Using Kriging, Cokriging, and GIS to Visualize Fe and Mn in Groundwater

Crystal Johnson
East Tennessee State University

Follow this and additional works at: http://dc.etsu.edu/etd

Part of the Environmental Monitoring Commons, Geochemistry Commons, Geographic Information Sciences Commons, Geology Commons, Hydrology Commons, and the Water Resource Management Commons

Recommended Citation

This Thesis - Open Access is brought to you for free and open access by Digital Commons @ East Tennessee State University. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ East Tennessee State University. For more information, please contact digilib@etsu.edu.
Using Kriging, Cokriging, and GIS to Visualize Fe and Mn in Groundwater

A thesis
presented to
the department of Geosciences
East Tennessee State University

in partial fulfillment
of the requirements for the degree
Masters of Science in Geosciences
with a concentration in Geospatial Analysis

by
Crystal Deanne Johnson
May 2015

Dr. Arpita Nandi, Chair
Dr. Timothy Andrew Joyner
Dr. Ingrid Luffman

Keywords: GIS, Kriging, Cokriging, Geospatial, Fe, Mn, Groundwater
ABSTRACT

Using Kriging, Cokriging, and GIS to Visualize Fe and Mn in Groundwater

by

Crystal Deanne Johnson

For aesthetic, economic, and health-related reasons, allowable concentrations of iron (Fe) and manganese (Mn) found present in drinking water are 0.3 mg/L and 0.05 mg/L, respectively. Water samples taken from private drinking wells in the rural communities within Buncombe County, North Carolina contain amounts of these metals in concentrations higher than the suggested limits. This study focused on bedrock geology, elevation, saprolite thickness, and well depth to determine factors affecting Fe and Mn. Using ArcGIS 10.2, spatial trends in Fe and Mn concentrations ranges were visualized, and estimates of the metal concentrations were interpolated to unmonitored areas. Results from this analysis were used to create a map that delineates the actual spatial distribution of Fe and Mn. The study also established a statistically significant correlation between Fe and Mn concentrations, which can be attributed to bedrock geology. Additionally, higher Fe in groundwater was concentrated in shallower wells and valley areas.
DEDICATION

This thesis is dedicated to my son, Lleyton, my mom, Judy, and my dad, Dean. You three, with the help of some of my own personal reasons, are the source of my crazy inspiration to attempt graduate school as a single mother. It was tough at times, and I sincerely thank each you for your support and guidance. I also want to dedicate this thesis to my advisor, Dr. Arpita Nandi; without you the entire idea would not have come to fruition, and I give you a sincere thank you for your kindness, patience and expertise.
ACKNOWLEDGEMENTS

Getting through graduate school and writing this thesis would not have been possible without the kind and loving support of my immediate and extended family. Thank you to my 9-year-old son for having the patience of an adult and giving me the space to complete a dream; I love you!! Thank you to my parents for always supporting me through every endeavor I have ever attempted; I can never express to you how important you are to me and how much I love you all!

To my school family: thank you! Thank you to Amy and Noah for always giving me a place to crash and always putting the fuzzy sheet on the inside of the covers. To Selena, my sister from Bristol, we really can do this! Thank you for always being there for Earth Fare chocolate runs, hanging out with me in the computer lab, and lending a listening ear and allowing me to confide in you. Thank you Jeremy for being my very first friend at ETSU on day one, and for staying up until 1am in the GEL to work on rough drafts; it would not have been fun without you! I cannot forget my carpool buddies. Nan and Billy, I will always remember sunset pictures, laughing loud and often, and, of course, roadside donut stands. (I will let you all know when it is open!)

Thank you to my committee members, Dr. Andrew Joyner and Dr. Ingrid Luffman. Your patience, encouragement, donuts, and coffee were just what I needed to get through writing this thesis. Most of all, a huge thank you goes to Dr. Arpita Nandi; your kindness, encouragement, and everlasting patience has inspired me in more ways than just writing a thesis. I am very blessed to have been placed around so many loving, supportive people; this process would have been so much more challenging without all of you with me!
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>2</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>4</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>5</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>8</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>9</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>10</td>
</tr>
<tr>
<td>Economic, Aesthetic, and Health-Related Reasons to Monitor Mn and Fe</td>
<td>11</td>
</tr>
<tr>
<td>Factors that Control Fe Concentrations in Groundwater</td>
<td>12</td>
</tr>
<tr>
<td>Factors that Control Mn Concentrations in Groundwater</td>
<td>13</td>
</tr>
<tr>
<td>Previous Studies Involving Fe and Mn in Groundwater</td>
<td>14</td>
</tr>
<tr>
<td>Geostatistical Prediction Models Used in Groundwater Contaminant Studies</td>
<td>15</td>
</tr>
<tr>
<td>Kriging and Cokriging</td>
<td>16</td>
</tr>
<tr>
<td>Model Cross-Validation and Validation</td>
<td>18</td>
</tr>
<tr>
<td>Research Goals</td>
<td>19</td>
</tr>
<tr>
<td>2. IRON AND MANGANESE IN GROUNDWATER: USING KRIGING AND GIS TO LOCATE HIGH CONCENTRATIONS IN BUNCOMBE COUNTY, NC</td>
<td>21</td>
</tr>
<tr>
<td>Abstract</td>
<td>21</td>
</tr>
<tr>
<td>1.0 Introduction</td>
<td>22</td>
</tr>
<tr>
<td>1.1 Background Information</td>
<td>22</td>
</tr>
<tr>
<td>1.2 Fe and Mn in North Carolina</td>
<td>24</td>
</tr>
<tr>
<td>1.3 Objective of study</td>
<td>26</td>
</tr>
<tr>
<td>2.0 Study Area</td>
<td>26</td>
</tr>
<tr>
<td>3.0 Research Methods</td>
<td>30</td>
</tr>
<tr>
<td>3.1 Data collection</td>
<td>31</td>
</tr>
<tr>
<td>3.2 Geocoding</td>
<td>33</td>
</tr>
<tr>
<td>3.3 Exploratory Non-spatial Statistics</td>
<td>34</td>
</tr>
<tr>
<td>3.4 Exploratory Spatial Statistics</td>
<td>35</td>
</tr>
<tr>
<td>3.5 Spatial statistics - Kriging and Cokriging</td>
<td>36</td>
</tr>
<tr>
<td>4.0 Results</td>
<td>37</td>
</tr>
<tr>
<td>4.1 Location of Fe and Mn Contaminated Wells</td>
<td>37</td>
</tr>
</tbody>
</table>
4.2 Exploratory Non-spatial Statistics Results .................................................................40
4.3 Exploratory Spatial Statistics Results .................................................................41
5.0 Kriging/Cokriging Results .................................................................................42
5.1 Kriging/Cokriging with Fe data .........................................................................43
5.2 Fe – Validation with 30% Test Dataset ...............................................................46
5.3 Mn Kriging and Cokriging .................................................................................47
5.4 Mn – Validation with 30% Test Dataset ...............................................................49
5.5 Percent Area Exceeding Acceptable Limit ........................................................50
6.0 Discussion ..............................................................................................................52
6.1 Exploratory Non-Spatial and Spatial Statistics ..................................................52
6.2 Final Spatial Statistics and Prediction Maps ......................................................53
6.3 Limitations ............................................................................................................61
7.0 Conclusion .............................................................................................................62
8.0 References .............................................................................................................63
3. DISCUSSION AND MAJOR FINDINGS .................................................................68
   Study Limitations ..................................................................................................70
   Future Research ....................................................................................................71
REFERENCES ..............................................................................................................72
VITA ..............................................................................................................................79
LIST OF TABLES

