with the co-variances, variances and sample mean are represented by  $a_i$ . The ordered sample values are  $x_i$  and  $\bar{x}$  is the sample mean. The Shapiro-Wilk W-test statistic value is always between zero and one. A W-test statistic value of 1 indicates that the univariate continuous data follow a normal distribution.

The Shapiro-Wilk test is the preferred test of normality due to the tests better power properties compared to alternative normality tests, such as the Kolmogorov-Smirnov and the chi-square goodness-of-fit tests (Shapiro, Wilk, & Chen, 1968). Generally, the Shapiro-Wilk test is the statistically preferred test of univariate normality compared to the Kolmogorov-Smirnov- and

the chi-square goodness-of-fit tests when evaluating geochemical and environmental data sets (Reimann and Filzmoser, 2000).

### **Location and Scale Estimates**

Rock (1988) studied non-robust and robust estimates of location and scale with geochemical data. Rock (1988) studied the following four non-robust estimates of location using geochemical data: (1) arithmetic mean (2) geometric mean (3) harmonic mean (4) midrange. The following estimates of location are in order of increasing robustness: midrange, arithmetic mean, geometric mean and harmonic mean (Rock, 1988). The arithmetic, geometric and harmonic means are described in the following equations, respectively, where  $x_i$  is each observation and  $n$  are the number of observations in the data set.

The arithmetic mean is defined using the following equation:

$$A = \left(\frac{1}{n}\right) * \sum_{i=1}^n x_i$$

The geometric mean is defined using the following equation:

$$GM = n\sqrt{\prod_{i=1}^n x_i}$$

The harmonic mean is defined using the following equation:

$$HM = 1 / \sum_{i=1}^n (1/x_i)$$

The midrange is defined as one-half of the sum of the minimum and the maximum values in a data set.

Means are considered non-robust average values that are sensitive to outliers and skewed distributions (Rock, 1988). The arithmetic mean is considered the worst estimate of location for data following a non-normal distribution or for data containing outlier values (Rock, 1988). For data that is skewed right, the arithmetic mean is too high and biased (Filzmoser et al., 2009).

Censored values below the lower limit of detection are overestimates, resulting in positive bias

of the arithmetic mean (Grunsky, 1999). Extreme values influence the estimate of the arithmetic mean (Reimann et al., 2005). When working with random samples from normal and lognormal distributions, the arithmetic and geometric means are considered maximum likelihood unbiased estimates (Rock, 1988). When data follows a lognormal distribution, the geometric mean is preferred over the arithmetic mean; however, most geochemical distributions are not lognormal (Rock, 1988). The geometric mean is not affected by skewed distributions; however, it is affected by data outliers (Rock, 1988). Reimann and Filzmoser (2000) state that the geometric mean can be used as an estimate of location, but this can be problematic. The midrange is entirely non-robust to data outliers (Rock, 1988).

Rock (1988) studied nineteen robust estimates of location using geochemical data; however, not all will be mentioned here. Rock (1988) studied the 10, 20 and 25% trimmed means as robust estimates of location. Trimmed means assumes that the estimate of mean is performed without outliers in the data (Grunsky, 2006). The 10, 20 and 25% trimmed means are calculated by first arranging the data values in ascending order and then taking the arithmetic average of the data after 10, 20 and 25% of the data is removed at both ends, respectively (Rock, 1988). The 50% trimmed mean, otherwise known as the median is a robust estimate of location studied by Rock (1988).

Rock (1988) studied the following combined L estimates as robust estimates of location: Gastwirth median and Trimean. Rock (1988) defines the Gastwirth median using the following equation:

$$\text{Gastwirth median} = (\text{median} * 0.4) + 0.3 * (\text{upper} + \text{lower tertiles}).$$

Rock (1988) defines the Trimean using the following equation:

$$(Q1 + 2 * Q2 + Q3) / 4.$$

Robust estimates are not as affected by extreme values or non-normally distributed data than estimates such as the mean and median (Rock, 1988). Robust estimates of location are not as affected by truncation compared to the arithmetic mean (Rock, 1988). Robust estimates yield values that are higher in accuracy and precision compared to the mean and median (Rock, 1988). Since the arithmetic mean is influenced by outliers and skewed distributions, the median is preferred as it is more robust (Reimann and Filzmoser, 2000). The median is robust against extreme values (Reimann et al., 2005). The median can be erroneous since this measure does not account for the relationships between variables (Reimann and Filzmoser, 2000). Reimann and Filzmoser (2000) state that the median is the preferred estimate of location with geochemical data; however, robust estimates of the mean can be used, such as Hampel or Huber estimates.

Rock (1988) considered robust estimates of location such as L, M and W estimates, in addition to parametric, robust and nonparametric estimates. Rock (1988) concluded that robust estimates of location, such as dominant cluster mode and shorth mean are unstable, while Hampel's series, trimmed means and combined L estimates are preferred when used with geochemical data (Rock, 1988). Rock (1988) suggests that when calculating the average of geochemical data sets that are truncated, the values outside the limits of detection can either be either omitted or replaced with an arbitrary value.

Rock (1988) makes the following three conclusions about what are considered to be the best estimates of location: (1) those in which the variance is minimized (2) those which are closest to the correct value for a data set with a known location (3) those in which yield the most consistent estimates using a range of real data sets. Rock (1988) states that the following five factors should be considered when evaluating the value of the robust estimate of location of real geochemical data: (1) distribution type, (2) modality, (3) continuity, (4) number of outliers,

(5) size of data set, and (6) source of error. Rock (1988) stated the common absolute order of location estimates with real geochemical data is the following: midrange  $\gg$  arithmetic mean  $>$  geometric mean  $\geq$  robust estimates of location  $\geq$  harmonic mean.

Rock (1988) studied the following six estimates of scale with real geochemical data: (1) standard deviation, (2) mean absolute deviation from the mean, (3) mean absolute deviation from the median, (4) median absolute deviation from the median, (5) semi interquartile range, and (6) a robust biweight A estimator based on the M estimates of location. The above-mentioned estimates of scale are arranged in increasing order of robustness (Rock, 1988).

The standard deviation is defined using the following equation:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}$$

The mean absolute deviation from the mean is defined as the average distance between the mean and each data value. The mean absolute deviation from the median is defined as the average distance between the median and each data value. The median absolute deviation from the median is defined as the median of the average distance between the median and each data value. The semi-interquartile range (semi-IQR) is defined as one-half the value of the IQR.

The standard deviation is based on the squared differences of each observation from the mean, thus, it is not a good estimation of the spread of the data (Reimann and Filzmoser, 2000). The standard deviation is influenced by censored values, outliers and data values from other populations (Reimann and Filzmoser, 2000). The standard deviation should not be used as an estimate of scale if data outliers are present (Reimann and Filzmoser, 2000). Data values from a secondary population influence the standard deviation estimate (Reimann and Filzmoser, 2000). Extreme values influence the estimate of the standard deviation (Reimann et al., 2005).



The standard deviation, mean absolute deviation from the mean and mean absolute deviation from the median provide larger estimates of scale compared to the IQR or median absolute deviation from the median (Rock, 1988). Robust estimates of scale including the median absolute deviation from the median and hinge spread are preferred estimates of scale (Reimann and Filzmoser, 2000). The median absolute deviation from the median is robust against extreme values and is considered a suitable replacement for the standard deviation (Reimann et al., 2005).

Filzmoser et al. (2009) suggests using the squared median absolute deviation from the median and the IQR as robust alternatives to the sample variance. Low values of the squared median absolute deviation from the median indicate stability (Reimann and Filzmoser, 2009). A large difference between the mean and median, and the standard deviation and squared median absolute deviation from the median indicates that the data does not follow a normal distribution (Reimann and Filzmoser, 2000). The range is defined as the difference between the minimum and maximum values and is considered a less robust estimate of scale than the standard deviation (Rock, 1988). Rock (1988) stated that the common absolute order of scale estimates with geochemical data is the following: standard deviation > mean absolute deviation from the mean > mean absolute deviation from the median >> median deviation from the median = semi-interquartile range

The arithmetic mean and standard deviation are not reliable estimates to describe the data distribution in geochemistry (Rock, 1988). When using the arithmetic mean and standard deviation to summarize a geochemical data set, a robust estimate of location and scale should also be used, in addition to a statement about the normality and homogeneity of the data distribution (Rock, 1988). The arithmetic mean and standard deviation are overestimates when

working with positively skewed geochemical data sets containing outliers with large values (Rock, 1988). It is common to remove data outliers prior to calculating the arithmetic mean and standard deviation; however, the removal of outliers is subjective (Reimann et al., 2005).

Robust estimates are preferred to the logarithmic transformation of geochemical data sets (Rock, 1988). Since geochemical data often is representative of more than one population, it is problematic to calculate the arithmetic mean and standard deviation after logarithmic transformation, as this does not correct for the influence of outliers (Reimann et al., 2005). Since a logarithmic transformation rarely results in a normal distribution, calculating the arithmetic mean from the logarithmically transformed data and then back transforming the result is misleading (Reimann and Filzmoser, 2000). If a logarithmic data transformation results in a near normal data distribution, the arithmetic mean can be calculated and then transformed to the original data scale for a less biased arithmetic mean (Filzmoser et al., 2009).

Robust estimates are not entirely unbiased (Rock, 1988). Robust estimates of location and scale are smaller and more consistent compared to non-robust estimates such as the arithmetic mean and standard deviation, respectively (Rock, 1988). Data closure must be achieved before visually looking at the data distribution using a histogram or calculating estimates such as the arithmetic mean and standard deviation (Reimann and Filzmoser, 2000). Robust estimates are not as affected as non-robust estimates when there is a large proportion of censored values below the lower limit of detection in the geochemical data (Rock, 1988). When the number of outliers in a data set increases, the differences between robust estimates of location and scale become more inconsistent, as the differences between the robust estimates are large (Rock, 1988).

## Correlation

Pearson's  $r$  correlation is a bivariate, correlational and parametric statistical test that measures the strength of the linear relationship between two variables. Pearson's  $r$  correlation assumes that the variables are measured at least at the interval scale and that these two variables are proportional, meaning they are linearly related (Hill and Lewicki, 2006). The assumptions of Pearson's  $r$  correlation are the following: (1) the units of measurement are the same for all variables, (2) there is a linear relationship between the two variables, (3) the variables are either normal or lognormal in distribution (Rollinson, 2014). Pearson's  $r$  correlation requires a multivariate normal distribution (Reimann et al., 2002).

The correlation coefficient can be misleading, as it is influenced by data outliers and skewed data distributions (Filzmoser and Hron, 2009). The Pearson correlation coefficient,  $r$ , is also known as the Pearson's product moment correlation coefficient. The coefficient of determination,  $r^2$ , is calculated by taking the square of the Pearson's correlation coefficient. The coefficient of determination is a measure of the amount of shared variance between two variables. When converted into a percentage, this measure indicates the proportion of shared variance between two variables.

Prior to Pearson's  $r$  correlation, a logarithmic transformation of the univariate data is recommended, as is a visual inspection of the univariate distribution (Reimann et al. 2008). Due to the Euclidean geometry of compositional data, correlation analysis using raw and logarithmically transformed data is discouraged (Filzmoser and Hron, 2009). The use of logarithmically transformed geochemical data with correlation analysis can lead to misleading results due to the bias in the relationship between two variables (Filzmoser and Hron, 2009). The constant sum problem affects the results of Pearson's  $r$  correlation by introducing negative bias

into the results (Rollinson, 2014). The closure of geochemical data affects the results of Pearson's r correlation by forcing two variables to be associated (Rollinson, 2014).

The use of data transformations, nonparametric correlation methods or robust correlation measurements account for data outliers and skewed data distributions (Filzmoser and Hron, 2009). Spearman's rho correlation is the nonparametric analog to the Pearson's r correlation. Spearman's correlation is advantageous in that this test can be applied to ranked data and this test is preferred to Pearson's r correlation when the data is not normally distributed or contains outliers (Rollinson, 2014). Spearman's rho correlation computes a correlation coefficient,  $r_s$ , based on ranked data. Spearman's rho correlation assumes that the variables are measured at least at the ordinal scale (Hill and Lewicki, 2006). Spearman's rho correlation is preferred over Pearson's r correlation when one or both variables are non-normally distributed or contain outliers (Cooksey, 2014). The results of the Spearman rho correlation for raw and logarithmically transformed geochemical data are the same; therefore, a logarithmic transformation of the geochemical data is not needed prior to Spearman rho correlation (Reimann et al., 2008).

A correlation coefficient includes both strength and direction. The strength of association between two variables can range from weak to very strong. The variables can either move in a positive or negative direction. A correlation matrix displays the correlation coefficients between variable pairs. A correlation coefficient is calculated using the equation:

$$r_{xy} = \text{covariance}(x,y) / ((\text{variance}(x) * \text{variance}(y))^{1/2})$$

A correlation coefficient ranges from -1 to +1, a perfect negative and perfect positive correlation, respectively. A correlation coefficient of zero corresponds to no correlation. A correlation coefficient between +/- 0.5 to 0.7 is good, while a correlation coefficient between +/-

0.7 to 1.0 is strong. A positive association is one in which both variables increase. A negative association occurs when one variable increases and the other variable decreases.

The interpretation of Pearson's  $r$  correlation and Spearman's  $\rho$  correlation tests are the same. The probability of a statistically significant relationship between the two variables,  $x$  and  $y$ , is denoted by the  $p$ -value. In a two-tailed test, the null hypothesis states that there is no correlation between  $x$  and  $y$  and the population correlation coefficient is equal to zero. The alternative hypothesis states that the population correlation coefficient is not equal to zero and there is a correlation between  $x$  and  $y$ .

A strong correlation, but non-linear relationship between two variables, can be misleading (Hill and Lewicki, 2006). Correlation between variables is often forced and negative bias is introduced into correlation due to the constant sum constraint of geochemical data (Rollinson, 1993). Closed data affects the results of correlation analysis (Rock, 1988). Rock (1988) suggests that when using closed data, a negative correlation is less significant, and a positive correlation is more significant compared to open data. Data transformation does not correct for data closure (Reimann et al., 2002). A significant difference between the Pearson's  $r$  and Spearman's  $\rho$  correlation coefficients suggests that the Pearson's correlation coefficient is influenced by data abnormalities (Cooksey, 2014).

### **Hierarchical Cluster Analysis**

Cluster analysis is a non-parametric, multivariate and correlational statistical test. Cluster analysis is a data mining technique with the purpose of discovering patterns in a set of data. The term cluster analysis was first used by Tryon (1939) to describe a set of algorithms that serve to partition a multivariate data sets into homogenous groups based on similarity. R-mode and Q-mode cluster analysis correspond to the clustering of the variables and observations, respectively.

Cluster analysis results in homogenous groups of either observations or variables which can then be interpreted (Templ et al., 2008).

There are two classes of cluster analysis, non-hierarchical (non-HCA) and hierarchical (HCA). In non-hierarchical cluster analysis, the number of clusters is subjectively determined prior to analysis and the assignment of a member to a cluster occurs in a single step based on a similarity measure. In agglomerative hierarchical cluster analysis (HCA), all possible numbers of clusters are looked at and a similarity measure is used to assess the members before assignment to a cluster. In the beginning of agglomerative HCA, there are the same number of clusters as there are members. The members are then combined or linked pairwise using an algorithm. This continues until assignment of all members finally into one larger cluster.

In cluster analysis, the distance between members is a measure of the similarity between members in a multidimensional space. Euclidean distance is a common distance measurement chosen when clustering observations of geochemical data (Templ et al., 2008). The Euclidean distance is often squared to put more weight on members which are located further apart. The squared Euclidean distances between the cluster centers is proportional to the initial distance between the individual clusters.

Ward's minimum variance algorithm developed by Ward in 1963 is a common linking algorithm in agglomerative HCA. Ward's method attempts to minimize the sum of the squared distances from the clusters. A centroid represents each cluster and Ward's minimum variance algorithm assigns each data point to a separate cluster and merges the clusters together until only one cluster remains. The proximity measure is the increase in the squared error when merging two clusters. This method finds and merges cluster pairs in which the result is a minimum increase in the total within cluster variance. Other values of linkage used in HCA include

complete, single, and average. A dendrogram illustrates the clusters and linkages. Horizontal and vertical lines on the dendrogram correspond to the linkage of the clusters and a distance measure corresponds to the height of the lines.

Cluster analysis using a distance coefficient, such as the squared Euclidean distance makes no assumptions and does not require a normal distribution (Templ et al., 2008). Templ et al. (2008) recommends transforming largely skewed data to a more symmetrical distribution. The most common standardization method used in cluster analysis is the z-score which accounts for differences in variability (Templ et al., 2008). Outliers in geochemical data can affect the results of cluster analysis (Templ et al., 2008). For example, outliers can be assigned to homogenous clusters or to their own clusters based on what clustering method is used (Templ et al., 2008). Outliers should be deleted, or the number of clusters chosen should be increased to minimize the effects of outliers in cluster analysis (Templ et al., 2008).

When variables are clustered, HCA is a preferred method (Templ et al., 2008). A variable with many censored values can affect the results of cluster analysis and should be excluded (Templ et al., 2008). Identical elemental concentrations such as censored data below the lower limit of detection can influence the results of cluster analysis (Templ et al., 2008). The addition or deletion of a variable in cluster analysis can significantly affect the results (Templ et al., 2008).

The results of cluster analysis of geochemical data depend on the data transformation and the clustering algorithm (Templ et al., 2008). Temple et al. (2008) suggests transforming geochemical data that is heavily skewed to a distribution that is more symmetrical, prior to cluster analysis. A logarithmic transformation is commonly used to approach symmetry prior to cluster analysis (Templ et al., 2008). Geochemical data should be standardized if there is a large

difference in variability between the variables (Templ et al., 2008). The most common standardization method chosen is the z-score transformation (Templ et al., 2008).

### **Principal Component Analysis**

Principal component analysis (PCA) is a type of exploratory factor analysis that is parametric, multivariate and correlational. PCA was introduced by Pearson (1901) and expanded by Hotelling (1933). PCA is an extraction method in which a principal component (PC) is extracted one by one in terms of variance. The purpose of PCA is to extract the fewest number of PCs that account for the highest proportion of variance out of the total available variance. This allows the number of variables in a data set to be reduced so that the variation within the data set can be explained. There are no statistical assumptions in PCA (Reimann and Filzmoser, 2002).

In PCA, the structure of the data and the relationships between the variables is explained using either a variance-covariance matrix or a correlation matrix. A variance-covariance matrix is used when all variables are measured using the same unit and when the difference in variance between the variables is important. A correlation matrix is used when some variables are measured using different units. A variance-covariance and correlation matrix are the same, except in a correlation matrix, the variables are standardized to a unit standard deviation and mean of zero. This standardization procedure does not allow the variables with the highest variance to dominate the PCs. Standardization to zero mean and unit standard deviation results in the equal influence of all variables in factor analysis (Reimann and Filzmoser, 2002).

Either a correlation or co-variance matrix is used to calculate eigenvectors and eigenvalues. Eigenvectors are geometrical representations of a matrix in a multidimensional space, thus, eigenvectors define axes within a multidimensional ellipsoid. Like any vector, an eigenvector contains a magnitude and direction. Eigenvectors are normalized to a unit length of



1. Each eigenvector is associated with an eigenvalue and represents the length of the semi-axis of an ellipsoid. Eigenvalues are indices that correspond to the amount of variance a component represents in the variables out of the total variables available. PCA extracts one PC at a time in decreasing order of eigenvalue.

The number of variables is equal to the number of PCs extracted in PCA; however, PCA is best applied when trying to reduce the number of variables. Extracting the same number of PCs as variables would defeat the purpose of this method. In PCA, the direction of maximum variance is found in a multidimensional cloud of data points. There are the same number of dimensions as there are variables.

A PC is an eigenvector of either a variance-covariance or a correlation matrix. In PCA, each eigenvector corresponds to a PC. PC1 is extracted first and the corresponding eigenvector shows the direction of largest variance. The corresponding eigenvector to each successive PC extracted is orthogonal and explains less variance than the previous. A decision rule is used in PCA to determine how many components to extract from the data. Only PCs that account for the largest amount of variance amongst the variables are extracted in PCA. The number of PCs to select is determined based on the eigenvalue greater than one rule. An eigenvalue greater than one rule states that the only acceptable PCs are those with an eigenvalue greater than one, meaning that an acceptable PC must explain the variance of more than one variable (Kaiser, 1960).

The simplest way to determine how many factors to extract is by using a scree plot with the lowest number of extracted factors being the most interpretable (Reimann and Filzmoser, 2002). In a scree plot, the x-axis represents the eigenvalues and the y-axis represents the number of components. A general rule-for selecting the number of components to interpret is the last

number located before the bend in the plot. The initial solution of PCA is rotated to obtain simple structure. In simple structure, one or a few variables aligns with one component. Rotation serves to align one or very few variables with a single component. Rotation makes the relationships between the variables and components more prominent, thus, the solution is more interpretable. A rotation is performed in a subspace. The new axes formed from rotation explain less variance than the original components; however, the same total amount of variance is explained. There are two main types of rotation: orthogonal and oblique.

A component loading is analogous to a correlation coefficient and is defined as the correlation between the component loading and original variables. A squared component loading value is the amount of variation in a variable that is explained by a component. A component loading is the value that is multiplied with the standardized original variable to obtain a component score. The component loadings can be interpreted geologically or geochemically (Grunsky, 2010). For example, each principal component can be interpreted as a geological or weathering process (Grunsky, 2010). In lake sediment surveys, the first and second components can reflect lithological variation (Grunsky, 2010). A component score is the value of the transformed variable of a data point. A component score denotes the position of an observation on the new coordinate system and is calculated using the linear combination of the original variables and the component loadings.

A bi-plot contains the first two PCs, on the x and y axes, respectively. Points represent observations and correspond to the component scores of PC1 and PC2. Points that are close to each other represent similar observations. Vectors correspond to the correlation of the variables with PC1 and PC2. The vector length is proportional to the variance of the variable. The cosine of the angle between the vectors on the bi-plot is equal to the correlation between the variables.

Vectors that point in the same general direction are more correlated to each other. Vectors that are at right angles to one another are uncorrelated.

Since major and trace elements are reported in different units of concentration, they should not be mixed in multivariate analyses such as PCA and cluster analysis (Rock 1988). Converting the major and trace elements to one unit of concentration is not a preferred solution (Templ et al., 2008). Variables with the highest magnitude correspond to the highest variance, thus, these variables have the most influence on the multivariate statistical method (Templ et al., 2008). Therefore, the major elements will have more influence than the trace elements on the multivariate statistical method. The addition or deletion of one variable in factor analysis can influence the results significantly (Reimann and Filzmoser, 2002). Reimann et al. (2002) states that most geochemical data sets should not be used with factor analysis due to the following problems encountered: closed data, presence of outliers and non-normal, and non-lognormal distribution.

Prior to factor analysis, the distribution of the variables must be looked at and a transformation such as a logarithmic transformation might be required to approach normality (Reimann and Filzmoser, 2002). Filzmoser et al. (2009) state that when factor analysis is used with compositional data, a data transformation is required. Reimann and Filzmoser (2002) suggest that all variables should come as close to a normal distribution as possible, prior to PCA. Data closure cannot be fixed with a conventional method of data transformation, thus, there are consequences for using closed data in correlation (Reimann and Filzmoser, 2002). A negative correlation is less significant, while a positive correlation is more significant when using closed data (Rock, 1988). Outliers should be removed prior to factor analysis; however, this is almost never done (Reimann et al., 2002).

Correlation and covariance estimates are sensitive to the presence of outliers (Grunsky, 2010). Censored values should be replaced or removed prior to PCA (Grunsky, 2010). A significant amount of identical values such as values near the lower limit of detection of a variable can influence correlation, thus these variables should be excluded from factor analysis (Reimann and Filzmoser, 2002). Large regional geochemical data sets should be divided into smaller data sets based on knowledge or cluster analysis prior to factor analysis (Reimann and Filzmoser, 2002). Factor analysis can be used in geochemistry to re-interpret a geological map or to identify processes that influence the elemental distribution regionally (Reimann et al., 2002). Grunsky (2010) states that in lake sediment surveys, the first and second PCs generally represent variations in lithology.

Principal factor analysis is preferred over more advanced factor analysis techniques when working with geochemical data (Reimann and Filzmoser, 2002). Invalid data points should be removed, and the geochemical data should be transformed prior to multivariate statistical analyses (Reimann et al., 2002). Transformation and/or standardization is preferred when performing a multivariate statistical method (Templ et al., 2008). When using PCA, the data should be transformed before standardization (Reimann and Filzmoser, 2002).

## 4. METHODOLOGY AND SAMPLING

### **Introduction**

The data for the statistical methods detailed in previous chapter which are the focus of this study originate with ICP-AES data from Al-Azraq AZ1 and AZ2 cored sediments.

### **Core Description**

The sediment classification scheme used by the Limnological Research Center, University of Minnesota is the basis for the initial sediment core description (Schnurrenberger, et. al., 2003). The following ideas make up this classification scheme: (1) core sediment is compositionally and potentially paleolimnologically variable; (2) sedimentary composition of the major lithologic units can determine what types of studies can be employed on the different units; and (3) the sedimentary unconformities and bedding structures have an influence on how the core is sampled (Schnurrenberger et al., 2003). The classification scheme involves describing both the macroscopic and microscopic components of the sediment; however, this study describes only the macroscopic core description procedure. The overall goal of the sediment macroscopic description is to get a basic idea of the overall lithology of the core to guide sampling. Following the macroscopic core description, digital images of the core are taken.

Following core splitting and cleaning, the macroscopic features of the sediment core are described by observation of the overall sediment lithology, color and sedimentary characteristics. The sediment core is described by bed using four characteristics: lithology and color, thickness and inclination, sedimentary structures and bedding planes, and core stratigraphy disturbance (Schnurrenberger et al., 2003). A macroscopic description of grain size, texture, and sediment color using Munsell color charts details the lithology and color of the bed. McKee and Weir (1953) define a bed and a laminate as larger than and less than one centimeter in thickness,

respectively. Beds are either very thin, thin, medium, thick or very thick bedded (Mckee and Weir, 1953). Boundaries between bedding planes are sharp, diffuse and indistinct. Sedimentary structures are primary or secondary and are either physical, chemical or biogenic. AZ1 cored sediments are described in previous work (Ahmad, 2013). AZ2 cored sediments extend from the surface to 35.5 meters.

### **Inductively Coupled Plasma- Atomic Emission Spectroscopy (ICP-AES)**

ICP-AES sample preparation included drying overnight at 45°C, then ground in mortar until a fine powder which is sieved at 80 mesh. The pan portion is weighed and submitted to ALS Global - Geochemistry Analytical Lab in Reno, NV for ICP-AES analysis. Tables 4.1. and 4.2. show the major and trace elements analyzed by ICP-AES and the corresponding units of concentration, respectively.

Table 4.1. ICP-AES major elements and corresponding units of concentration.

Element	Concentration
	Reported
Al	wt. %
Ca	wt. %
Fe	wt. %
K	wt. %
Mg	wt. %
Na	wt. %
S	wt. %
Ti	wt. %

Table 4.2. ICP-AES trace elements and corresponding units of concentration.

Element	Concentration Reported
Ag	ppm
As	ppm
B	ppm
Ba	ppm
Be	ppm
Bi	ppm
Cd	ppm
Co	ppm
Cr	ppm
Cu	ppm
Ga	ppm
Hg	ppm
La	ppm
Mn	ppm
Mo	ppm
Ni	ppm
P	ppm
Pb	ppm
Sb	ppm
Sc	ppm



Sr	ppm
Th	ppm
Tl	ppm
U	ppm
V	ppm
W	ppm
Zn	ppm

Ahmad (2013) collected two hundred fifty-three sediment samples for ICP-AES analysis from AZ1. Ahmad (2013) addresses the ICP-AES results in previous work. The AZ1 ICP-AES results are used in statistical analyses. One hundred sixty-three samples were collected from AZ2 at major color or lithological changes for ICP-AES analysis. The trace elements Sb and Sc are included in AZ2, but not included in AZ1. Table 4.3. illustrates the total number of samples and elements analyzed by ICP-AES in each sediment core.

Table 4.3. Number of samples and elements analyzed by ICP-AES in each sediment core.

Core	No. of Samples	Total No. of Elements Analyzed
AZ1	253	33
AZ2	163	35

### Statistical Methodology

The statistical methods used in this study considered AZ1 and AZ2 data separately. Variables containing more than 5% censored values were excluded from all statistical analyses. For those variables containing less than 5% censored values, a replacement method based on Rock (1988) substituted a value in between zero and the lower limit of detection. For values less than the lower limit of detection, a value of one-half the lower limit of detection was chosen as a replacement value based on (Reimann and Filzmoser, 2000). In addition, elements containing a high number of duplicate values were excluded from all statistical analyses. Figure 4.1 details data preprocessing methods used in this study.

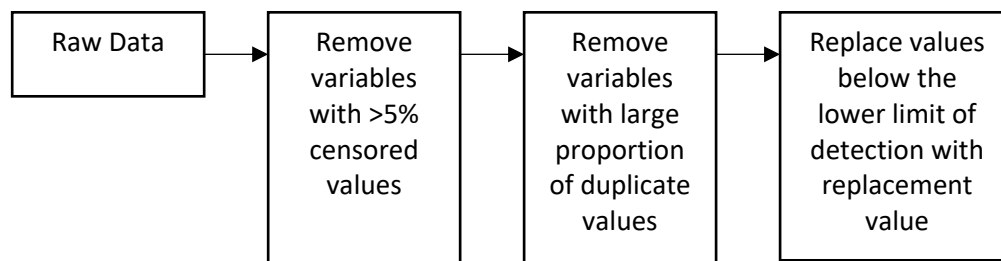


Figure 4.1 Flow chart of data preprocessing methods.

Four graphical methods used to subjectively determine if both raw and ln transformed univariate data from AZ1 and AZ2 follow a normal or lognormal distribution include: (1) histogram, (2) quantile-quantile (QQ) plot, (3) de-trended quantile-quantile (QQ) plot, and (4)

box plot. Two statistical methods used to determine if univariate data from AZ1 and AZ2 follow a normal distribution include (1) skewness and kurtosis statistics with the corresponding standard error (S.E.) values, and (2) the Shapiro-Wilk test. Skewness and kurtosis z-scores were computed by dividing the skewness and kurtosis values by the corresponding standard error value. This method was used only for raw univariate data from AZ1 and AZ2. The Shapiro-Wilk test was performed with and without outliers removed based on work by Reimann et al. (2005). Table 4.4. lists the graphical and statistical methods of detecting a normal distribution used in this study. Figure 4.2. details the methods of detecting a normal distribution used in this study.

Table 4.4. Graphical and statistical methods used to detect normality with univariate data.

<b>Graphical</b>	<b>Statistical</b>
Histogram	Skewness & Kurtosis with S.E.
Q-Q Plot	Shapiro-Wilk
Detrended Q-Q Plot	
Box Plot	

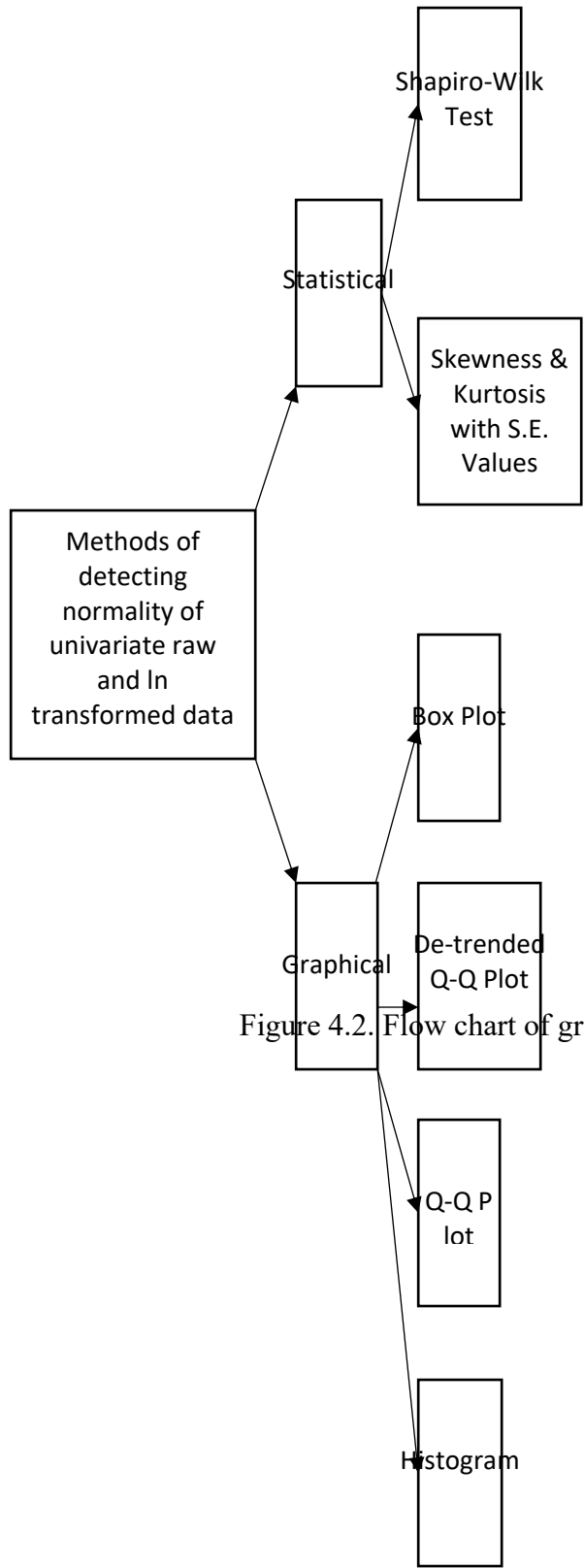


Figure 4.2. Flow chart of graphical and statistical methods used to de

The following two commonly used outlier detection rules based on Reimann et al. (2005) was used with univariate geochemical data from AZ1 and AZ2: (1) median  $\pm$  2 standard deviation, and (2) the boxplot inner fence. If the number of outliers in the univariate data is below 10% or between 15 – 50%, then the boxplot inner fence and median  $\pm$  2 median absolute deviation rules are used, respectively, as these methods are the most informative (Reimann et al., 2005). Initially, outliers are detected by finding mild and extreme outlier values in the univariate box plot.

In this study, non-robust and robust estimates of location and scale are considered. Four non-robust estimates of location include: (1) arithmetic mean, (2) geometric mean, (3) harmonic mean, and (4) midrange. Robust estimates of location include: (1) the 10, 20, and 25 % trimmed means, (2) median, (3) Gastwirth median, and (4) Trimean. Rock (1988) stated that dominant cluster mode and shorth mean are unstable robust location estimates, so these methods were omitted. The estimates of robust (1- x) and non-robust (x-x) estimates of scale include: (1) range, (2) standard deviation, (3) mean absolute deviation from the mean, (4) mean absolute deviation from the median, (5) median absolute deviation from the median, and (6) semi interquartile range. Table 4.5. details the estimates of location and scale used in this study.

Table 4.5. Estimates of Location and Scale used with univariate data of AZ1 and AZ2.

<b>Non-Robust Location Estimates</b>	<b>Robust Location Estimates</b>	<b>Scale Estimates</b>
Arithmetic Mean	10% Trimmed Mean	Range
Geometric Mean	20% Trimmed Mean	Std.
Harmonic Mean	25% Trimmed Mean	Mean Abs. Dev. From Mean
Midrange	Median	Mean Abs. Dev. From Median
	Gastwirth Median	Median Abs. Dev. From Median
	Trimean	Semi-IQR

Pearson's r correlation and the Spearman's rho correlation, correspond?? to the parametric and non-parametric correlation methods used to determine the relationship between each variable set, respectively. Pearson's r correlation was performed using the ln transformed data. Spearman's rho correlation was performed using the raw data.

Two methods used to determine the relationship between the variables include: Hierarchical Cluster Analysis (HCA) and Principal Components Analysis (PCA). HCA and PCA are non-parametric and parametric statistical tests, respectively. Prior to HCA and PCA, geochemical data was transformed and standardized. The natural log (ln) transformed the data. The z-score standardized the data. HCA analysis using Ward's method and squared Euclidean distance was chosen. PCA using the correlation matrix and varimax rotation was chosen univariate. Figure 4.3 details the multivariate methods used in this study and the order in which...

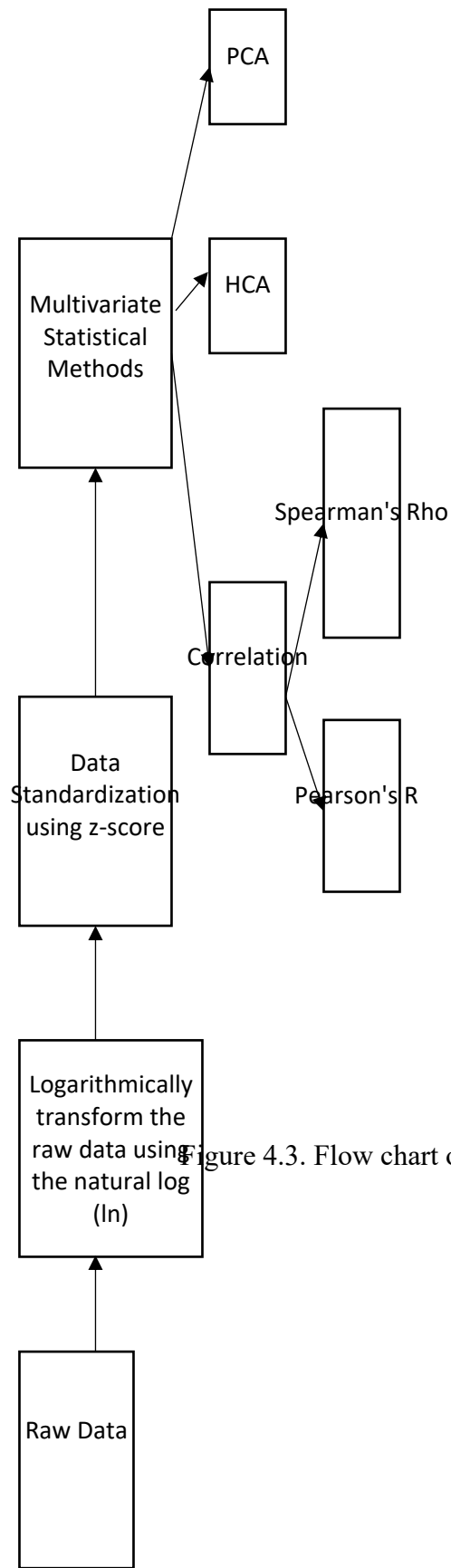


Figure 4.3. Flow chart of multivariate statistical methods used

Methods used with excel include: (1) histograms, (2) box plots using raw data and (3) location and scale estimates. Methods used with SPSS 24 include: (1) descriptive statistics, (2) quantile-quantile plots, (3) de-trended quantile-quantile plots, (4) correlation analyses, (5) hierarchical cluster analysis, and (6) principal component analysis. Methods used with Minitab 17 include ln transformed box plots. The level of significance chosen for all statistical methods is 0.05.



## 5. AZ2 RESULTS

### **AZ2 Sediment Core Description**

AZ2 extends from the surface to 35.5 meters. Meters are used to describe the core. From 0 – 2 m, the sediment is light brown to light greenish gray. From 2 – 3.5 m, there was little recovery. From 3.5 – 4.5 m, the sediment is light yellowish brown. From 4.5 – 5.5 m, there was little recovery. The upper 5.5 m of sediment is wet, sticky and plastic clay. From 5.5 – 6.5 m, the sediment is light yellowish brown with oxidation present. From 6.5 – 11 m, the sediment is light brownish gray with oxidation present. From 5.5 – 11 m, the sediment mainly consists of sandy silt with clayey silt and silty sand lenses. From 11 to 12.5 m, there is little recovery. From 12.5 – 14 m, the sediment is mainly silty sand and dark gray, with oxidation and mottling present. From 14 – 15.5 m, the sediment is dark olive gray and consists of silty sand and silt. From 15.5 – 17.5 m, the sediment is dark olive gray, friable and contains organics. The sediment consists of silt and sandy silt. From 17.5 – 19 m, the sediment is black and dark gray and consists of organics, mainly sandy silt with some sand, and diatoms are present. From 19 – 20.5 m, there is little recovery. From 21.5 – 22 m, the sediment is green with laminations and contains diatoms. The sediment is mainly sandy silt with clayey silt lenses. From 22 – 22.5 m, the sediment is dark gray and consists of sandy silt, with diatoms present. From 22.5 – 24 m, the sediment is greenish gray, mottled and laminated. The sediment is mainly silty sand with sandy silt and silt lenses. From 24 – 25 m, the sediment is light gray and mainly silty sand. From 25 – 26 m, the sediment is greenish gray and contains quartz sand and clay. The sediment mainly is silt with a sandy silt lens. From 26 – 31.5 m, the sediment consists of quartz sand. From 31.5 – 34.5 m, there was little recovery. From 34.5 – 35.5 m, the sediment is pale olive to light gray. The sediment

consists of mainly sandy silt with silt and clayey silt lenses. There is sand located near the bottom of the sediment core.

### **Geochemical Concentration by Depth**

The distribution of major elements in AZ2 demonstrates significant fluctuations with depth throughout AZ2. The major constituents in AZ2 are Ca (0.08-23.9 wt. %), Fe (0.66-8.36 wt. %), Mg (0.33-7.75 wt. %) and Na (0.13-6.01 wt. %). The AZ2 samples are depleted in S (0.03-3.75 wt. %), Al (0.67-3.12 wt. %), K (0.24-1.40 wt. %) and Ti (0.01-0.05 wt. %).

The Al and K concentrations mirror one another throughout the majority of AZ2. From 34.50 to 35.54 m, Al and K mirror each other's fluctuations. From 5.58 to 25.76 m, Al and K mirror each other's fluctuations, with the exception of two places. Al and K do not mirror each other at 6.42 m, where Al contains a major peak, while K does not. Al and K do not mirror each other at 20.67 m, where K contains a major peak, while Al does not. The Ca and Mg concentrations mirror one another throughout the majority of AZ2. From 34.50 to 35.54 m, Ca and Mg do not mirror each other's fluctuations, as the concentration of Ca is high, while the concentration of Mg is low. From 5.58 to 25.76 m, Ca and Mg mirror each other's fluctuations, with the exception of one place. Ca and Mg do not mirror each other at 6.42 m. Ca contains a major peak at 6.42 m, while Mg does not. The S and Fe concentrations mirror one another from 17.46 to 23.25 m. The Fe and Na concentrations mirror one another to a lesser extent from 5.58 to 14.50 m. Figure 5.1. presents graphs of the major elements with depth for AZ2.

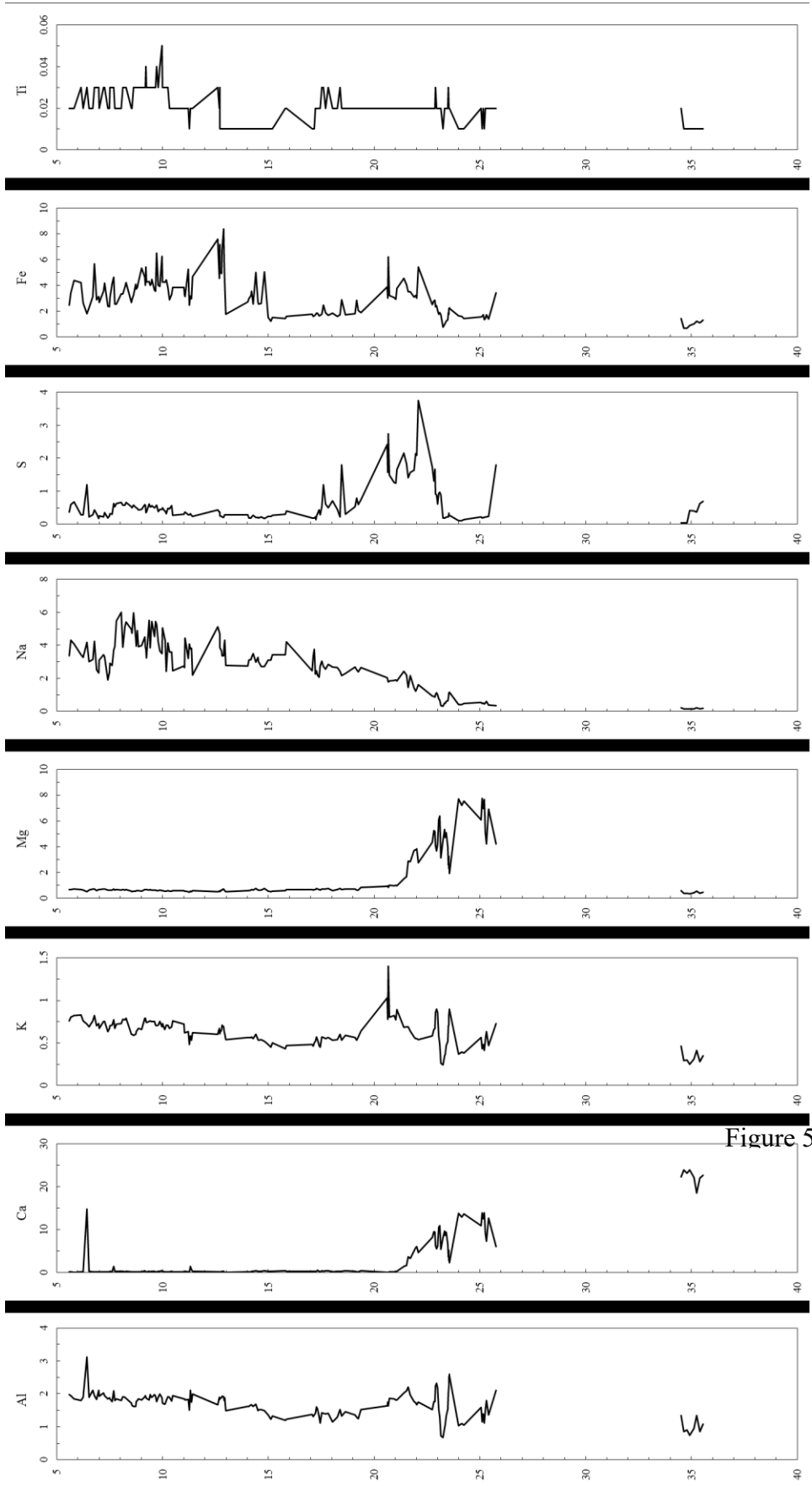


Figure 5.1. Graphs of Major Elements with D

The concentration of Al fluctuates throughout AZ2; however, fluctuates dramatically throughout the lower part of AZ2. From 34.50 to 35.54 m, Al is low to moderate, ranging from 0.74 to 1.34 wt. %. From 23.35 to 25.76 m, Al is moderate to high, ranging from 1.03 to 2.59 wt. %. Four minor peaks occur at 22.95, 23.54, 23.56 and 23.57 m, which correspond to Al concentrations of 2.32, 2.49, 2.59 and 2.55 wt. %, respectively. From 23.15 to 23.32 m, Al is low, ranging from 0.67 to 0.97 wt. %. From 5.58 to 23.10 m, Al is moderate to high, ranging from 1.12 to 2.32 wt. %, excluding the major peak that occurs during this interval. One major peak occurs at 6.42 m and corresponds to 3.12 wt. %.

The concentration of Ca remains low and relatively stable throughout the upper part of AZ2; however, fluctuates dramatically in the lower part of AZ2. From 34.40 to 35.54 m, Ca is highest and is a major peak, ranging between 18.5 to 23.9 wt. %. From 21.41 to 25.76 m, Ca fluctuates dramatically between moderate to high values, ranging from 1.37 to 13.9 wt. %. Seven major peaks occur at 24.00, 24.13, 24.24, 25.11, 25.16, 25.21 and 25.42 m, which correspond to Ca concentrations of 13.8, 12.9, 13.6, 13.9, 12.5, 13.9 and 12.6 wt. %, respectively. Five minor peaks occur at 22.81, 22.88, 23.05, 23.10 and 23.32 m, which correspond to Ca concentrations of 9.5, 9.4, 10.5, 10.9 and 9.7 wt. %, respectively. From 5.58 to 21.07 m, Ca remains very low and stable, ranging from 0.08 to 1.46 wt. %, excluding the major peak that occurs during this interval. One major peak occurs at 6.42 m and corresponds to a Ca concentration of 14.7 wt. %. From 5.58 to 21.07 m, two minor increases occur at 7.68 and 11.31 m, which correspond to a Ca concentration of 1.46 and 1.41 wt. %, respectively.

The concentration of K fluctuates throughout AZ2; however, fluctuates dramatically throughout the lower part of AZ2. From 34.50 to 35.54 m, K is low, ranging from 0.29 to 0.46 wt. %. From 20.62 to 25.76 m, K fluctuates dramatically, with values from low to high, ranging

from 0.24 to 1.4 wt. %. One major peak occurs at 20.67 m and corresponds to 1.4 wt. %. Five minor peaks occur at 22.91, 22.95, 23.00, 23.56 and 23.57 m and correspond to 0.86, 0.9, 0.85, 0.9 and 0.89 wt. %, respectively. From 12.98 to 19.37 m, K is low to moderate, ranging from 0.43 to 0.64 wt. %. From 5.58 to 12.93 m, K is moderate to high, ranging from 0.48 to 0.83 wt. %.

The concentration of Mg remains low and relatively stable throughout the upper and lower parts of AZ2; however, fluctuates dramatically in the middle part of AZ2. From 34.50 to 35.54 m, Mg is low, ranging from 0.33 to 0.60 wt. %. From 21.41 to 25.76 m, Mg increases and fluctuates between moderate to high values, ranging from 1.52 to 7.75 wt. %. Six major peaks occur at 24.00, 24.13, 24.24, 25.11, 25.21 and 25.42 m, which correspond to Mg concentrations of 7.72, 7.20, 7.54, 7.75, 7.65 and 6.93 wt. %, respectively. Five minor peaks occur at 22.81, 22.88, 23.05, 23.10 and 23.32 m, which correspond to Mg concentrations of 5.27, 5.15, 6.09, 6.36 and 5.35 wt. %, respectively. From 5.58 to 21.07 m, Mg is low, ranging from 0.47 to 1.02 wt. %.

The concentration of Na decreases with increasing depth throughout the length of AZ2; however, fluctuates dramatically throughout. From 34.50 to 35.54 m, the concentration of Na is lowest, ranging from 0.13 to 0.22 wt. %. From 21.88 to 25.76 m, Na is low to moderate, ranging from 0.31 to 1.62 wt. %. From 17.22 to 21.71 m, Na fluctuates between low to moderate values, ranging from 1.44 to 3.04 wt. %. From 5.58 to 17.16 m, Na fluctuates dramatically between moderate to high values, ranging from 1.9 to 6.01 wt. %. Two major peaks occur at 8.04 and 8.64 m, which correspond to 6.01 and 5.97 wt. %. Minor peaks occur at 7.80, 8.28, 9.36, 9.49 and 9.68 m, which correspond to 5.49, 5.42, 5.51, 5.44 and 5.44 wt. %, respectively.

The concentration of S remains relatively low and stable throughout the upper part of AZ2; however, fluctuates dramatically in the lower part of AZ2. From 34.92 to 35.54 m, S is low to moderate, ranging from 0.37 to 0.68 wt. %. From 34.50 to 34.80 m, the concentration of S is lowest, ranging from 0.03 to 0.04 wt. %. From 23.25 to 25.76 m, S is low, ranging from 0.10 to 0.33 wt. %, excluding the minor peak that occurs at 25.76 m and corresponds to 1.80 wt. %. From 20.62 to 22.88 m, S is moderate to high, ranging from 1.24 to 2.42 wt. %, excluding the two major peaks which occurs during this interval. Two major peaks occur at 20.67 and 22.09 m and correspond to 2.74 and 3.75 wt. %, respectively. From 17.52 to 19.37 m, S is low to moderate, ranging from 0.21 to 0.79 wt. %, excluding the two minor peaks which occur during this interval. Two minor peaks occur at 17.59 and 18.45 m, which correspond to 1.2 and 1.8 wt. %, respectively. From 5.58 to 17.46 m, S is low, ranging from 0.14 to 0.67 wt. %, excluding the minor peak which occurs during this interval. One minor peak occurs at 6.42 m, which corresponds to 1.19 wt. %.

The concentration of Fe fluctuates throughout AZ2. From 34.50 to 35.54 m, Fe is low, ranging from 0.66 to 1.42 wt. %. From 23.00 to 25.76 m, Fe is low, ranging from 0.77 to 2.24 wt. %, excluding the increased value which occurs at 25.76 m. From 18.45 to 22.95 m, Fe fluctuates dramatically from low to high values, ranging from 1.6 to 6.21 wt. %. One minor peak occurs at 20.67 m, which corresponds to 6.21 wt. %. From 14.99 to 17.52 m, Fe is low, ranging from 1.21 to 1.86 wt. %. From 12.98 to 14.81 m, Fe fluctuates from low to high values, ranging from 1.74 to 5.06 wt. %. From 5.58 to 12.93 m, Fe fluctuates dramatically from low to high values, ranging from 1.79 to 8.36 wt. %. One major peak occurs at 12.88 m, which corresponds to 8.36 wt. %. Five minor peaks occur at 9.73, 9.98, 12.61, 12.69 and 12.81, which correspond to 6.51, 6.24, 7.57, 7.17 and 6.76 wt. %, respectively.

The concentration of Ti remains relatively low and stable throughout AZ2. From 10.00 to 35.54 m, Ti is low, with concentrations between 0.01 to 0.03 wt. %. From 9.21 to 9.98 m, Ti is moderate to high, ranging between 0.03 to 0.05 wt. %. From 5.58 to 9.18 m, Ti is low to moderate, ranging between 0.02 to 0.03 wt. %. One major peak occurs at 9.98 m. Three minor peaks occur at 9.21, 9.73 and 9.86 m.

The distribution of trace elements in AZ2 demonstrates significant fluctuations with depth throughout AZ2. The minor constituents in AZ2 are Mn (49-1445 ppm), Sr (24-759 ppm) and Ba (10-400 ppm). The Ba and Sr concentrations mirror one another from 5.58 to 17.22 m; however, there are two places where the concentrations do not mirror each other. B and Sr do not mirror each other at 8.04 m, where B increases, while Sr does not. B and Sr do not mirror each other at 11.31 m, where Sr increases, while Ba does not. Ni and Pb mirror one another from 5.58 to 21.88 m; however, there are three main places, where the concentrations do not mirror one another. At 6.42, 8.23 and 11.31 m, the concentration of Ni increases, while Pb does not. Ni and Zn mirror one another from 14.99 to 35.54 m; however, there is one place, where the concentrations do not mirror one another. Ni and Zn mirror one another to a lesser extent from 5.58 to 14.81 m; however, at 8.28 m, the concentration of Zn increases, while Ni does not. Cr and Cu mirror one another from 5.58 to 12.81 m. Co and Cu mirror one another to a lesser extent. Mn and Sr mirror one another from 5.58 to 20.65 m; however, there are two places where the concentrations do not mirror one another. At 7.68 and 11.31 m, the concentration of Sr increases, while Mn does not. V and Zn mirror one another from 23.00 to 35.54 m. Figure 5.2. presents graphs of the trace elements with depth for AZ2.

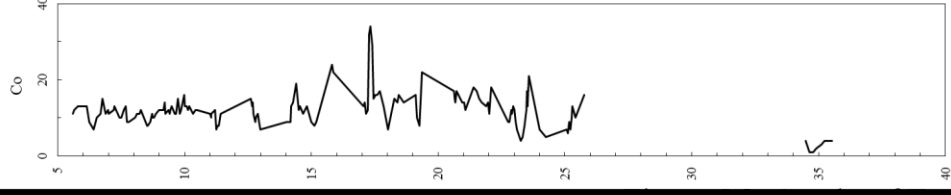
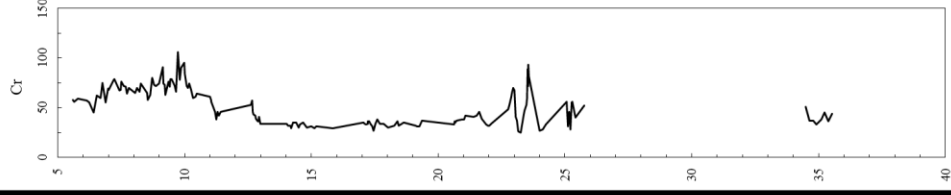
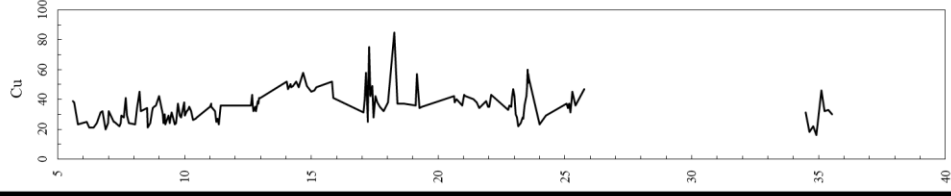
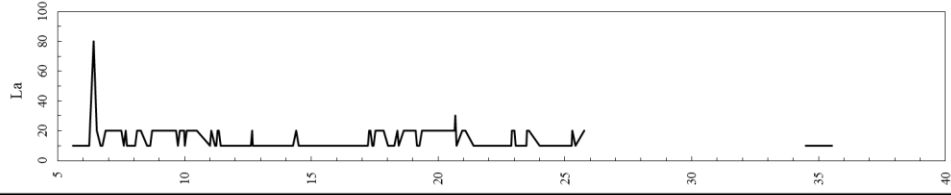
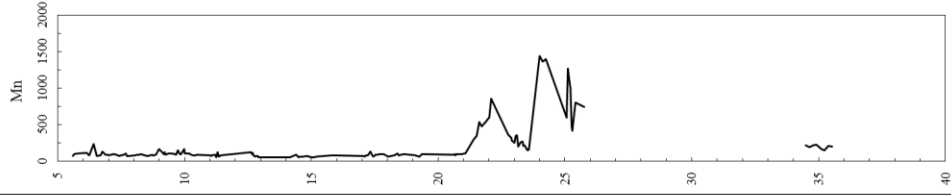
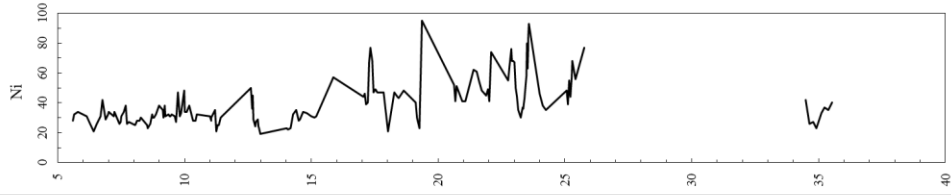
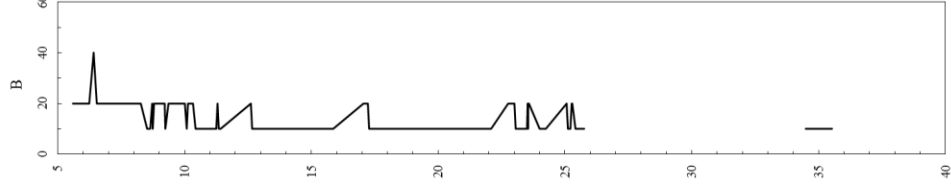
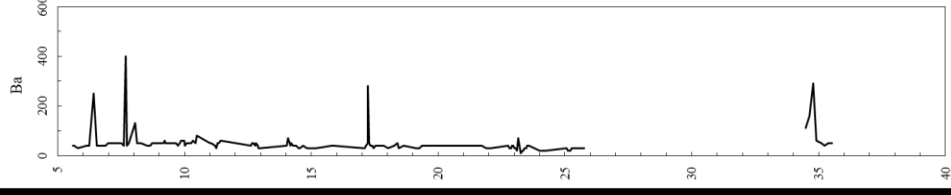


Figure 3.2. Graphs of Trace Elements with Depth of





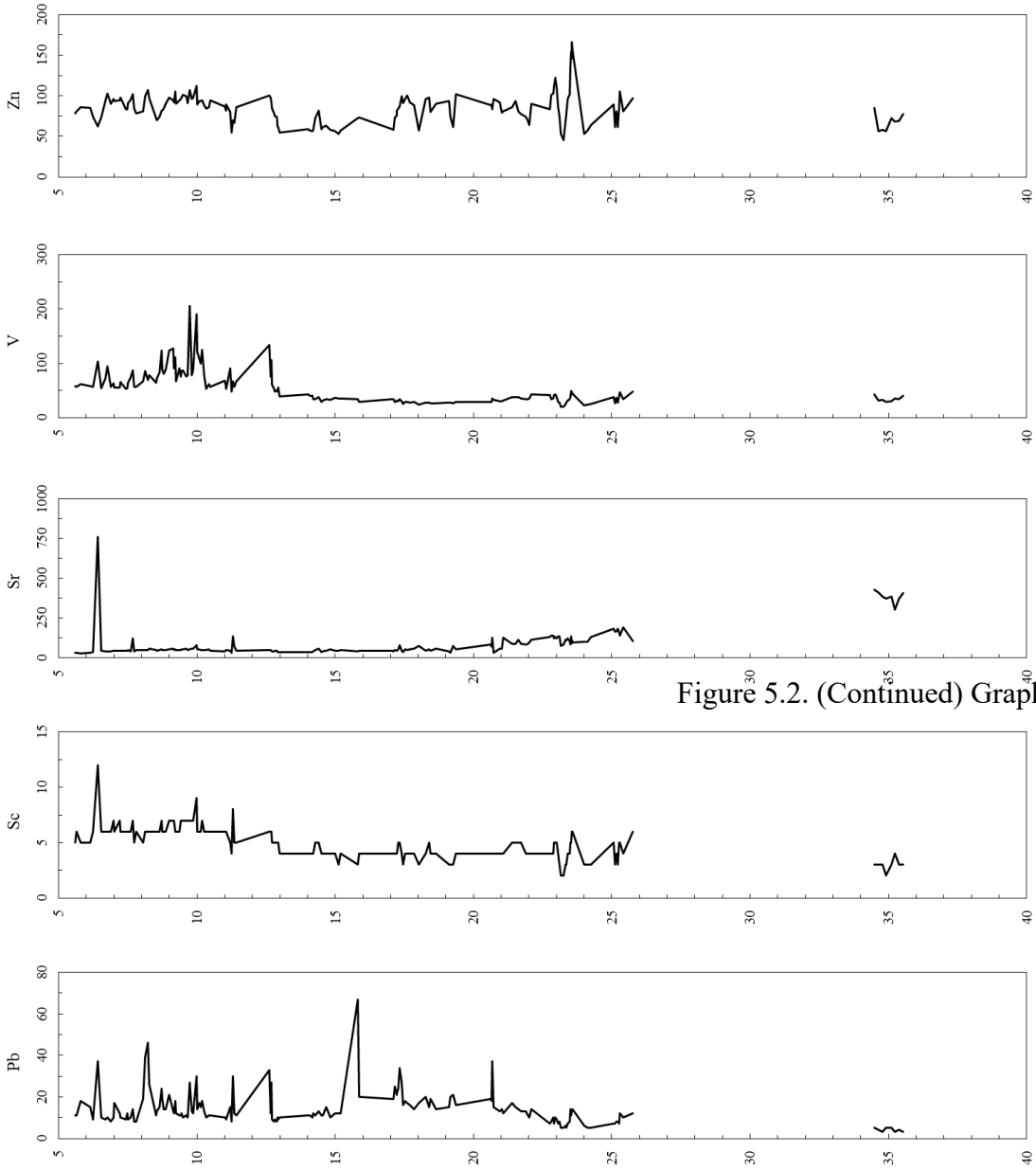


Figure 5.2. (Continued) Graphs of Trace Elements with

The concentration of B fluctuates between low and high values, 10 and 20 ppm, respectively, throughout AZ2. From 10.43 to 35.54 m, B is low, where values of 10 ppm dominate the samples. From 5.58 to 10.33 m, B is high, where values of 20 ppm dominate the samples. B contains one major peak at 6.42 m.

The concentration of Ba remains relatively low and stable throughout the length of most of AZ2; however, there are several major and minor peaks. From 34.92 to 35.53 m, the concentration of Ba is low. From 34.50 to 34.80 m, Ba fluctuates, with values ranging from 40-290 ppm. Three major peaks occur at 34.50, 34.64 and 34.80 m. From 8.13 to 25.76 m, Ba is low, with values ranging from 10-80 ppm after the exclusion of a major peak that occurs at 17.24 m. From 5.58 to 8.04 m, Ba is low, ranging from 30 to 50 ppm after the exclusion of one minor peak and two major peaks. One minor peak occurs at 8.04 m and two major peaks occur at 6.42 and 7.68 m.

The concentration of Co fluctuates dramatically throughout AZ2. From 34.50 to 35.54 m, Co is low, ranging from 1-4 ppm. From 11.42 to 25.76 m, Co fluctuates dramatically between low and high values, ranging from 4 to 34 ppm. Three major peaks occur at 17.28, 17.32 and 17.40 m. Four minor peaks occur at 15.81, 15.86, 19.37 and 23.57 m. From 5.58 to 11.36 m, Co is moderate, ranging from 7 to 16 ppm.

The concentration of Cr remains relatively stable throughout the middle of AZ2; however, fluctuates dramatically towards the top and bottom parts of AZ2. From 34.50 to 35.54 m, Cr is low, with values ranging from 33 to 51 ppm. From 22.81 to 25.76 m, Cr fluctuates dramatically between low and high values, ranging from 25 to 93 ppm. Two major peaks occur at 23.52 and 23.56 m. From 11.20 to 22.75 m, Cr remains low and stable, ranging from 27 to 57

ppm. From 5.58 to 11.06 m, Cr fluctuates dramatically between moderate and high values, ranging from 45 to 106 ppm. A major peak occurs at 9.73 m.

The concentration of Cu fluctuates dramatically throughout AZ2. From 14.02 to 35.54 m, Cu fluctuates dramatically between low and high values, ranging between 16-85 ppm. Two major peaks occur at 17.28 and 18.27 m. From 5.58 to 12.98 m, Cu fluctuates between low to moderate values, ranging between 20-45 ppm.

The concentration of La fluctuates between low and high values throughout AZ2. From 21.41 to 35.54 m and from 12.69 to 17.24 m, La is low, where values of 10 ppm dominate the samples. From 17.28 to 21.07 m and from 5.58 to 12.67 m, La is high, where values of 20 ppm dominate the samples. One major and one minor peak occurs at 6.42 and 20.67 m, respectively.

The concentration of Mn remains relatively low and stable throughout the upper length of AZ2; however, fluctuates dramatically toward the middle of AZ2. From 34.50 to 35.54 m, Mn is low, ranging from 147 to 225 ppm. From 21.41 to 25.76 m, Mn fluctuates dramatically between low to high values, ranging from 150 to 1445 ppm. Six major peaks occur at 24.00, 24.13, 24.24, 25.11, 25.16 and 25.21 m. One minor peak occurs at 22.09 m. From 5.58 to 21.07 m, Mn is low, ranging from 49 to 231 ppm.

The concentration of Ni fluctuates dramatically throughout AZ2. The concentration of Ni fluctuates dramatically, between low to high values from 11.20 to 35.54 m, ranging from 19-95 ppm. Two major peaks occur at 19.37 and 23.57 m. Eight minor peaks occur at 17.32, 22.09, 22.88, 23.50, 23.54, 23.56, 23.57 and 25.76 m. The concentration of Ni fluctuates between low and moderate values from 5.58 to 11.06 m, ranging from 21-48 ppm.

The concentration of Pb fluctuates dramatically throughout AZ2. From 34.50 to 35.54 m, the concentration of Pb is low, ranging from 3-5 ppm. From 20.72 to 25.76 m, Pb is low, ranging

from 5-17 ppm. From 15.81 to 20.67 m, Pb fluctuates dramatically between low to high values from 14-67 ppm. One major peak occurs at 15.81 m. Two minor peaks occur at 17.32 and 20.67 m. From 5.58 to 15.20 m, Pb fluctuates between low to moderate values, ranging between 8-46 ppm. Four minor peaks occur at 6.42, 8.13, 8.23 and 12.61 m.

The concentration of Sc fluctuates dramatically throughout the upper part of AZ2; however, remains low in most of the lower part of AZ2. From 11.36 to 35.54 m, Sc is low, ranging from 2-6 ppm. From 5.58 to 11.31 m, Sc fluctuates dramatically between low and high values, ranging from 4-12 ppm. One major peak occurs at 6.42 m. Two minor peaks occur at 9.98 and 11.31 m.

The concentration of Sr remains relatively low and stable throughout AZ2. From 34.50 to 35.54 m, the Sr is high, ranging from 301-427 ppm. From 20.67 to 25.76 m, Sr is moderate to high, ranging from 28-189 ppm. From 5.58 to 20.65 m, Sr is low to moderate, ranging from 24-132 ppm, with the exception of the major peak which occurs at 6.42 m.

The concentration of V remains relatively low and stable throughout most of the lower part of AZ2; however, fluctuates dramatically throughout the upper part of AZ2. From 12.72 to 35.54 m, V is low, ranging from 20-60 ppm. From 5.58 to 12.69 m, V fluctuates dramatically between moderate to high values, ranging from 47-206 ppm. Two major peaks occur at 9.73 and 9.98 m. Six minor peaks occur at 8.73, 9.00, 9.15, 10.00, 10.18, and 12.61 m.

The concentration of Zn fluctuates dramatically throughout AZ2. From 12.67 to 35.54 m, the concentration of Zn fluctuates dramatically between low and high values, ranging from 45-166 ppm. Five major peaks occur at 23.50, 23.52, 23.54, 23.56 and 23.57 m. One minor peak

occurs at 22.95 m. From 5.58 to 12.61 m, Zn fluctuates between low to moderate values, ranging from 55-112 ppm.

### **Chemical Index of Alteration**

The CIA increases with depth throughout the length of AZ2. From 34.50 to 35.54 m, the CIA is highest, ranging from 1.66 to 3.74. From 21.88 to 25.76 m, the CIA fluctuates dramatically from low to high values, ranging from 3.79 to 26.49. From 5.58 to 21.81 m, the CIA is low and stable, ranging from 13.17 to 30.87, excluding the major peak which occurs at 6.42 m. At 6.42 m, the CIA is 8.33.

### **Elemental Ratios**

Elemental molar ratios such as Na/Al, Na/Ti and Na/K are also useful in determining the degree of chemical weathering in the sediments (Roy et al., 2006). These ratios combined with the CIA and evaluation of the depletion and enrichment of certain elements aid in the determination of the degree of chemical weathering (Roy et al., 2006). Low values for the above mentioned elemental molar ratios indicate increased chemical weathering (Roy et al., 2006).

The Na/Al molar ratio decreases with depth throughout the length of AZ2. From 34.50 to 35.54 m, Na/Al is lowest, ranging from 0.24-0.29. From 21.88 to 25.76 m, Na/Al is low, ranging from 0.27-1.52. From 5.58 to 21.81 m, Na/Al fluctuates dramatically from low to high values, ranging from 1.08-6.06.

The Na/Ti molar ratio decreases with depth throughout the length of AZ2. From 34.50 to 35.54 m, Na/Ti is lowest, ranging from 13.53-28.35. From 21.88 to 25.76 m, Na/Ti is low, ranging from 21.91-104.38. From 5.58 to 21.81 m, Na/Ti fluctuates dramatically from 89.94-554.10.

The Na/K molar ratio decreases with depth throughout the length of AZ2. From 34.50 to 35.54 m, Na/K is lowest, ranging from 0.69-0.82. From 21.88 to 25.76 m, Na/K is low, ranging from 0.71 to 4.56. From 5.58 to 21.81 m, Na/K is low to high, ranging from 1.92-15.38.

The Na/Al and Na/K molar ratios mirror one another throughout the length of AZ2. A slight decrease is seen in the Na/K molar ratio at 20.67 m, corresponding to 1.92, while no increase is seen at this depth for Na/K. The Na/Al and Na/Ti molar ratios mirror one another to a slightly lesser extent. Figure 5.3. presents graphs of the CIA and elemental ratios with depth of AZ2.

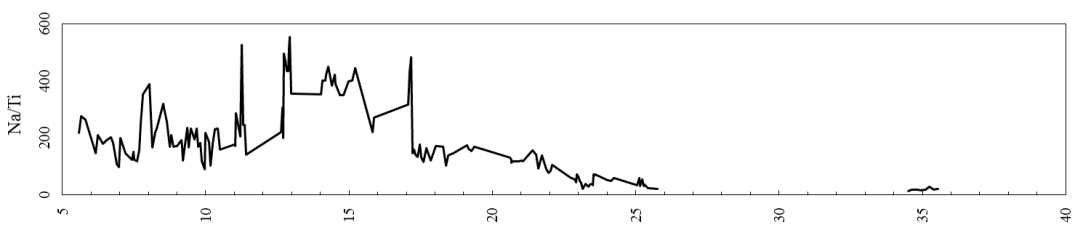
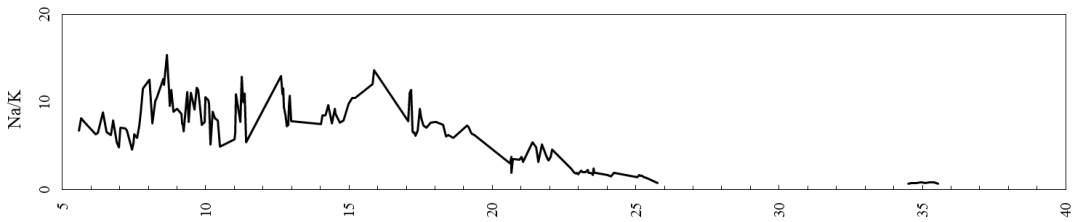
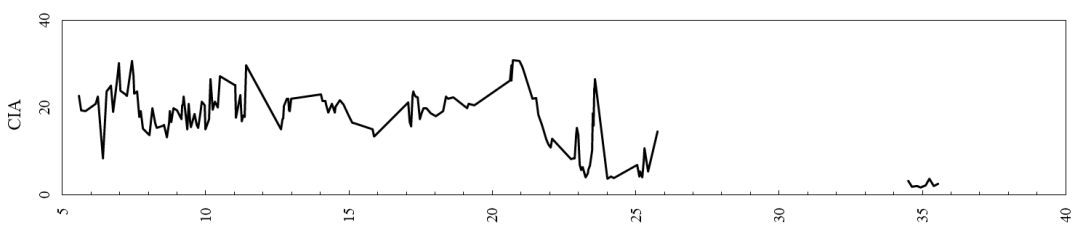
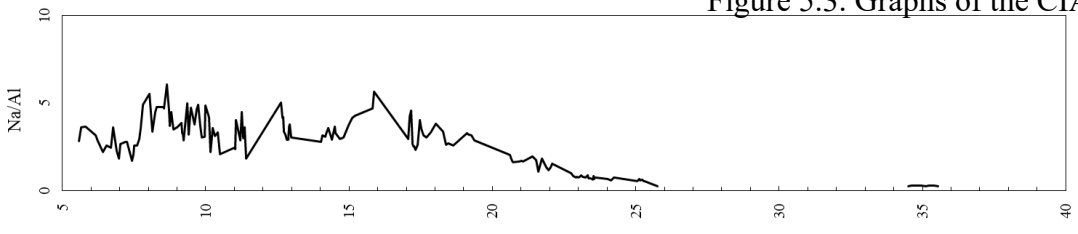


Figure 5.3. Graphs of the CIA and Elemental Ratios w



## 6. AZ2 DISCUSSION

### **Introduction**

McLennan et al. (1993) considers the following four sedimentary processes which affect the distribution and behavior of major elements in sediments: weathering, diagenesis, sorting and sedimentary recycling. The distribution of elements in lakes is related to both climatic and watershed processes (Cohen, 2003). Weathering and atmospheric deposition are two processes which bring major ions into lakes (Cohen, 2003). Climate determines the major residence times of these major ions (Cohen, 2003). The inorganic component of sediment consists of mineral particles washed or blown into the drainage basin, in addition to inorganic ions which are attributed to inflow or regional precipitation (Engstrom and Wright, 1984). The elemental composition of lacustrine sediments reflects both the composition of the earth's crust, as well as the solubility of the elements in lake water (Cohen, 2003).

The inorganic component of lake sediments is not easily distinguishable in terms of mineralogy; however, elemental analysis is a way of environmental reconstruction (Engstrom and Wright, 1984). Even though the chemical constituents of the inorganic portion of lake sediments are interdependent, relative changes in elemental concentrations yield information related to the composition of source material (Engstrom and Wright, 1984). Conclusions regarding lake and catchment history generally cannot be made using chemical analysis of lake sediments because it is difficult to determine the source and the diagenetic processes that occurred post deposition; however, the distribution of detrital minerals in lake sediment provides information as to weathering and erosion in the catchment (Engstrom and Wright, 1984).



## Geochemical Elements

Major elements and uncensored trace elements that occur in the AZ2 sediments include lithophile, siderophile and chalcophile elements. Lithophile elements show an affinity for silicate phases. Lithophile elements are concentrated in the silicate portion of the earth's crust and mantle. Lithophile elements include Al, Ca, K, Mg, Na, Ti, B, Ba, Be, Br, Cr, Mn, Sc, Sr and V. Siderophile and chalcophile elements show an affinity for metallic liquid and sulfide liquid phases, respectively. Siderophile elements are concentrated in the earth's core. Chalcophile elements are concentrated in the earth's core and are depleted in the silicate portion of the earth. Siderophile elements include Fe, Co and Ni. Chalcophile elements include Cu, Pb, S and Zn.

Al is considered a poor metal, while Si is a metalloid. Si and Al are the second and third most abundant elements in earth's crust, after O. Al occurs mainly in silicate igneous rock minerals. In sedimentary environments, clays are the most common Al enriched mineral. Clays are found in natural water environments, most soils and hydrolysate sedimentary rocks. Generally, Al and Si are the two most abundant inorganic elements in lacustrine sediments (Engstrom and Wright, 1984). A large proportion of Al and Si supplied to lacustrine sediments are contributed by clastic silicate minerals (Engstrom and Wright, 1984).

Al is considered primarily allochthonous in lacustrine sediments (Engstrom and Wright, 1984). Al is resistant to chemical weathering, as it is resistant to erosion and soil leaching (Engstrom and Wright, 1984). Al and Si are relatively unreactive and insoluble, and are concentrated in soil profiles during periods of increased weathering (Cohen, 2003). Increased terrigenous Al in lacustrine sediments indicates an increase in weathering intensity, suggesting warm and wet climatic conditions (Cohen, 2003). A large proportion of Si contained in lake sediments is biogenic and attributed to siliceous algae (Cohen, 2003).

Alkali metals are Na and K. Alkaline earth metals are Mg and Ca. Additional alkaline earth metals include Ba, Be and Sr. Generally, the alkali and alkaline earth metals comprise silicate minerals. Na and K are the sixth and seventh most abundant elements in earth's crust, respectively. Na sources include mainly feldspar minerals, but Na is also present in unaltered mineral grains or as sodium salt crystals deposited or precipitated in saline water. In sediments, K commonly occurs in unaltered feldspar, mica or illite. On clay mineral surfaces, Na is retained by adsorption due to the high cation exchange capacity of clay minerals; therefore, Na tends to remain in solution. In natural systems, Na and K behave differently. Na tends to remain in solution after it is freed from silicate mineral structures. It is more difficult to free K from silicate mineral structures. There is a strong tendency for K to become reincorporated into weathering products, generally clays.

Na and K are precipitated in authigenic minerals in hypersaline and closed basin lakes (Cohen, 2003). Na and K are considered highly soluble; thus, the past salinity of a lake is difficult to determine without mineralogical data (Cohen, 2003). A large proportion of Na and K occurs in crystalline and clastic sedimentary bedrock, and in related watersheds as part of silicate mineral grains (Cohen, 2003). Under arid conditions, Na and K salts accumulate in soils, as Na and K are soluble weathering products (Cohen, 2003). Na is easily removed from weathering products (Nesbitt et al., 1980). Ca behaves similarly to Na during chemical weathering (Nesbitt et al., 1980).