Table 2.1: Summary of data types used in the study ..............................................................33
Table 2.2: Descriptive statistics results for primary variables and covariates .........................40
Table 2.3: Global Moran’s I test statistic and significance for Fe and Mn .................................42
Table 2.4: Accuracy metrics produced from the kriging and cokriging of Fe and covariates .........................................................................................................................................44
Table 2.5: Accuracy metrics created from 30% testing Fe dataset validation ............................46
Table 2.6: Accuracy metrics produced from the kriging and cokriging of Mn and covariates .............................................................................................................................................49
Table 2.7: Accuracy metrics created from 30% testing Mn dataset validation ..........................50
Table 2.8: Percent area exceeding acceptable Fe concentrations ..............................................51
Table 2.9: Percent area exceeding acceptable Mn concentrations ..............................................51
LIST OF FIGURES

Figure 1.1: Semivariogram example exhibiting range and nugget ...........................................17

Figure 2.1: Location map of Buncombe County, in Western North Carolina, USA. The study area is located in the eastern United States (A) in North Carolina (B) in the western portion of the state in Buncombe County (C) ..................................................27

Figure 2.2: Geologic map of Buncombe County, NC ................................................................28

Figure 2.3: Fe and Mn research project steps flow chart ..........................................................31

Figure 2.4: Fe contaminated well locations .............................................................................38

Figure 2.5: Mn contaminated well locations .............................................................................39

Figure 2.6: GeoDa local cluster analysis results. LISA cluster map of Fe concentrations (A) and LISA significance map of Fe concentrations (B), LISA cluster map of Mn (C) and LISA significance map of Mn (D) .............................................................................42

Figure 2.7: Kriging/cokriging surfaces for Fe and covariates: Fe (A), Fe/saprolite (B), Fe/well depth (C), Fe/geology (D), Fe/elevation (E), Fe/geology/elevation (F), Fe/saprolite/elevation (G), Fe/geology/saprolite (H) all shown in concentrations of mg/L ................................................................................................................45

Figure 2.8: Kriging/cokriging surfaces for Mn and covariates: Mn (A), Mn/saprolite (B), Mn/well depth (C), Mn/geology (D), Mn/elevation (E), Mn/geology/elevation (F), Mn/saprolite/elevation (G), Mn/geology/saprolite (H) all shown in concentrations of mg/L ........................................................................................................48

Figure 2.9: Final prediction map of Fe concentrations cokriged with well depth ......................56

Figure 2.10: Standard error prediction map of Fe/well depth cokriged surface ......................57

Figure 2.11: Comparison of 1961 Fe concentration data with Fe/well depth cokriged surface ...............................................................................................................................................58

Figure 2.12: Final prediction map of Mn concentrations cokriged with geology ....................60

Figure 2.13: Standard error prediction map of Mn/geology cokriged surface .........................61
Iron (Fe) and manganese (Mn) are naturally occurring metals found in groundwater which originate from the weathering of Fe and Mn-rich bedrock (ATSDR 2012). Both minerals are trace elements and in the proper dosage, are vital for the production of necessary processes within the body, but at high and unsafe doses can cause serious health implications. Identifying and monitoring spatial areas of high concentrations of Mn and Fe in groundwater is important for economic, aesthetic, and adverse health-related reasons. With the establishment of the 1974 Safe Drinking Water Act (SDWA), the United States Environmental Protection Agency (USEPA) assumed responsibility for developing standards and monitoring water quality in public drinking water systems (Tiemann 2014). USEPA was tasked with creating unenforceable guidelines Maximum Contaminant Level Goals (MCLGs) for various constituents in drinking water that either pose an immediate threat to human life, or cause aesthetic or economic negative impacts on drinking water piping systems or fixtures. Furthermore, USEPA divides water quality contaminants into two categories: primary and secondary, and sets enforceable Maximum Contaminant Levels (MCLs) for these two categories (Tiemann 2014; USEPA 2013a). Secondary maximum contaminant levels (SMCLs) set forth by USEPA dictate Fe concentrations exceeding 0.3 mg/L and Mn concentrations exceeding 0.05 mg/L will become insoluble and precipitate from solution, causing aesthetic, economic, and health related issues.
Fe and Mn are essential elements required for proper body function, are natural constituents in food, and are therefore absorbed in some concentration daily (WHO 2011; Farina et al. 2013). However, under or overexposure to these metals can cause health related issues. The brain is very susceptible to excess amounts of Fe and Mn and different neurodegenerative diseases may occur (Spangler and Spangler 2010). Occasionally, Fe deficiency causes an increase in Mn uptake leading to excess Mn deposition in the brain. The transition between ferric (Fe III) and ferrous (Fe II) oxidative states in a body’s cellular stage can cause neuronal oxidative stress disorder. Degenerative diseases from excess Fe in the human system present symptoms similar to Parkinson’s disease, Alzheimer’s disease, and Huntington’s disease (Farina et al. 2013).

Inhalation of dust containing high concentrations of Mn induces a neurotoxic effect resembling Parkinson’s disease-like symptoms and is referred to as manganism (Farina et al. 2013; Bouchard et al. 2007). The health effects of inhaling Mn-containing dust are well documented whereas the health effects of ingesting Mn orally is still under investigation. In contrast, some cases have been documented where high levels of soluble Mn from irrigation water contaminated with sewage showed decreased performance of short-term memory and visual identification (Bouchard et al. 2007). Ingestion of high concentrations of Mn has also been correlated to hyperactivity and learning disability in small children (Graziano 2010), and infant mortality (Spangler and Spangler 2010). In addition, pregnant women with late term anemia have a tendency to absorb greater amounts of Mn due to the absence of Fe in the body, affecting the fetus (Spangler and Spangler 2010).

Fe and Mn contaminants present in concentrations higher than recommended alter the aesthetic or mechanical properties of potable water. Aesthetic effects are those that affect the taste, odor, or color quality of the drinking water. Mechanical effects are those that cause
physical damage to piping and distribution systems. Mn, when oxidized, produces a black precipitate that forms on piping and fixtures often when the metal ion is present in concentrations greater than 0.05 mg/L. Fe, similar to Mn, when present in concentrations greater than 0.3 mg/L, can oxidize to the insoluble form to create an orange-red precipitate that also stains fixtures and piping systems. Over time, these precipitates can build up in the pipelines, reducing the quantity of the water supply. Additionally, aggressive chemical weathering of water distribution systems can release insoluble Fe ions and create an excess of Fe in drinking water (WHO 2006). Thus, when both Fe and Mn are present in elevated levels, special considerations should be taken to locate the sources and spatial distribution of these metals in groundwater.