Ca is the fifth most abundant element in the earth's crust. Ca makes up mainly pyroxene, amphibole and feldspars. In sedimentary environments, calcium is found mainly in carbonate minerals. Sources of Ca in the AZ2 sediment include the precipitation of  $\text{CaCO}_3$  and aeolian inputs. Since the bedrock of the Al-Azraq basin is made up of limestone, specifically the Umm

Rijam Chert-Limestone and the overlying Wadi Shallala Formations, determining the source of Ca is difficult. Limestone is made up of mainly calcite; however, magnesium carbonate and other impurities commonly occur. In sedimentary environments, magnesium is present with calcium carbonate. Dolomite occurs if the Mg:Ca molar ratio approaches 1:1. Other minerals which contain Ca include sulphates, such as gypsum and anhydrite. Mg is the eighth most abundant element in the earth's crust. Ca and Mg exhibit similar cation exchange behavior, in which both ions are strongly adsorbed by clay minerals.

In watersheds, generally Ca and Mg comprise the bedrock (Cohen, 2003). In the detrital sediment fraction, Ca and Mg behave similarly to Na and K (Cohen, 2003). During periods of increased erosion, the concentrations of Ca and Mg increase in the sediment (Cohen, 2003). Mackereth (1966) investigated the chemical stratigraphy of English lakes and concluded that sedimentary profiles of the alkali and alkaline earth elements can be used to infer about in-situ leaching of mineral constituents and erosion of soil particles. Mackereth (1966) concluded that during active erosion, the concentrations of Na, K and Mg increase in the sediment. Also, during periods of deep weathering, the concentration of Na, K and Mg become depleted in the sediment (Mackereth, 1966). Mackereth (1966) concluded that Na, K and Mg in the sediment reflect erosion within the catchment and the degree of weathering, as these elements are associated with detrital minerals.

In lake sediments, alkali and alkaline earth metals commonly occur in allogenic clastics eroded from rocks and catchment soils (Engstrom and Wright, 1984). Engstrom and Wright (1984) considers Na<sup>+</sup>, K<sup>+</sup> and Mg<sup>+</sup> in surface waters to be conservative, as these ions are generally not sedimented to a significant degree through biological uptake or chemical precipitation. Na, K and Mg are incorporated into sediment through exchange across the

sediment-water interface, sorption onto suspended particles or exchange processes which occur in the soil profile (Engstrom and Wright, 1984). Ca precipitates as authigenic carbonates in alkaline lakes, as do Mg and Fe to a lesser extent (Engstrom and Wright, 1984). In saline environments, Ca, Mg and Fe may precipitate as sulphates (Engstrom and Wright, 1984).

Fe is the fourth most abundant element in earth's crust. Both the chemical behavior and the solubility of Fe in water depends upon oxidation intensity and pH. If excessive amounts of Fe occur in water, oxyhydroxide precipitates form. Fe occurs in igneous rock minerals such as pyroxenes, amphiboles, biotite, magnetite and olivine. When igneous rock minerals are exposed to water, Fe is released and reprecipitated as sedimentary species. In reducing environments, ferrous polysulfides such as pyrite and marcasite may occur. In the earth's crust, there is one fifteenth as much manganese as there is Fe. Manganese makes up a significant portion of basalt, olivine, pyroxene and amphibole. Manganese substitutes for Fe, magnesium or calcium in silicate structures. During weathering, Fe and Mn are a part of redox processes. Small amounts of manganese are commonly found in carbonates, as manganese substitutes for calcium.

Fe and Mn are supplied to lake sediments as mineral particles, and the concentration of these elements reflects erosion within the catchment similarly to the alkali and alkaline earth elements (Engstrom and Wright, 1984). The concentration of Fe and Mn in lake sediment is influenced by the conditions of the lake and the catchment; however, these elements are difficult to interpret (Engstrom and Wright, 1984). In lake sediments, Fe and Mn are found in authigenic clastics and comprise authigenic oxides, sulphides, carbonates, as well as organic complexes (Engstrom and Wright, 1984). Authigenic Fe and Mn are considered labile, thus, those processes controlling preservation of these elements should be considered (Engstrom and Wright, 1984). Fe and Mn exhibit a low solubility under oxidizing conditions; therefore, oxidized Fe and Mn

can travel to a lake through groundwater (Engstrom and Wright, 1984). Fe and Mn are easily released from the soil under reducing conditions (Engstrom and Wright, 1984). Mn has a higher solubility compared to Fe, thus, Mn is released from the soil more easily (Engstrom and Wright, 1984). Increased Fe and Mn in lake sediments indicates warmer climatic conditions; however, the redox environment of the lake should be considered (Cohen, 2003). Fe and Mn can also be supplied to lake sediments as oxides, colloids or organic complexes (Cohen, 2003). The mobility of Fe and Mn increases at redox boundaries, therefore, the concentrations of these elements can be used to determine redox history of a lake (Cohen, 2003).

Trace metals are supplied to a lake through erosion of bedrock or atmospheric sources (Cohen, 2003). The concentration of trace metals in lake sediments can be influenced by many processes including watershed erosion, lake transport and diagenesis (Cohen, 2003). The concentrations of the following trace metals are influenced by redox conditions: Cd, Co, Cr, Cu, Ni, Pb (Cohen, 2003). Ti, Cr and Cu are elements with a low solubility, thus, these elements are less mobile during chemical weathering (Yanes et al., 2006). Ti and Cr are transported in refractory minerals such as quartz, ilmenite, chromite and zircon (Yanes et al., 2006). Cu is associated with both clay minerals and organic matter (Yanes et al., 2006). Al, P, Sc, V and Pb are elements concentrated in clays and organic matter (Yanes et al., 2006). The sedimentary signal of Fe, Mn and P is mixed, thus, these signals are hard to interpret (Engstrom and Wright, 1984). P occurs in P bearing igneous and sedimentary rocks (Cohen, 2003). P and Mn behave similarly to Ti during chemical weathering (Nesbitt and Marcovics, 1997). Inorganic compounds containing P exhibit a low solubility, as P is a nutrient for biota. The solubility of P is controlled by coprecipitation and adsorption, and also uptake by biota. S is supplied to a lake through groundwater, stream and atmospheric sources (Cohen, 2003). Reduced forms of sulfur occur in

sulfide minerals and are considered relatively immobile. In sedimentary environments, S occurs as metallic sulfides.

Uncensored transition metals in AZ2 sediments include Fe, Ti, Co, Cr, Cu, Mn, Ni, Sc, V and Zn. Transition metal behavior during chemical weathering is not well known (Wei et al., 2004). Fe, Ti, Co, Cr, Cu, Mn, Ni, V and Zn are considered to be sensitive to water column and sedimentary redox conditions (Mackereth, 1966; Davison, 1993). Sc is considered a conservative element during chemical weathering (Nesbitt and Marcovics, 1997).

### **Chemical Index of Alteration**

The geochemistry of major and trace elements is affected by chemical weathering of siliciclastic sediments (Roy et al., 2008). The CIA is an index which measures changes in major elements in the siliclastic fractions due to sediment-water interaction (Roy et al., 2010). The degree of chemical weathering; thus, the composition of the siliclastic sediments is controlled by climate and tectonism (Roy et al., 2008). Chemical weathering includes those interactions between precipitation and rocks and their weathering products (Nesbitt and Young, 1989). The amount of acids in precipitation ultimately determines the extent of chemical weathering (Singer, 1980; Nesbitt and Young, 1984).

The degree of chemical weathering is an indication of moisture availability and average precipitation of the region (Roy et al., 2008). Chemical weathering is enhanced in a warm and wet climate, as chemical weathering is controlled by both temperature and moisture (Nesbitt and Young, 1982). White and Blum (1995) state that during warm periods with high precipitation, chemical weathering intensity is increased, while the opposite occurs during dry periods with decreased precipitation.

Using molecular proportions, the CIA can be used to determine the degree of chemical weathering (Nesbitt and Young, 1982). The CIA is defined by the following equation:

$$[\text{Al}_2\text{O}_3 / (\text{Al}_2\text{O}_3 + \text{CaO} + \text{Na}_2\text{O} + \text{K}_2\text{O})] \times 100$$

CIA values less than 60, between 60-80 and greater than 80 indicate low, moderate and extreme chemical weathering, respectively (Fedo et al., 1995). The chemical composition of the sediments is different compared to that of the parent rocks because during the chemical weathering process, elements are mobilized and redistributed in the sediment (Nesbitt et al., 1979, Nesbitt et al., 1980, Nesbitt and Marcovics, 1997). The average CIA values for shale ranges between 70 to 75 (Nesbitt and Young, 1982). Changes in the CIA reflect proportional changes in feldspar and other clay minerals in the sediment (Nesbitt and Young, 1982).

Wedepohl (1969) estimated that quartz, plagioclase and potassium feldspar make up 21, 41 and 21% by volume of the earth's crust, respectively. Nesbitt et al. (1980) studied the behavior of alkali and alkaline earth metals in the Toorongro Granodiorite weathering profile and determined that the following two processes influence the behavior of these elements: (1) the degradation of primary minerals and the release of elements into solutions, and (2) fixation by exchange, as well as, adsorption and incorporation into clay structural sites.

During periods of increased sediment-water interaction, primary detrital minerals, such as plagioclase and amphibole, are degraded and secondary clay minerals, such as kaolinite, are formed (Nesbitt and Young, 1982). Labile cations such as  $\text{Ca}^{2+}$ ,  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Fe}^{2+}$  and  $\text{Mg}^{2+}$  are leached in preference to hydrolysates such as  $\text{Al}^{3+}$  and  $\text{Ti}^{4+}$  (Nesbitt and Young, 1982). Alkali and alkaline earth elements are leached in preference to Al, Fe and Ti from the weathering profile (Nesbitt and Young, 1989). Fe behaves similarly to Ti during chemical weathering (Nesbitt and Marcovics, 1997). Under humid conditions, chemical weathering removes Na, Ca

and K from feldspars, while Al and Ti are enriched in clay minerals and sheet silicates (Mason and Moore, 1982; Nesbitt and Young, 1982). Thus, the ratio of the alkaline metal, Na, to the hydrolysates Al and Ti would be low under humid conditions (Nesbitt and Young, 1982). High values of the CIA indicate increased moisture availability in the catchments and deposition under humid conditions, while low CIA values indicate the decreased moisture availability in the catchments and deposition under arid conditions (McLennan and Taylor, 1991; Andersson et al., 2004). When precipitation increases, the concentrations of the following soluble elements tend to increase in lake sediments: K, Be, Mg, Ca, Ba and Sr (Engstrom and Wright, 1984).

Elemental concentrations and elemental ratios aid in the determination of the degree of chemical weathering that has occurred in the sediment, in addition to paleoenvironmental history reconstruction in the sediment source region (Nesbitt and Young, 1982). The molar ratios of Na/Al and Na/Ti indicate the degree of chemical weathering of the detrital sediment fraction (Roy et al., 2006). The Na/K molar ratio indicates the abundance of Ca-Na feldspars relative to K-feldspars (Roy et al., 2008). Low values of the Na/Al, Na/Ti and Na/K molar ratios indicate increased chemical weathering (Roy et al., 2006). Under humid conditions, Na<sup>+</sup> is displaced by water, while K<sup>+</sup> is adsorbed in the clay minerals, thus K<sup>+</sup> is retained in the sediments (Roy et al., 2008).

### **Weathering of AZ2 Sediments**

Chemical weathering indicates the availability of moisture and the average precipitation in the Al-Azraq basin, as well as changes in environment. The enrichment or depletion of elements, in addition to the CIA and elemental ratios, such as Na/Al, Na/Ti and Na/K indicate the degree of chemical weathering in the AZ2 sediment. AZ2 data indicates periods of humid and arid conditions, as well as periods of increased and decreased moisture availability.



CIA values of AZ2 are low throughout the entire length of the sediment core; however, values decrease with depth, ranging from 1.66 to 30.87. The range of CIA values of the AZ2 sediment is much lower than the shale average (70 – 75), suggesting that clay minerals perhaps do not dominate the AZ2 sediments (Das et al., 2008). Low CIA values of the sediments of AZ2 indicate low chemical weathering based on Fedo et al. (1995). The AZ2 sediments are divided into three zones based on the enrichment and depletion of selected elements, molar ratios and the CIA.

CIA values of Zones 1, 2 and 3 are all less than 60, which indicate low overall chemical weathering in the AZ2 sediment (Fedo et al., 1995). In terms of relative CIA values, Zone 1 exhibits the highest CIA values. In Zone 2, the CIA values are slightly less compared to Zone 1. In Zone 3, the CIA values are lowest compared to the other zones. Roy et al. (2006) states that low values of the Na/Al, Na/Ti and Na/K molar ratios indicate increased chemical weathering; however, the lowest molar ratios are exhibited by Zone 3 followed by Zones 2 and 1, respectively. Perhaps due to the degree of low chemical weathering in the AZ2 sediments, differences in the CIA and the relative changes in the Na/Al, Na/Ti and Na/K molar ratios are an indication of low chemical weathering.

Zone 1 is from 5.58 to 21.81 m, characterized by relatively high CIA values, with high values of the molar ratios of Na/Al, Na/Ti and Na/K. Zone 2 is from 21.88 to 25.76 m, characterized by intermediate CIA values, with low values of the molar ratios of Na/Al, Na/Ti and Na/K. Zone 3 is from 34.50 to 35.54 m, characterized by the lowest CIA values, with the lowest values of the molar ratios of Na/Al, Na/Ti and Na/K. Zone 1 corresponds to a period of the highest chemical weathering in the AZ2 sediments, as indicated by the CIA, followed by Zones 2 and 3, respectively.

Zone 1 is characterized by relatively low concentrations of Ca, Mg and K and relatively high concentrations of Al, Ti and Fe, indicating that during chemical weathering these elements were removed and retained, respectively. Low concentrations of Ca, Mg, K, Ba and Sr indicate a period of decreased precipitation (Engstrom and Wright, 1984). In Zone 1, low concentrations of Ca, Mg and Sr indicate little to no carbonate precipitation (Engstrom and Wright, 1984). High concentrations of Na in Zone 1 indicates a period of aridity with increased evaporation and perhaps increased aeolian input.

Zones 2 and 3 are characterized by an increase in precipitation due to the increase in the concentrations of Ca, Mg, K and Sr observed in the AZ2 sediments (Engstrom and Wright, 1984). Zones 2 and 3 are characterized by increased Ca values and low Na values, corresponding to a period of greater humidity with decreased evaporation and perhaps decreased aeolian input. In Zone 2, periods of increased K, could be an indication of lacustrine deposition. Although, the sediments of AZ2 experience low chemical weathering (<60); Zone 3 is characterized by relatively low values of Na, Mg and K, indicating that these elements were removed during chemical weathering. Ca was not removed during chemical weathering in Zone 3 and this possibly indicates carbonate precipitation and perhaps increased aeolian input (Engstrom and Wright, 1984).

Comparison of the AZ2 sediments to the mean upper continental crust (UCC) composition reported by Taylor and McLennan (1985) yields interesting results. The concentration values of all samples of AZ2 for Al, Fe and Ti are depleted relative to the UCC values reported by Taylor and McLennan (1985) of 15.9, 9.1 and 0.9, respectively. The concentration values of the mobile, soluble elements including Ca, K and Mg are depleted in almost all AZ2 samples compared to the UCC values determined by Taylor and McLennan

(1985) of 7.4, 1.1 and 5.3, respectively. Perhaps there is a period of increased Aeolian input in Zone 3, specifically from 22.75 m through the bottom of the AZ2 sediment core, as the majority of Ca concentration values are enriched relative to the UCC value (Taylor and McLennan, 1985). The concentration values of the majority of AZ2 samples in Zone 1, specifically from the 5.58 to 17.16 m are enriched in Na relative to the UCC value for Na determined by Taylor and McLennan (1985), of is 3.1. Perhaps, there is a period of increased aridity in Zone 1, specifically from 5.58 to 17.16 m. From 17.22 m through the bottom of the AZ2 sediment core, all the Na concentration values are depleted relative to the UCC value (Taylor and McLennan, 1985), perhaps indicating more humid environments in Zones 2 and 3 compared to Zone 1.

## 7. AZ1 & AZ2 STATISTICAL RESULTS

### **Introduction**

This chapter shows the results of methods used to remove and replace censored values, as well as detect outliers in the data. Graphical methods of detecting data normality include histograms, Q-Q plots, de-trended Q-Q plots, and box plots. Statistical methods of detecting data normality include skewness and kurtosis z-scores, and the Shapiro-Wilk Test. Results of robust and non-robust estimates of location and scale are presented. Finally, results of correlation, principal components analysis, and hierarchical cluster analysis are presented.

### **Censored Values**

In AZ1 and AZ2, those elements with less than 5% censored values less than the lower limit of detection were replaced with one-half the censored value, while censored values greater than the upper limit of detection were replaced with the censored value. In AZ1, the following trace elements contain less than 5 % censored values, therefore, a replacement method is used before statistical analysis: Be, La, Mo, P, Pb and Sr. In AZ2, the following trace elements contain less than 5 % censored values, therefore, a replacement method is used before statistical analysis: Be and P. Table 7.1. presents the number and percentage of censored values in AZ1 and AZ2.

Table 7.1. Number and Percentage of Censored Values in the trace elements of AZ1 and AZ2.

AZ1			AZ2		
Element	No. of Values	% of Values	Element	No. of Values	% of Values
Ag	236	93.28	Ag	130	79.75
As	76	30.04	As	31	19.02
Be	6	2.37	Be	7	4.29
Bi	243	96.05	Bi	125	76.69
Cd	124	49.01	Cd	107	65.64
Ga	49	19.37	Ga	31	19.02
Hg	227	89.72	Hg	157	96.32
La	3	1.19	-	-	-
Mo	1	0.4	Mo	40	24.54
P	3	1.19	P	1	0.61
Pb	1	0.4	-	-	-
-	-	-	Sb	147	90.18
Sr	2	0.79	-	-	-
Th	249	98.42	Th	163	100
Tl	253	100	Tl	162	99.39
U	236	93.28	U	143	87.73
W	253	100	W	163	100

## Outliers

In AZ1 and AZ2, the removal of outliers was determined using the box plot inner fence and the median  $\pm$  2 MAD methods. If the proportion of outliers in univariate data is below 10 % or between 15-50 %, then the boxplot inner fence or median  $\pm$  2 MAD methods was used, respectively, based off of Reimann et al. (2005). No variables in AZ1 contained more than 10 % outliers, thus, the box plot inner fence method was the only method used with AZ1 data. A few variables in AZ2 contained a high proportion of data outliers, therefore, the median  $\pm$  2 MAD method was used in addition to the box plot inner fence method. Table 7.2. presents the number of outliers and extreme outliers for AZ1 and AZ2 data and ln transformed data using outlier removal methods. Appendix C and D show outlier values with associated depths for the raw and

ln transformed elements of AZ1 and AZ2, respectively, using the box plot inner fence rule.

Appendix E shows outlier values with associated depths for the raw and ln transformed elements of AZ2 using the median +/- 2 MAD method.

Table 7.2. Number of outliers and extreme outliers for AZ1 and AZ2 data and ln transformed data using outlier removal methods.

Element	AZ1				AZ2			
	Raw		Ln Transformed		Raw		Ln Transformed	
	No. Outliers	No. Extreme Outliers	No. Outliers	No. Extreme Outliers	No. Outliers	No. Extreme Outliers	No. Outliers	No. Extreme Outliers
Al	0		3		2		7	
Ca	0		0		25	13	0	
K	0		1		4	1	8	
Mg	0		0		37	35	35	17
Na	0		1		0		9	
S	6	6	6	1	22	11	6	
Fe	5	1	2		2		0	
Ti	1		0		0		37	
B	0		0		1		0	
Ba	10		0		7	6	8	4
Be	6		6	6	10		12	7
Co	5	2	3	1	10	3	12	3
Cr	9		10	2	0		0	
Cu	6	2	12	3	3	1	3	
La	1		0		1	1	1	
Mn	16	5	0		23	16	12	
Mo	19	5	0		9		2	
Ni	2	1	2	1	8	3	3	
P	8	4	10		12	3	6	
Pb	20	11	29	9	1		4	
Sr	18	6	5		12	9	8	
V	15	5	7	1	9	2	1	
Zn	2		1	1	5	1	5	

## Geochemical Data Distribution

### *Histograms*

In AZ1, the major element histograms of Al, Ca, K, Mg, Na, S and Fe are skewed right, while Ti is skewed left. The histograms of Al, K and Na are non-symmetrical and bi-modal. The S, Fe and Ti histograms contain outliers. The major element histograms of Al, Ca, K, Mg and Na have a negative kurtosis and are platykurtic. The major element histograms of S, Fe and Ti have a positive kurtosis and are leptokurtic.

The Al, K and Na histogram look similar in shape; however, K and Na look the most similar. The modes of the first and second peaks of the Al, K and Na histograms correspond to a large and smaller number, respectively. The histograms of Ca and Mg are very similar in shape.

In AZ1, the ln transformed major element histograms of Al, K, Mg, Na and S are skewed right, while Ca, Fe and Ti are skewed left. The histograms of Ca, K and Na are non-symmetrical and bi-modal. The Al, K, Na, S and Fe histograms contain outliers. The ln transformed major element histograms of Ca, K, Mg, Na and Ti have a negative kurtosis and are platykurtic. The ln transformed major elements of Al, S and Fe have a positive kurtosis and are leptokurtic.

In AZ1, the trace element histograms of B, Ba, Co, Cu, La, Mn, Mo, Ni, P, Pb, Sr, V and Z are skewed right, while Be and Cr are skewed left. The histogram of Zn appears normally distributed; however, it is skewed to the right and contains one outlier. The histograms of Co, Cu, La, Mn, Mo, Ni, P, Pb, Sr and V contain outliers. The histogram of B and Be have a negative kurtosis and are platykurtic. The histograms of Ba, Co, Cr, Cu, Mn, Mo, Ni, P, Pb, Sr and V have a positive kurtosis and are leptokurtic. The histogram of La has a positive kurtosis and is platykurtic. The histogram of Zn is mesokurtic and has a kurtosis value close to zero.

In AZ1, the ln transformed trace elements of Be, Cr, Cu, Mn, P, V and Zn are skewed left, while B, Ba, Co, La, Mo, Ni, Pb and Sr are skewed right. The Ba and Mn histograms are non-symmetrical and bi-modal. The histogram of V appears normally distributed; however, contains outliers. The ln transformed histograms of Be, Co, Cu, La, Mo, Ni, P, Pb, Sr and Zn histograms contain outliers. The ln transformed histograms of B, Ba, La, Mn and Mo have a negative kurtosis and are platykurtic. The ln transformed trace element histograms of Co, Cr, Cu, Ni, Pb, Sr, V and Zn have a positive kurtosis and are leptokurtic. The ln transformed histogram of Be and P have a positive kurtosis and are platykurtic. The histograms of raw and ln transformed major element and trace element data from AZ1 are in appendix F and H, respectively.

In AZ2, the major element histograms of Ca, Mg, S, Fe and Ti are skewed right, while Al, K and Na are skewed left. The histograms of Al, Na and Fe are non-symmetrical and bi-modal. The Al, Ca, K, Mg, S and Ti histograms contain outliers. The major element histograms of Na has a negative kurtosis and is platykurtic. The major element histograms of Al, Ca, K, Mg, S, Fe and Ti have a positive kurtosis and are leptokurtic.

In AZ2, the ln transformed major element histograms of Al, K, Na, Fe and Ti are skewed left, while Ca, Mg and S are skewed right. The histogram of Fe is non-symmetrical and bi-modal. The K, S and Ti histograms contain outliers. The ln transformed major element histograms of Fe and Ti have a negative kurtosis and are platykurtic. The ln transformed major element histograms of Al, Ca, K, Mg, Na and S have a positive kurtosis and are leptokurtic.

In AZ2, the uncensored trace element histograms of B, Ba, Be, Cr, Cu, La, Mn, Ni, P, Pb, Sr, V and Zn are skewed right. The histograms of Co and Sc are skewed left. The histogram of Cr is non-symmetrical and bi-modal. The B, Ba, Co, Cr, Cu, La, Ni, Pb, Sc, Sr, and V histograms



contain outliers. The uncensored trace element Cr has a negative kurtosis and is platykurtic. The histogram of Be has a positive kurtosis and is platykurtic. The uncensored trace elements of B, Ba, Cu, Ln, Mn, Ni, P, Pb, Sc, Sr, V and Zn have a positive kurtosis and are leptokurtic.

In AZ2, the ln transformed trace elements of B, Ba, Cr, La, Mn, Ni, P, Pb, Sr and V are skewed right, while Be, Co, Cu, Sc and Zn are skewed left. The histogram of Cr is non-symmetrical and bi-modal. The Cu histogram appears normally distributed. The Co, La, Sc, Sr and V histograms contain outliers. The ln transformed uncensored trace elements of B, Cr, Ni and V have a negative kurtosis and are platykurtic. The ln transformed trace element of P has a positive kurtosis and are platykurtic. The ln transformed uncensored trace elements of Ba, Be, Co, La, Mn, Pb, Sc, Sr and Zn have a positive kurtosis and are leptokurtic. The ln transformed Cu is mesokurtic and has a kurtosis value close to zero. The histograms of raw and ln transformed major element and trace element data from AZ2 are in appendix G and I, respectively.

#### *Quantile-Quantile Plots*

The interpretation of the Q-Q plots of major and trace element data from AZ1 are similar to the interpretation of the corresponding histograms. All of the raw and ln transformed major element Q-Q plots in AZ1 show data points that deviate from the 45 degree line and represent either a skewed distribution or a skewed and bi-modal distribution. All of the Q-Q plots of the trace elements in AZ1 deviate from normality, with the exception of Be and Zn. The Be and Zn Q-Q plots show a normal distribution with nearly all of the data points on or close to the 45 degree line. All of the Q-Q plots of the ln transformed trace elements in AZ1 deviate from normality. The Q-Q and detrended Q-Q plots of raw and ln transformed major element and trace element data from AZ1 are in appendix J.

The interpretation of the Q-Q plots of major and uncensored trace element data from AZ2 are similar to the interpretation of the corresponding histograms. All of the raw and ln transformed major element Q-Q plots in AZ2 show data points that deviate from the 45 degree line and represent either a skewed distribution or a skewed and bi-modal distribution. All of the Q-Q plots of the uncensored trace elements in AZ2 deviate from normality. All of the Q-Q plots of the ln transformed uncensored trace elements in AZ2 deviate from normality with the exception of Cu. The ln transformed Cu Q-Q plot shows a normal distribution with nearly all of the data points on or close to the 45 degree line. The Q-Q and de-trended Q-Q plots of raw and ln transformed major element and uncensored trace element data from AZ2 are in appendix K.

#### *Box Plots*

Appendix L presents box plots of data from AZ1 and AZ2. Appendix M presents box plots of ln transformed data from AZ1 and AZ2. NEED MORE HERE

### **Statistical Methods of Determining Data Normality**

#### *Skewness and Kurtosis*

In AZ1, both the skewness and kurtosis z-scores for Be and Zn are within +/- 1.96 and are normally distributed. Both the skewness and kurtosis z-scores for ln K and ln Ni are within +/- 1.96 and are normally distributed.

The skewness z-scores for the following ln transformed major elements are within +/- 1.96: ln Ca and ln Mg. The skewness z-scores for the following ln transformed trace elements are within +/- 1.96: ln Ba, ln La, ln Mn and ln V. All other major and uncensored trace elements have skewness z-scores outside the range of +/- 1.96.

The kurtosis z-scores for the following major elements are within +/- 1.96: Al, K, ln Na and Ti. The kurtosis z-scores for the following trace elements are within +/- 1.96: B and ln P.

The kurtosis z-scores for B and ln P is -0.55 and 1.84, respectively. All other major and trace elements have kurtosis z-scores outside the range of +/- 1.96. Table 7.3. presents skewness and kurtosis z-scores for raw and ln transformed elements in AZ1.

Table 7.3. Skewness and Kurtosis z-scores of AZ1. Bolded values are within +/- 1.96.

Element	Skewness z-score	Kurtosis z-score	Element	Skewness z-score	Kurtosis z-score
Al	2.69	<b>-0.95</b>	ln Al	-4.28	3.80
Ca	3.05	-4.16	ln Ca	<b>-1.80</b>	-5.37
K	4.74	<b>-1.31</b>	<b>ln K</b>	<b>-0.38</b>	<b>-1.48</b>
Mg	4.87	-2.68	ln Mg	<b>0.84</b>	-5.34
Na	2.42	-3.90	ln Na	-2.53	<b>-0.93</b>
S	37.60	177.26	ln S	5.46	9.17
Fe	7.51	13.42	ln Fe	-4.31	2.84
Ti	4.52	<b>0.84</b>	ln Ti	-2.62	-2.15
B	6.31	<b>-0.55</b>	ln B	2.03	-3.83
Ba	9.34	4.81	ln Ba	<b>1.91</b>	-4.17
Be	<b>0.40</b>	<b>-0.49</b>	ln Be	-7.96	7.92
Co	20.57	69.41	ln Co	-2.93	21.73
Cr	4.45	4.17	ln Cr	-5.81	7.57
Cu	27.67	127.01	ln Cu	-2.55	12.42
La	93.96	717.50	ln La	<b>-0.02</b>	-3.88
Mn	14.73	21.51	ln Mn	<b>0.67</b>	-3.36
Mo	20.43	43.03	ln Mo	3.57	-3.15
Ni	7.31	4.62	<b>ln Ni</b>	<b>0.30</b>	<b>1.03</b>
P	25.20	66.96	ln P	-2.52	1.84
Pb	93.90	717.45	ln Pb	17.74	54.90
Sr	55.85	254.76	ln Sr	8.95	11.56
V	25.89	95.97	ln V	<b>1.29</b>	5.78
Zn	<b>1.65</b>	<b>1.12</b>	ln Zn	-7.40	15.83

In AZ2, both the skewness and kurtosis z-scores for Be are within +/- 1.96 and are normally distributed. Skewness and kurtosis z-scores for Be are 0.21 and 1.47, respectively. Both

the skewness and kurtosis z-scores for ln Fe, ln Cu, ln P, ln Sc and ln Zn are within  $\pm 1.96$  and are normally distributed.

The skewness z-scores for the following major elements are within  $\pm 1.96$ : Al and Na. The skewness z-scores for Al and Na are -0.45 and -0.20, respectively. The skewness z-scores for the following ln transformed major elements are within  $\pm 1.96$ : ln S. The skewness z-scores for the following trace elements are within  $\pm 1.96$ : ln Cr, ln P and ln Pb. All other major and trace elements have skewness z-scores outside the range of  $\pm 1.96$ .

The kurtosis z-scores for the following major elements are within  $\pm 1.96$ : Ti. The kurtosis statistics for the following ln transformed major elements are within  $\pm 1.96$ : ln Mg and ln Ti. The kurtosis z-scores for the following trace elements are within  $\pm 1.96$ : Cr. The kurtosis statistics for the following ln transformed trace elements are within  $\pm 1.96$ : ln Ni and ln V. All other major and trace elements have kurtosis z-scores outside the range of  $\pm 1.96$ . Table 7.4. presents skewness and kurtosis z-scores for raw and ln transformed elements in AZ2.

Table 7.4. Skewness and Kurtosis z-scores of AZ2. Bolded values are within +/- 1.96

Element	Skewness z-score	Kurtosis z-score	Element	Skewness z-score	Kurtosis z-score
Al	<b>-0.45</b>	3.32	ln Al	-5.57	4.62
Ca	11.58	11.24	ln Ca	4.79	-2.02
K	2.25	7.98	ln K	-4.82	4.74
Mg	10.41	7.45	ln Mg	7.05	<b>0.70</b>
Na	<b>-0.20</b>	-2.25	ln Na	-7.14	2.75
S	13.47	21.16	ln S	<b>0.62</b>	2.92
Fe	4.58	2.17	ln Fe	<b>-1.65</b>	<b>-0.77</b>
Ti	2.10	<b>0.72</b>	ln Ti	-2.54	<b>-1.74</b>
B	5.07	3.65	ln B	2.66	-3.74
Ba	28.83	89.02	ln Ba	9.87	21.62
Be	<b>0.21</b>	<b>1.47</b>	ln Be	-8.78	10.29
Co	7.03	13.68	ln Co	-9.92	19.77
Cr	3.03	<b>-1.86</b>	ln Cr	<b>0.86</b>	-3.38
Cu	6.26	8.37	ln Cu	<b>0.57</b>	<b>0.63</b>
La	24.41	105.05	ln La	4.27	2.33
Mn	17.13	29.27	ln Mn	7.71	4.04
Ni	7.15	4.36	ln Ni	3.06	<b>-0.54</b>
P	15.60	37.24	ln P	<b>-0.20</b>	<b>1.85</b>
Pb	14.02	30.32	ln Pb	<b>0.40</b>	2.64
Sc	4.05	6.55	ln Sc	<b>-1.77</b>	<b>0.83</b>
Sr	20.47	50.82	ln Sr	7.49	4.91
V	10.37	14.60	ln V	3.26	<b>-0.73</b>
Zn	5.47	8.26	ln Zn	<b>0.18</b>	<b>1.46</b>

*Shapiro-Wilk Test*

The raw and ln transformed major element concentrations in AZ1 are all non-normally distributed ( $p < 0.05$ ). The raw and ln transformed trace element concentrations in AZ1 are all non-normally distributed ( $p < 0.05$ ), with the exception of Zn ( $p > 0.05$ ). The concentration of Zn is normally distributed with  $W = 0.99$ ,  $p = 0.076$ . Table 7.5. presents Shapiro-Wilk Test Statistics and p-values for raw and ln transformed element data in AZ1. NEED REF

Table 7.5. Shapiro-Wilk Test Statistics of AZ1.

Element	Shapiro-Wilk Statistic	p-value	Element	Shapiro-Wilk Statistic	p-value
Al	0.98	0.001	ln Al	0.97	0.000
Ca	0.84	0.000	ln Ca	0.84	0.000
K	0.93	0.000	ln K	0.97	0.000
Mg	0.84	0.000	ln Mg	0.86	0.000
Na	0.93	0.000	ln Na	0.94	0.000
S	0.61	0.000	ln S	0.96	0.000
Fe	0.93	0.000	ln Fe	0.96	0.000
Ti	0.92	0.000	ln Ti	0.91	0.000
B	0.83	0.000	ln B	0.87	0.000
Ba	0.80	0.000	ln Ba	0.92	0.000
Be	0.98	0.005	ln Be	0.91	0.000
Co	0.78	0.000	ln Co	0.92	0.000
Cr	0.97	0.000	ln Cr	0.95	0.000
Cu	0.75	0.000	ln Cu	0.95	0.000
La	0.70	0.000	ln La	0.71	0.000
Mn	0.76	0.000	ln Mn	0.95	0.000
Mo	0.63	0.000	ln Mo	0.86	0.000
Ni	0.91	0.000	ln Ni	0.98	0.001
P	0.65	0.000	ln P	0.97	0.000
Pb	0.11	0.000	ln Pb	0.74	0.000
Sr	0.20	0.000	ln Sr	0.88	0.000
V	0.72	0.000	ln V	0.98	0.000
Zn	0.99	0.076	ln Zn	0.94	0.000

The raw and ln transformed major element concentrations in AZ2 are all non-normally distributed ( $p < 0.05$ ). The raw and ln transformed trace element concentrations in AZ2 are all non-normally distributed ( $p < 0.05$ ) with the exception of ln Cu ( $p > 0.05$ ). The ln concentration of Cu is normally distributed with  $W = 0.99$ ,  $p = 0.468$ . Table 7.6. presents Shapiro-Wilk Test Statistics and p-values for raw and ln transformed major element data in AZ2.

Table 7.6. Shapiro-Wilk Test Statistics of AZ2.

Element	Shapiro-Wilk Statistic	p-value	Element	Shapiro-Wilk Statistic	p-value
Al	0.96	0.000	ln Al	0.92	0.000
Ca	0.59	0.000	ln Ca	0.81	0.000
K	0.96	0.000	ln K	0.94	0.000
Mg	0.60	0.000	ln Mg	0.73	0.000
Na	0.97	0.001	ln Na	0.84	0.000
S	0.70	0.000	ln S	0.95	0.000
Fe	0.95	0.000	ln Fe	0.98	0.043
Ti	0.84	0.000	ln Ti	0.82	0.000
B	0.64	0.000	ln B	0.65	0.000
Ba	0.39	0.000	ln Ba	0.77	0.000
Be	0.97	0.002	ln Be	0.85	0.000
Co	0.90	0.000	ln Co	0.84	0.000
Cr	0.92	0.000	ln Cr	0.94	0.000
Cu	0.93	0.000	ln Cu	0.99	0.468
La	0.51	0.000	ln La	0.66	0.000
Mn	0.53	0.000	ln Mn	0.83	0.000
Ni	0.88	0.000	ln Ni	0.97	0.000
P	0.73	0.000	ln P	0.97	0.003
Pb	0.78	0.000	ln Pb	0.98	0.011
Sc	0.91	0.000	ln Sc	0.93	0.000
Sr	0.54	0.000	ln Sr	0.85	0.000
V	0.81	0.000	ln V	0.95	0.000
Zn	0.92	0.000	ln Zn	0.96	0.000

### Estimates of Location

In AZ1, the values of the non-robust estimates of location for all major elements decreased in the following order: midrange > arithmetic mean > geometric mean > harmonic mean. In AZ1, the values of the non-robust estimates of location for all trace elements, with the exception of La, decreased in the following order: midrange > arithmetic mean > geometric mean > harmonic mean. In AZ1, the non-robust estimates of location for La decreased in the following order: midrange > geometric mean > harmonic mean > arithmetic mean. In AZ1, the

highest and lowest values of the estimate of location for all major elements and all trace elements with the exception of La, are the midrange and harmonic mean, respectively. Table 7.7. presents non-robust estimates of location for elements in AZ1.

Table 7.7. Non-Robust Estimates of Location for Elements in AZ1.

Element	Arithmetic Mean	Geometric Mean	Harmonic Mean	Midrange
Al	2.176	2.046	1.902	2.370
Ca	6.036	1.872	0.434	9.530
K	0.997	0.906	0.821	1.255
Mg	3.435	2.217	1.475	5.645
Na	2.916	2.497	2.081	3.315
S	0.335	0.306	0.285	1.245
Fe	2.949	2.682	2.390	5.100
Ti	0.029	0.026	0.022	0.045
B	43.083	35.545	29.947	65.000
Ba	167.470	104.290	69.710	395.000
Be	1.070	1.008	0.924	1.075
Co	12.794	12.237	11.649	23.000
Cr	53.680	50.346	46.468	69.000
Cu	29.312	27.576	25.804	75.500
La	12.770	13.899	12.980	22.500
Mn	398.016	255.270	168.724	1361.000
Mo	6.630	3.145	1.870	35.750
Ni	45.573	42.015	38.749	71.500
P	1454.070	1058.873	726.360	5045.000
Pb	12.771	8.555	7.537	334.000
Sr	290.964	130.796	94.636	5020.500
V	76.897	69.122	62.526	241.500
Zn	79.506	76.514	72.911	87.500

In AZ1, the values of the robust estimates of location for Al, Ni, P and Pb decreased in the following order: 10 % trimmed mean > Trimean > 20 % trimmed mean > 25 % trimmed mean > Gastwirth median > median. The values of the robust estimates of location for K, Mg, B



and Ba decreased in the following order: 10 % trimmed mean > 20 % trimmed mean > Trimean > 25 % trimmed mean > Gastwirth median > median. The values of the robust estimates of location for S and Mo decreased in the following order: 10 % trimmed mean > Trimean > 20 % trimmed mean > Gastwirth median > 25 % trimmed mean > median. The values of the robust estimates of location for Co and La decreased in the following order: 10 % trimmed mean > 20 % trimmed mean > 25 % trimmed mean > Gastwirth median > Trimean > median. For the following elements, the order of decreasing robust estimates of mean was unique in AZ1: Ca, Na, Fe, Ti, Be, Cr, Cu, Mn, Sr, V and Zn.