Factors that Control Fe Concentrations in Groundwater

The presence of Fe is controlled by both natural and anthropogenic factors. Fe, the fourth most abundant element in the earth’s crust, migrates into the groundwater due to the chemical weathering of iron-bearing minerals like ferromagnesian silicates and sulfide minerals such as olivine, pyroxenes, amphiboles, and iron-containing micas, magnetite, ilmenite, and pyrite (USEPA 2013b). In groundwater, Fe primarily remains in two oxidation states; ferric or Fe (III) and ferrous or Fe (II). The solubility of Fe (III) and Fe (II) depends on pH and Eh of the local groundwater aquifer. Fe (III) has less solubility in aerated water (high dissolved oxygen creating a high Eh situation) and at neutral to alkaline conditions when the pH is at least 7. Fe (II) is highly soluble and mobile in acidic, anoxic water, i.e., the groundwater has pH less than 7 and is low in dissolved oxygen. Iron species can also become more soluble when paired with organic ligands, forming complexes, and iron hydroxide may disperse as a colloid in water (Hem 1972).
Fe present in the environment as a result of human activities can be attributed to a variety of sources. The significant contribution of Fe in the environment is due in part to the availability of Fe-containing wastes from mining processes. Bauxite and iron ore mining wastes often contain high concentrations of Fe that is donated to the environment. These processes convert raw Fe-containing minerals to industry products. In addition, wastes from steel production also contain Fe. Sulfide-bearing rock, exposed as a result of disturbance from construction and mining activities, can produce sulfidic acid when the exposed rock comes in contact with groundwater or precipitation. This process creates acid mine drainage. The acidic water favors the chemical release of Fe into the environment. Not all Fe comes from the process of production of iron metal ores; some Fe in the environment is contributed by the rusting and weathering of exposed manufactured Fe products (USEPA 1978).

Factors that Control Mn Concentrations in Groundwater

Mn is found in 0.1% of the earth’s crust (WHO 2011). It does not naturally occur in the environment as a pure metal, but instead, exists as a chemical component in various Fe-bearing minerals. Some of the most common minerals are: ferromagnesian micas, pyrolusite, rhodocrosite, olivine, rhodonite, hausmannite. Additionally, Mn is found in clastic (shale), chemical (limestone and dolostone) and metasedimentary rocks and soils (Briel 1997; Chapman et al. 2013). Mn can substitute for Fe, Mg (magnesium), and Ca (calcium) in many common minerals (Briel 1997) and is released into the groundwater from the chemical weathering of the bedrock. In comparison to Fe, Mn is considerably more stable in a soluble condition for a wide range of pH-Eh conditions (Hem 1972; Howe et al. 2004). Mn exists in soluble states as Mn (II) or Mn (IV). Mn (II) is more commonly present in groundwater because of the lower pH and
redox potential found in deeper groundwater. Furthermore, microbes can also have an effect on the conversion of soluble Mn (II) to Mn (IV).

Mn found in the environment due to human activity can be traced to byproducts of the smelting process and industrial mining. However, no substantial amount of Mn ore has been mined in the United States since 1978; therefore, most Mn ore used in industrial production is imported (Corathers 2007). Ferromanganese is used in steel production to improve the structural qualities of the steel. Mn compounds are also used in dry cell batteries, plant fertilizers, animal feeds, matches, fireworks and as a glass bonding agent (ATSDR 2012).

Previous Studies Involving Fe and Mn in Groundwater

Groundwater concentrations of Mn and Fe have been studied in various parts of the world. Kondikas et al. (1989) observed high concentrations of Mn found in drinking water in areas of Greece. In Central Mexico, Santos-Burgoa et al. (2000) began analyses on the correlation between Mn exposure and health effects in the general population. Katsoyiannis and Zouboulis (2006) studied the coexistence of arsenic (As), Fe, and Mn in groundwater as well as oxidative states of the metals to determine the most efficient means of removal for remediation of groundwater. Several other studies emphasized the coexistence of Fe, Mn, and As in groundwater and surface water as a result of the regional geology (Klein and Hurlbut 1985; Ayotte et al. 1999; Hinkle and Polette 1999), groundwater chemistry (Anderson and Bruland 1991), and seasonal variation (Agbaire et al. 2009; Farmer et al. 2000). A handful of studies examined metal concentrations with ground residence time (Edmunds and Smedley 2000) and groundwater recharge-discharge areas (Richardson et al. 1992). Studies found that the saprolite layer above the bedrock could be a contributor of Fe oxides in several areas of the United States.
(Anand and Paine 2002). Lytle et al. (1994) documented the existence of significant Mn concentrations along Utah roadways as a result of air pollution from Mn additives in gasoline. The North Carolina Geological Survey headed by Jeffrey Reid (1993) documented the occurrence of Fe and Mn in streams in North Carolina as a component in the National Uranium Resource Evaluation (NURE). More groundwater quality studies in NC include work done by Spangler and Reid (2010) who examined groundwater Mn concentrations in well water and found associations between these concentrations and increased infant and cancer-related mortality rates. In addition, Spangler (2012) identified a positive correlation between high concentrations of air-borne Mn off- and on-road and high liver disease mortalities among residents of NC at the county level. Polizzotto et al. (2013) noted high levels of Mn concentrations in the NC Piedmont area associated with bedrock weathering and groundwater redox conditions; however, they realized the need of geospatial analyses to better understand the causes for spatial variability in well water concentrations, and Giese et al. (1986) indicated the presence of Fe and Mn in crystalline aquifers of North Carolina.

**Geostatistical Prediction Models Used in Groundwater Contaminant Studies**

Geostatistical analysis using Geographic Information Systems (GIS) has been employed extensively in multiple studies to determine groundwater quality in disparate areas. Specifically interpolation, a primary form of Geostatistical analysis, is often used to predict values in unobserved locations based on the values in observed locations (Zhu et al. 2001; He and Jia 2004; Hu et al. 2005; Lui et al. 2006; Ahmadi and Sedghamiz 2007; Sanders et al. 2012;). Nas (2009) used Ordinary Kriging (OK), a type of interpolation within the Geostatistical Analyst extension in ArcGIS Desktop to determine spatial patterns of water quality in rural areas in
Turkey that rely on groundwater for drinking water, and Hu et al. (2005) also used OK to track nitrate-contaminated groundwater in the North China Plain. Lui et al. (2006) tracked heavy metal concentrations in soil in rice paddy fields in China using interpolation methods. Additionally, GIS techniques using kriging and other forms of interpolation mapped high radon levels in bedrock in southern Belgium (Zhu et al. 2001). He and Jia (2004) used geostatistical methods to track and interpolate fuel additives in the groundwater in California. Sanders et al. 2012 used OK to determine arsenic (As) distribution in groundwater in North Carolina.

**Kriging and Cokriging**

Kriging is a geostatistical technique that is used to interpolate a surface from a scattered set of known points (X,Y plots) in which a continuous surface of values can be predicted between the known locations. For example, creating a continuous surface of water quality contaminant concentration values based on measured concentrations at specific points. Kriging is based on an assumption that spatial autocorrelation exists among the measured points in a given dataset. The method statistically weights the values from data points clustered around a central primary variable in order to predict a continuous surface that estimates unmeasured values at other locations (Oliver and Webster 1990). The statistical weights are derived from a semivariogram, which is a model that assesses the degree of spatial correlation as a function of distance between the data points (Cressie 1985; Cressie 1991; Salih et al. 2002; Ahmadi and Sedghamiz 2007). Spatial autocorrelation is quantified by the positions (coordinates) of observation points and the correlation between observations, by calculating the differences squared of the variable values. If positive spatial autocorrelation is present, data points located closer together will have smaller squared differences which is indicated by the placement of the data points in the semivariogram.
(Figure 1.1). The range signifies the distance at which the data points will no longer be autocorrelated; data points closer to the range exhibit stronger autocorrelation and data points farther away become less, or not, autocorrelated. The nugget represents the measurement error among the data points.

![Semivariogram example exhibiting range and nugget](image)

Figure 1.1: Semivariogram example exhibiting range and nugget

This geostatistical tool is used to determine the autocorrelation among the values in a dataset described by clusters of values with comparable values of differences squared (Eldeiry 2012). Once autocorrelation is confirmed, the kriging process creates an interpolated surface for unmeasured values by determining a weighted average of the distances between neighboring values (Eldeiry 2012). However, the weighted average is determined by the semivariogram and affects the way the weights are assigned to each set of data points in a cluster (Salih et al. 2002). In addition, kriging also provides an estimate of uncertainty in the prediction surface.
Several different types of kriging have been developed with each method serving to effectively produce a predicted continuous surface per each unique dataset. These methods include: ordinary, simple, universal, indicator, probability and disjunctive (Goovaerts 1997). The general mathematical expression of kriging is described as:

\[ Z(s) = \mu(s) + \epsilon(s) \]

where the mean \( \mu(s) \) along with the errors \( \epsilon(s) \) compose the deterministic trend \( Z(s) \) in which \( \mu(s) \) and \( \epsilon(s) \) are determined at each location \( s \) of each variable (Goovaerts 1997; Joyner 2013). For each different method of kriging, the general kriging formula is modified to account for each of the different types of surfaces that the model formula predicts (Goovaerts 1997).

Cokriging is also an interpolation technique that relies on the same premise as kriging; the technique uses autocorrelation and cross correlation to create an interpolated surface that predicts values at unmeasured locations. Additionally, cokriging includes a secondary variable in the interpolation model (Queiroz et al. 2008), and assumes that some autocorrelation exists between a primary and secondary variable; stronger the autocorrelation among the multivariates results in greater accuracy for the prediction of the primary variable in the cokriging model (Salih, et al., 2002). Using two or more variables, cokriging produces the best model based on eliminating bias between the estimated value and the true value and minimizing the variance among estimations. Using cokriging, the accuracy of the interpolated surface can be increased when compared to univariate kriging.

Model Cross-Validation and Validation

To validate the interpolated surfaces produced from the kriging and cokriging models, two methods of quality assurance and quality control are performed; cross-validation and validation.
Cross-validation is performed with the full (100%) dataset using an \( n-1 \) method, while validation is performed using subsets of the data for model training and testing purposes. Geostatistical Analyst toolbar in ArcGIS Desktop offers automatic cross-validation of the kriging and cokriging models created when the model parameters are optimized using the Geostatistical Wizard. Cross-validation is the process of systematically removing each point in the data, predicting the value at the surface in which the point is missing, and then comparing the predicted value to the actual value. Accuracy metrics are produced in the cross-validation process and are used to determine the efficacy and accuracy of the model. Cross-validation is performed on the entire dataset in every model produced.

The validation method involves dividing the dataset into two unequal subsets; the training dataset contains the majority of the data, while the test dataset contains the remaining data points. The training dataset is used to develop the kriging and cokriging model, and accuracy metrics are produced by running the model on the test dataset and comparing the results to the test data. The accuracy of the models produced from the subset data give an indication of the accuracy of the model overall (Isaaks and Srivastava 1989; Goovaerts 1997; Kitanidis 1997).

**Research Goals**

Previous studies indicated that elevated amounts of Mn and Fe exist in private wells in Western North Carolina, including Buncombe County. This has raised concerns about the need to identify spatial locations of elevated Mn and Fe concentrations in the region for proper groundwater remediation and management. Therefore, the goals of this research are to: (1) quantitatively analyze the spatial distribution of Fe and Mn in Buncombe County, North Carolina, and (2) evaluate the natural factors controlling the variability of Fe and Mn in
groundwater across the county. To evaluate the Fe and Mn distribution in the county, geocoding and interpolation using OK and cokriging techniques were used within ArcGIS. This same procedure can be used to estimate the trend of other contaminants in the groundwater in different states and counties. Finally, this study estimates Fe and Mn concentration values in unmonitored locations where direct measurements were not feasible due to surface topography, time, or resources and can aid in making educated decisions when sourcing new groundwater supplies in unmonitored areas.
CHAPTER 2
IRON AND MANGANESE IN GROUNDWATER: USING KRIGING AND GIS TO LOCATE HIGH CONCENTRATIONS IN BUNCOMBE COUNTY, NC
Crystal D. Johnson, Dr. Arpita Nandi, Dr. T. Andrew Joyner and Dr. Ingrid Luffman
Department of Geosciences, College of Arts and Sciences, East Tennessee State University
Johnson City, Tennessee 37614, United States

Abstract – For aesthetic, economic, and health-related reasons, allowable concentrations (as suggested by the United States Environmental Protection Agency) of the secondary contaminants manganese (Mn) and iron (Fe) found present in drinking water are 0.05 mg/L and 0.3 mg/L, respectively. Water samples taken from private drinking wells in the rural communities within Buncombe County, North Carolina contain amounts of these metals in concentrations higher than the suggested limits. To evaluate the factors affecting Fe and Mn concentrations in well water, this study focused on bedrock geology, ground elevation, saprolite thickness, and drinking water well depth. Using ArcGIS 10.2 and conducting statistical analyses with SPSS and GeoDa, spatial trends in the ranges of Fe and Mn concentrations were visualized and estimated. The results from this analysis were used to create a map that delineates the actual spatial distribution of Fe and Mn in Buncombe County. The study also established a statistically significant correlation between Fe and Mn concentrations, which can be attributed to bedrock geology. Additionally, higher Fe in groundwater was concentrated in shallower wells and valley areas. Finally, the study generated a prediction model that estimates the presence of Fe and Mn in groundwater where direct measurements are not possible.
1.0 Introduction

Many rural residents living outside the range of public drinking water supply infrastructure obtain their water supply from private groundwater wells. Potentially, detection of contamination of drinking water sources by those using groundwater wells may go unnoticed due to lack of current mandatory regulations as well as absence of direct sampling methods. Therefore, installation of water wells in areas prone to high concentrations of contamination could occur without the contractor or landowner being informed of the issue. Identification of areas likely to contain high concentrations of contaminants is useful in prevention of unknowingly accessing contaminated water, and also provides professionals with information about areas needing to be monitored for Mn and Fe concentrations.