In AZ1, the highest values of the robust estimates of location was the 10 % trimmed mean for the following elements: Al, Ca, K, Mg, Na, S, B, Ba, Co, Cu, La, Mo, Ni, P, Pb, Sr, V and Zn. The values of the median and Trimean were highest and tied for Ti and Be. The value of the Gastwirth median was highest for Fe. The value of the median was highest for Cr and Mn. The lowest values of the robust estimates of location was the median for the following elements: Al, K, Mg, Na, S, B, Ba, Co, La, Mo, Ni, P, Pb, V and Zn. The value of the 25 % trimmed mean was lowest for Ca. The value of the Trimean was lowest for Fe. The value of the Gastwirth median was lowest for Ti. The value of the 10 % trimmed mean was lowest for Be. The value of the 20 % trimmed mean was lowest for Cr. The values of the median and Gastwirth median were lowest and tied for Cu. The value of the 25 % trimmed mean was lowest for Mn. The value of the Gastwirth median was lowest for Sr. The second lowest value of the robust estimates of location was the Gastwirth median for the many elements including: Al, K, Mg, Na, B, Ba, Be, Ni, P and Pb. Table 7.8. presents robust estimates of location for elements in AZ1.

Table 7.8. Robust Estimates of Location for Elements in AZ1.

Core Element	AZ1					
	10%	20%	25%	Median	Gastwirth	Trimean
Al	2.146	2.108	2.088	2.030	2.073	2.114
Ca	5.467	4.973	4.754	4.760	4.856	5.344
K	0.957	0.924	0.908	0.830	0.896	0.911
Mg	3.064	2.732	2.595	2.070	2.508	2.711
Na	2.832	2.756	2.721	2.480	2.644	2.703
S	0.312	0.305	0.302	0.300	0.303	0.309
Fe	2.877	2.900	2.923	2.930	2.943	2.853
Ti	0.028	0.028	0.028	0.030	0.027	0.030
B	39.064	35.695	34.409	30.000	33.000	35.000
Ba	137.241	118.675	111.969	80.000	104.000	115.000
Be	1.068	1.073	1.077	1.100	1.070	1.100
Co	12.502	12.483	12.480	12.000	12.300	12.250
Cr	52.547	52.530	52.693	54.000	53.100	53.000
Cu	28.463	28.199	28.134	28.000	28.000	28.125
La	14.729	14.636	14.567	10.000	13.000	12.500
Mn	317.158	300.987	298.016	325.000	302.548	318.250
Mo	4.478	3.517	3.181	3.000	3.300	4.000
Ni	43.374	41.987	41.339	40.000	41.200	42.250
P	1255.813	1212.252	1185.433	1110.000	1152.960	1225.000
Pb	8.291	8.205	8.165	8.000	8.000	8.250
Sr	143.133	131.238	127.024	129.000	124.392	136.750
V	71.123	70.364	70.378	70.000	70.348	70.125
Zn	79.010	78.695	78.717	78.000	78.348	78.250

In AZ2, the values of the non-robust estimates of location for the major and trace elements decreased in the following order: midrange > arithmetic mean > geometric mean > harmonic mean. In AZ2, the highest and lowest values of the estimate of location for all major elements and all trace elements are the midrange and harmonic mean, respectively. Table 7.9. presents non-robust estimates of location for elements in AZ2.

Table 7.9. Non-Robust Estimates of Location for Elements in AZ2.

Element	Arithmetic Mean	Geometric Mean	Harmonic Mean	Midrange
Al	1.672	1.626	1.574	1.895
Ca	3.144	0.629	0.275	11.990
K	0.623	0.602	0.578	0.820
Mg	1.550	0.979	0.770	4.040
Na	2.697	2.026	1.125	3.070
S	0.566	0.404	0.290	1.890
Fe	3.022	2.676	2.339	4.510
Ti	0.021	0.019	0.018	0.030
B	14.172	13.296	12.563	25.000
Ba	49.080	42.262	38.801	205.000
Be	1.143	1.063	0.948	1.225
Co	11.755	10.722	9.058	17.500
Cr	50.865	47.696	44.832	65.500
Cu	35.724	34.352	33.061	50.500
La	14.785	13.674	12.839	45.000
Mn	191.920	126.656	102.633	747.000
Ni	39.877	37.460	35.459	57.000
P	1450.613	1105.015	833.851	5055.000
Pb	13.847	12.087	10.587	35.000
Sc	4.939	4.732	4.521	7.000
Sr	85.822	64.985	55.572	391.500
V	53.061	46.686	42.011	113.000
Zn	84.221	82.090	80.022	105.500

In AZ2, the values of the robust estimates of location for Ca, Mn, Ni, P and Sr decreased in the following order: 10 % trimmed mean > Trimean > 20 % trimmed mean > 25 % trimmed mean > Gastwirth median > median. The values of the robust estimates of location for Mg, S and V decreased in the following order: 10 % trimmed mean > 20 % trimmed mean > Trimean > 25 % trimmed mean > Gastwirth median > median. The values of the robust estimates of location for B and La decreased in the following order: 10 % trimmed mean > 20 % trimmed mean > 25

% trimmed mean > Gastwirth median > Trimean > median. For the following elements, the order of decreasing robust estimates of mean was unique in AZ2: Al, K, Na, Fe, Ti, Ba, Be, Co, Cr, Cu, Pb, Sc and Zn.

In AZ2, the highest values of the robust estimates of location was the 10 % trimmed mean for the following elements: Ca, Mg, S, Fe, B, Be, Cr, La, Mn, Ni, P, Pb, Sr and V. The value of the median was highest for Al. The value of the 25 % trimmed mean was highest for Na. The value of the Trimean was highest for K and Ti. The value of the Gastwirth median was highest for Ba. The value of the median was highest for Co and Cu. The values of the median, Gastwirth median and Trimean were tied and highest for Sc. The value of the 25 % trimmed mean was highest for Zn.

The lowest values of the robust estimates of location was the median for the following elements: Ca, Mg, S, B, Cr, La, Mn, Ni, P, Sr and V. The value of the 10 % trimmed mean was lowest for Al. The value of the 10 % trimmed mean was lowest for K. The value of the 10 % trimmed mean and Trimean were tied and lowest for Na. The value of the Gastwirth median was lowest for Fe. The values of the median and Gastwirth median were lowest and tied for Ti. The values of the 25 % trimmed mean were lowest for Co, Cu and Sc. The values of the median and Trimean were lowest and tied for Ba. The values of the median and Gastwirth median were lowest and tied for Be and Pb. The value of the 10 % trimmed mean was lowest for Zn. The second lowest value of the robust estimates of location was the Gastwirth median for many elements in AZ2, including: Ca, Mg, S, Cr, Mn, Ni, P, Sr and V. The second lowest estimates for K are tied values between the 25 % trimmed mean and the Gastwirth median. Table 7.10. presents robust estimates of location for elements in AZ2.

Table 7.10. Robust Estimates of Location for Elements in AZ2.

Core Element	AZ2					
	10%	20%	25%	Median	Gastwirth	Trimean
Al	1.684	1.706	1.718	1.760	1.715	1.705
Ca	1.715	0.816	0.478	0.280	0.308	1.100
K	0.626	0.628	0.627	0.630	0.627	0.633
Mg	1.101	0.750	0.682	0.660	0.666	0.728
Na	2.680	2.747	2.786	2.730	2.771	2.680
S	0.443	0.394	0.389	0.370	0.385	0.390
Fe	2.887	2.848	2.836	2.870	2.810	2.873
Ti	0.021	0.021	0.021	0.020	0.020	0.023
B	13.817	13.402	13.086	10.000	13.000	12.500
Ba	41.450	41.959	42.346	40.000	43.000	40.000
Be	1.142	1.140	1.137	1.100	1.100	1.125
Co	11.496	11.485	11.481	12.000	11.700	11.750
Cr	49.427	48.361	47.802	46.000	47.176	48.250
Cu	34.870	34.691	34.679	35.000	34.700	34.750
La	14.198	13.918	13.704	10.000	13.000	12.500
Mn	126.557	107.227	101.580	96.000	97.770	109.500
Ni	37.641	36.526	36.210	35.000	35.876	36.750
P	1251.908	1128.454	1059.012	990.000	1030.380	1173.750
Pb	12.580	12.237	12.099	12.000	12.000	12.250
Sc	4.885	4.897	4.877	5.000	5.000	5.000
Sr	64.740	58.351	55.494	50.000	55.304	59.750
V	48.092	45.536	44.691	42.000	43.800	45.250
Zn	83.214	84.216	84.420	84.000	84.276	83.750

**Estimates of Scale**

In AZ1, the values of the estimates of scale for Al, K, S, Fe, B, Ba, Co, Cu, Mo, Ni, P, Pb, Sr and V decreased in the following order: standard deviation > mean absolute deviation from the mean > mean absolute deviation from the median > semi-IQR > median absolute deviation from the median. The values of the estimates of scale for Ca and Na decreased in the following order: standard deviation > semi-IQR > mean absolute deviation from the mean >

mean absolute deviation from the median > median absolute deviation from the median. The values of the estimates of scale for Cr, Mn and Zn decreased in the following order: standard deviation > mean absolute deviation from the mean > mean absolute deviation from the median > median absolute deviation from the median > semi-IQR. The values of the estimates of scale for Mg decreased in the following order: standard deviation > mean absolute deviation from the mean > semi-IQR > mean absolute deviation from the median > median absolute deviation from the median.

The values of the estimates of scale for Ti contain the following two sets of tied values: mean absolute deviation from the mean and mean absolute deviation from the median, and the median absolute deviation from the median and semi-IQR, respectively. The highest estimate of scale for Ti is the standard deviation. The lowest estimate of scale for Ti is the median absolute deviation from the median and semi-IQR. The values of the estimates of scale for Be contain tied values between the median absolute deviation from the median and semi-IQR, in which both these values are the lowest. The highest value of estimates of scale for Be is the standard deviation followed by the mean absolute deviation from the mean and the mean absolute deviation from the median, respectively. The values of the estimates of scale for La contain tied values between the mean absolute deviation from the median and semi-IQR. The highest value of estimates of scale for La is the standard deviation followed by the mean absolute deviation from the mean. The lowest estimate of scale for La is the median absolute deviation from the median.

In AZ1, the highest estimate of scale for all elements is the standard deviation. In AZ1, the lowest estimate of scale for Al, Ca, K, Mg, Na, S, Fe, B, Ba, Co, Cu, La, Mo, Ni, P, Pb, Sr and V is the median absolute deviation from the median. The lowest estimate of scale for Ti and

Be is the median absolute deviation from the median and the semi-IQR, in which these values are tied. The lowest estimate of scale for Cr, Mn and Zn is the semi-IQR. Table 7.11. presents estimates of scale for elements in AZ1.

Table 7.11. Estimates of Scale for Elements in AZ1.

Element	Std.	Mean Abs. Dev. Mean	Mean Abs. Dev. Med.	Med. Abs. Dev. Med.	Semi-IQR
Al	0.740	0.602	0.591	0.460	0.538
Ca	6.040	5.472	5.423	4.590	5.673
K	0.441	0.375	0.352	0.260	0.343
Mg	2.970	2.617	2.474	1.340	2.553
Na	1.526	1.362	1.343	1.200	1.390
S	0.190	0.109	0.104	0.070	0.078
Fe	1.256	0.937	0.936	0.800	0.855
Ti	0.015	0.011	0.011	0.010	0.010
B	27.617	23.194	22.095	10.000	20.000
Ba	166.299	133.184	122.174	50.000	100.000
Be	0.383	0.277	0.276	0.200	0.200
Co	4.214	2.785	2.763	2.000	2.500
Cr	18.564	14.137	14.130	12.000	11.500
Cu	11.626	7.299	7.225	5.000	5.250
La	42.809	5.212	5.000	0.000	5.000
Mn	405.461	280.058	275.300	215.000	208.000
Mo	9.630	6.400	5.453	2.000	4.000
Ni	19.151	15.437	14.577	11.000	12.500
P	1393.810	845.559	802.648	500.000	630.000
Pb	42.809	8.909	6.344	1.000	1.500
Sr	990.288	289.142	225.482	74.000	87.500
V	42.797	26.349	25.316	18.000	18.250
Zn	21.070	17.073	17.032	16.000	15.500

In AZ2, the values of the estimates of scale of Ca, Mg, S, Co, Cu, La, Mn, Ni, P, Pb, Sr and V decreased in the following order: standard deviation > mean absolute deviation from the mean > mean absolute deviation from the median > semi-IQR > median absolute deviation from

the median. The values of the estimates of scale of Na, Fe, B and Cr decreased in the following order: standard deviation > semi-IQR > mean absolute deviation from the mean > mean absolute deviation from the median > median absolute deviation from the median. The values of the estimates of scale of Be and Zn decreased in the following order: standard deviation > mean absolute deviation from the mean > mean absolute deviation from the median > median absolute deviation from the median > semi-IQR. The values of the estimates of scale of Al decreased in the following order: semi-IQR > standard deviation > mean absolute deviation from the mean > mean absolute deviation from the median > median absolute deviation from the median.

The values of the estimates of scale of K is highest for the semi-IQR followed by the standard deviation. The values of the mean absolute deviation from the mean and the mean absolute deviation from the median contain tied values for K. The value of the median absolute deviation from the median is lowest for K. The values of the estimates of scale of Ti is highest for the semi-IQR and the median absolute deviation from the median, in which these values are tied. The value of the standard deviation is lower compared to these tied values, followed by the values of the mean absolute deviation from the mean and mean absolute deviation from the median, respectively. The values of the estimates of scale of Ba and Sc is lowest for the median absolute deviation from the median and semi-IQR, in which these values are tied. The highest values of the estimates of scale of Ba and Sc is the standard deviation, followed by the mean absolute deviation from the mean and the mean absolute deviation from the median.

In AZ2, the highest estimate of scale for Ca, Mg, Na, S, Fe, B, Ba, Be, Co, Cr, Cu, La, Mn, Ni, P, Pb, Sc, Sr, V and Zn is the standard deviation. The highest estimate of scale for Al and K is the semi-IQR. The highest estimate of scale for Ti is the semi-IQR and the median absolute deviation from the median, in which these values are tied. In AZ2, the lowest estimate



of scale for Al, Ca, K, Mg, Na, S, Fe, Ti, B, Co, Cr, Cu, La, Mn, Ni, P, Pb, Sr and V is the median absolute deviation from the median. The lowest estimate of scale for Be and Zn is the semi-IQR. The lowest estimate of scale for Ba and Sc is the median absolute deviation from the median and the semi-IQR, in which these values are tied. Table 7.12. presents estimates of scale for elements in AZ2.

Table 7.12. Estimates of Scale for Elements in AZ2.

Element	Std.	Mean Abs. Dev. Mean	Mean Abs. Dev. Med.	Med. Abs. Dev. Med.	Semi-IQR
Al	0.372	0.289	0.282	0.190	0.880
Ca	5.738	4.164	2.959	0.130	0.140
K	0.158	0.121	0.121	0.090	0.315
Mg	1.871	1.396	0.968	0.090	0.330
Na	1.520	1.249	1.248	1.130	1.365
S	0.563	0.370	0.329	0.160	0.185
Fe	1.470	1.172	1.165	1.120	1.435
Ti	0.008	0.006	0.005	0.010	0.010
B	5.307	4.965	4.172	0.000	5.000
Ba	44.870	17.889	15.828	10.000	10.000
Be	0.383	0.285	0.284	0.200	0.150
Co	4.730	3.204	3.202	2.000	2.500
Cr	18.489	16.259	15.908	14.000	16.500
Cu	10.394	7.754	7.706	6.000	6.500
La	7.230	5.402	4.785	0.000	5.000
Mn	256.609	153.767	116.325	26.000	45.000
Ni	15.290	11.791	11.172	8.000	8.500
P	1257.753	877.039	764.663	330.000	622.500
Pb	8.306	5.466	5.147	3.000	3.500
Sc	1.443	1.149	1.141	1.000	1.000
Sr	94.095	54.896	44.092	11.000	26.500
V	30.515	22.737	21.491	14.000	17.500
Zn	19.582	14.406	14.405	12.000	11.500

## Correlation

A Pearson's correlation was performed to determine the relationship between the variables. There was a strong correlation (0.7-1.0) between twenty one variable pairs in AZ1.

Figure 7.1. presents the Pearson's correlation matrix of AZ1.

The coefficient of determination,  $r^2$ , indicates the amount of shared variance between the variables. For the strongly correlated variables in AZ1, the value of  $r^2$  was the highest for Al and K followed by Ca and Mg. The  $r^2$  value of Al and K is the largest ( $r = 0.869$ ) and indicates that 87 % of the variance is shared between these two variables. The  $r^2$  value is second highest ( $r = 0.810$ ) for Ca and Mg ( $r = 0.810$ ), indicating that 81 % of the variance is shared between these two variables. Figure 7.1. presents the Pearson's correlation matrix of AZ1.

A Spearman's Rank correlation was performed to determine the relationships between the variables. There was a strong correlation between 26 variable pairs in AZ1. For the strongly correlated variables in AZ1, the value of  $r$  was the highest for Al and K ( $r = 0.929$ ,  $p < 0.001$ ) followed by Ca and Mg ( $r = 0.876$ ,  $p < 0.001$ ). Figure 7.2. presents the Spearman's correlation matrix of AZ1.

	Al	Ca	K	Mg	Na	S	Fe	Ti	B	Ba	Be	Co	Cr	Cu	Mn	Mo	Ni	P	Pb	Sr	V	Zn
Al	1.000																					
Ca	0.092	1.000																				
K	0.932*	0.315	1.000																			
Mg	-0.048	0.900*	0.208	1.000																		
Na	0.048	-0.847*	-0.172	-0.801*	1.000																	
S	-0.266	-0.514	-0.361	-0.459	0.563	1.000																
Fe	0.659	-0.476	0.476	-0.539	0.504	0.050	1.000															
Ti	0.695	0.381	0.757*	0.289	-0.375	-0.406	0.383	1.000														
B	0.648	0.596	0.826*	0.554	-0.405	-0.401	0.060	0.598	1.000													
Ba	0.432	0.637	0.592	0.550	-0.573	-0.346	-0.075	0.569	0.665	1.000												
Be	0.662	-0.500	0.437	-0.596	0.555	0.120	0.842*	0.272	0.056	-0.092	1.000											
Co	0.530	0.235	0.483	0.208	-0.083	-0.186	0.397	0.425	0.379	0.308	0.394	1.000										
Cr	0.570*	-0.017	0.801*	-0.089	0.130	-0.090	0.623	0.706*	0.543	0.361	0.666	0.479	1.000									
Cu	0.572	-0.029	0.485	-0.127	0.150	0.028	0.379	0.306	0.300	0.187	0.513	0.505	0.591	1.000								
Mn	0.021	0.886*	0.234	0.887*	-0.795*	-0.575	-0.375	0.383	0.516	0.537	-0.890	0.525	-0.094	-0.163	1.000							
Mo	0.082	0.534	0.226	0.529	-0.435	-0.137	-0.108	0.304	0.430	0.408	-0.187	0.375	0.162	0.099	0.533	1.000						
Ni	0.760*	0.452	0.801*	0.376	-0.270	-0.252	0.273	0.625	0.737*	0.585	0.340	0.734*	0.730*	0.615	0.328	0.480	1.000					
P	0.384	0.592	0.445	0.446	-0.330	-0.119	-0.011	0.361	0.550	0.459	0.080	0.487	0.370	0.332	0.436	0.318	0.671	1.000				
Pb	-0.128	-0.012	-0.206	-0.036	0.037	0.278	0.013	-0.064	-0.307	-0.077	0.042	0.275	-0.026	0.163	-0.035	0.301	0.084	0.176	1.000			
Sr	0.165	0.750*	0.342	0.604	-0.757*	-0.270	-0.303	0.429	0.494	0.644	-0.361	0.067	0.043	0.009	0.620	0.289	0.339	0.418	-0.118	1.000		
V	0.684	0.050	0.671	-0.005	0.013	-0.317	0.570	0.695	0.426	0.324	0.429	0.332	0.666	0.385	0.041	0.168	0.465	0.151	-0.129	0.140	1.000	
Zn	0.684	-0.032	0.594	-0.071	0.149	-0.234	0.609	0.466	0.402	0.219	0.648	0.608	0.727*	0.487	0.004	0.315	0.627	0.375	-0.098	-0.098	0.503	1.000

Figure 7.1. Pearson's Correlation Matrix of AZ1 Variables

	Al	Ca	K	Mg	Na	S	Fe	Ti	B	Ba	Be	Co	Cr	Cu	Mn	Mo	Ni	P	Pb	Sr	V	Zn
Al	1.000																					
Ca	0.057	1.000																				
K	0.929*	0.232	1.000																			
Mg	0.021	0.876*	0.232	1.000																		
Na	-0.131	-0.847*	-0.281	-0.808*	1.000																	
S	-0.292	-0.626	-0.398	-0.547	0.754*	1.000																
Fe	0.547	-0.578	0.404	-0.527	0.437	0.112	1.000															
Ti	0.754*	0.261	0.798*	0.243	-0.428	-0.463	0.333	1.000														
B	0.662	0.530	0.813*	0.563	-0.467	-0.438	-0.024	0.621	1.000													
Ba	0.516	0.540	0.614	0.515	-0.588	-0.434	-0.069	0.581	0.651	1.000												
Be	0.546	0.548	0.364	-0.556	0.467	0.171	0.805*	0.240	0.065	-0.063	1.000											
Co	0.582	0.198	0.550	0.263	-0.244	-0.271	0.350	0.475	0.460	0.413	0.374	1.000										
Cr	0.867*	-0.061	0.823*	-0.006	-0.059	-0.174	0.537	0.760*	0.381	0.463	0.530	0.527	1.000									
Cu	0.523	-0.017	0.452	-0.059	0.033	0.001	0.247	0.366	0.368	0.219	0.379	0.453	0.540	1.000								
Mn	0.168	0.821*	0.326	0.860*	-0.822*	-0.629	-0.324	0.413	0.557	0.549	-0.413	0.396	0.088	-0.056	1.000							
Mo	0.117	0.475	0.258	0.593	-0.492	-0.265	-0.125	0.312	0.457	0.426	-0.193	0.383	0.218	0.111	0.563	1.000						
Ni	0.759*	0.398	0.806*	0.427	-0.417	-0.312	0.147	0.663	0.753*	0.622	0.219	0.743*	0.708*	0.574	0.424	0.486	1.000					
P	0.474	0.490	0.531	0.503	-0.387	-0.143	-0.028	0.405	0.629	0.512	0.067	0.523	0.490	0.405	0.436	0.580	0.768*	1.000				
Pb	-0.116	-0.129	-0.186	-0.149	0.062	0.091	0.120	-0.070	-0.332	-0.141	0.138	0.172	-0.082	0.160	-0.107	0.154	-0.006	0.039	1.000			
Sr	0.345	0.838*	0.472	0.734*	-0.779*	-0.516	-0.330	0.499	0.618	0.688	-0.341	0.291	0.199	0.155	0.725*	0.428	0.549	0.528	-0.077	1.000		
V	0.626	0.037	0.686	0.032	-0.148	-0.345	0.495	0.761*	0.448	0.568	0.393	0.347	0.675	0.325	0.133	0.189	0.456	0.185	-0.014	0.320	1.000	
Zn	0.601	-0.097	0.554	-0.022	-0.011	-0.184	0.567	0.449	0.382	0.300	0.588	0.612	0.693	0.388	0.094	0.346	0.546	0.433	0.186	0.061	0.450	1.000

Figure 7.2. Spearman's Correlation Matrix of AZ1 Variables

A Pearson's correlation was performed to determine the relationship between the variables. There was a strong correlation (0.7-1.0) between 18 variable pairs in AZ2. Figure 7.3. presents the Pearson's correlation matrix of AZ2.

For the strongly correlated variables in AZ2, the value of  $r^2$  decreases in the following order: Al and K, Ca and Sr, Ca and Na. The  $r^2$  value of Al and K is the largest ( $r^2 = 0.790$ ) and indicates that 79 % of the variance is shared between these two variables. The  $r^2$  values of Ca and Sr, Ca and Na, and Al and Sc are 0.767, 0.755, 0.706, respectively, and correspond to 77, 76, and 71 % shared variance, respectively.

A Spearman's Rank correlation was performed to determine the relationship between the variables. There was a strong correlation between 12 variables pairs. For the strongly correlated variables in AZ2, the value of  $r$  was the highest for Al and K ( $r = 0.855$ ,  $p < 0.001$ ), followed by Cr and Sc ( $r = 0.836$ ,  $p < 0.001$ ). Figure 7.4. presents the Spearman's correlation matrix of AZ2.

	Al	Ca	K	Mg	Na	S	Fe	Ti	Ba	Be	Co	Cr	Cu	Mn	Ni	P	Pb	Se	Sr	V	Zn
Al	1.000																				
Ca	-0.366	1.000																			
K	0.889*	-0.514	1.000																		
Mg	-0.109	0.692	-0.140	1.000																	
Na	0.568	-0.869*	0.606	-0.557	1.000																
S	0.294	-0.061	0.402	0.131	0.179	1.000															
Fe	0.677	-0.593	0.708*	-0.293	0.682	0.356	1.000														
Ti	0.480	-0.247	0.551	-0.070	0.381	0.365	0.407	1.000													
Ba	0.232	-0.108	0.156	-0.503	0.151	-0.104	0.129	0.242	1.000												
Be	0.752*	-0.649	0.756*	-0.271	0.798*	0.203	0.746*	0.437	0.108	1.000											
Co	0.551	-0.471	0.607	-0.011	0.532	0.385	0.493	0.414	-0.147	0.726*	1.000										
Cr	0.645	-0.195	0.604	-0.146	0.320	0.064	0.455	0.674	0.295	0.413	0.150	1.000									
Cu	0.120	-0.058	0.123	0.105	0.060	0.104	0.028	-0.195	-0.207	0.191	0.467	-0.224	1.000								
Mn	-0.194	0.822*	-0.278	0.833*	-0.672	0.119	-0.280	-0.042	-0.280	-0.404	-0.212	-0.109	-0.071	1.000							
Ni	0.086	0.420	0.102	0.562	-0.330	0.272	-0.145	0.196	-0.226	0.045	0.475	-0.015	0.432	0.496	1.000						
P	-0.029	0.765*	-0.122	0.503	-0.577	0.127	-0.324	0.052	0.078	-0.318	-0.127	-0.013	0.051	0.592	0.550	1.000					
Pb	0.368	-0.543	0.451	-0.309	0.641	0.300	0.402	0.437	0.125	0.555	0.680	0.093	0.317	-0.348	0.146	-0.198	1.000				
Se	0.840*	-0.429	0.725*	-0.314	0.629	0.121	0.650	0.608	0.354	0.692	0.402	0.790*	-0.105	-0.261	-0.111	-0.134	0.378	1.000			
Sr	-0.286	0.876*	-0.381	0.426	-0.796*	0.013	-0.500	-0.153	0.133	-0.620	-0.538	-0.096	-0.101	0.665	0.288	0.763*	-0.418	-0.291	1.000		
V	0.590	-0.408	0.514	-0.414	0.548	0.047	0.675	0.586	0.365	0.509	0.156	0.815*	-0.238	-0.253	-0.265	-0.244	0.290	0.829*	0.829*	1.000	
Zn	0.603	-0.058	0.630	0.095	0.162	0.251	0.322	0.679	0.147	0.471	0.525	0.675	0.188	0.026	0.538	0.270	0.276	0.528	-0.033	0.387	1.000

Figure 7.3. Pearson's Correlation Matrix of AZ2 Variables.

	Al	Ca	K	Mg	Na	S	Fe	Ti	Ba	Be	Co	Cr	Cu	Mn	Ni	P	Pb	Se	Sr	V	Zn
Al	1.000																				
Ca	-0.275	1.000																			
K	0.855*	-0.421	1.000																		
Mg	0.025	0.435	0.084	1.000																	
Na	0.369	-0.627	0.392	-0.531	1.000																
S	0.241	-0.088	0.402	0.076	0.131	1.000															
Fe	0.622	-0.566	0.642	-0.199	0.642	0.311	1.000														
Ti	0.462	-0.246	0.564	-0.033	0.414	0.397	0.441	1.000													
Ba	0.376	-0.211	0.335	-0.513	0.372	0.053	0.340	0.402	1.000												
Be	0.577	-0.431	0.558	-0.156	0.607	0.130	0.689	0.454	0.405	1.000											
Co	0.309	-0.153	0.398	0.245	0.144	0.274	0.353	0.342	0.018	0.482	1.000										
Cr	0.703*	-0.219	0.696	-0.158	0.415	0.175	0.480	0.684	0.485	0.450	0.088	1.000									
Cu	-0.035	0.030	-0.038	0.170	-0.136	0.020	-0.031	-0.246	-0.297	0.024	0.438	-0.276	1.000								
Mn	0.026	0.665	0.000	0.539	-0.526	0.207	-0.149	0.194	-0.128	-0.238	0.034	0.139	-0.096	1.000							
Ni	-0.043	0.460	0.007	0.657	-0.478	0.212	-0.182	0.144	-0.320	-0.027	0.586	-0.054	0.403	0.610	1.000						
P	-0.033	0.804*	-0.109	0.493	-0.642	0.107	-0.352	-0.090	-0.179	-0.316	0.016	-0.072	0.110	0.702*	0.550	1.000					
Pb	0.133	-0.386	0.247	-0.140	0.416	0.288	0.312	0.403	0.170	0.396	0.632	0.040	0.288	-0.219	0.184	-0.266	1.000				
Se	0.789*	-0.395	0.705*	-0.277	0.643	0.136	0.648	0.664	0.522	0.645	0.224	0.836*	-0.210	-0.073	-0.194	-0.258	0.251	1.000			
Sr	-0.129	0.740*	-0.148	0.369	-0.569	0.191	-0.319	0.015	-0.061	-0.353	-0.165	0.031	-0.018	0.778*	0.439	0.797*	-0.236	-0.193	1.000		
V	0.672	-0.425	0.597	-0.467	0.680	0.127	0.696	0.582	0.541	0.531	0.042	0.823*	-0.255	-0.094	-0.310	-0.337	0.163	0.875*	-0.225	1.000	
Zn	0.555	-0.036	0.630	0.221	0.133	0.303	0.339	0.698	0.249	0.492	0.563	0.652	0.111	0.300	0.486	0.165	0.268	0.535	0.107	0.401	1.000

Figure 7.4. Spearman's Correlation Matrix of AZ2 Variables.

## **Hierarchical Cluster Analysis**

In AZ1, HCA of the major elements resulted in three main groups. Group 1 contains Al, K, Ti and Fe. In group 1, Al and K are grouped together first, followed by Ti and then Fe. Group 2 contains Ca and Mg, which are located close together. Group 3 contains Na and S, which are located close together. In AZ1, HCA of the trace elements resulted in three main groups. Group 1 contains B, Ba, Mo, Mn, Ni, P and Sr. Group 1 can be divided into two subgroups. In the first subgroup, B and Ni are grouped together first, followed by P and then Mo. In the second subgroup, Ba and Sr are grouped together first, followed by Mn. Group 2 contains Be, Co, Cr, Cu, Pb, V and Zn. Group 2 can be divided into two subgroups. In the first subgroup, Co and Cu are grouped together. In the second subgroup, Cr and Zn are grouped together first, followed by Be and then V. Group 3 contains Pb. Figures 7.5. and 7.6. present the HCA dendograms of major and trace elements in AZ1, respectively.



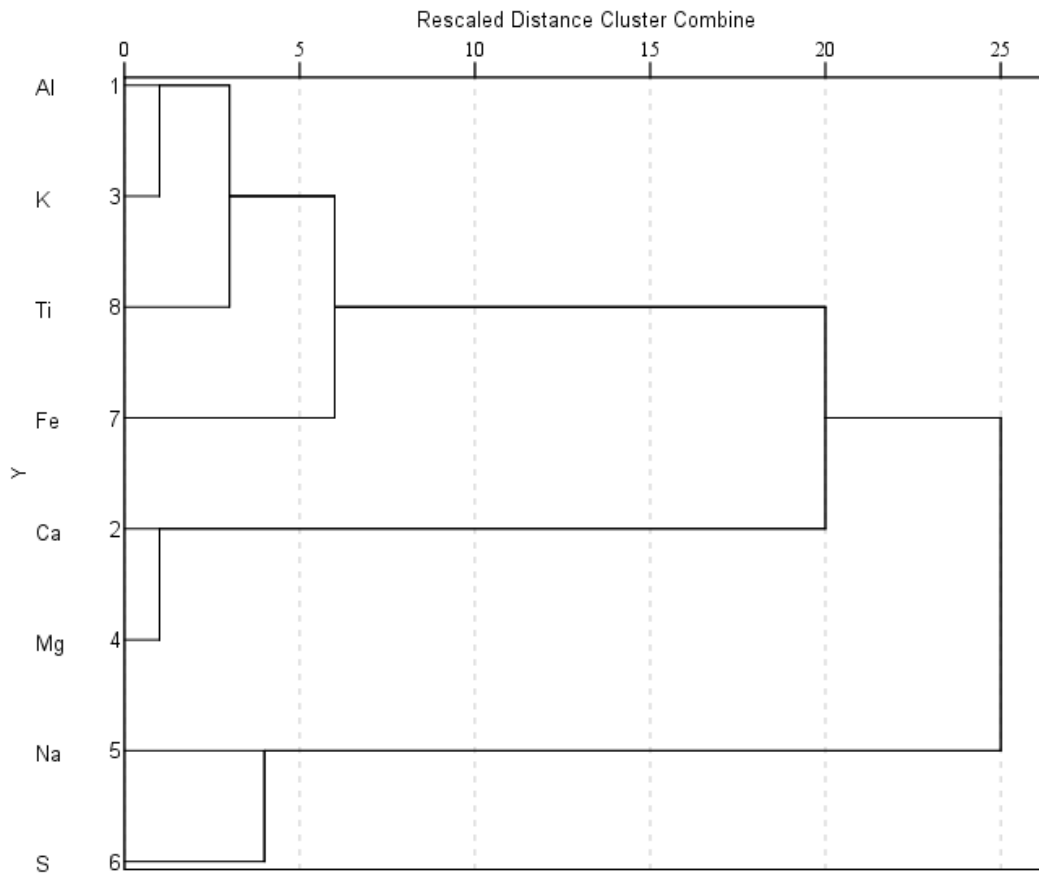


Figure 7.5. HCA Dendrogram of Major Elements in AZ1.

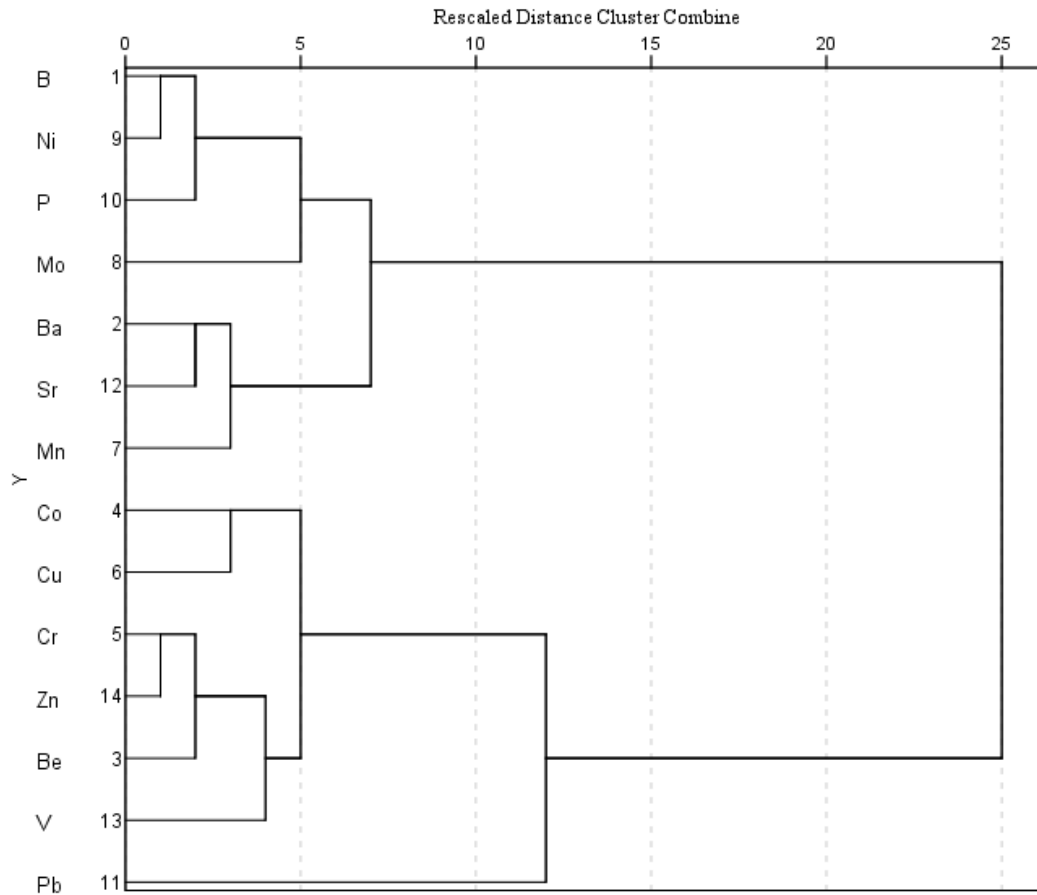


Figure 7.6. HCA Dendrogram of Trace Elements in AZ1.

In AZ2, HCA of the major elements resulted in two main groups. Group 1 contains Al, K, Na, Fe, S and Ti. Group 1 can be divided into two subgroups. In subgroup 1, Al and K are grouped together first, followed by Na and Fe. In subgroup 2, S and Ti are located close together. Group 2 contains Ca and Mg, which are located close together. In AZ2, HCA of the trace elements resulted in three main groups. Group 1 contains Ba, Cr, Sc and V. Group 1 can be divided into two subgroups. In subgroup 1, Sc, V and Cr are located close together. Subgroup 2 contains Ba. Group 2 contains Be, Co, Cu, Ni, Pb and Zn. Group 2 can be divided into two subgroups. In subgroup 1, Be and Co are grouped together first, followed by Pb. In subgroup 2, Ni and Zn are grouped together first, followed by Cu. Group 3 contains Mn, P and Sr. In group 3,

P and Sr are grouped together first, followed by Mn. Figures 7.7. and 7.8. present the HCA dendograms of major and trace elements in AZ1, respectively.

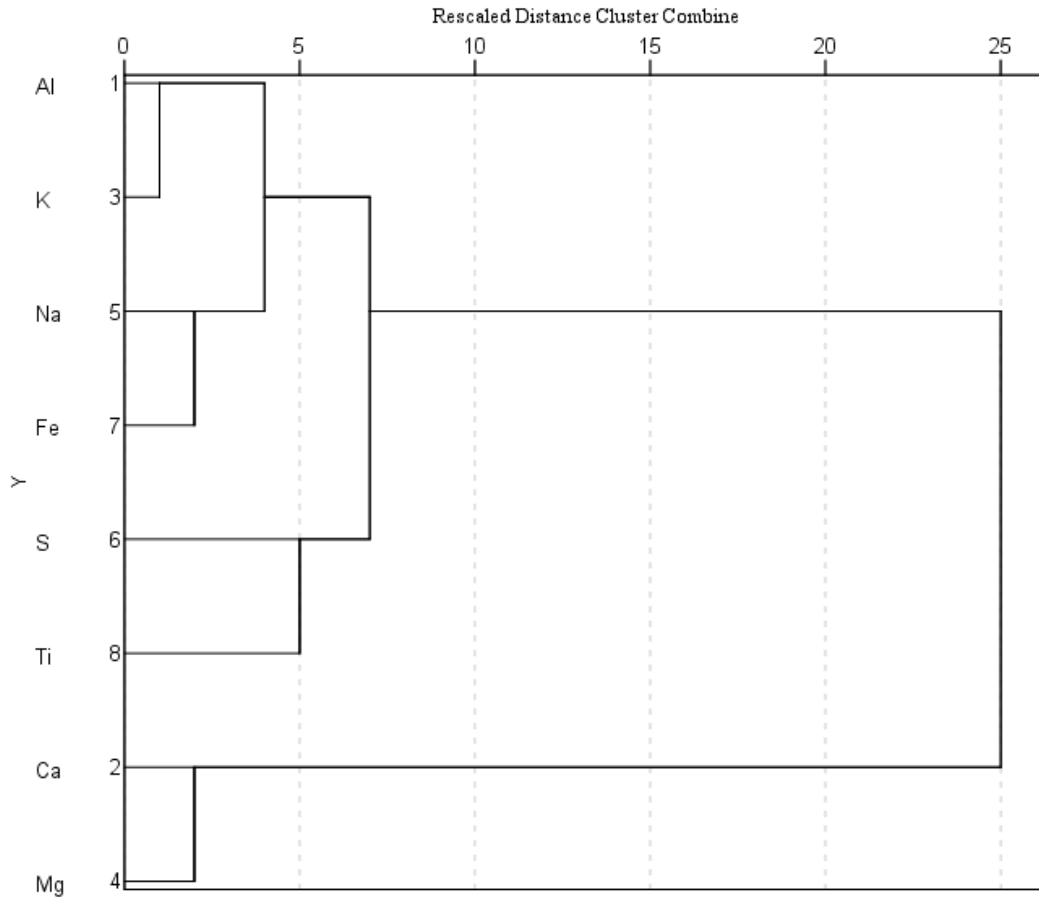


Figure 7.7. HCA Dendrogram of Major Elements in AZ2.

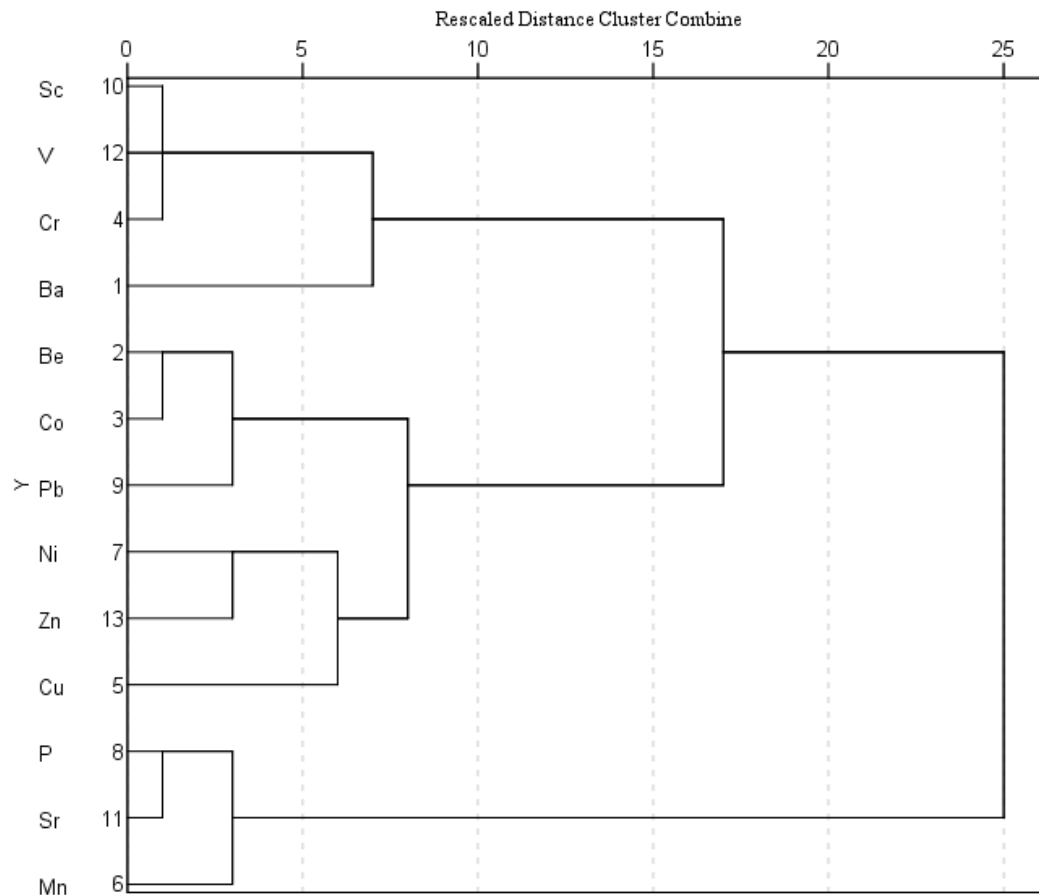


Figure 7.8. HCA Dendrogram of Trace Elements in AZ2.

### Principal Component Analysis

In AZ1, PCA of the variables extracted four components. A high factor loading ( $>0.5$ ) indicates that the variable is well explained (Reimann et al., 2002). The four components account for 79.414 % of the total variance. Factors 1 and 2 explain 65.816 % of the total variance. Factor 1 explains 38.623 % of the total variance. Factor 1 is dominated by the following elements: Al, K, Fe, Ti, Be, Cr, Cu, Ni, V and Zn. Factor 2 explains 27.193 % of the total variance. Factor 2 is dominated by the following elements: Ca, Mg, Na, B, Ba, Mn, P and Sr. Factor 3 explains 8.777 % of the total variance. Factor 3 is dominated by Pb, Co and Mo. Factor 4 explains 4.820 % of

the total variance. Factor 4 is dominated by S. Tables 7.13. and 7.14. present PCA factors and eigenvalues and PCA factor loadings, respectively, for the variables of AZ1.

Table 7.13. PCA Factors and Eigenvalues of AZ1 Variables.

Core	Variables	Factor	Eigenvalue	% Var.	% Cumulative Var.
AZ1	22	1	8.497	38.623	38.623
		2	5.982	27.193	65.816
		3	1.931	8.777	74.593
		4	1.060	4.820	79.414

Table 7.14. PCA Factor Loadings of AZ1 Variables.

Variable	AZ1			
	Factor 1	Factor 2	Factor 3	Factor 4
Al	0.963			
Cr	0.924			
K	0.901	0.325	-0.109	
Ni	0.765	0.431	0.330	0.139
Zn	0.762	-0.113	0.384	-0.183
V	0.756			-0.288
Ti	0.727	0.378		-0.245
Be	0.724	-0.568		
Fe	0.721	-0.540		-0.137
Cu	0.638		0.254	0.359
Ca		0.953	0.133	-0.113
Mg		0.906	0.167	-0.192
Na	0.108	-0.877		0.250
Mn		0.846	0.194	-0.381
Sr	0.104	0.823	-0.176	0.120
Ba	0.415	0.716		
B	0.620	0.646		
P	0.378	0.544	0.425	0.285
Pb	-0.130	-0.110	0.788	0.161
Co	0.568	0.141	0.614	-0.105
Mo	0.128	0.521	0.599	
S	-0.222	-0.428	0.111	0.745

In AZ2, PCA of the variables extracted four components. The four components account for 78.545 % of the total variance. Factors 1 and 2 explain 59.692 % of the total variance. Factor 1 explains 40.546 % of the variance. Factor 1 is dominated by the following elements: Al, K, Fe, Ti, Be, Cr, Sc, V and Zn. Factor 2 explains 19.146 % of the variance. Factor 2 is dominated by the following elements: Ca, Na, Mn, P and Sr. Factor 3 explains 12.929 % of the total variance. Factor 3 is dominated by the following elements: S, Co, Cu, Ni and Pb. Factor 4 explains 5.924 % of the total variance. Factor 4 is dominated by the following elements: Mg and Ba. Tables 7.15. and 7.16. present PCA factors and eigenvalues and PCA factor loadings, respectively, of AZ2 variables.

Table 7.15. PCA Factors and Eigenvalues of AZ2 Variables.

Core	Variables	Factor	Eigenvalue	% Var.	% Cumulative Var.
AZ2	21	1	8.515	40.546	40.546
		2	4.021	19.146	59.692
		3	2.715	12.929	72.621
		4	1.244	5.924	78.545

Table 7.16. PCA Factor Loadings of AZ2 Variables.

AZ2				
Variable	Factor 1	Factor 2	Factor 3	Factor 4
Cr	0.885		-0.193	-0.138
Sc	0.882	-0.239		-0.172
Al	0.832	-0.179	0.240	
V	0.805	-0.265	-0.252	-0.216
K	0.792	-0.274	0.314	
Ti	0.779		0.138	
Zn	0.725	0.253	0.430	
Fe	0.676	-0.490		0.141
Be	0.629	-0.521	0.382	
Ca	-0.248	0.920	-0.149	0.137
Sr	-0.142	0.891	-0.172	-0.138
P		0.872	0.180	
Na	0.456	-0.805	0.204	
Mn		0.777		0.480
Co	0.393	-0.317	0.777	0.217
Cu	-0.185		0.774	
Ni	0.101	0.577	0.676	0.249
Pb	0.287	-0.392	0.627	-0.175
S	0.318		0.340	0.324
Ba	0.287			-0.838
Mg		0.634		0.680





## 8. AZ1 & AZ2 STATISTICAL DISCUSSION

### **Geochemical Data Distribution**

Ahrens (1953), Rock (1988) and Reimann and Filzmoser (2000) stated that geochemical data is commonly not normally or lognormally distributed; and overall the majority of elements in AZ1 and AZ2 are neither normal nor lognormal. In AZ1 and AZ2, one element and one ln transformed element is considered normally distributed according to the following visual and statistical methods: histogram, Q-Q plot, critical values of the skewness and kurtosis statistic and the results of the Shapiro-Wilk test. In AZ1, Zn is normally distributed. In AZ2, ln transformed Cu is normally distributed.

In AZ1, the critical values of the skewness and kurtosis statistics indicated that Be follows a normal distribution; however, the histogram, Q-Q plot and Shapiro-Wilk test did not indicate this. In AZ1, the critical values of the skewness and kurtosis statistics indicated that ln transformed K and ln transformed Ni follow a normal distribution; however, the histogram, Q-Q plot and Shapiro-Wilk test did not indicate this. In AZ2, the critical values of the skewness and kurtosis statistics indicated that ln transformed Fe, P, Sc and Zn follow a normal distribution; however, the histogram, Q-Q plot and Shapiro-Wilk test did not indicate this. In AZ2, the critical values of the skewness and kurtosis statistics indicated that Be follows a normal distribution; however, the histogram, Q-Q plot and Shapiro-Wilk test did not indicate this. Rock (1988) stated that geochemical data is commonly skewed right and leptokurtic; and overall the majority of elements in AZ1 and AZ2 are skewed right and leptokurtic.

## **Performance of Location and Scale Estimates**

### *Non-Robust Location Estimates*

Non-robust estimates of location, such as the arithmetic, geometric, harmonic means, and the midrange are sensitive to outliers, extreme values, censored values and skewed data distributions (Rock, 1988). Non-robust estimates of location used with geochemical data of AZ1 and AZ2 reflect the influence of outliers, extreme values, censored values and skewed data distributions. In AZ1 and AZ2, the values of the non-robust estimates of location for all major and trace elements, with the exception of La in AZ1, decrease in the following order: midrange > arithmetic mean > geometric mean > harmonic mean. The absolute order of the non-robust estimates of location used with the majority of geochemical data of AZ1 and AZ2 decrease with increasing order of robustness. In AZ1, the non-robust estimates of location for the trace element of La decrease in the following order: midrange > geometric mean > harmonic mean > arithmetic mean. The difference observed in the absolute order of non-robust estimates of location for La in AZ1 is attributed to outliers, censored values and skewed data distribution. The distribution of La in AZ1 contains 1 outlier value, 3 censored values and is skewed right. The distribution of La contains a large proportion of duplicate data values, which perhaps also influences the absolute order of non-robust estimates of location.

The midrange contained the highest values primarily due to the presence of outliers in the geochemical data of AZ1 and AZ2, as Rock (1988) stated that the midrange is entirely non-robust to outliers. Significant differences are observed between the two most non-robust estimates of location, the midrange and arithmetic mean. The arithmetic mean contains the second highest values, with the exception of La in AZ1, primarily due to the presence of outliers and skewed data distributions in the geochemical data of AZ1 and AZ2, as Rock (1988) stated

that the arithmetic mean is an overestimate that is affected by the presence of large data outliers and skewed data distributions. The geometric mean contains lower values, with the exception of La in AZ1, compared to the arithmetic mean, as the geometric mean is less affected by skewed distributions than the arithmetic mean (Rock, 1988). The harmonic mean contains the lowest estimates, with the exception of La in AZ1, as the harmonic mean is less sensitive to non-normal data distributions, outliers, extreme values and censored values compared to the midrange, arithmetic and geometric means (Rock, 1988).

### *Robust Location Estimates*

Robust estimates of location are less sensitive to outliers, extreme values, censored values and skewed data distributions (Rock, 1988). Robust estimates of location used with geochemical data of AZ1 and AZ2 reflect the decreased influence of outliers, extreme values, censored values and skewed data distributions. In AZ1 and AZ2, the values of the robust estimates of location for all major and trace elements, did not decrease in a general order such as what is observed when the non-robust estimates of location are used. According to Rock (1988), the following five factors should be taken into account when looking at the differences between robust estimates of location of real geochemical data, as these characteristics of a set of data could influence the results: (1) distribution type, (2) modality, (3) continuity, (4) number of outliers, (5) size of data set and (6) source of error.

For the majority of elements in AZ1 and AZ2, the value of the 10 % trimmed mean is the highest. The value of the 50 % trimmed mean, also known as the median is the lowest for the majority of elements in AZ1 and AZ2, primarily due to the robustness of this estimate against extreme values and the decreased influence of outliers and skewed data distributions when compared to the arithmetic mean (Reimann and Filzmoser, 2000; Reimann et al., 2005).

Rock (1988) prefers robust estimates of location over the median. Reimann and Filzmoser (2000) stated that estimates of the median can be erroneous since the relationships between the variables is not taken into account. Estimates of the median are highest in several trace elements of AZ1 and AZ2, as this can be attributed to the presence of outliers, extreme values, censored values and skewed data distributions (Rock, 1988). Rock (1988) stated that robust estimates of location are higher in accuracy and precision, and use more information in the data compared to the median. In AZ1 and AZ2, the values of the robust estimates of location such as the trimmed means and combined L estimates are larger in value compared to the median for the majority of elements. The values of the trimmed means and combined L estimates are preferred estimates of location when using real geochemical data according to Rock (1988).

The absolute order of location estimates with the majority of elements in AZ1 and AZ2 is similar to the common absolute order of location estimates determined by Rock (1988) with real geochemical data and is the following: midrange  $\gg$  arithmetic mean  $>$  geometric mean  $\geq$  robust  $\geq$  harmonic mean. In AZ1, Zn is considered normally distributed according to histogram and Q-Q plots, the critical values of the skewness and kurtosis statistics and the results of the Shapiro-Wilk test; however, there is not a clear distinguishable difference between the absolute order of estimates of location compared to those determined by Rock (1988). In AZ1, Zn is skewed right, unimodal, leptokurtic, contains 2 outliers and contains no censored values.

Rock (1988) states that when the number and intensity of outliers increases in a data distribution, the observable differences between non-robust estimates of location will be larger compared to observable differences between robust estimates of location. The observable differences between the non-robust estimates of location are larger compared to the observable differences between robust estimates of location for all geochemical data in AZ1 and AZ2. Rock

(1988) stated that skewed and leptokurtic data distributions yield larger values of non-robust estimates of location compared to robust estimates of location. The non-robust estimates of location for those elements in AZ1 and AZ2 that are skewed and leptokurtic are larger compared to robust estimates of location.

### *Scale Estimates*

The absolute order of scale estimates with the majority of elements in AZ1 and AZ2 is the following: standard deviation > mean absolute deviation from the mean > mean absolute deviation from the median > semi-IQR > median absolute deviation from the median. The absolute order of scale estimates determined by Rock (1988) with geochemical data is the following: standard deviation > mean absolute deviation from the mean > mean absolute deviation from the median >> median absolute deviation from the median = semi-IQR. According to Rock (1988), the standard deviation, mean absolute deviation from the mean and the mean absolute deviation from the median provide larger estimates of scale compared to the semi-IQR or median absolute deviation from the median, as this is demonstrated by the geochemical data of AZ1 and AZ2. For the majority of elements in AZ1 and AZ2, the value of the semi-IQR is larger than and not equal to the value of the median absolute deviation from the median resulting in a slightly different absolute order of estimates of scale than the one determined by Rock (1988) with real geochemical data. Reimann and Filzmoser (2000) consider the median absolute deviation from the median and the semi-IQR to be preferred estimates of scale. Reimann et al. (2005) stated that the median absolute deviation from the median is considered a suitable replacement for the standard deviation, as this estimate is robust against extreme values.

The value of the standard deviation is highest for all the elements in AZ1 and the

majority of the elements in AZ2. In AZ1, the highest estimate of scale is the standard deviation. In AZ2, the highest estimate of scale with the majority of elements is the standard deviation. Rock (1988) stated that the standard deviation is an overestimate and considered the least robust estimate of scale after the range. The standard deviation is sensitive to censored values, outliers and data values from other populations (Reimann and Filzmoser, 2000). The standard deviation is also influenced by extreme data values (Reimann et al., 2005).

### **Mutually Consistent Estimates of Location and Scale**

The determination of the mutual consistency of robust estimates of location and scale should be accompanied by graphical summaries of the geochemical data when making the decision as to whether or not summary statistics are meaningful (Rock, 1988). By inspecting the mutual consistency of estimates of location and scale while providing visual summaries of the geochemical data, such as a histogram or box plot, will influence the decision on whether or not summary statistics are meaningful with the geochemical data of AZ1 and AZ2. The best estimate of location is determined considering those in which are most consistent, as this is a widely used method by Rock (1988). According to Rock (1988), robust estimates are smaller and more consistent; therefore, this criterion is used to determine the best estimate of scale.

#### *Non-Robust Location Estimates*

The midrange contains the largest values for the entirety of real geochemical data in AZ1 and AZ2, and is considered to be the least consistent estimate of location compared to the other non-robust estimates including the arithmetic, geometric and harmonic means. The harmonic mean contains the smallest values for the majority of the real geochemical data in AZ1 and AZ2, and is considered to be the most consistent non-robust estimate of location compared to the other non-robust estimates including the arithmetic and geometric means and the midrange. Overall,

the four non-robust estimates of location can be ranked in the following order of mutual consistency, which corresponds to decreasing order of robustness, as demonstrated by the use of these estimates with geochemical data from AZ1 and AZ2: harmonic mean > geometric mean > arithmetic mean > midrange.

### *Robust Location Estimates*

In AZ1 and AZ2, the values of the robust estimates of location for the majority of elements is highest for the 10 % trimmed mean. The 10 % trimmed mean is considered a robust estimate of location; however, due to the skewed nature and outliers present in the majority of elements, this is likely not an optimal robust estimate of location. In AZ1 and AZ2, the values of the robust estimate of location for the majority of elements is lowest for the median. The median values are the smallest estimates of location when used with the elements of AZ1 and AZ2; however, the median is not considered to be the most consistent. The 20 and 25 % trimmed mean values of the elements are slightly lesser in value than the 10 % trimmed mean. The values of the Trimean and Gastwirth median are slightly more consistent and lower than the values of the trimmed means. The 20 and 25 % trimmed means and combined L estimates, including the Trimean and Gastwirth median, are preferred robust estimates of location as they are more mutually consistent overall.

### *Scale Estimates*

The three least robust estimates of scale, the standard deviation, mean absolute deviation from the mean and the mean absolute deviation from the median, provide the largest estimates for the majority of geochemical data of AZ1 and AZ2 and are the least mutually consistent. The median absolute deviation from the median and the semi-IQR are considered more robust estimates of scale and when these estimates are used with the geochemical data of AZ1 and AZ2,

they are more mutually consistent. The semi-IQR is the most robust of the five estimates of scale; therefore, the semi-IQR should contain smaller values when used with real geochemical data according to Rock (1988); however, when this estimate is used with the geochemical data of AZ1 and AZ2, this is not true. The semi-IQR yielded slightly larger values compared to the median absolute deviation from the median; therefore, the semi-IQR is considered to be the second most mutually consistent estimate of scale with the geochemical data of AZ1 and AZ2. The median absolute deviation from the median provides the lowest estimates of scale and is likely the most mutually consistent estimate with the geochemical data of AZ1 and AZ2.

## **Multivariate Statistical Methods**

### *Correlation*

Cooksey (2014) stated that a significant difference between the Pearson and Spearman's rho correlation coefficients suggests that the Pearson's correlation coefficient is influenced by data abnormalities. Reimann et al. (2002) stated that Pearson's correlation requires a multivariate normal distribution; however, the majority of the geochemical data of AZ1 and AZ2 are neither normal nor lognormal in distribution. Filzmoser and Hron (2009) stated that the Pearson's correlation coefficient is influenced by data outliers and skewed data distributions, as reflected in the results of Pearson's correlation with the geochemical data of AZ1 and AZ2. Visually, the difference observed in the magnitude the Pearson's versus Spearman's rho correlation coefficients was not significant; however, there was a difference in terms of how many variable pairs were considered significant with each type of correlation.

Considering the strongly correlated variables from both Pearson's and Spearman's correlations in both AZ1 and AZ2, the relationships between several elements can be explained. Al, Be and K are elements which are associated with clay minerals. Mg, Fe, Ti, Cr, Cu and Zn



are also elements associated with clay minerals to a lesser extent. The strong association between the above mentioned elements is an indication of clay minerals in the Al-Azraq basin sediment. A strong association was observed with some clay mineral associated elements and transition metals, such as Al and Ni, K and Ni, and K and Sc, Be and Co. Ca and Mg are elements which are associated with carbonates. Mn and Sr are also elements associated with carbonates to a lesser extent. The strong association between carbonate elements was observed. A strong correlation was observed between several transition element pairs including those transition metals associated with sensitive redox conditions. Strong correlation was observed between immobile and mobile element pairs. For example, there is a strong correlation between the following immobile element pairs: Al and Ti, Al and Sc, Ti and Cr, Mn and P. Ti and Cr are elements transported in refractory minerals and a strong correlation was observed. Considering mobile element pairs, a strong correlation was observed between Ca and Na, and Ca and Mg. The strong correlation between Al and Ni, K and B, and K and Ni, K and Sc, Be and Co indicates the association of elements commonly associated with clay minerals and those of transition and metalloid elements.

#### *HCA and PCA*

Similarly, the resulting groups from HCA and PCA of the elements from AZ1 and AZ2 can explain relationships between elements. Factors 1 and 2 explained by PCA reflects the composition of the bedrock and clay content of the AZ1 and AZ2 sediments, respectively. HCA of major elements resulted in three main groups for AZ1 and two main groups for AZ2. In both AZ1 and AZ2, Ca and Mg cluster together into one group and reflect the composition of the bedrock. In AZ1 and AZ2, Al and K are located close together, reflecting the clay content of the Al-Azraq sediments.

## 9. CONCLUSIONS

### **Conclusions**

Assessing a geochemical variable for normality using visual and statistical methods allows for an understanding of the data distribution prior to the decision on the use of parametric or non-parametric statistical analyses. Ahrens (1953), Rock (1988) and Reimann and Filzmoser (2000) stated geochemical data is commonly not normally or lognormally distributed. Overall the majority of geochemical elements in Al-Azraq sediment cores AZ1 and AZ2 are neither normal nor lognormal. Different graphical and statistical methods to determine data distribution largely yielded slightly different results. Careful consideration should be taken when choosing multivariate statistical methods to be used with geochemical data. Ultimately, a non-parametric test may be better suited to use with data that does not follow a normal nor lognormal distribution. It is likely that non-parametric and parametric multivariate statistical methods, such as correlation, PCA and HCA, used with geochemical data both yield geologically interpretable results. However, the use of non-parametric methods may yield a better interpretation of the data.

Non-robust estimates of location and scale used with AZ1 and AZ2 geochemical data corresponded to higher values and less consistent values compared to more robust estimates. Smaller differences were observed between robust estimates of location and scale compared to their non-robust counterparts. Since multivariate statistical methods are broken down into parametric and non-parametric, considering non-robust and robust estimates of location and scale, of which these statistical methods are based is important for selecting statistical tests. Additionally, by understanding of mutually consistent estimates of location and scale, of which non-parametric and parametric statistical tests are based, may lead to a better choice when selecting multivariate statistical tests.

This research adds to the paucity of studies addressing geochemical data distribution, methods of taking into account censored values and identifying outliers, data standardization and transformation methods commonly used with geochemical data. This study used two large data sets to assess methods commonly used with geochemical data. This study provides insight into the importance of the selection of statistical method by assessing location and scale estimates in which parametric and non-parametric statistical methods are based. In this study, statistics aided in an environmental interpretation of the geochemical data. ICP-AES data from the AZ2 sediments indicated periods of humidity and aridity, increased and decreased precipitation, and low and high chemical weathering. The ICP-AES data from the AZ2 sediments supported interpretations of past environments. This research adds to the previous work in the Al-Azraq basin by using ICP-AES data as a main indicator of past environmental conditions represented by the geochemistry of the AZ2 sediments.

APPENDIX A  
AZ1 Inductively Coupled Plasma-Atomic Emission Spectroscopy (ICP-AES) Data







48301	1.86	5.50	0.61	5.79	1.44	0.88	3.49	0.02	-0.2	5	20	110	0.8	<2	23.5	24	30	25	10	<1	10	312	5	94	2410	61	140	<10	10	44	<10	70
48302	1.86	5.50	0.61	5.79	1.44	0.88	3.49	0.02	-0.2	5	20	110	0.8	<2	23.5	24	30	25	10	<1	10	312	5	94	2410	61	140	<10	10	44	<10	70
48303	1.50	13.40	0.53	6.79	1.39	0.77	1.54	0.02	-0.2	<2	30	140	0.7	<2	9.8	11	37	33	10	<1	10	465	10	41	3090	39	155	<10	<10	32	<10	75
48304	1.42	10.90	0.56	6.06	1.44	0.31	1.29	0.02	-0.2	3	20	60	0.7	<2	10.9	12	48	31	10	<1	10	354	18	40	2370	23	145	<10	10	36	<10	106
48684	1.11	12.90	0.45	7.08	1.07	0.26	1.07	0.02	-0.2	6	20	60	0.6	2	<0.5	10	39	22	<10	<1	10	477	10	48	2250	5	161	<10	10	33	<10	53
48685	1.57	12.40	0.57	6.06	1.07	0.26	1.07	0.02	-0.2	6	20	60	0.6	2	<0.5	10	39	22	<10	<1	10	477	10	48	2250	5	161	<10	10	33	<10	53
48686	1.57	12.40	0.57	6.06	1.07	0.26	1.07	0.02	-0.2	6	20	60	0.6	2	<0.5	10	39	22	<10	<1	10	477	10	48	2250	5	161	<10	10	33	<10	53
48977	1.71	7.06	0.64	4.25	1.77	0.34	1.54	0.03	-0.2	4	20	300	0.8	<2	8.5	13	59	36	10	<1	10	252	16	50	2700	21	177	<10	10	50	<10	133
50224	1.29	12.40	0.50	7.04	1.08	0.28	1.23	0.02	-0.2	2	20	70	0.6	<2	<0.5	8	48	35	<10	<1	10	433	3	36	2560	4	169	<10	10	40	<10	60
50680	0.91	10.20	0.58	8.95	0.96	0.27	1.08	0.01	-0.2	10	20	20	0.5	<2	<0.3	7	30	46	<10	<1	10	522	5	31	1430	5	184	<10	<10	31	<10	45



APPENDIX B  
AZ2 Inductively Coupled Plasma-Atomic Emission Spectroscopy (ICP-AES) Data





APPENDIX C  
AZ1 Box Plot Inner Fence Rule Outlier Charts

Al		Ca		K		Mg		Na		S		Fe		Ti	
Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth
										1.05*	4.09	6.92	40.308	0.08	2.82
										1.11*	4.2	6.76	42.6615		
										2.37*	48.0384	8.11	43.1865		
										0.98*	48.2816	9.79*	44.835		
										0.88*	48.4096	6.9	46.7018		
										0.85*	48.4864				
B		Ba		Be		Co		Cr		Cu		La		Mn	
Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth
		570	0.01	0.25	4.09	2	4.09	112	8.83	6	4.09	40	35.82	1335	20.61
		730	2.54	0.25	7.65	43*	47.98	106	11.99	61	12.95			2670*	20.67
		650	4.51	0.25	8.46	44*	48.28	103	18.20	73*	15.01			1150	21.65
		620	5.03	0.25	24.83	24	48.91	101	19.46	50	20.1156			1195	21.90
		670	5.61	0.25	25.57	25	49.92	128	19.64	58	47.9083			1240	22.85
		570	8.83	0.25	25.69			97	20.12	145*	47.9845			1755	23.10
		770	9.69					99	20.26					1635	23.33
		610	10.16					102	20.28					1145	23.46
		730	22.26					102	20.97					1365	23.81
		580	23.46											1520	23.94
														1800*	24.83
														1455	25.23
														2100*	25.57
														1860*	25.69
														1610	30.58
														1805*	31.52
Mo		Ni		P		Pb		Sr		V		Zn			
Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth		
22	5.32	102	18.48	10000*	33.28	17	2.20	752	0.01	219*	2.79	13	4.09		
26	11.66	135*	49.92	9840*	33.77	1	4.09	686	0.37	217*	2.82	162	49.92		
51*	12.54			5500	33.99	15	33.14	645	0.50	169	3.09				
22	13.71			3870	35.20	15	33.47	712	0.85	196	3.16				
23	14.06			5570	35.31	21*	35.82	685	0.87	173	4.81				
23	14.28			10000*	35.82	19	46.7	697	1.06	160	5.32				
28	14.32			4190	37.82	19	47.81	631	1.25	146	5.59				
26	14.40			10000*	46.70	33*	47.98	699	1.57	171	5.61				
53*	15.01					667*	48.04	508	1.67	200*	5.75				
54*	15.17					18	48.28	600	1.86	471*	8.83				
29	15.66					87*	48.41	10000*	4.09	233*	33.14				
32	18.48					73*	48.49	10000*	4.20	174	33.47				
22	18.69					72*	48.61	517	4.40	127	42.11				
27	19.26					111*	48.77	534	4.51	158	42.66				
30	33.14					61*	48.91	2410*	17.11	166	43.19				
71*	48.04					55*	49.00	6200*	18.16						
24	48.41					19	49.28	3420*	18.20						
38*	49.82					23*	49.50	832*	46.70						
24	49.92					18	49.92								
						22*	49.97								

Al		Ca		K		Mg		Na		S		Fe		Ti	
Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth
-0.84	4.09			-1.51	4.09			-1.51	4.09	0.05	4.09	-0.89	4.09		
-0.34	24.83									0.10	4.20	2.28	44.84		
-0.54	25.57									0.86*	48.04				
										-0.02	48.28				
										-0.13	48.41				
										-0.16	48.49				
B		Ba		Be		Co		Cr		Cu		La		Mn	
Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth
				-1.39*	4.09	0.69*	4.09	2.30*	4.09	1.79*	4.09				
				-1.39*	7.65	3.76	47.98	3.04	7.65	4.11	12.95				
				-1.39*	8.46	3.78	48.28	4.85	19.64	4.29	15.01				
				-1.39*	24.83			3.04	21.65	2.48	21.59				
				-1.39*	25.57			3.04	23.10	2.30	23.10				
				-1.39*	25.69			3.00	23.81	2.30	23.94				
								2.64	24.83	2.08*	24.83				
								2.40*	25.57	2.20	25.23				
								2.89	25.69	2.30	25.69				
								3.00	31.52	4.06	47.91				
										4.98*	47.98				
										2.56	48.61				
Mo		Ni		P		Pb		Sr		V		Zn			
Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth
		2.08	4.09	5.01	32.06	0*	4.09	9.21	4.09	5.39	2.79	2.56*	4.09		
		4.91*	49.92	9.21	33.28	1.39	4.20	9.21	4.20	5.38	2.82				
				4.94	33.49	1.39	4.81	7.79	17.11	5.28	3.16				
				9.19	33.77	1.10	5.89	8.73	18.16	2.48	4.09				
				5.01	34.26	1.10	18.20	8.14	18.20	5.30	5.75				
				4.94	34.45	1.10	20.61			6.15*	8.83				
				9.21	35.82	1.10	20.77			5.45	33.14				
				4.87	44.05	1.10	20.87								
				4.50	44.13	1.39	20.97								
				9.21	46.70	1.10	22.26								
						1.39	22.85								
						1.10	26.26								
						3.04	35.82								
						2.94	46.70								
						2.94	47.81								
						3.50*	47.98								
						6.50*	48.04								
						2.89	48.28								
						4.47*	48.41								
						4.29*	48.49								
						4.28*	48.61								
						4.71*	48.77								
						4.11*	48.91								
						4.01*	49.00								
						2.94	49.28								
						3.14	49.50								
						2.89	49.92								
						3.09	49.97								
						1.39	50.24								

APPENDIX D  
AZ2 Box Plot Inner Fence Rule Outlier Charts

Al		Ca		K		Mg		Na		S		Fe		Ti	
Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth
3.12	6.42	14.70*	6.42	1.03	20.62	1.69	21.53			1.19	6.42	7.57	12.61		
0.67	23.25	9.50	22.81	1.40*	20.67	2.90*	21.61			1.20	17.59	8.36	12.88		
		9.40	22.88	0.24	23.25	2.86*	21.71			1.80*	18.45				
		10.50	23.05	0.25	34.92	3.72*	21.88			2.42*	20.62				
		10.90	23.10			3.74*	21.95			1.56	20.65				
		9.70	23.32			3.85*	22.02			2.74*	20.67				
		8.70	23.35			2.74*	22.09			1.49	20.72				
		9.40	23.41			4.34*	22.75			1.25	20.94				
		13.80*	24.00			5.27*	22.81			1.24	21.01				
		12.90	24.13			5.15*	22.88			1.65	21.07				
		13.60*	24.24			4.08*	22.91			2.15*	21.41				
		10.90	25.06			3.66*	22.95			1.82*	21.53				
		13.90*	25.11			4.20*	23.00			1.40	21.61				
		12.50	25.16			6.09*	23.05			1.56	21.71				
		13.90*	25.21			6.36*	23.10			1.63	21.88				
		9.20	25.26			3.12*	23.15			2.14*	21.95				
		12.60	25.42			4.49*	23.25			2.09*	22.02				
		22.20*	34.50			5.35*	23.32			3.75*	22.09				
		23.80*	34.64			4.69*	23.35			1.73*	22.75				
		23.10*	34.80			5.07*	23.41			1.30	22.81				
		23.90*	34.92			3.97*	23.47			1.67*	22.88				
		22.10*	35.12			2.60*	23.50			1.80*	25.76				
		18.50*	35.23			3.27*	23.52								
		22.00*	35.39			2.28*	23.54								
		22.60*	35.54			2.34*	23.56								
						1.93	23.57								
						7.72*	24.00								
						7.20*	24.13								
						7.54*	24.24								
						6.06*	25.06								
						7.75*	25.11								
						6.95*	25.16								
						7.65*	25.21								
						5.09*	25.26								
						4.23*	25.30								
						6.93*	25.42								
						4.23*	25.76								

B		Ba		Be		Co		Cr		Cu		La		Mn	
Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth
40	6.42	250*	6.42	2.10	12.61	24	15.81			75	17.28	80*	6.42	348	21.53
		400*	7.68	2.20	12.81	22	15.86			85*	18.27			533*	21.61
		130*	8.04	2.10	12.88	32*	17.28			60	23.52			472*	21.71
		280*	17.24	0.25	23.15	34*	17.32							541*	21.88
		110	34.50	0.25	23.25	29*	17.40							579*	21.95
		160*	34.64	0.25	34.64	22	19.37							600*	22.02
		290*	34.80	0.25	34.80	21	23.57							855*	22.09
				0.25	34.92	1	34.64							361	22.75
				0.25	35.12	1	34.80							342	22.81
				0.25	35.39	2	34.92							320	22.88
														347	23.05
														358	23.10
														1445*	24.00
														1365*	24.13
														1400*	24.24
														600*	25.06
														1265*	25.11
														1120*	25.16
														1005*	25.21
														472*	25.26
														418	25.30
														799*	25.42
														744*	25.76

Ni		P		Pb		Sc		Sr		V		Zn	
Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth
77	17.32	10000*	6.42	37	6.42	12	6.42	759*	6.42	123	8.73	136	23.50
95	19.37	6160*	7.68	39*	8.13			180	25.06	123	9.00	154	23.52
74	22.09	5970*	11.31	46*	8.23			183	25.21	127	9.15	146	23.54
76	22.88	3710	21.61	27	9.73			189	25.42	111	9.21	166*	23.56
80	23.50	3750	22.95	30	9.98			427*	34.50	206*	9.73	157	23.57
78	23.54	4790	23.52	30	11.31			410*	34.64	191*	9.98		
84	23.56	4620	23.56	33	12.61			386*	34.80	121	10.00		
93	23.57	3680	23.57	27	12.69			371*	34.92	125	10.18		
77	25.76			67*	15.81			384*	35.12	133	12.61		
				34	17.32			301*	35.23				
				27	17.4			372*	35.39				
				37	20.67			404*	35.54				



Al		Ca		K		Mg		Na		S		Fe		Ti	
Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth
1.14	6.42			0.34	20.67	1.06	21.61	-1.17	23.25	0.88	20.62			-4.61	11.25
-0.31	23.15			-1.35	23.15	1.05	21.71	-1.56	34.50	1.01	20.67			-4.61	12.72
-0.40	23.25			-1.43	23.25	1.31	21.88	-1.97	34.64	1.32	22.09			-4.61	12.79
-0.15	34.64			-1.24	34.64	1.32	21.95	-1.90	34.80	-3.51	34.50			-4.61	12.81
-0.11	34.80			-1.20	34.80	1.35	22.02	-2.04	34.92	-3.51	34.64			-4.61	12.88
-0.30	34.92			-1.39	34.92	1.01	22.09	-1.90	35.12	-3.22	34.80			-4.61	12.90
-0.16	35.39			-1.17	35.12	1.47	22.75	-1.51	35.23					-4.61	12.93
				-1.27	35.39	1.66*	22.81	-1.90	35.39					-4.61	12.98
						1.64*	22.88	-1.83	35.54					-4.61	14.02
						1.41	22.91							-4.61	14.07
						1.30	22.95							-4.61	14.18
						1.44	23.00							-4.61	14.19
						1.81*	23.05							-4.61	14.28
						1.85*	23.10							-4.61	14.41
						1.14	23.15							-4.61	14.50
						1.50*	23.25							-4.61	14.54
						1.68*	23.32							-4.61	14.67
						1.55*	23.35							-4.61	14.81
						1.62*	23.41							-4.61	14.99
						1.38	23.47							-4.61	15.11
						0.96	23.50							-4.61	15.20
						1.18	23.52							-4.61	17.06
						0.82	23.54							-4.61	17.11
						0.85	23.56							-4.61	17.16
						2.04*	24.00							-4.61	23.25
						1.97*	24.13							-4.61	24.00
						2.02*	24.24							-4.61	24.13
						1.80*	25.06							-4.61	24.24
						2.05*	25.11							-4.61	25.11
						1.94*	25.16							-4.61	25.21
						2.03*	25.21							-4.61	34.64
						1.63*	25.26							-4.61	34.80
						1.44	25.30							-4.61	34.92
						1.94*	25.42							-4.61	35.12
						1.44	25.76							-4.61	35.23
														-4.61	35.39
														-4.61	35.54