1.1 Background Information

High concentrations of iron (Fe) and manganese (Mn) in groundwater may cause serious problems in public and private water supply systems including economic, aesthetic and adverse health concerns. Health concerns surrounding over-exposure to Fe can result in oxidative stress, and dyshomeostasis of metal homeostasis in the body (Farina et al. 2013). Degenerative diseases from excess Fe in the human system present symptoms similar to Parkinson’s disease, Alzheimer’s disease, and Huntington’s disease (Farina et al. 2013). High levels of Mn in the body are associated with decreased performance of short-term memory and visual identification (Bouchard et al. 2007), are correlated to hyperactivity and learning disability in small children (Graziano 2010), and infant mortality (Spangler and Spangler 2009). In addition, pregnant
women with late term anemia have a tendency to absorb greater amounts of Mn due to the absence of Fe in the body, affecting the fetus (Spangler and Spangler 2009). Furthermore, Fe and Mn can cause technical damage in household and industrial water supply distribution systems and plumbing fixtures by forming precipitates, which constrict and corrode the pipelines. Insoluble Fe and Mn have a negative aesthetic effect on drinking water by altering the taste, odor, or color of the water.

Both Fe and Mn are naturally occurring metals in rocks and soils, and bedrock chemical weathering and leaching from soils may lead to the presence of the metals at elevated concentrations in groundwater (Drever 2002). Fe and Mn generally coexist in groundwater as they have similar chemical properties (Davidson 1993; Hem 1972), and Fe and Mn can be found in both the reduced and oxidized states in groundwater. In an anoxic and acidic environment, Fe and Mn become soluble, allowing the metals to interact with the groundwater. Additionally, the geochemistry of the aquifer, in particular pH, Eh (redox potential), and dissolved oxygen, also contribute to controlling the concentrations of these metals in ground water (Hem 1972). Industrial and agricultural activities such as materials production of manganese-iron alloys through the smelting process, fertilizers, fungicides, livestock feed, batteries, paints, and gasoline additives may also contribute to the occurrence of the metals in ground water (Howe et al. 2004; Lytle 1995). Both Fe and Mn are listed under the United States Environmental Protection Agency (USEPA) National Secondary Drinking Water Regulations as secondary contaminants in drinking water at an allowable concentration level of 0.3 mg/L for Fe and 0.05 mg/L for Mn (USEPA 2013a). These regulations are non-enforceable standards that have been created by the USEPA; however, different states have the option to adopt the standards as minimum requirements for drinking water. Several studies have pointed out the coexistence of Fe, Mn, and
arsenic (As) in groundwater and surface water due to the regional geology (Klein and Hurlbut 1985; Hinkle and Polette 1999; Ayotte et al. 1999), groundwater chemistry (Anderson and Bruland 1991), and seasonal variation (Agbaire et al. 2009; Farmer et al. 1999). However, few studies have examined the metal concentrations with respect to groundwater residence time, i.e., the length of time the water spent in contact with the host bedrock (Edmunds and Smedley 2000), and groundwater recharge-discharge areas (Richardson et al. 1992). Additional studies found higher Mn concentrations in soil and plants along heavily trafficked roadways in central Utah as a result of the increased usage of manganese oxides in fuel as anti-knock agents (Lytle et al. 1994).

1.2 Fe and Mn in North Carolina

About 14% (42 million people) in the Unites States rely on unregulated private domestic water wells (Kenny et al. 2009). An estimated 2.3 million individuals in North Carolina rely on private well water for drinking water. This population is the fourth largest of state populations in the U.S. who rely on private well water for drinking water. (Kenny et al. 2009; Sanders et al. 2012). These wells are predominately located in the rural areas of North Carolina where residences are located outside municipal water supply infrastructure. The testing of well water is not mandated by neither state nor federal regulations; therefore, well water is not monitored consistently. Water treatment facilities, as well as private well owners, are not subjected to federal water quality standards regarding secondary contaminants (Herman 1996). Maximum ideal concentrations of secondary contaminants in drinking water developed by the USEPA are merely guidelines which residents or municipalities are given the option to follow. In lieu of mandatory well water testing by the state, many homeowners choose to test their well water.
through private means. The large number of private wells in North Carolina that contain elevated amounts of Mn and Fe has raised concerns about the spatial distribution and the sources of the contamination in the wells. The North Carolina Department of Health and Human Services (NCDHHS) in conjunction with North Carolina Public Health (NCPH) used the State Laboratory of Public Health dataset to map the distribution of private well water contaminants from 1998-2010 across the state. Additionally, NCDHHS evaluated the overall groundwater quality and potential public health concerns related to groundwater contamination in each individual county (NCDHHS 2014).

Spangler and Reid (2010) examined groundwater Mn concentrations in well water in NC and found associations with increased infant and cancer-related mortality rates. In addition, Spangler (2012) identified a positive correlation between high concentrations of air-borne Mn off- and on-road and high liver disease mortalities among residents of NC at the county level. Polizzotto et al. (2013) noted high levels of Mn concentrations in the NC Piedmont area associated with bedrock weathering and groundwater redox conditions; however, they realized the need of geospatial analyses to better understand the causes for spatial variability in well water concentrations. Giese et al. (1986) indicated the presence of Fe and Mn in crystalline aquifers of North Carolina. Evidence of high concentrations of Mn and Fe in groundwater in the crystalline aquifers of mountain region of western NC was documented by Campbell (2013). Campbell analyzed 63 private wells and 4 springhouses in Haywood County, NC for water quality. Among the constituents tested, Fe and Mn concentrations were found to be elevated, although the sources of the elevated concentrations were not identified conclusively. In a study in Bent Creek Experimental Forest watershed in Buncombe County, North Carolina, Campbell (2011) reported that schist rocks contain significantly higher iron oxide and the iron sulfide mineral pyrrhotite.
Dissolved manganese concentrations were below the detection limit in the Bent Creek study area.

1.3 Objective of study

The large number of private wells in western NC that contain elevated amounts of Mn and Fe has raised concerns about the source or sources of the contamination of the wells. Given the health risk and other economic problems associated with Fe and Mn in the groundwater, the objectives of this study are to: (1) quantitatively analyze the spatial distribution of Fe and Mn in Buncombe County, North Carolina, and (2) evaluate the natural factors controlling the variability of Fe and Mn in groundwater across the county.

2.0 Study Area

Buncombe County is located in the Blue Ridge physiographic province of the Appalachian Highlands in Western North Carolina. The county is bounded by Madison and Yancey Counties to the north, Henderson County to the south, McDowell and Rutherford Counties to the east and Haywood County to the west. It covers 1056.8 km² (656.67 mi²) and is home to 247,912 residents (United States Census Bureau 2015). Figure 2.1 shows the location of Buncombe County, in Western North Carolina, USA. The detailed map of Buncombe County (C) shows the locations of the major municipalities’ boundaries/jurisdictions in the county (colored polygons) as well as the major cities (C).
Figure 2.1: Location map of Buncombe County, in Western North Carolina, USA. The study area is located in the eastern United States (A) in North Carolina (B) in the western portion of the state in Buncombe County (C).