B		Ba		Be		Co		Cr		Cu		La		Mn	
Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth
		5.52*	6.42	0.69	10.33	3.47	17.28			4.32	17.28	4.38	6.42	6.36	21.95
		5.99*	7.68	0.74	12.61	3.53	17.32			4.44	18.27			6.40	22.02
		4.87	8.04	0.79	12.81	3.37	17.40			2.77	34.92			6.75	22.09
		5.63*	17.24	0.74	12.88	1.39	23.25							7.28	24.00
		2.30	23.25	-1.39*	23.15	1.39	34.50							7.22	24.13
		4.70	34.50	-1.39*	23.25	0*	34.64							7.24	24.24
		5.08	34.64	-0.69	23.32	0*	34.80							6.40	25.06
		5.67*	34.80	-0.51	23.35	0.69*	34.92							7.14	25.11
				-0.51	24.13	1.10	35.12							7.02	25.16
				-0.51	24.24	1.39	35.23							6.91	25.21
				-0.51	25.11	1.39	35.39							6.68	25.42
				-0.51	25.21	1.39	35.54							6.61	25.76
				-0.51	34.50										
				-1.39*	34.64										
				-1.39*	34.80										
				-1.39*	34.92										
				-1.39*	35.12										
				-0.69	35.23										
				-1.39*	35.39										
				-0.69	35.54										

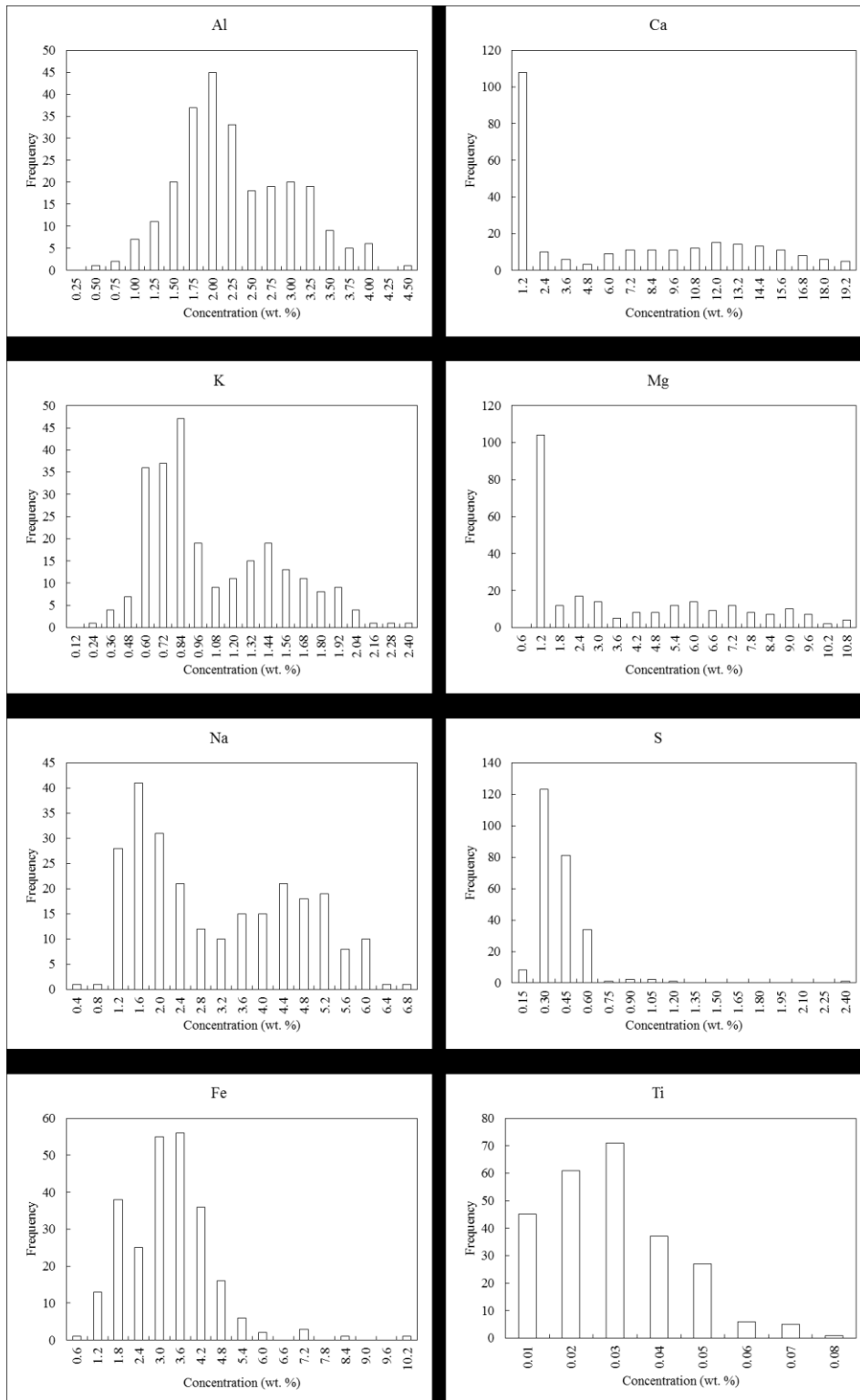
  

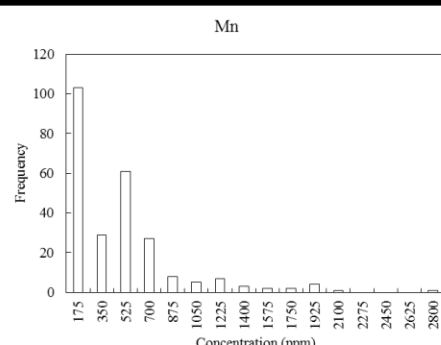
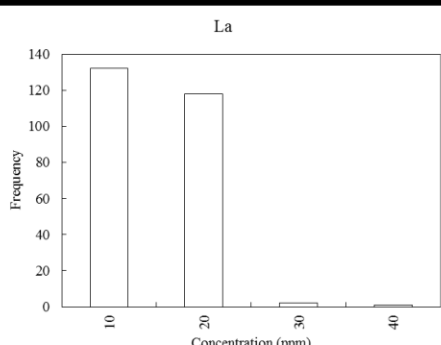
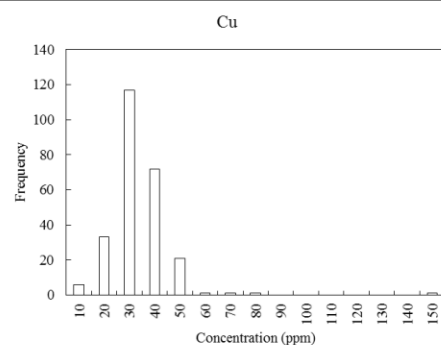
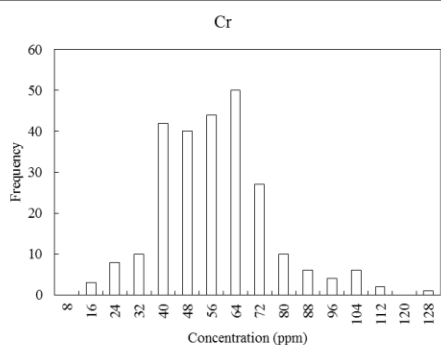
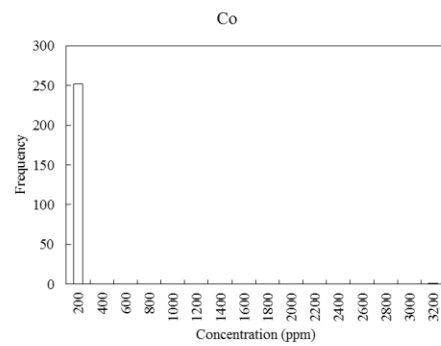
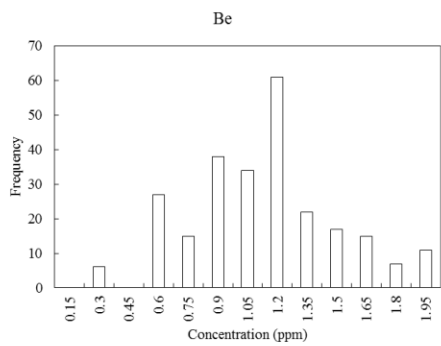
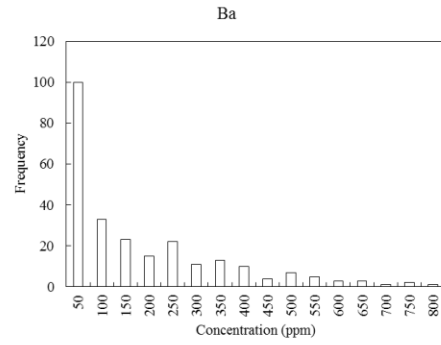
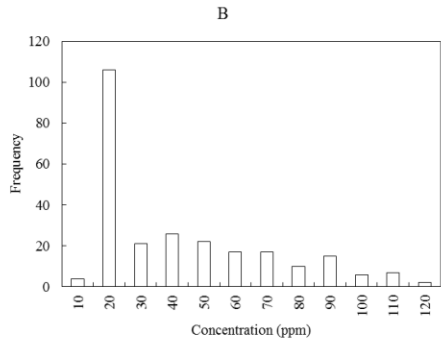
Ni		P		Pb		Sc		Sr		V		Zn	
Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth	Outlier	Depth
4.55	19.37	4.79	12.93	3.66	8.13	2.48	6.42	6.63	6.42	5.33	9.73	3.81	23.25
4.53	23.57	4.70	12.98	3.83	8.23	0.69	23.15	6.06	34.50			5.04	23.52
		5.08	14.02	4.20	15.81	0.69	23.25	6.02	34.64			4.98	23.54
				1.10	34.80	0.69	34.92	5.96	34.80			5.11	23.56
				1.10	35.23			5.92	34.92			5.06	23.57
				1.10	35.54			5.95	35.12				
								5.92	35.39				
								6.00	35.54				

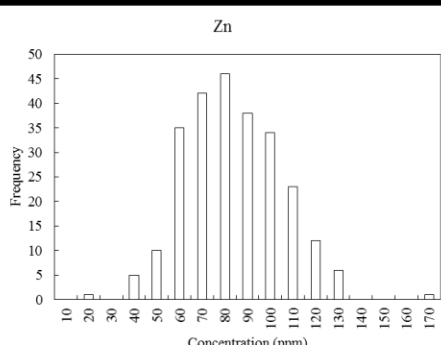
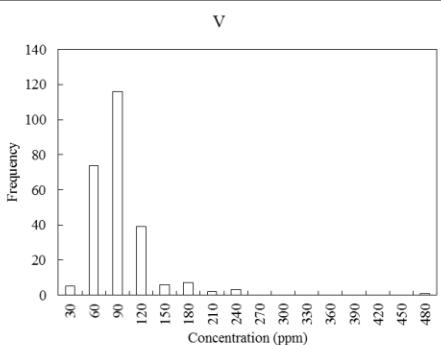
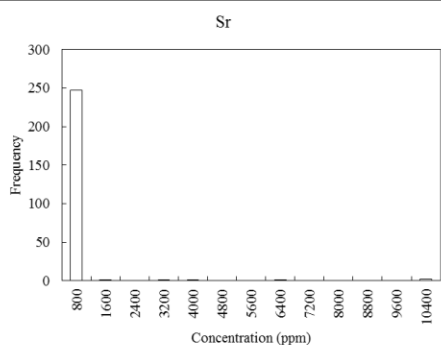
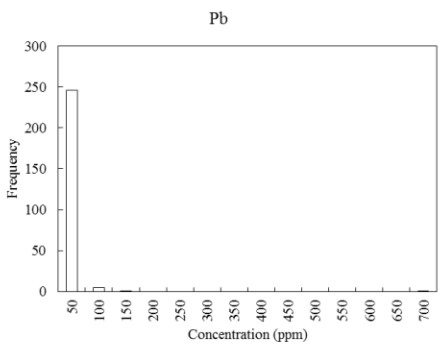
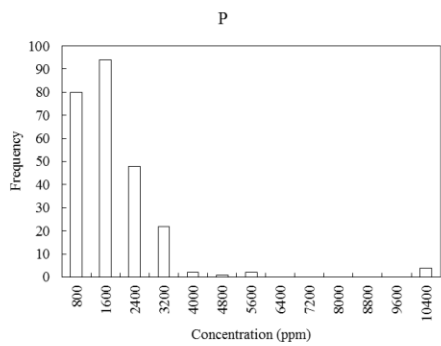
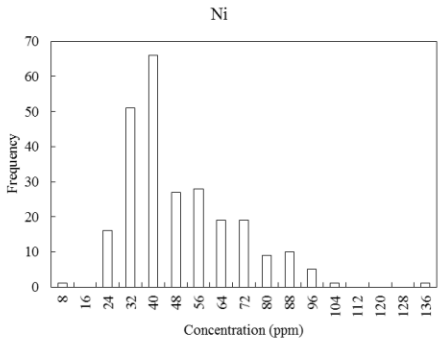
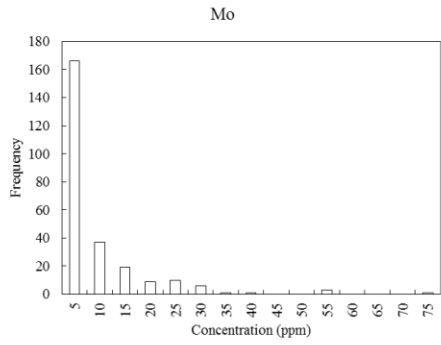
APPENDIX E  
AZ2 Median  $\pm$  2 MAD Rule Outlier Chart

Ca		Mg		ln Mg		ln Ti	
Outlier	Depth	Outlier	Depth	Outlier	Depth	AZ1	AZ2
14.70	6.42	0.47	11.25	-0.67	6.42		
1.46	7.68	0.91	20.62	-0.76	11.25		
1.41	11.31	0.96	20.65	-0.67	12.61		
0.66	11.36	0.86	20.67	-0.69	12.98		
0.60	17.32	1.02	20.72	-0.09	20.62		
1.37	21.41	0.99	20.94	-0.04	20.65		
1.62	21.53	1.01	21.01	-0.15	20.67		
3.65	21.61	0.99	21.07	0.02	20.72		
3.34	21.71	1.52	21.41	-0.01	20.94		
5.02	21.88	1.69	21.53	0.01	21.01		
5.84	21.95	2.90	21.61	-0.01	21.07		
6.03	22.02	2.86	21.71	0.42	21.41		
4.69	22.09	3.72	21.88	0.52	21.53		
8.20	22.75	3.74	21.95	1.06	21.61		
9.50	22.81	3.85	22.02	1.05	21.71		
9.40	22.88	2.74	22.09	1.31	21.88		
6.30	22.91	4.34	22.75	1.32	21.95		
5.50	22.95	5.27	22.81	1.35	22.02		
6.13	23.00	5.15	22.88	1.01	22.09		
10.50	23.05	4.08	22.91	1.47	22.75		
10.90	23.10	3.66	22.95	1.66	22.81		
5.46	23.15	4.20	23.00	1.64	22.88		
8.20	23.25	6.09	23.05	1.41	22.91		
9.70	23.32	6.36	23.10	1.30	22.95		
8.70	23.35	3.12	23.15	1.44	23.00		
9.40	23.41	4.49	23.25	1.81	23.05		
6.58	23.47	5.35	23.32	1.85	23.10		
3.69	23.50	4.69	23.35	1.14	23.15		
5.13	23.52	5.07	23.41	1.50	23.25		
2.82	23.54	3.97	23.47	1.68	23.32		
3.17	23.56	2.60	23.50	1.55	23.35		
2.33	23.57	3.27	23.52	1.62	23.41		
13.80	24.00	2.28	23.54	1.38	23.47		
12.90	24.13	2.34	23.56	0.96	23.50		
13.60	24.24	1.93	23.57	1.18	23.52		
10.90	25.06	7.72	24.00	0.82	23.54		
13.90	25.11	7.20	24.13	0.85	23.56		
12.50	25.16	7.54	24.24	0.66	23.57		
13.90	25.21	6.06	25.06	2.04	24.00		
9.20	25.26	7.75	25.11	1.97	24.13		
7.30	25.30	6.95	25.16	2.02	24.24		
12.60	25.42	7.65	25.21	1.80	25.06		
6.06	25.76	5.09	25.26	2.05	25.11		
22.20	34.50	4.23	25.30	1.94	25.16		
23.80	34.64	6.93	25.42	2.03	25.21		
23.10	34.80	4.23	25.76	1.63	25.26		
23.90	34.92	0.38	34.64	1.44	25.30		
22.10	35.12	0.39	34.80	1.94	25.42		
18.50	35.23	0.33	34.92	1.44	25.76		
22.00	35.39	0.42	35.12	-0.97	34.64		
22.60	35.54	0.38	35.39	-0.94	34.80		
		0.47	35.54	-1.11	34.92		
				-0.87	35.12		
				-0.97	35.39		
				-0.76	35.54		

APPENDIX F  
AZ1 Histograms of Raw ICP-AES Data

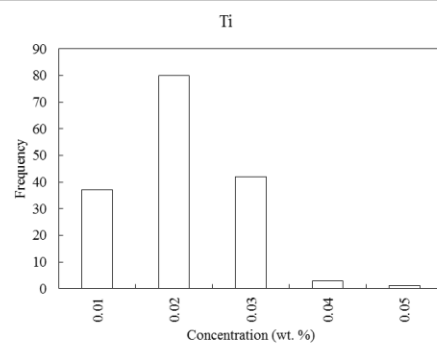
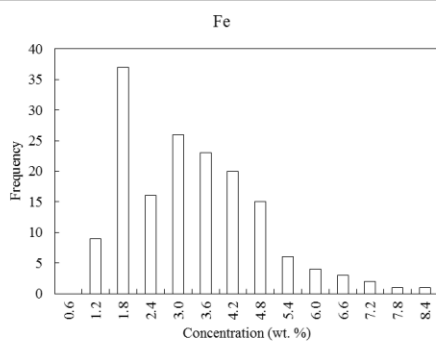
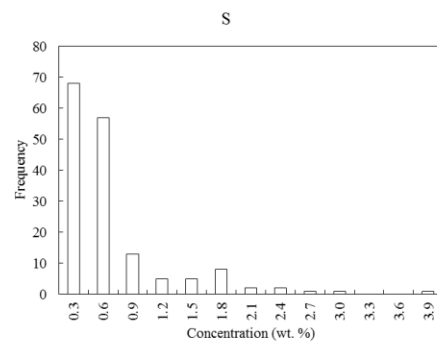
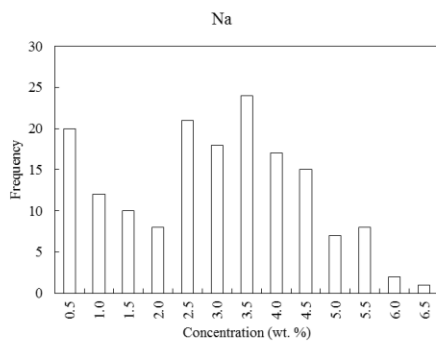
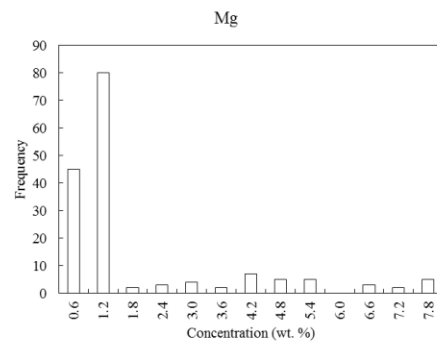
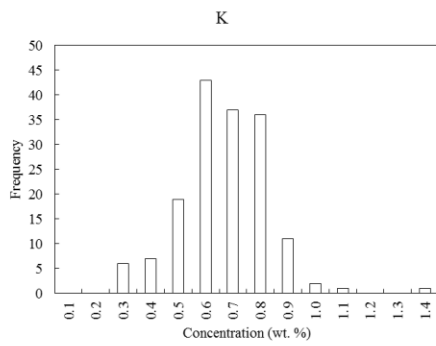
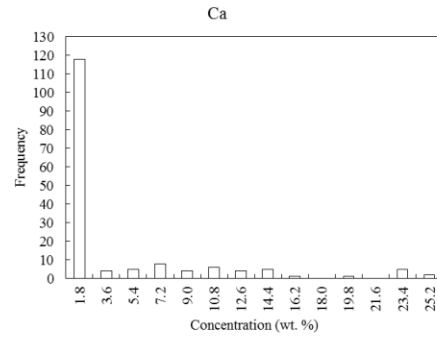
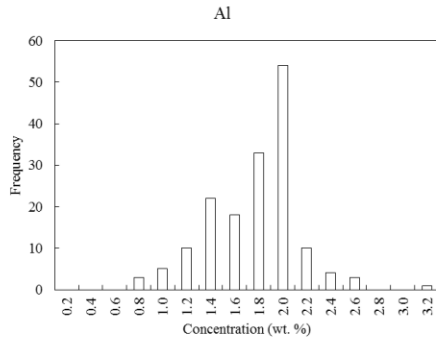


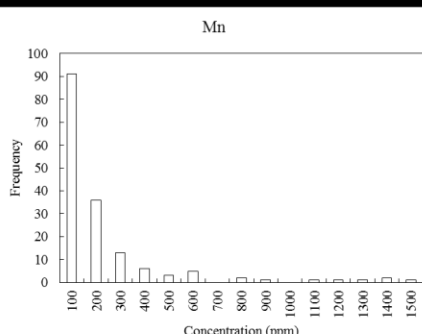
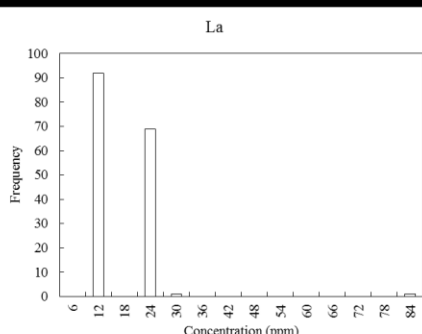
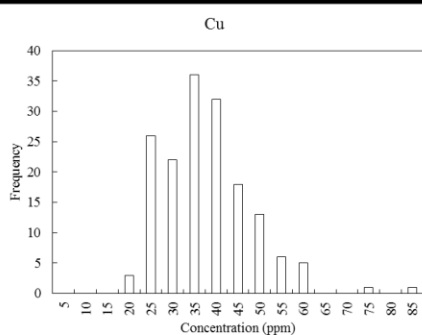
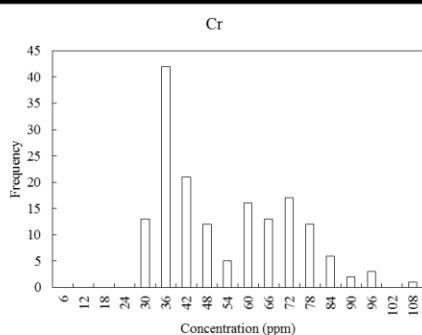
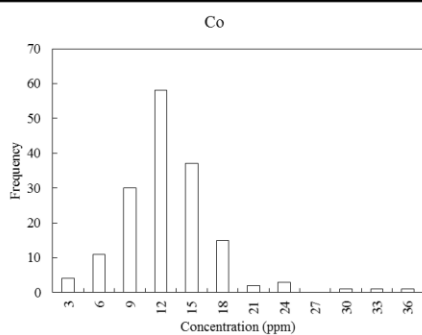
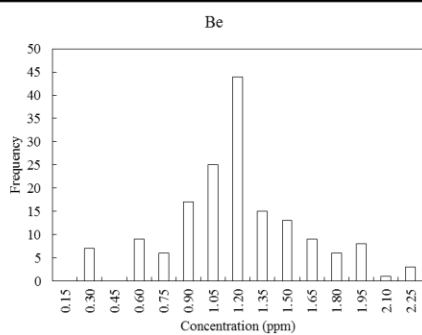
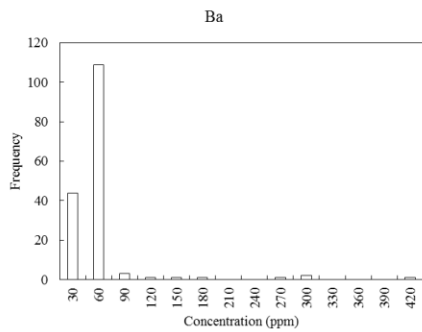
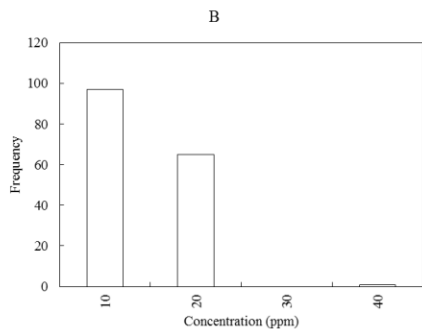


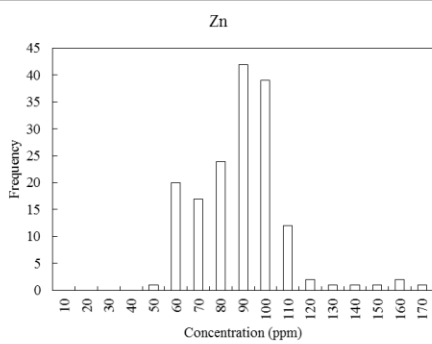
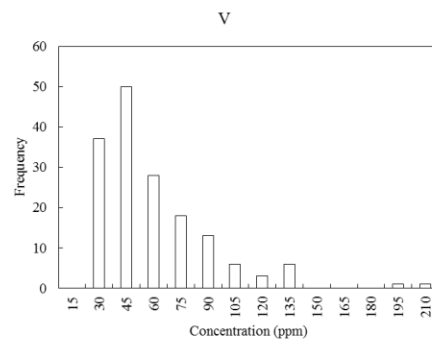
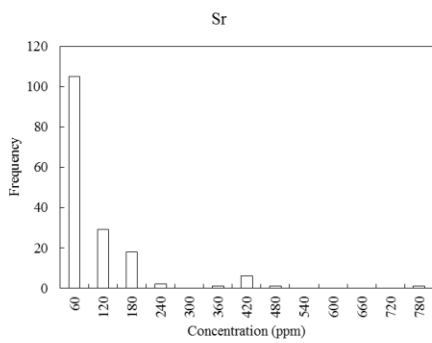
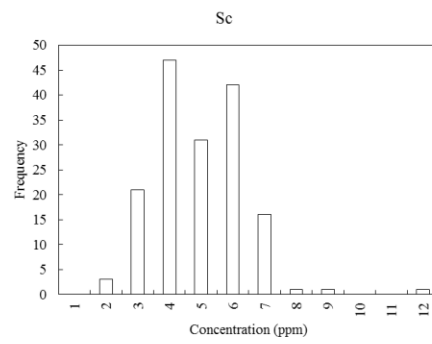
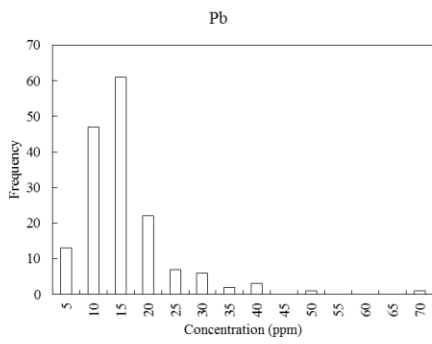
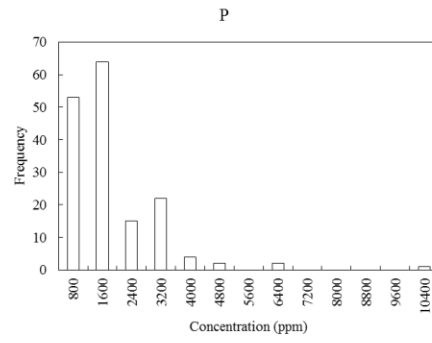
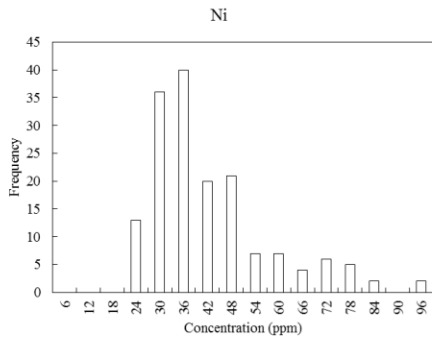


APPENDIX G  
AZ2 Histograms of Raw ICP-AES Data

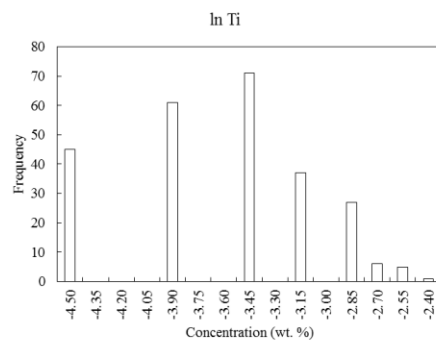
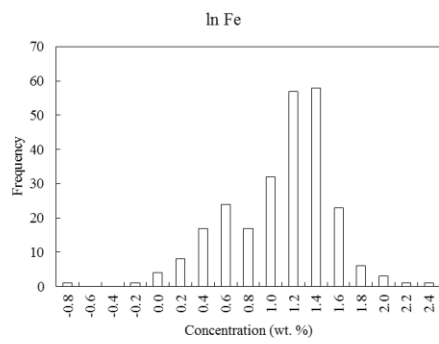
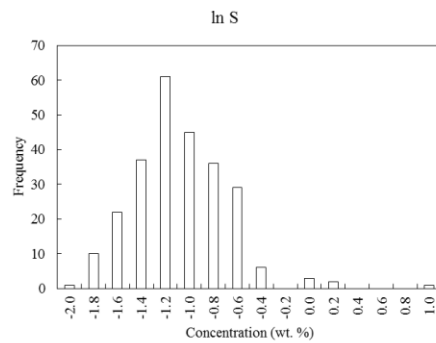
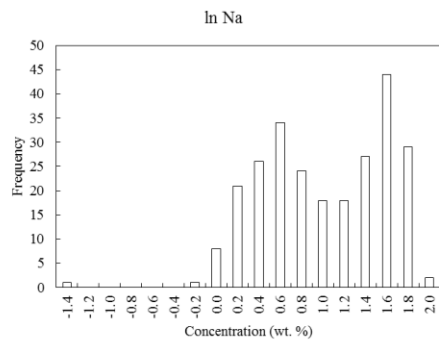
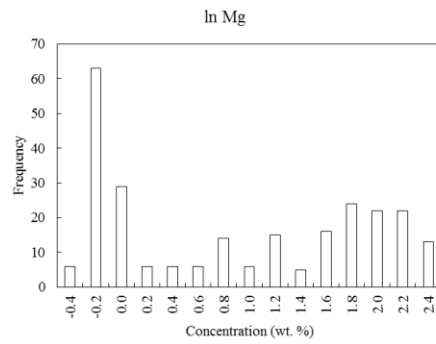
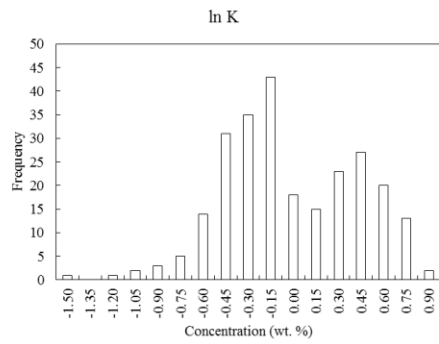
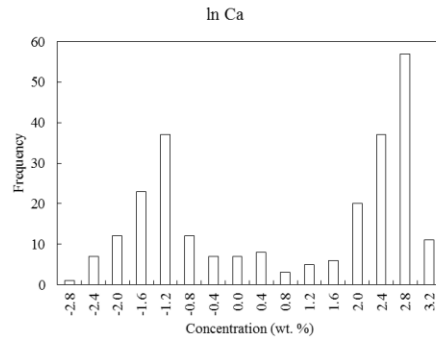
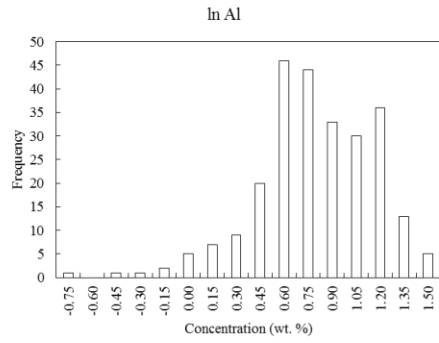


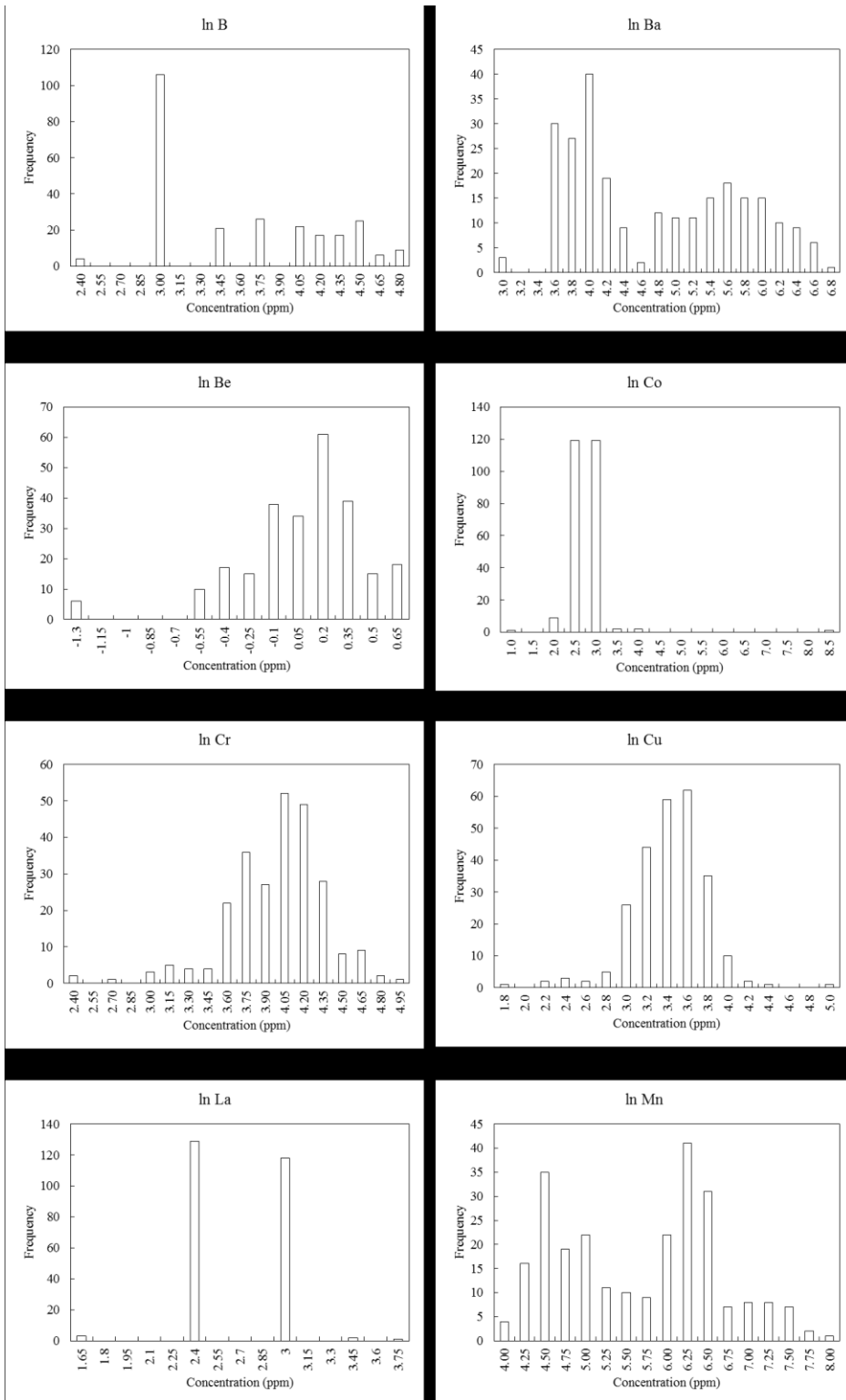


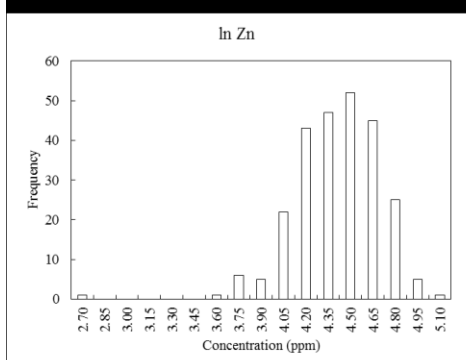
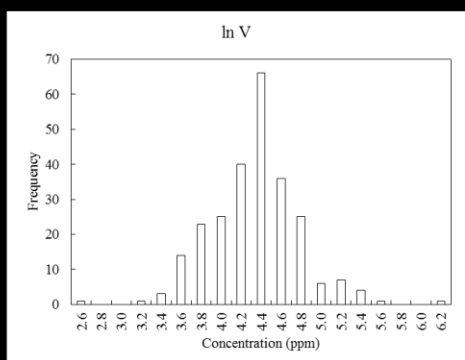
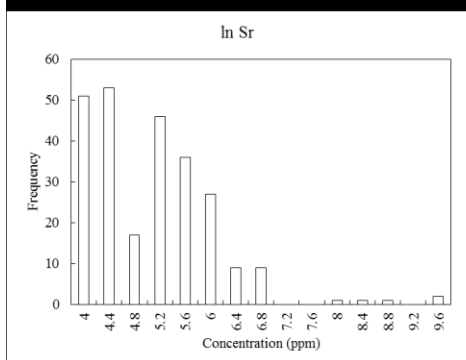
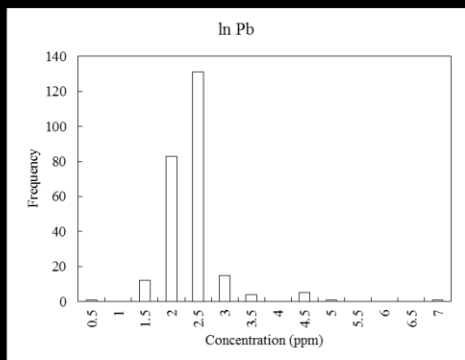
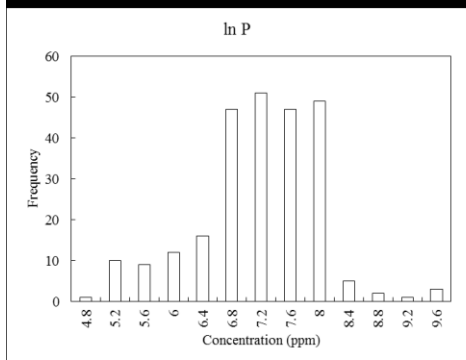
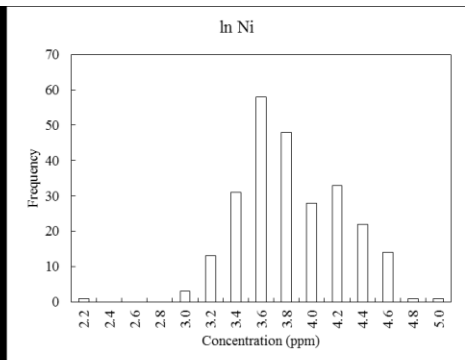
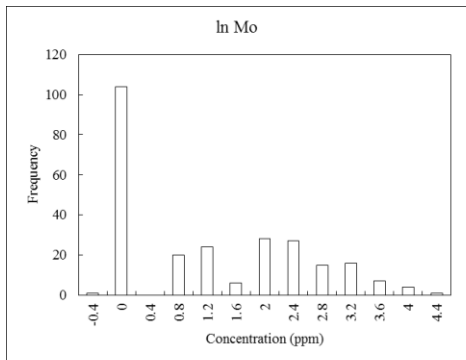