The county is bordered by the Appalachian Mountains to the west and the Black Mountains to the east and sits on a central dissected plateau known as the Asheville Plateau (Trapp 1970), a broad valley through which the French Broad and Swannanoa Rivers flow. Most of Buncombe County is located in the French Broad River drainage basin with the exclusion of the southeast tip of the county which drains to the Broad River drainage basin. The City of Asheville, located in the center of the county, is also the location of the widest area of the Asheville Plateau at 670.6 m (2,200 ft) elevation. The lowest areas in the County 524.3 m (1,720 ft) are in the
northeast corner along the French Broad River and in the southeast along the Broad River. The highest point in Buncombe County is found in the northwest corner at Potato Knob, elevation 1956.5 m (6,419 ft).

The Blue Ridge mountain belt consists of a mixture of meta-sedimentary and meta-igneous rocks that occurs southeast of the Valley and Ridge and northwest of the inner Piedmont province (Figure 2.2).

---

**Figure 2.2:** Geologic map of Buncombe County, NC
The Blue Ridge consists of Grenville (∼1.0–1.2 Ga) basement rocks in a series of northeast and northwesterly oriented thrust sheets bounded by shear zones. The major rock units occur as northeast trending belts, and are primarily composed of gneiss and schist, with minor amounts of granitic rocks and mafic intrusives. The central part of the county along the Asheville Plateau consists largely of garnet-muscovite schist and biotite gneiss with alternating, discontinuous elongated hornblende gneiss. This rock belongs to the Ashe Metamorphic Suite and the Tallulah Falls Formation dated to the Late Proterozoic. To the northwest of this muscovite-biotite gneiss found in the center of the county the formations exist in bands running southwest to northcentral and consist largely of migmatitic biotite-hornblende gneisses, amphibolites and meta-ultramafic rock which are mostly metamorphic and intrusive rock dating Late to Middle Proterozoic. (Dicken et al. 2005). Pegmatite dikes and quartz veins are common intrusions into the existing rock throughout the county and are part of the Carolina Gneiss (Keith 1904) which is mostly metasedimentary rock. The high mountains in the northeast corner along the Madison County line consist of predominately hornblende gneiss along with biotite gneiss and granitic body intrusions composing the Roan Gneiss (Trapp 1970; Keith 1904). Some ultramafic bodies exist in pockets in the area.

Opposite to the Roan Gneiss in the west, the Brevard Schist Belt lies in the southeast corner of the county, and is composed of quartzite, phyllite, argillite, slate, garnet-muscovite schist, graphite schist, marble, fine-grained gneiss and pyroclasts. Farther southeast of the Brevard Schist Belt lies the Henderson Gneiss. This formation is composed of granite gneiss, biotite gneiss and quartzite with some feldspar porphyroblasts. These formations also include bands of amphibolite and meta-ultramafic rock and belong to metamorphic/intrusive rock with some formations belonging to the Alligator Back Formation. These rocks date to the Late Proterozoic
and Paleozoic periods. Regionally, the bedrock dips 40 degrees to the southeast until Mt. Pisgah (located in the southwest corner of the county) where the dip is reversed in direction northeast. This ridge, called the Pisgah Ridge forms the border of an anticline that dips to the northeast (Trapp 1970).

Most crystalline metamorphic and igneous rocks serve as aquifers in the area (Trap and Horn 1997). The main types of crystalline rocks are coarse-grained gneiss and schist with various mineral compositions. In crystalline-rock, the fractures in the bedrock serve as the primary areas of groundwater storage. Where bedrock fractures have preferred directions of orientation, groundwater tends to flow more readily in the direction of the fractures. The porosity of the bedrock is generally low and ranges from only about 0.01 to 2 percent (Drever 2002). Wells located in valleys have higher yields as these areas commonly coincide with fracture zones, and the water table is closer to the surface in topographic lows. Some crystalline rocks contain minerals which, when weathered, can contribute to elevated iron and manganese in groundwater. In contrast, water in deep crystalline aquifers is found to be soft water with higher a sodium concentration (Rosenshein and Hunn 1968).

3.0. Research Methods

The research methods are divided into five steps: (i) data collection, (ii) geocoding, (iii) exploratory non-spatial statistics, (iv) exploratory spatial statistics, and (v) spatial statistical methods, kriging and cokriging (Figure 2.3).
3.1 Data collection

The Environmental Inorganic Chemistry Unit in the State Laboratory of Public Health under the North Carolina Division of Public Health and the Department of Health and Human Services (DHHS) performed the domestic well water analysis for home owners. First, the Buncombe
County Environmental Health Services collected and received water samples from private
drinking water wells using strict sampling protocols set forth by USEPA. The samples were
delivered to the Environmental Inorganic Chemistry Unit for analysis (Buncombe County
Environmental Health Services 2015). The results from the analysis of inorganic testing of water
samples were made available by DHHS at
http://slph.ncpublichealth.com/EnvironmentalSciences/inorganic/default.asp. Each individual
report lists results from the chemical analyses of inorganic constituents along with the
corresponding well owner’s address, GPS coordinates, collection date of the sample, sampling
point, sampling source, well permit number, and well owner’s identification number. In this
study, 1167 wells (sampled January 2010 - December 2013) were considered for the analysis. An
Excel database was created including: StarLiMS number (personal identification number per
individual well owner), well permit number, first and last name, address including city, state, zip
code, GPS coordinates, collection date, and sampling point. The inorganic constituents of the
tested water samples include: iron (Fe), and manganese (Mn). The areas within the national
forest in the northeastern and southwestern parts of the county as well as the urban areas in the
middle of the county did not have private drinking water well records.

The total well depth and the casing depth of the wells were also noted from the Groundwater
Log (GW1) records stored in the North Carolina Department of the Environment and Natural
Resources (NCDENR), Asheville Regional Office. A casing is installed during well construction
to prevent loose and disintegrated saprolite material (on top of the bedrock) from collapsing into
the well. Each well is drilled to the depth at which bedrock begins. From here after, the casing
depth is referred to as saprolite depth. A detailed spatial bedrock geology map of Buncombe
County was accessed from the NCDENR. Digital Elevation Model (DEM) data available in a
raster format at 6 meter accuracy were collected from Buncombe County GIS, located in Buncombe County, NC and managed by the North Carolina Flood Mapping Program (2006). All data were projected using the projected coordinate system NAD_1983_StatePlane_North_Carolina_FIPS_3200_Feet (Table 2.1).

<table>
<thead>
<tr>
<th>Data Name</th>
<th>Data Type</th>
<th>Scale</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe</td>
<td>Vector - Point</td>
<td>1:250,000</td>
<td>DHHS</td>
</tr>
<tr>
<td>Mn</td>
<td>Vector - Point</td>
<td>1:250,000</td>
<td>DHHS</td>
</tr>
<tr>
<td>Well Depth</td>
<td>Raster</td>
<td>1:250,000</td>
<td>DHHS</td>
</tr>
<tr>
<td>Saprolite Depth</td>
<td>Raster</td>
<td>1:250,000</td>
<td>DHHS</td>
</tr>
<tr>
<td>LiDAR - DEM</td>
<td>Raster</td>
<td>6 meter resolution</td>
<td>Buncombe County GIS</td>
</tr>
<tr>
<td>Geology</td>
<td>Vector - Polygon</td>
<td>1:250,000</td>
<td>NCDENR</td>
</tr>
</tbody>
</table>

3.2 Geocoding

Each well record contained either postal address information or latitude and longitude from a Cartesian coordinate system. The data were divided into records with latitude/longitude identification information (regardless of the presence of address information) and those with only address information. Some records contained neither address nor latitude/longitude coordinates. The records with missing location information were discarded from the dataset. Each record was geocoded and displayed as a point in ArcGIS 10.2 for Desktop.