APPENDIX H  
AZ1 Histograms of Ln Transformed ICP-AES Data

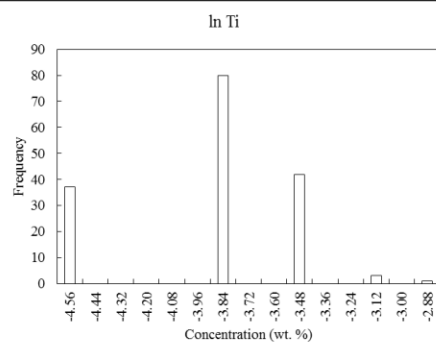
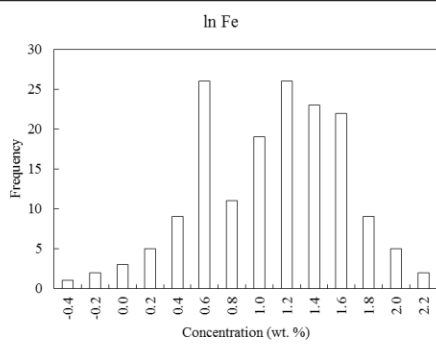
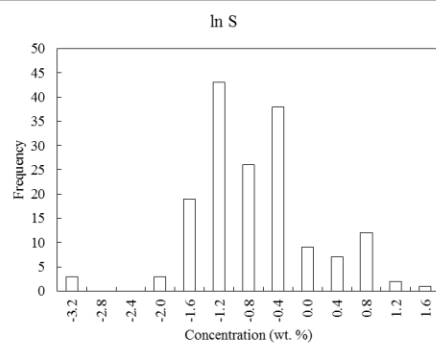
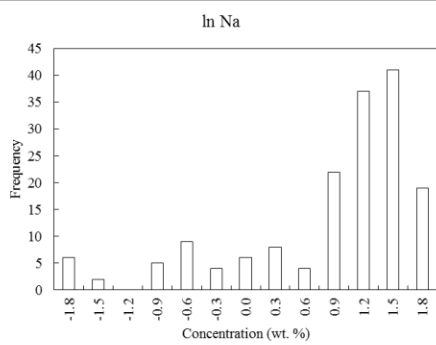
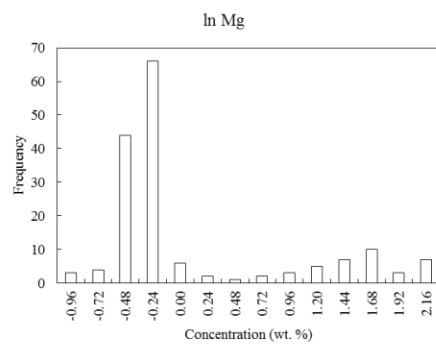
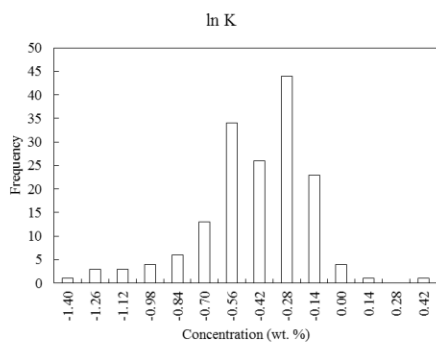
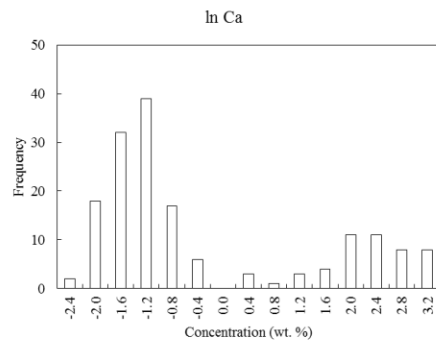
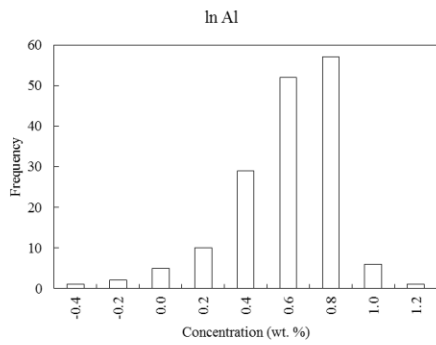


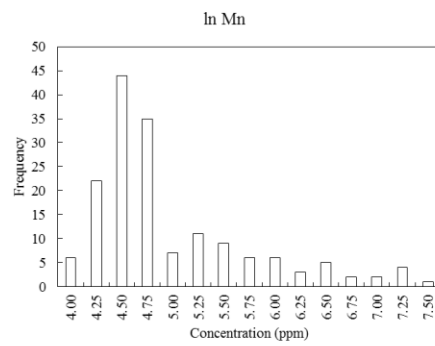
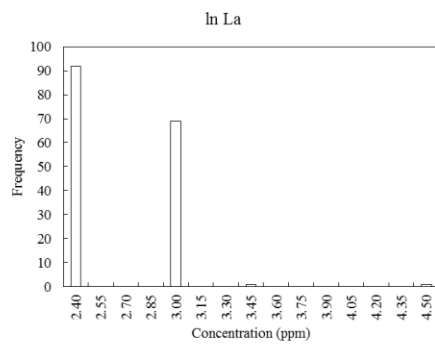
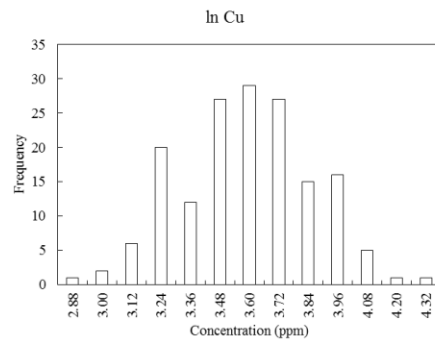
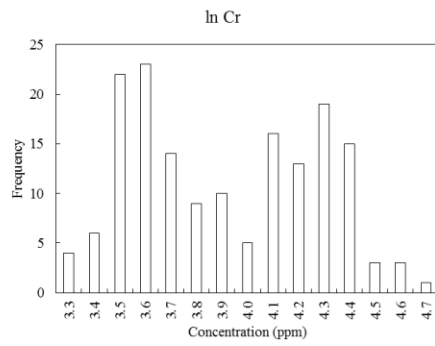
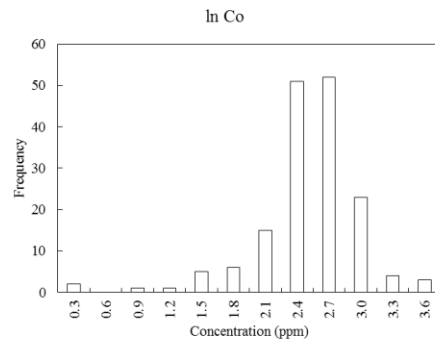
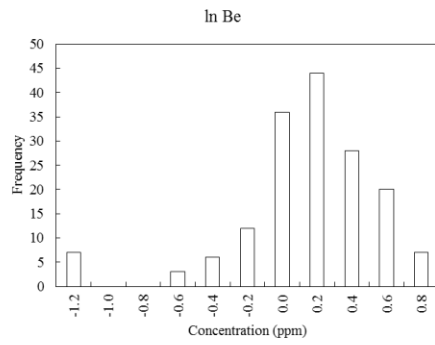
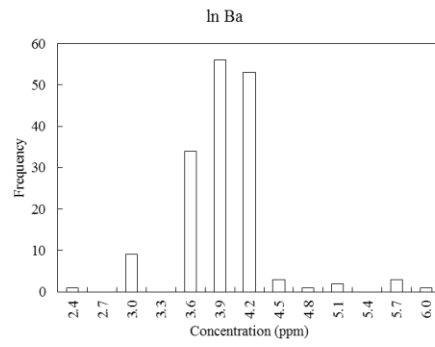
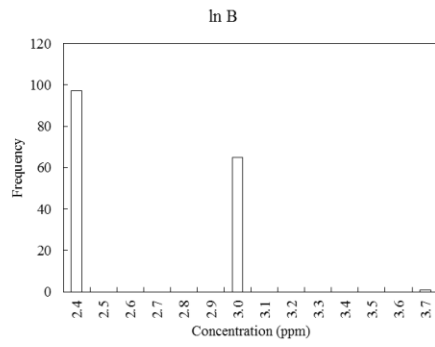


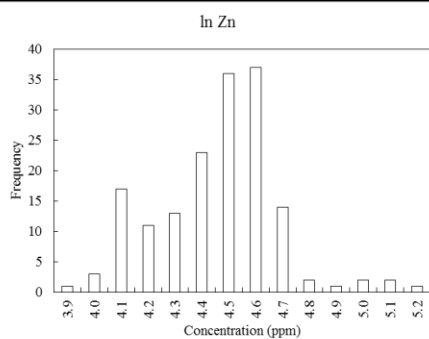
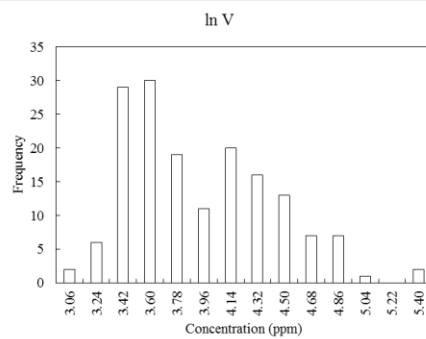
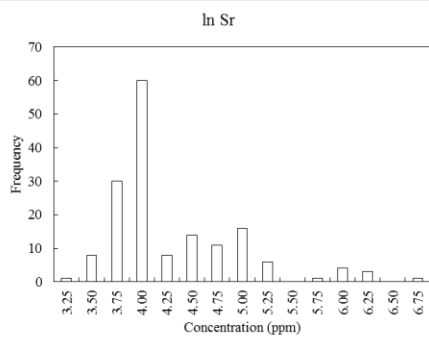
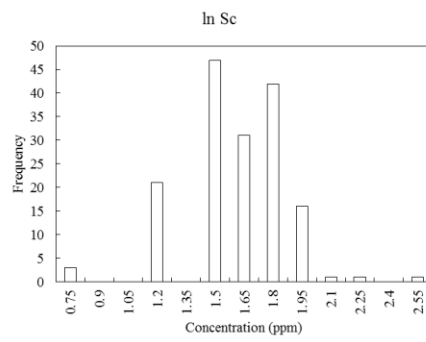
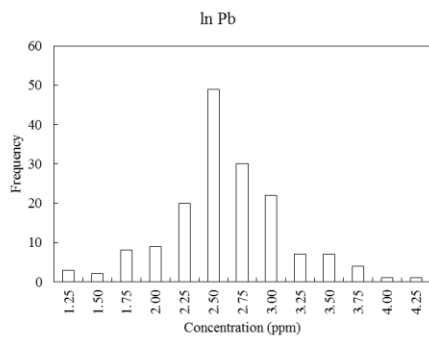
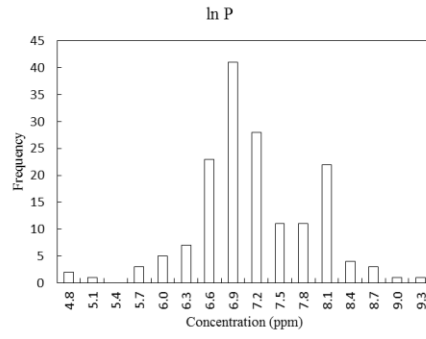
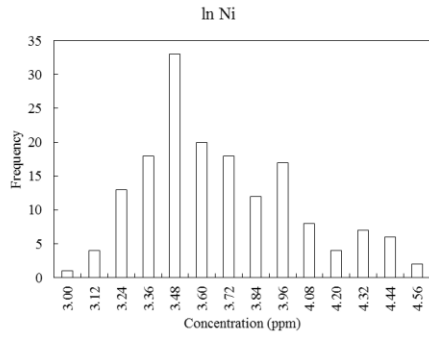


APPENDIX I  
AZ2 Histograms of Ln Transformed ICP-AES Data

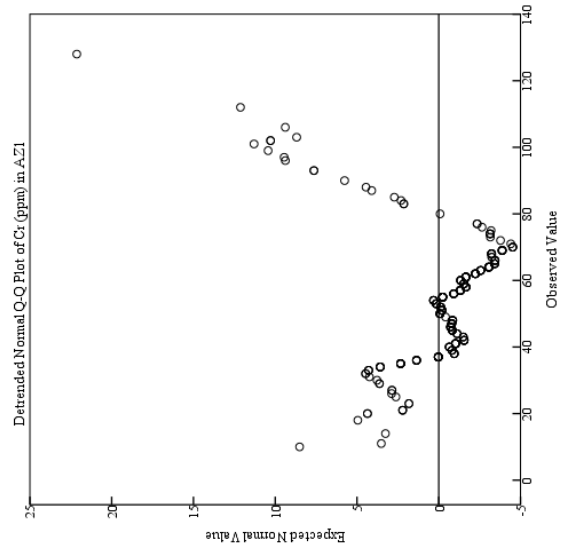
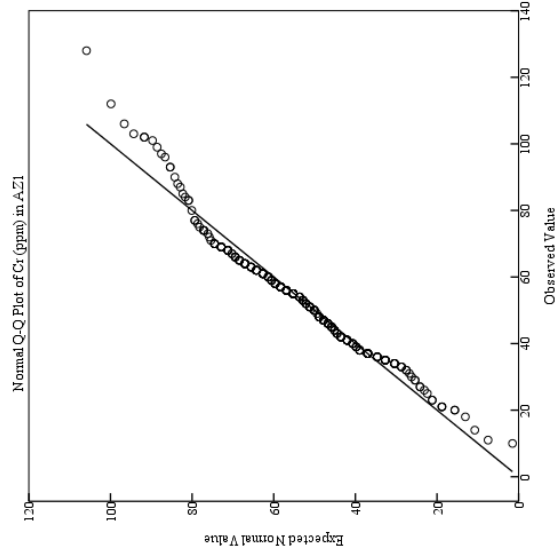
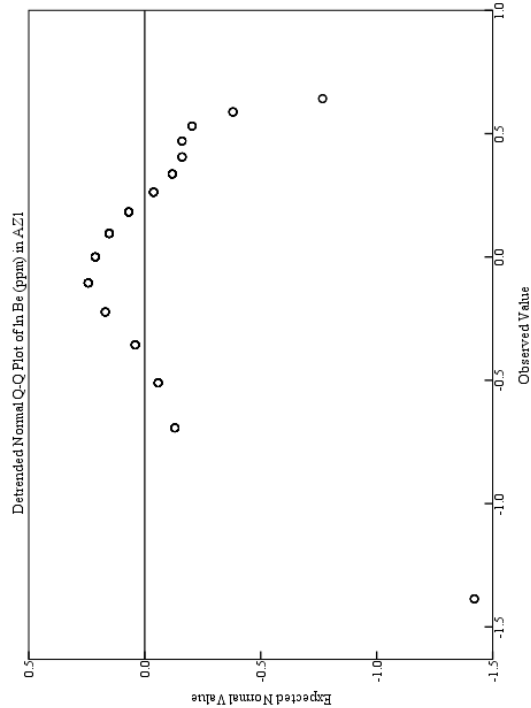
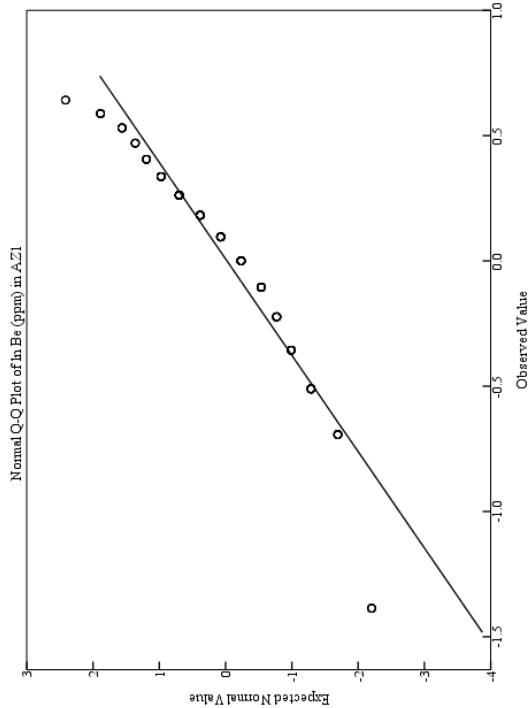


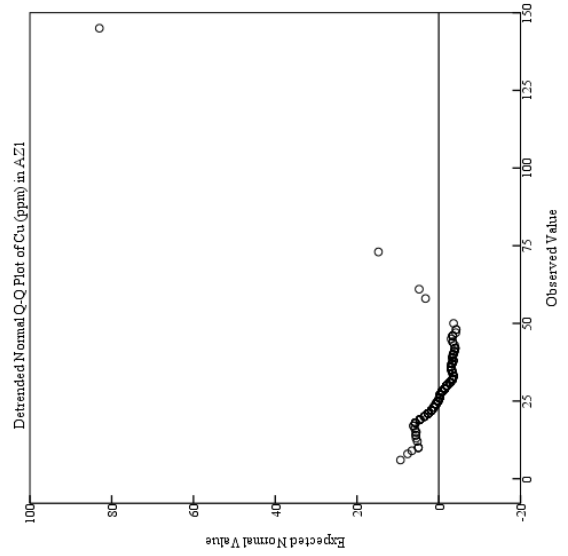
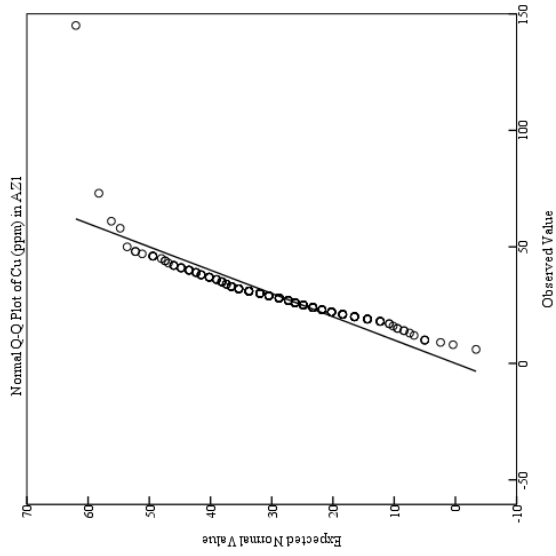
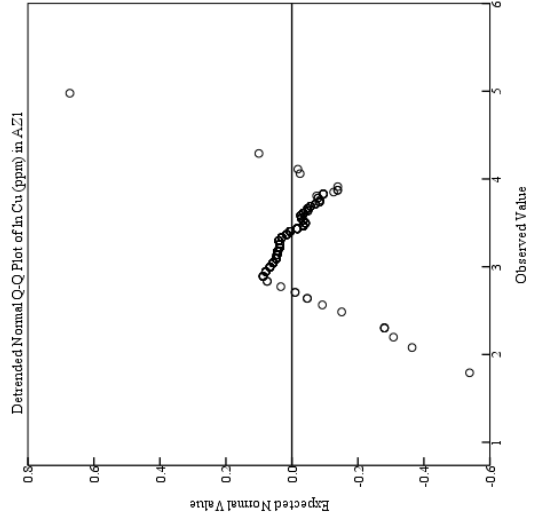
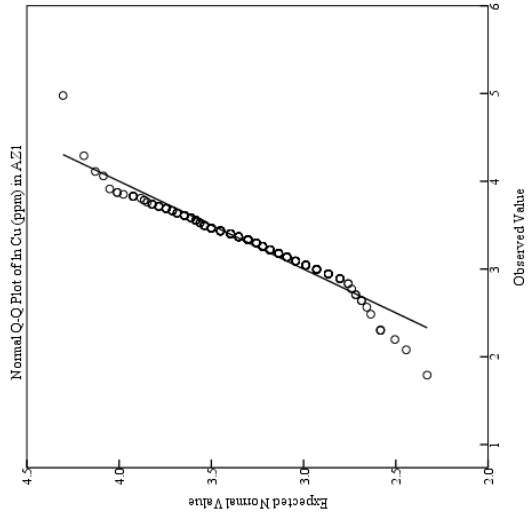


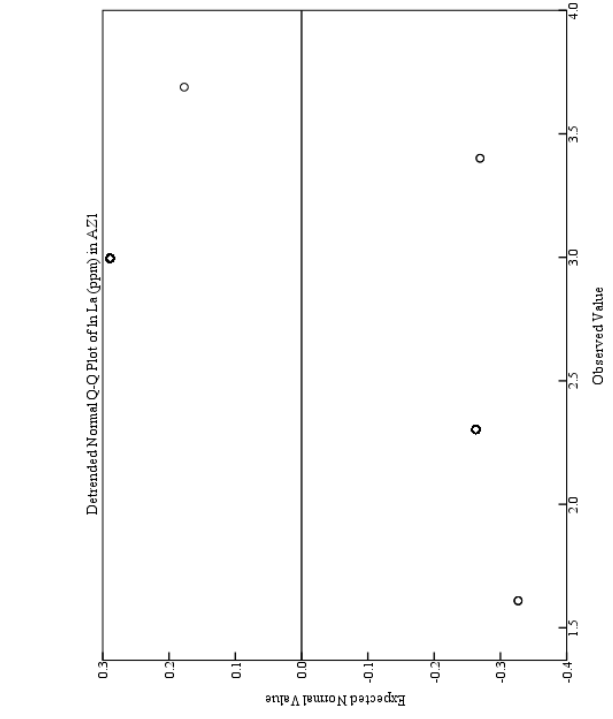
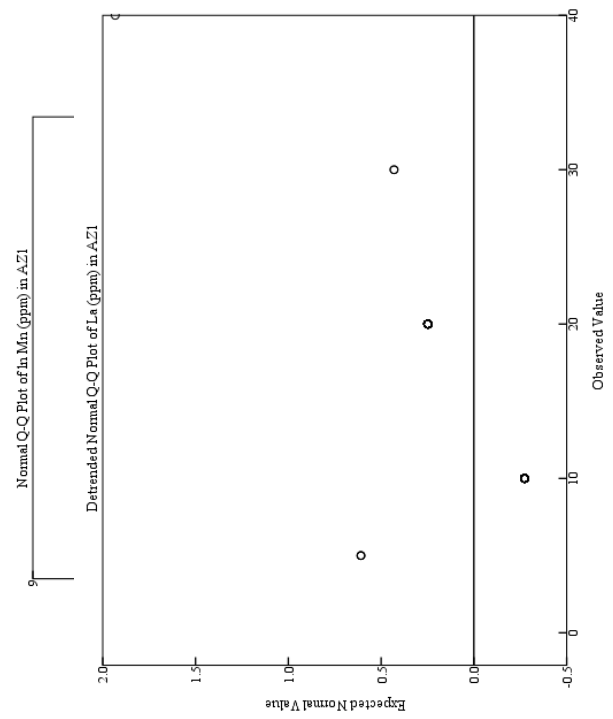
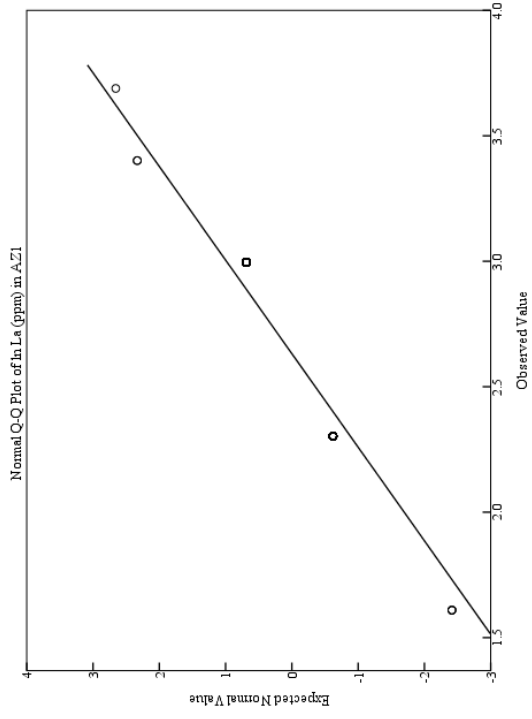
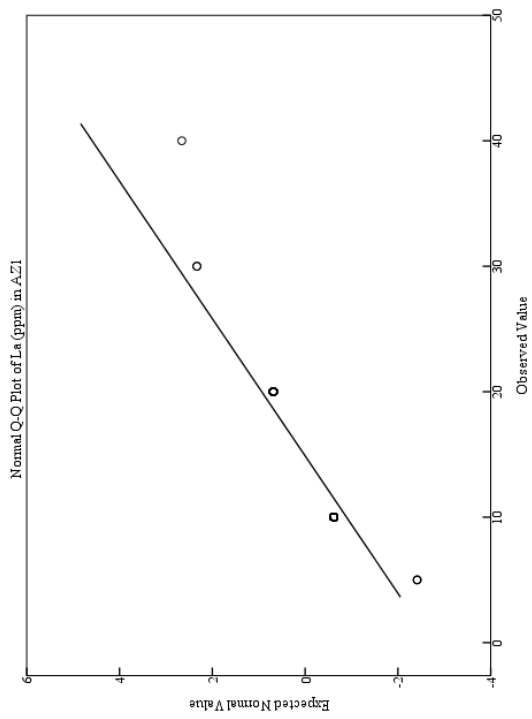


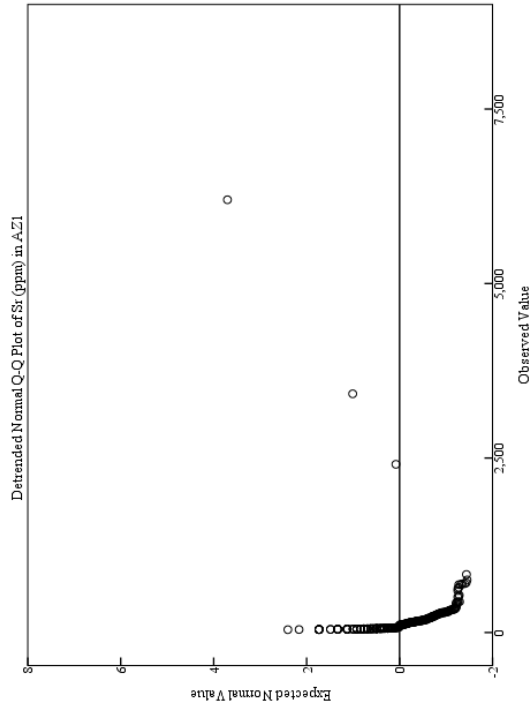
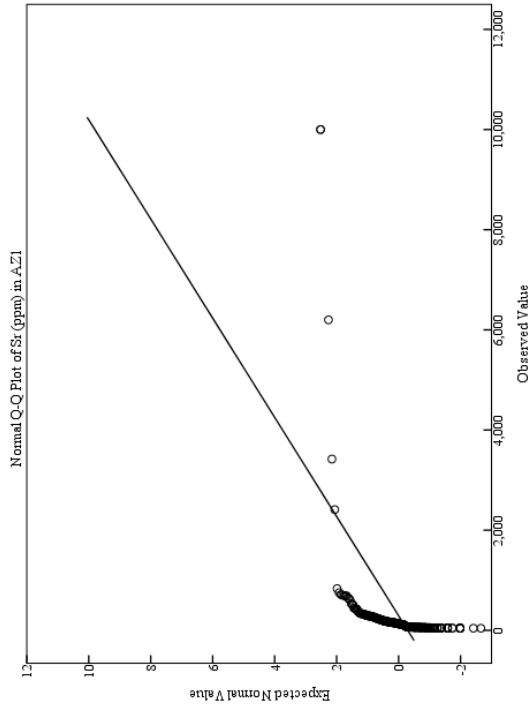
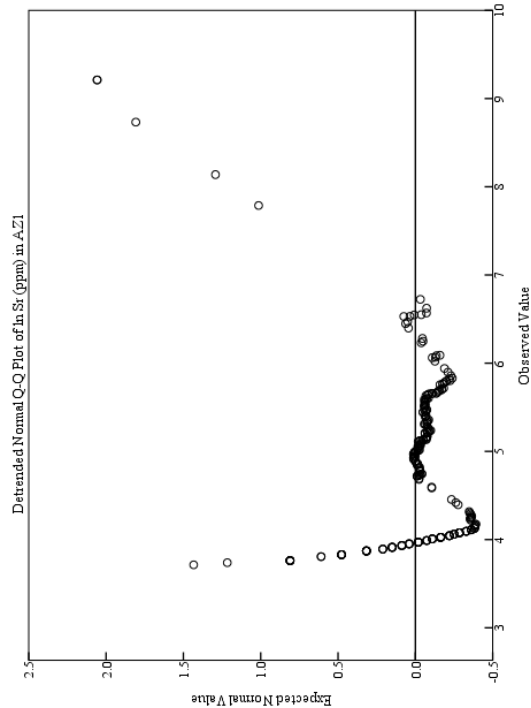
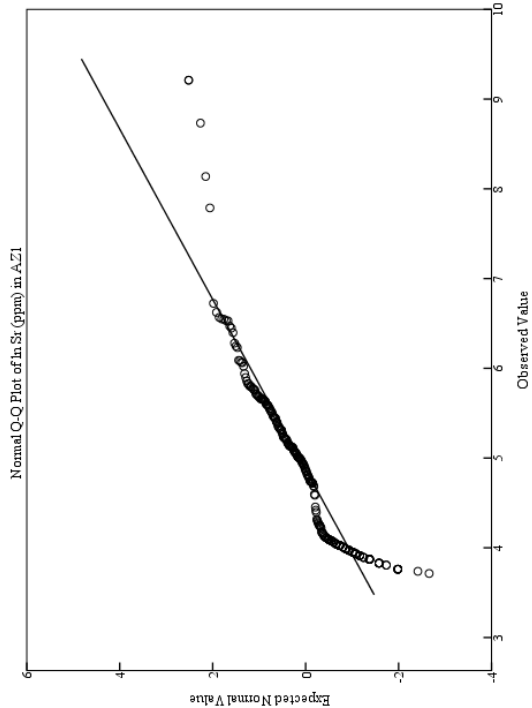


APPENDIX J  
AZ1 Q-Q and Detrended Q-Q Plots of ICP-AES Data



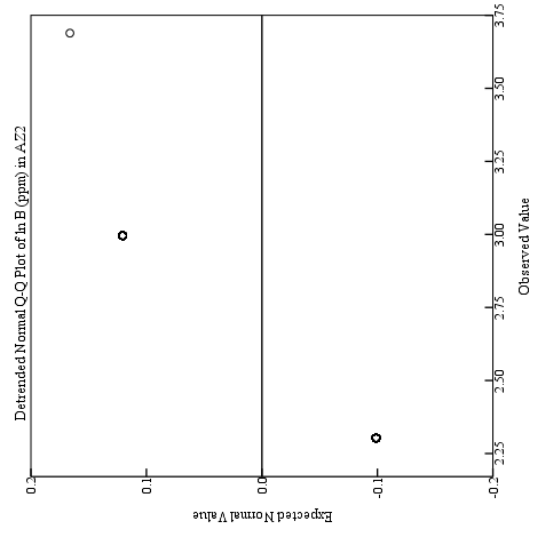
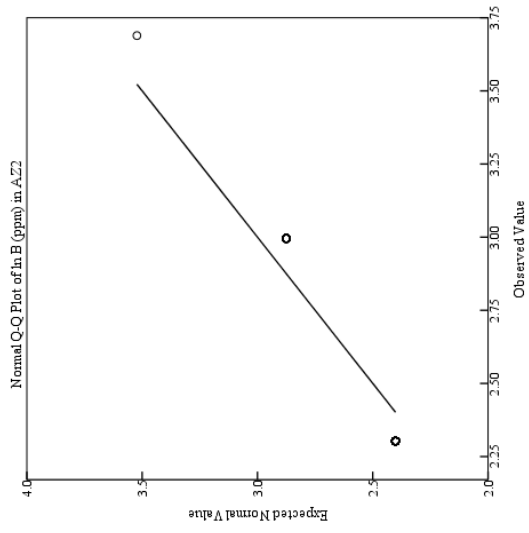
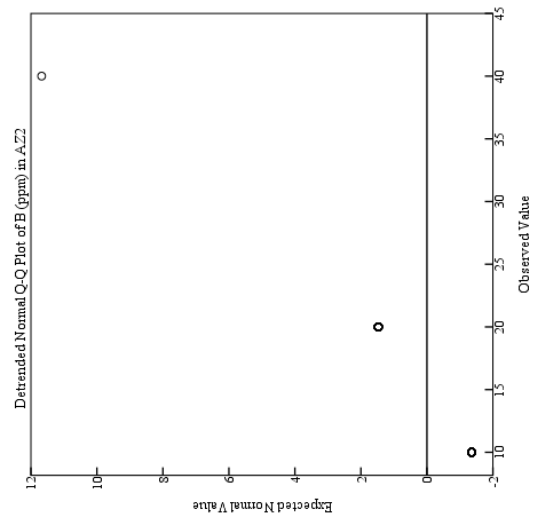
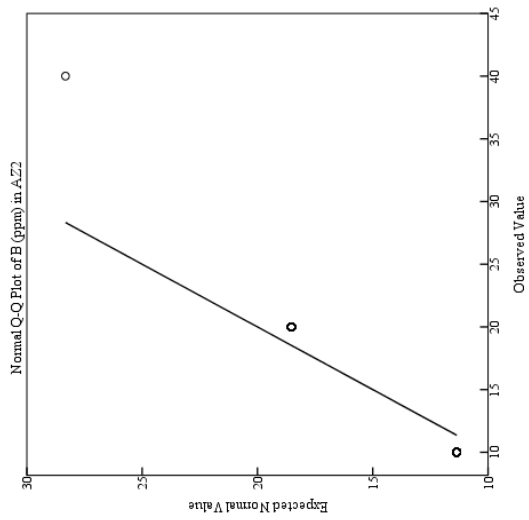


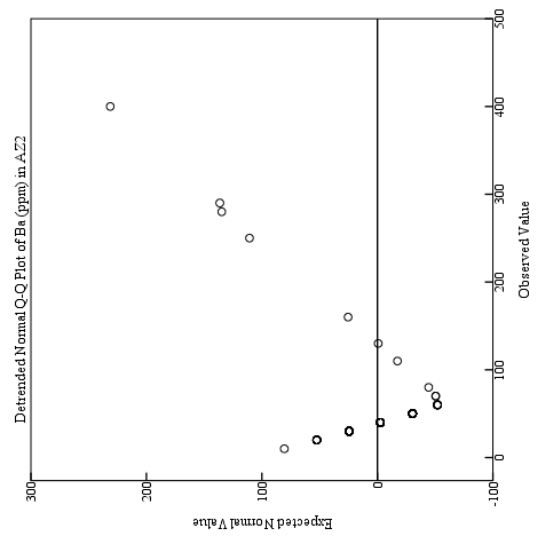
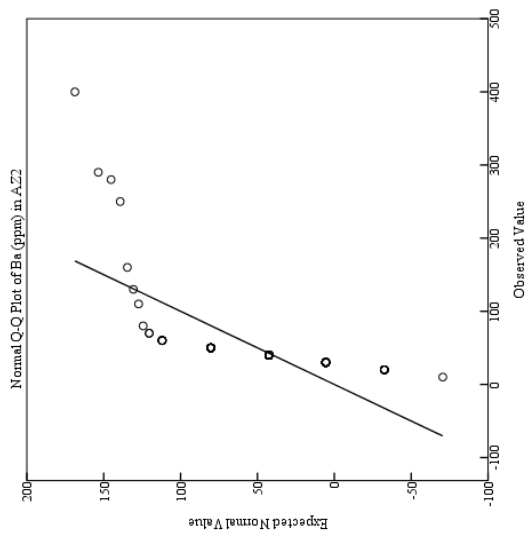
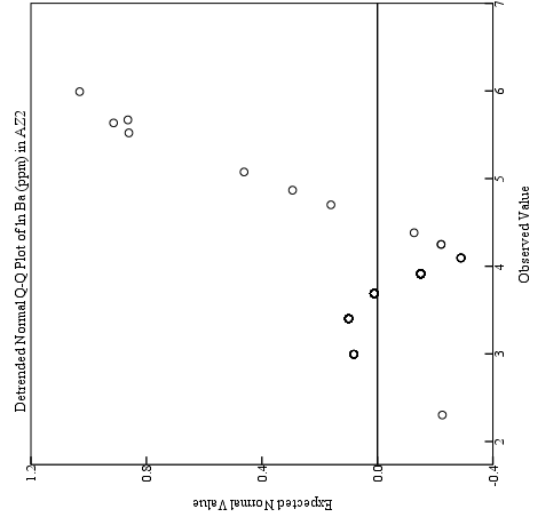
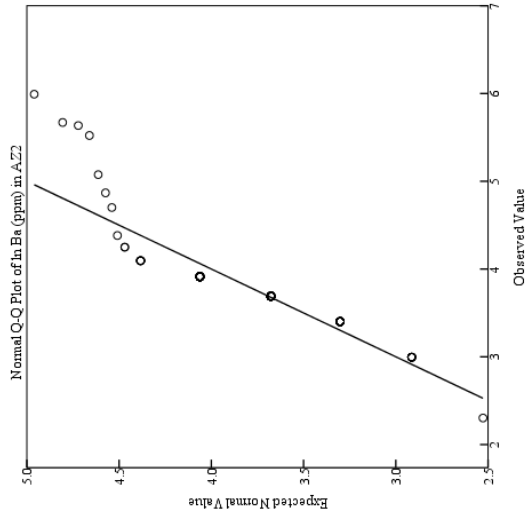


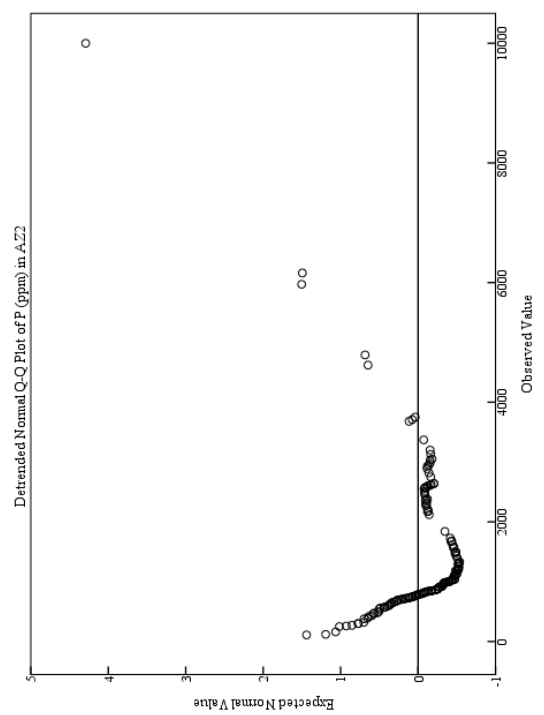
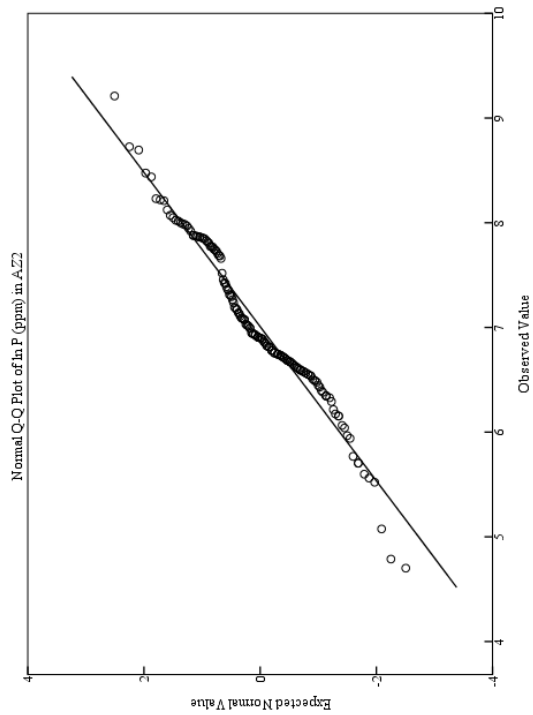
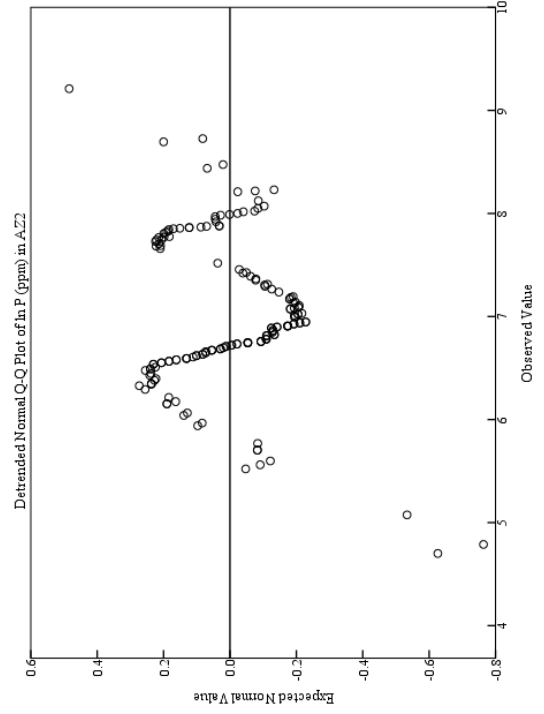
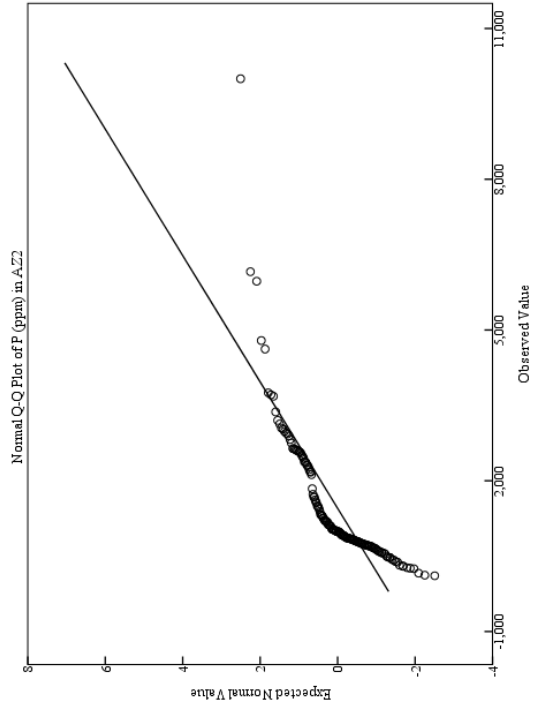




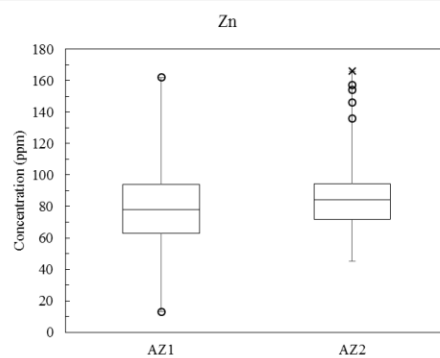
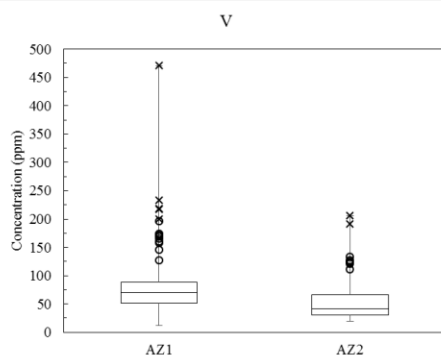
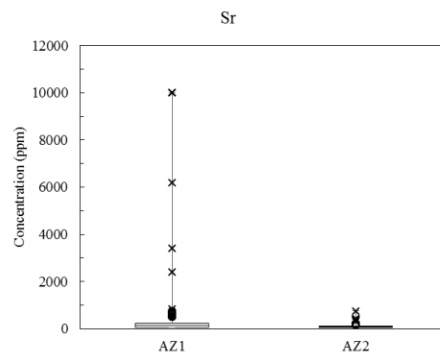
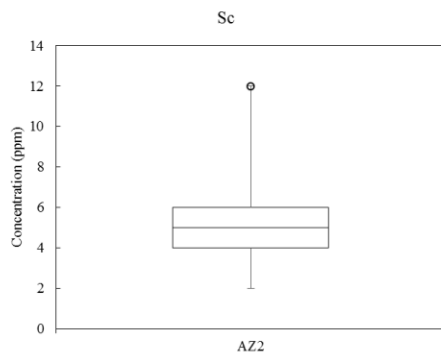
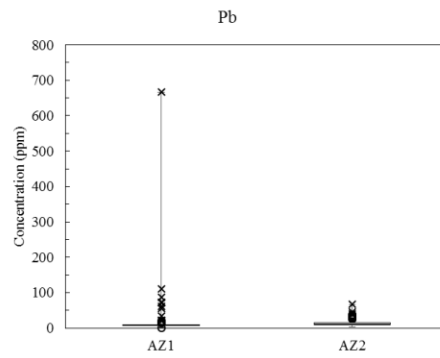
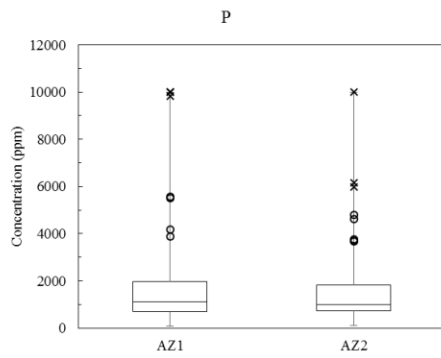
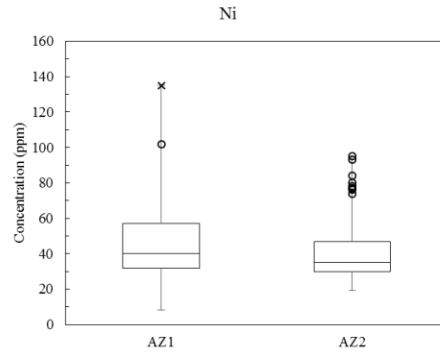
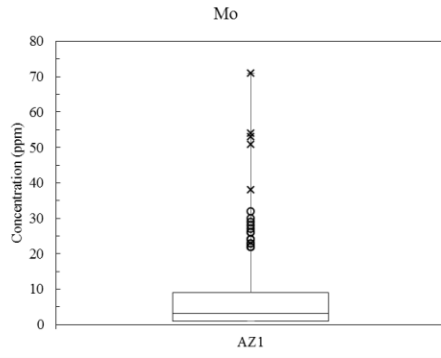
APPENDIX K  
AZ2 Q-Q and Detrended Q-Q Plots of ICP-AES Data



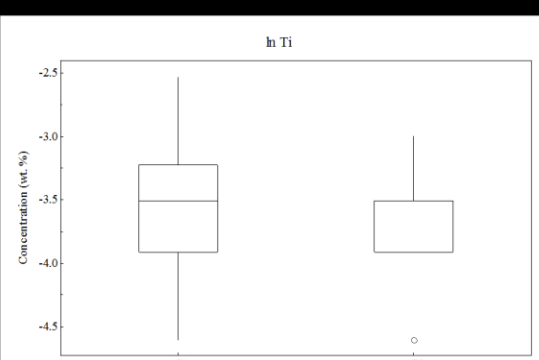
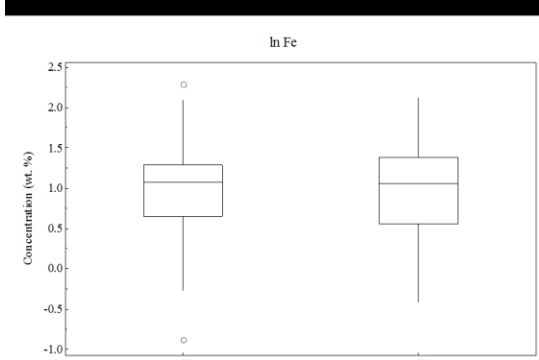
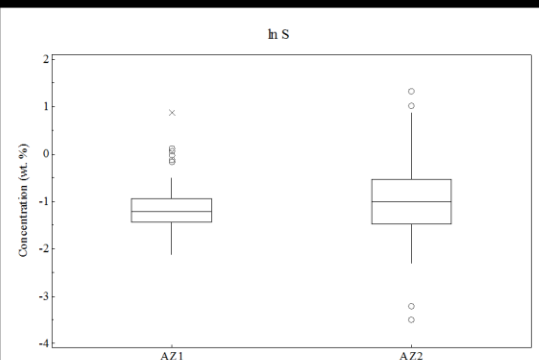
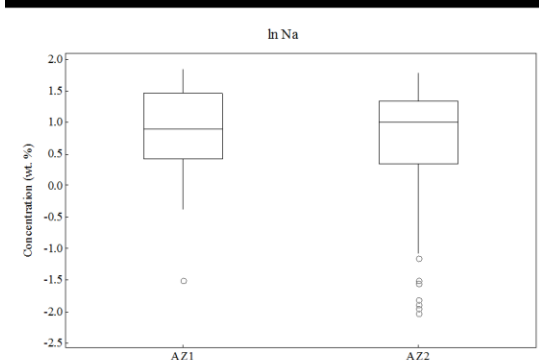
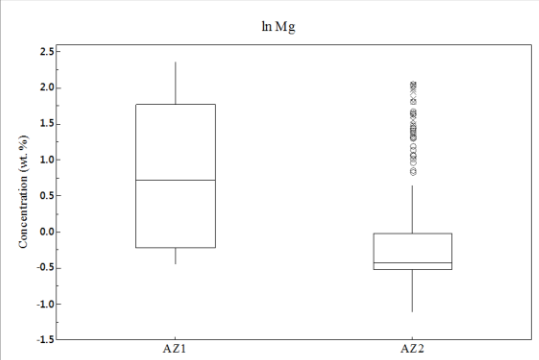
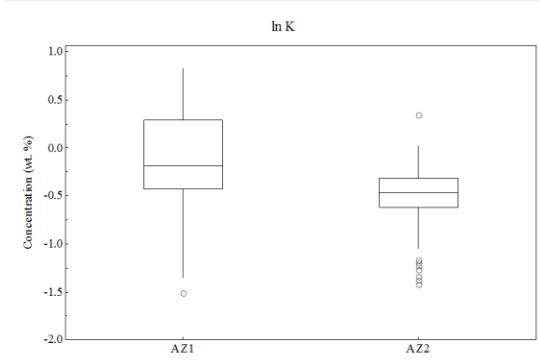
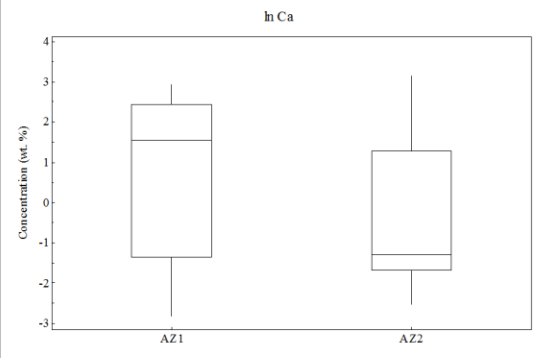
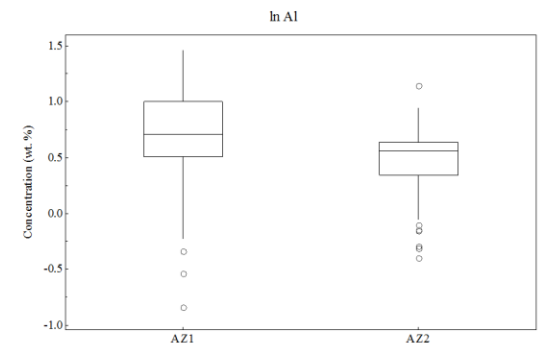




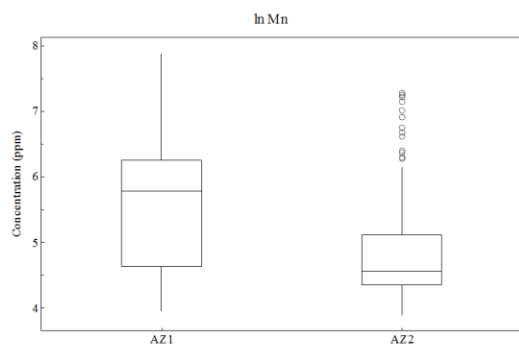
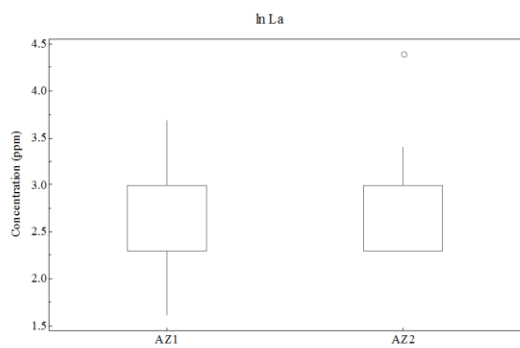
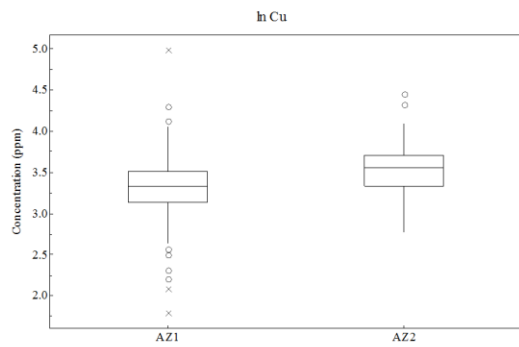
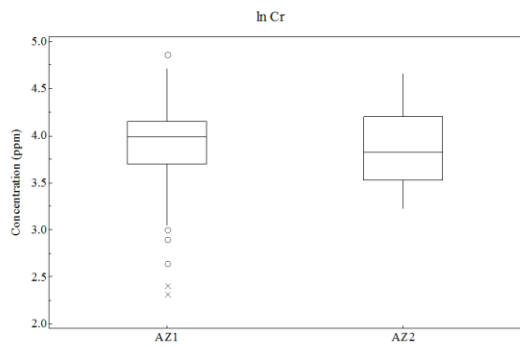
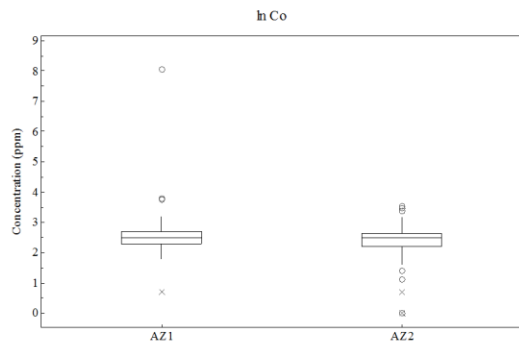
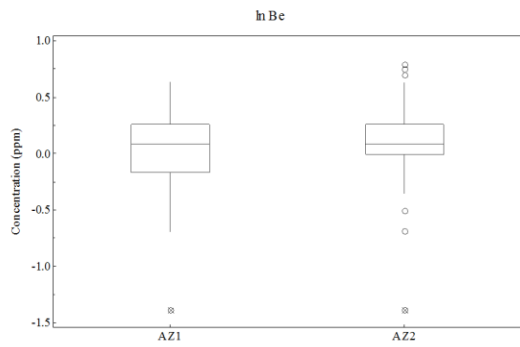
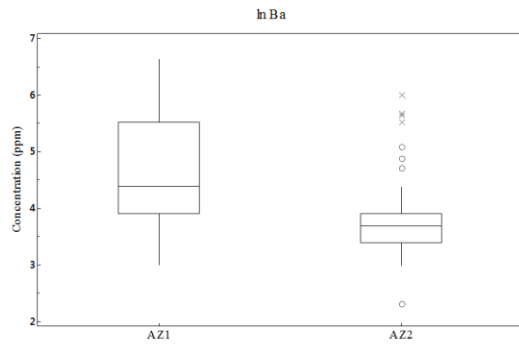
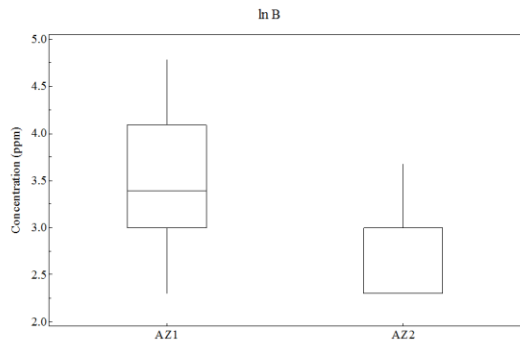
APPENDIX L  
Box Plots of Raw ICP-AES Data from AZ1 and AZ2

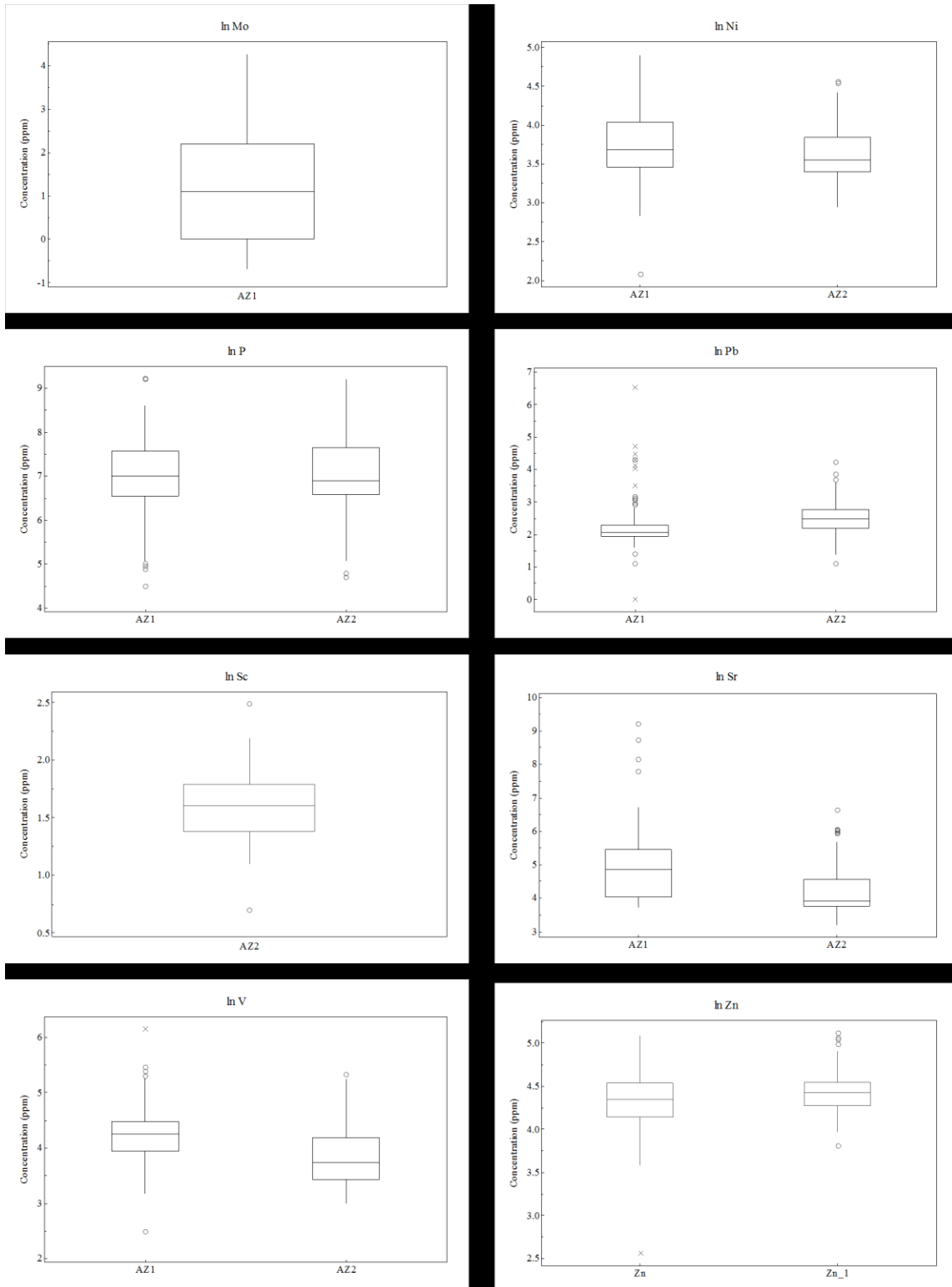


APPENDIX M  
Box Plots of Ln Transformed ICP-AES Data from AZ1 and AZ2









## REFERENCES

- Abed, A.M., Yasin, S., Sadaqa, R., Al-Hawari, Z., 2008. The paleoclimate of the eastern desert of Jordan during marine isotope stage 9. *Quat. Res.* 69, 458–468.  
<https://doi.org/10.1016/j.yqres.2008.02.006>
- Ahmad, K., Davies, C., 2017. Stable isotope ( $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$ ) based interpretation of organic matter source and paleoenvironmental conditions in Al-Azraq basin, Jordan. *Appl. Geochemistry* 78, 49–60. <https://doi.org/10.1016/j.apgeochem.2016.12.004>
- Ahmad, K.I., 2010. Organic chemistry of Al-Azraq basin, Jordan, an interpretation of paleoenvironment and paleoclimate using bulk organic matter.
- Ahmad, K.I., 2000. Biogeochemistry and Inorganic Geochemistry As Indicators of the Paleoenvironment and Paleohydrology of the Al-Azraq Basin.
- Ahrens, L.H., 1953. A fundamental law of geochemistry. *Nature* 172, 1148.  
<https://doi.org/10.1038/1721148a0>
- Andersson, P.O.D., Worden, R.H., Hodgson, D.M., Flint, S., 2004. Provenance evolution and chemostratigraphy of a Palaeozoic submarine fan-complex: Tanqua Karoo Basin, South Africa. *Mar. Pet. Geol.* 21, 555–577. <https://doi.org/10.1016/j.marpetgeo.2004.01.004>
- Bajjali, W., Al-Hadidi, K., 2006. Recharge origin, overexploitation, and sustainability of water resources in an arid area from Azraq basin, Jordan: Case study. *Nord. Hydrol.* 37, 277–292.  
<https://doi.org/10.2166/nh.2006.013>
- Bajjali, W., Al-Hadidi, K., 2005. Hydrochemical Evaluation of Groundwater in Azraq Basin, Jordan Using Environmental Isotopes and GIS Techniques. *Proceeding 25th Annu. ESRI Int. User Conf. San Diego Calif. July 25-29 2005* 1–19.
- Barberi, F., Capaldi, G., GA SPARINI, P., Marinelli, G., Santacroce, R., Scandone, R., Treuil, M., Varet, J., 1980. Recent basaltic volcanism of Jordan and its implications on the geodynamic hi-story of the Dead Sea shear zone.
- Bender, F., 1968. Geologie von Jordanien. *Beiträge zur Reg. Geol. der Erde* 7, 230.
- Burdon, D., 1959. Handbook of the geology of Jordan to accompany and explain the three sheets of the 1:250,000 geological map east of the rift by A.M. Quennell. Govt, of Jordan, 82 P.
- Cohen, A., 2003. Paleolimnology: the history and evolution of lake systems, *Choice Reviews Online*.  
<https://doi.org/10.5860/choice.41-3461>
- Cooksey, R.W., 2020. Illustrating Statistical Procedures: Finding Meaning in Quantitative Data. *Illus. Stat. Proced. Find. Mean. Quant. Data*. <https://doi.org/10.1007/978-981-15-2537-7>
- Das, S.K., Routh, J., Roychoudhury, A.N., Val Klump, J., 2008. Major and trace element geochemistry in Zeekoeflei, South Africa: A lacustrine record of present and past processes. *Appl. Geochemistry* 23, 2496–2511. <https://doi.org/10.1016/j.apgeochem.2008.02.011>
- Davies, C., 2007. Past Environments of the Jordan Plateau from the Paleolakes of the Eastern Desert. *Crossing Jordan North Am. Contrib. to Archaeol. Jordan*.
- Davies, C.P., 2005. Quaternary paleoenvironments and potential for human exploitation of the Jordan

- Plateau desert interior. *Geoarchaeology* 20, 379–400. <https://doi.org/10.1002/gea.20055>
- Davies, C.P., 2000. Reconstruction of Palaeoenvironments from Lacustrine Deposits of the Jordan Plateau.
- Davison, W., 1993. Iron and manganese in lakes. *Earth Sci. Rev.* 34, 119–163. [https://doi.org/10.1016/0012-8252\(93\)90029-7](https://doi.org/10.1016/0012-8252(93)90029-7)
- Deutsche Welle, n.d. Syrian refugees build a city in Jordan desert [WWW Document]. 2016. URL <https://p.dw.com/p/1JL19>
- Edgell, H.S., 2006. Arabian deserts: Nature, origin, and evolution. *Arab. Deserts Nature, Orig. Evol.* 1–590. <https://doi.org/10.1007/1-4020-3970-0>
- El-Naqa, A., Al-Momani, M., Kilani, S., Hammouri, N., 2007. Groundwater deterioration of shallow groundwater aquifers due to overexploitation in northeast Jordan. *Clean - Soil, Air, Water* 35, 156–166. <https://doi.org/10.1002/clen.200700012>
- Engstrom, D.R., Wright, H.E., 1984. Chemical stratigraphy of lake sediments as a record of environmental change. *Lake sediments Environ. Hist.* 11–67.
- Fedo, C.M., Wayne Nesbitt, H., Young, G.M., 1995. Unraveling the effects of potassium metasomatism in sedimentary rocks and paleosols, with implications for paleoweathering conditions and provenance. *Geology* 23, 921. [https://doi.org/10.1130/0091-7613\(1995\)023<0921:uteopm>2.3.co;2](https://doi.org/10.1130/0091-7613(1995)023<0921:uteopm>2.3.co;2)
- Filzmoser, P., Garrett, R.G., Reimann, C., 2005. Multivariate outlier detection in exploration geochemistry. *Comput. Geosci.* 31, 579–587. <https://doi.org/10.1016/j.cageo.2004.11.013>
- Filzmoser, P., Hron, K., Reimann, C., 2009a. Univariate statistical analysis of environmental (compositional) data: Problems and possibilities. *Sci. Total Environ.* 407, 6100–6108. <https://doi.org/10.1016/j.scitotenv.2009.08.008>
- Filzmoser, P., Hron, K., Reimann, C., Garrett, R., 2009b. Robust factor analysis for compositional data. *Comput. Geosci.* 35, 1854–1861. <https://doi.org/10.1016/j.cageo.2008.12.005>
- Filzmoser, P., Reimann, C., 2002. Robust Multivariate Methods in Geostatistics 429–436. [https://doi.org/10.1007/978-3-642-55991-4\\_46](https://doi.org/10.1007/978-3-642-55991-4_46)
- Filzmoser, P., Reimann, C., 1999. Normal and lognormal data distribution in geochemistry : death of a myth . Consequences for the statistical treatment of geochemical and environmental data. *Environ. Geol.* 39, 1001–1014.
- Frances, A., 2015. Jordan’s refugee crisis. Carnegie Endowment for International Peace. [WWW Document]. URL <http://carnegieendowment.org/2015/09/21/jordan-s-refugee-crisis-pub-61338>
- Grünfeld, K., 2005. Dealing with outliers and censored values in multi-element geochemical data - A visualization approach using XmdvTool. *Appl. Geochemistry* 20, 341–352. <https://doi.org/10.1016/j.apgeochem.2004.08.006>
- Grunsky, E.C., 2010. The interpretation of geochemical survey data. *Geochemistry Explor. Environ. Anal.* 10, 27–74. <https://doi.org/10.1144/1467-7873/09-210>
- Grunsky, E.C., 2006. The evaluation of geochemical survey data: data analysis and statistical methods

- using geographic information systems. *GIS Earth Sci. Geol. Assoc. Canada Spec. Publ.* 44, 229–283.
- Grunsky, E.C., Smee, B.W., 1999. The differentiation of soil types and mineralization from multi-element geochemistry using multivariate methods and digital topography. *J. Geochemical Explor.* 67, 287–299. [https://doi.org/10.1016/S0375-6742\(99\)00054-0](https://doi.org/10.1016/S0375-6742(99)00054-0)
- H., Wayne, Nesbitt, and, G., Markovics, 1997. Weathering of granodioritic crust, long-term storage of elements in weathering profiles, and petrogenesis of siliciclastic sediments. *Geochim. Cosmochim. Acta.*
- Hegazi F, 2016. Providing water for Syrian refugees in Jordan.
- HF., K., 1960. The application of electronic computers to factor analysis. *Educ. Psychol. Meas.* 20, 141–51.
- Hofste, R., Kuzma, S., Walker, S., Sutanudjaja, E., Bierkens, M., Kuijper, M., Faneca Sanchez, M., Van Beek, R., Wada, Y., Galvis Rodríguez, S., Reig, P., 2019. Aqueduct 3.0: Updated Decision-Relevant Global Water Risk Indicators. *WRI Publ.* <https://doi.org/10.46830/writn.18.00146>
- Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* 24, 417–441. <https://doi.org/10.1037/h0071325>
- Hugh R. Rollinson, 1993. *Using Geochemical Data: Evaluation, Presentation, Interpretation, Soil Science.* <https://doi.org/10.1097/00010694-199411000-00010>
- Ibrahim, K.M., 1993. The geological framework for the Harrat Ash-Shaam Basaltic Super-Group and its volcanotectonic evolution, *Bulletin* 25 33.
- Kaudse, T., 2014. Noble gases in groundwater of the Azraq Oasis, Jordan, and along the central Dead Sea Transform.
- Krumbein, W.C., Graybill, F.A., 1965. *An Introduction to Statistical Models in Geology.*
- Lewicki, P., Hill, T., 2006. *Statistics : Methods and Applications - A comprehensive reference for science, industry and data mining,* StatSoft Inc.
- Mackereth, F.J.H., 1966. Some chemical observations on post-glacial lake sediments. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 250, 165–213. <https://doi.org/10.1098/rstb.1966.0001>
- Mason, B.B., and Sons, J.W., 1958. *Principles of Geochemistry, Soil Science.* <https://doi.org/10.1097/00010694-195810000-00021>
- Mckee, E.D., Weir, G.W., 1953. Terminology for stratification and cross-stratification in sedimentary rocks. *Bull. Geol. Soc. Am.* 64, 381–390. [https://doi.org/10.1130/0016-7606\(1953\)64\[381:TFSACI\]2.0.CO;2](https://doi.org/10.1130/0016-7606(1953)64[381:TFSACI]2.0.CO;2)
- McLennan, S.M., Taylor, S.R., 1991. Sedimentary rocks and crustal evolution: tectonic setting and secular trends. *J. Geol.* 99, 1–21. <https://doi.org/10.1086/629470>
- Naylor, D., Al-Rawi, M., Clayton, G., Fitzpatrick, M.J., Green, P.F., 2013. Hydrocarbon potential in Jordan. *J. Pet. Geol.* 36, 205–236. <https://doi.org/10.1111/jpg.12553>
- Nesbitt, H.W., 1979. Mobility and fractionation of rare earth elements during weathering of a granodiorite. *Nature* 279, 206–210. <https://doi.org/10.1038/279206a0>

- Nesbitt, H.W., Markovics, G., Price, R.C., 1980. Chemical processes affecting alkalis and alkaline earths during continental weathering. *Geochim. Cosmochim. Acta* 44, 1659–1666. [https://doi.org/10.1016/0016-7037\(80\)90218-5](https://doi.org/10.1016/0016-7037(80)90218-5)
- Nesbitt, H.W., Young, G.M., 1989. Formation and diagenesis of weathering profiles. *J. Geol.* 97, 129–147. <https://doi.org/10.1086/629290>
- Nesbitt, H.W., Young, G.M., 1984. Prediction of some weathering trends of plutonic and volcanic rocks based on thermodynamic and kinetic considerations. *Geochim. Cosmochim. Acta* 48, 1523–1534. [https://doi.org/10.1016/0016-7037\(84\)90408-3](https://doi.org/10.1016/0016-7037(84)90408-3)
- Nesbitt, H.W., Young, G.M., 1982. Early proterozoic climates and plate motions inferred from major element chemistry of lutites. *Nature* 299, 715–717. <https://doi.org/10.1038/299715a0>
- Pearson, K., 1901. LIII. On lines and planes of closest fit to systems of points in space. London, Edinburgh, Dublin Philos. Mag. J. Sci. 2, 559–572. <https://doi.org/10.1080/14786440109462720>
- Powell, J.H., 1989. Stratigraphy and Sedimentation of the Phanerozoic rocks in Central and South Jordan. *Geol. Mapp. Div. Bull.* 11b 130.
- Pukelsheim, F., 1994. The Three Sigma Rule. *Am. Stat.* 48, 88. <https://doi.org/10.2307/2684253>
- Rawlins, B.G., Lister, T.R., Mackenzie, A.C., 2002. Trace-metal pollution of soils in northern England. *Environ. Geol.* 42, 612–620. <https://doi.org/10.1007/s00254-002-0564-5>
- Reimann, C., Filzmoser, P., Garrett, R.G., 2005. Background and threshold: Critical comparison of methods of determination. *Sci. Total Environ.* 346, 1–16. <https://doi.org/10.1016/j.scitotenv.2004.11.023>
- Reimann, C., Filzmoser, P., Garrett, R.G., 2002. Factor analysis applied to regional geochemical data: Problems and possibilities. *Appl. Geochemistry* 17, 185–206. [https://doi.org/10.1016/S0883-2927\(01\)00066-X](https://doi.org/10.1016/S0883-2927(01)00066-X)
- Reimann, C., Filzmoser, P., Garrett, R.G., Dutter, R., 2008. *Statistical Data Analysis Explained. Stat. Data Anal. Explain.* <https://doi.org/10.1002/9780470987605>
- Rock, N.M.S., 1988. Summary statistics in geochemistry: A study of the performance of robust estimates. *Math. Geol.* 20, 243–275. <https://doi.org/10.1007/BF00890256>
- Rollefson, G., Quintero, L., Low, R., Schnurrenberger, D., Watson, R., 1996. Paleolithic Hunters in the Azraq Oasis. *Biblic. Archaeol.* 59, 239–240. <https://doi.org/10.2307/3210567>
- Rollinson, H.R., 2014. *Using Geochemical Data. Using Geochemical Data.* <https://doi.org/10.4324/9781315845548>
- Roy, P.D., Smykatz-Kloss, W., Morton, O., 2008. Geochemical zones and reconstruction of late Holocene environments from shallow core sediments of the Pachapadra paleo-lake, Thar Desert, India. *Chemie der Erde* 68, 313–322. <https://doi.org/10.1016/j.chemer.2006.01.006>
- Roy, P.D., Smykatz-Kloss, W., Sinha, R., 2006. Late Holocene geochemical history inferred from Sambhar and Didwana playa sediments, Thar Desert, India: Comparison and synthesis. *Quat. Int.* 144, 84–98. <https://doi.org/10.1016/j.quaint.2005.05.018>
- Royston, J.P., 1982. An Extension of Shapiro and Wilk's W Test for Normality to Large Samples. *Appl.*

- Stat. 31, 115. <https://doi.org/10.2307/2347973>
- Schnurrenberger, D., Russell, J., Kelts, K., 2003. Classification of lacustrine sediments based on sedimentary components. *J. Paleolimnol.* 29, 141–154. <https://doi.org/10.1023/A:1023270324800>
- Shapiro, S.S., Wilk, M.B., Chen, H.J., 1968. A Comparative Study of Various Tests for Normality. *J. Am. Stat. Assoc.* 63, 1343. <https://doi.org/10.2307/2285889>
- Singer, D.A., Kouda, R., 2001. Some simple guides to finding useful information in exploration geochemical data. *Nat. Resour. Res.* 10, 137–147. <https://doi.org/10.1023/A:1011552810482>
- Taylor, S.R., McLennan, S.M., 1985. The Continental Crust: its Composition and Evolution. An Examination of the Geochemical Record Preserved in Sedimentary Rocks. *Cont. Crust its Compos. Evol. An Exam. Geochemical Rec. Preserv. Sediment. Rocks.*
- Templ, M., Filzmoser, P., Reimann, C., 2008. Cluster analysis applied to regional geochemical data: Problems and possibilities. *Appl. Geochemistry* 23, 2198–2213. <https://doi.org/10.1016/j.apgeochem.2008.03.004>
- TRYON, R.C., 2009. Cluster Analysis: Correlation Profile and Orthometric (factor) Analysis for the Isolation of Unities in Mind and Personality. Edwards Brother. Inc. lithoprinters Publ. 122.
- Tukey, J.W., 1977. Box-and-Whisker Plots. *Explor. Data Anal.* 39–43.
- Turner, B.R., Makhlof, I., 2005. Quaternary sandstones, northeast Jordan: Age, depositional environments and climatic implications. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 229, 230–250. <https://doi.org/10.1016/j.palaeo.2005.06.024>
- Wedepohl, K.H., 1969. *Handbook of Geochemistry*. Springer.
- Wei, G., Li, X.H., Liu, Y., Shao, L., Liang, X., 2006. Geochemical record of chemical weathering and monsoon climate change since the early Miocene in the South China Sea. *Paleoceanography* 21. <https://doi.org/10.1029/2006PA001300>
- White, A.F., Blum, A.E., 1995. Effects of climate on chemical weathering in watersheds. *Geochim. Cosmochim. Acta* 59, 1729–1747. [https://doi.org/10.1016/0016-7037\(95\)00078-E](https://doi.org/10.1016/0016-7037(95)00078-E)
- Wilk, M.B., Shapiro, S.S., 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52, 591–611.
- Yanes, C., Alvarez, H., Jaffé, R., 2006. Geochemistry of a tropical lake (Lake Leopoldo) on pseudo-karst topography within the Roraima Group, Guayana Shield, Venezuela. *Appl. Geochemistry* 21, 870–886. <https://doi.org/10.1016/j.apgeochem.2006.01.003>

## CURRICULUM VITAE

### Jennifer Knipper

---

Department of Geosciences  
University of Missouri – Kansas City

#### Work Experience

September 2019-Present	Staff Geologist Geotechnical Consultants, Inc. Lake Forest, CA
August 2017-September 2019	Project Geologist ENGEQ, Inc. Irvine, CA

#### Internship

June 2014-June 2016	United States Geological Survey Student Services Contractor Position USGS - Missouri Water Science Center Lee's Summit, MO
---------------------	--

#### Education

December 2012	University of Missouri Kansas City Major 1: Biology, B.S. Major 2: Geology, B.S. Minor: Chemistry
---------------	--

#### Abstracts and Poster Presentations

Knipper, J. and **Davies, C.** 2015. Quantitative analysis of major and trace elements of lacustrine sediment: a comparison of two marginal environments within the Al-Azraq Basin, Jordan using inductively coupled plasma mass spectrometry (ICPMS). Association of American Geographers, Chicago, IL. April 23, 2015.

Hirsch, J., **Knipper, J.**, Ahmad, K., and Davies, C. 2013. Characterizing paleolake environments of the Al-Azraq Basin, Jordan using stable isotopes, ICPMS, scanning electron microscopy, and magnetic susceptibility. Association of American Geographers, Los Angeles, CA April 9-13, 2013.

Hirsch, J., **Knipper, J.**, Lima, O., Petrillo, A., Smith, A., and Davies, C., 2012. Evidence of past environmental conditions during the evolution of an ancient saline lake in the Al-Azraq Basin, Jordan using sedimentology and geochemistry. Geological Society of America Abstracts with Programs, Vol. 44, No. 7, p. 438.

#### Professional Memberships

2014- Present	Association of Environmental and Engineering Geologists
2013- Present	Association for Women Geoscientists
2013- Present	Association of American Geographers
2012- Present	Geological Society of America

#### Scholarships, Research Grants, and Awards

2014	University of Missouri Women's Graduate Assistance Fund Grant \$2000
2014	Richard J. Gentile Scholarship \$1600
2014	The Association of Earth Science Clubs of Greater Kansas City Scholarship \$1000
2013	AAG Paleoenvironmental Student Poster Presentation of the Year 2013
2010	Stephen Bufton Memorial Education Fund Grant \$2000
2009	Stephen Bufton Memorial Education Fund Grant \$1500