Geocoding is the process of converting postal address data into geographic coordinates such as northings and eastings or latitude and longitude (Goldberg 2011). The procedure includes parsing the address to its individual components and assigning each component to a category. The categories are then compared to a reference layer which is used to match the addresses, and
the result is assigned XY coordinates on a map that correspond to the location of the address (Zandbergen 2009). Assigning the most accurate geographic coordinates to the postal address is dependent upon the reference data layers to which the parsed addresses are compared. The reference layer used for geocoding in this study was taken from the Buncombe County GIS department downloadable digital data website (Buncombe County GIS 2014) and contained primary and secondary street centerlines for Buncombe County. Street centerlines were created from 2002 aerial imagery and documented by L. Robert Kimball Associates and maintained daily in house by Buncombe County GIS using aerial imagery flown by the state of North Carolina. The end result of geocoding generates X, Y coordinates for each point; therefore, each record with existing latitude/longitude information was excluded from the geocoding process and was plotted on the map using a XY conversion process. Of 1168 original records, 1099 well records with Fe data, 1102 well records with Mn data, 315 well records with saprolite depth data, and 324 well records with well depth data remained for the analysis.

3.3 Exploratory Non-spatial Statistics

For all reported Fe concentrations below the minimum detection limit, standard USEPA statistical protocol was used; all values below 0.1 mg/L (the detection limit for Fe) mg/L were selected and half the values were assigned 0.1 mg/L while the remaining were assigned half the concentration value (0.05 mg/L). This same procedure was used for Mn concentrations; half the values below the detection limit for Mn (0.03mg/L) were assigned 0.03mg/L while the remaining data were assigned 0.015 mg/L (USEPA 2009). Descriptive statistical analysis was performed using the Statistical Package for the Social Sciences (SPSS). The mean, standard deviation, and range were calculated for each variable. Tests for normality and correlation among the variables
were also estimated using non-spatial tools/methods. The variable pairs that were significantly correlated in the non-spatial environment were analyzed in the open source statistical program GeoDa for exploratory spatial analysis (Anselin *et al.* 2006).

### 3.4 Exploratory Spatial Statistics

GeoDa was used to examine whether spatial dependence existed in the Fe and Mn data. The software has six main functions: data visualization and management, variable creation and transformation, mapping, exploratory statistics, spatial autocorrelation exploration, and spatial regression (mapping and visualization). In this study, the spatial autocorrelation exploration tool in GeoDa was used to first determine the presence or absence of any spatial relationship in the dataset. This provided the logical background for further analysis in ArcGIS. First, distance-based spatial weights of 10,000 ft were created. The spatial weights files smoothed out the autocorrelation connectivity by detecting problems in spatially discontinuous data, and this distance was manually selected to ensure every well would have a neighbor. Second, global and local Moran’s I test statistics were calculated. Global Moran’s I detects the presence of clustering in the dataset while local Moran’s I detects the location of the clusters of spatially autocorrelated data (Anselin *et al.* 2006). The global test produces a scatterplot of the data and creates a regression line which correlates to the Moran’s I test statistic. The significance of the clustering was estimated using a Monte Carlo approach to generate a pseudo $p$-value. The visual representation of the local Moran’s I test statistic is represented by LISA cluster and LISA significance maps. LISA cluster maps categorize the clusters by types of associations, i.e.; low values near low values, high values near high values and low values near high values. LISA significance maps show locations of statistically significant data clusters.
3.5 Spatial statistics - Kriging and Cokriging

Kriging and cokriging analysis were performed to produce prediction maps of Fe and Mn concentrations in Buncombe County, NC. Kriging relies on the assumption that autocorrelation exists among the data. In geography, it is assumed that data points that are located closer together are more similar than data points located farther apart (Miller 2007; Tobler 1970). However, this assumption was validated using GeoDa, explained in the previous section.

Ordinary kriging is widely used in soil science, groundwater quality assessment studies, and environmental interpolation mapping (Eldeiry and Garcia, 2012; Hu et al. 2005; Liu et al. 2005). Ordinary kriging is an interpolation technique that is useful in accounting for anisotropy in the dataset. This study uses ordinary kriging to produce a continuous surface map for predicting groundwater contamination concentrations in areas where sample data are not available. The geostatistical model that describes ordinary kriging is similar to the general kriging model.

Ordinary kriging is described using the following formula:

\[ Z(s) = \mu + \epsilon(s) \]

The formula follows the same convention as the general kriging formula with the exception that \( \mu \) is an unknown constant. Ordinary cokriging mathematically is almost identical to ordinary kriging. It is described as:

\[ Z_1(s) = \mu_1 + \epsilon_1(s) \]
\[ Z_2(s) = \mu_2 + \epsilon_2(s) \]
\[ Z_n(s) = \mu_n + \epsilon_n(s) \]

Where \( \mu_1 \ldots \mu_n \) are constants, \( \epsilon_n \) are multiple errors at location \( s \) which explain the predicted variable \( Z_n \) at location \( s \).
The empirical semivariogram was used to predict the spatial autocorrelation of datasets. The best model was selected on the basis of cross-validation and validation. The Mean Standardized Error (ME), Root Mean Square Standardized Error (RMSE), Root Mean Square (RMS), and Average Standard Error (ASE) were used to estimate the performance of the developed model. RMS indicates how closely the model predicts the observed data, with smaller RMS values being optimal. A better fitted model has the least difference between RMS and ASE. ME is close to 0 and RMSE is close to 1 if the prediction errors are valid. ME and RMSE also give indication of over- or underestimation of the variability of the errors in the model. Models with ME values greater than 0, and RMSE value less than 1, produce interpolated surfaces that are overestimates of the errors while models with ME values less than 0 and RMSE value greater than 1 are underestimates of the errors. In addition to using accuracy metrics to determine the accuracy of the model, standard prediction error maps were produced to quantify the variation in the difference between the true value and the predicted value. Standard prediction error maps were created for the final maps of Fe and Mn with their respective covariates.

4.0 Results

4.1 Location of Fe and Mn Contaminated Wells

Samples taken from wells containing Fe (excluding outliers) ranged in concentrations from 0.1 to 25 mg/L (Figure 2.4). The majority of the wells contained Fe concentrations below the EPA established 0.30 mg/L limit (710 wells). These wells were located in two main bands stretching southwest to northeast on either sides of the Asheville Plateau accounting for 64.6% of all the Fe samples. The remaining 389 wells 35.4% produced water samples that contained
concentrations greater than the EPA limit. Five wells (0.4%) produced concentrations of Fe greater than 15.0 mg/L, and were mostly located in the center of the county with the exception of one well located in the north central area of the county.

Figure 2.4: Fe contaminated well locations

Water samples containing significant levels of Mn concentrations exhibited a similar overall distribution pattern to the locations of wells with Fe concentration (Figure 2.5). The wells (excluding the outliers) ranged in concentrations 0.03 – 1.50 mg/L.
A total of 235 wells (21.3%) showed Mn concentrations less than the EPA limit of 0.05 mg/L. The remaining wells with concentrations greater than the EPA limit was 867 (78.7%) and the greatest concentrations (0.72 – 1.50 mg/L) were found in 2 wells located in the far southwest and southernmost borders (0.2%).
4.2 Exploratory Non-spatial Statistics Results

Mean, range, and standard deviation were calculated using SPSS for the original raw data for variables Mn, Fe, saprolite depth, and well depth. SPSS determined the presence of outliers in the Fe concentrations dataset. These values were reevaluated using the map. A substantial difference between the values identified as outliers and the surrounding values indicated the high concentrations of the outliers was more likely due to an issue with the existence distribution infrastructure than from natural sources. Subsequently, these outliers were removed, and the descriptive statistics procedure was performed again using Fe concentrations. Table 1 summarizes the descriptive statistics for these four variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Range</th>
<th>Standard Deviation</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fe (mg/L)</strong></td>
<td>0.680</td>
<td>0.1-25.00</td>
<td>1.714</td>
<td>1161</td>
</tr>
<tr>
<td><strong>Mn (mg/L)</strong></td>
<td>0.042</td>
<td>0.03-1.50</td>
<td>0.097</td>
<td>1167</td>
</tr>
<tr>
<td><strong>Saprolite Depth (ft)</strong></td>
<td>78.19</td>
<td>21-178</td>
<td>32.71</td>
<td>329</td>
</tr>
<tr>
<td><strong>Well Depth (ft)</strong></td>
<td>365.4</td>
<td>105-1005</td>
<td>173.7</td>
<td>329</td>
</tr>
</tbody>
</table>

As part of the exploratory statistics, test of normality on all variables were performed using the Kolmogorov-Smirnov normality test. None of the variables were found to be normally distributed. Significance of the test statistics for Fe, Mn and well depth were \( p = 0.000 \), and significance of saprolite depth test statistic was \( p = 0.036 \).

Two-tailed Spearman correlations were significant between the primary variables Fe and Mn (0.450, \( p = 0.01 \)), and also the combination of the primary variable Mn and covariate saprolite depth (-0.185, \( p = 0.01 \)).
4.3 Exploratory Spatial Statistics Results

Cluster analysis was performed on Fe and Mn concentrations to determine spatial autocorrelation. Globally, Fe concentrations were not spatially clustered (Global Moran’s I = 0.003; pseudo p-value = 0.115 at 999 permutations), however, local cluster analysis using LISA identified nine wells with high Fe concentrations were located next to similar high concentration wells and 46 wells with low concentrations were located near other low concentration wells (Figure 2.6A). Clusters of low-low values were found in the southeast corner of the county and were significant at $p = 0.05$ value (Figure 2.6A, 2.6B).

Significant spatial clusters were found among the Mn data (Figure 2.6C). Global Moran’s I was higher than Fe and significant at 999 permutations (Global Moran’s I = 0.104; pseudo p-value = 0.001). Locally, 185 wells with high concentrations were located near other wells with high concentrations, and 328 wells with low concentrations were located near other wells also with low concentrations. Three significant clusters of values were found among the Mn data. Clusters of wells with low concentrations were identified along the western and southeastern areas of the county and a cluster of wells with high concentrations were found in the north central area of the county (Figure 2.6C, 2.6D).
Figure 2.6: GeoDa local cluster analysis results. LISA cluster map of Fe concentrations (A) and LISA significance map of Fe concentrations (B), LISA cluster map of Mn (C) and LISA significance map of Mn (D)

Table 2.3: Global Moran’s I test statistic and significance for Fe and Mn

<table>
<thead>
<tr>
<th>Variable</th>
<th>Moran's I</th>
<th>Significance (α = 0.95)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe</td>
<td>0.003</td>
<td>0.115</td>
</tr>
<tr>
<td>Mn</td>
<td>0.104</td>
<td>0.001</td>
</tr>
</tbody>
</table>

5.0 Kriging/Cokriging Results

Kriging and cokriging methods were used to determine the optimum Fe and Mn concentration prediction models for Buncombe County groundwater.
5.1 Kriging/Cokriging with Fe data

Using the kriging method, the full dataset was used to create interpolation maps from the primary variable, i.e., Fe data (Figure 2.4). Cokriging was then used to evaluate the contribution of different covariates when estimating the spatial distribution of Fe concentration. The different covariates used in this study are saprolite depth, total well depth, geology, and elevation.

Kriging model accuracy was determined by intramodel metrics where stable semivariogram and true anisotropy selections produced the optimal model to predict Fe concentrations (in mg/L) of the kriged surface (Figure 2.7). The cross-validation matrix produced a ME value centered around 0 (ME = 0.000511), a low RMSE value (1.155), and the least difference among RMS (RMS = 2.1685) and ASE (ASE = 1.8777). Table 2.4 summarizes the accuracy metrics for Fe concentrations. However, with ME > 0, the model is an overestimate of the variability in the errors. In addition, ASE < RMS indicates the model’s underestimation in determining the variability between the predicted value from the true value. Each cokriging analysis was also performed using a stable semivariogram with anisotropy, to keep the models consistent. In this step, Fe concentrations were paired with the covariates: well depth, saprolite depth, geology, and elevation, and the combination of different covariates: well depth/elevation, well depth/elevation/geology, well depth/elevation/ geology/saprolite depth, elevation/saprolite depth, geology/saprolite depth, etc. Table 2.4 summarizes the accuracy metrics for the kriged surface created from Fe concentrations as well as the metrics from cokriged surfaces created from the remaining variables.
Table 2.4: Accuracy metrics produced from the kriging and cokriging of Fe and covariates

<table>
<thead>
<tr>
<th>Cokriging</th>
<th>ME</th>
<th>ASE</th>
<th>RMS</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe</td>
<td>0.00051</td>
<td>1.878</td>
<td>2.169</td>
<td>1.155</td>
</tr>
<tr>
<td>Fe/Well Depth</td>
<td>0.00009</td>
<td>2.090</td>
<td>2.171</td>
<td>1.0384</td>
</tr>
<tr>
<td>Fe/Geology/Elevation</td>
<td>0.00056</td>
<td>2.024</td>
<td>2.173</td>
<td>1.0738</td>
</tr>
<tr>
<td>Fe/Saprolite/Elevation</td>
<td>0.00067</td>
<td>2.023</td>
<td>2.173</td>
<td>1.0743</td>
</tr>
<tr>
<td>Fe/Elevation</td>
<td>0.00057</td>
<td>2.023</td>
<td>2.173</td>
<td>1.0744</td>
</tr>
<tr>
<td>Fe/Geology</td>
<td>0.00035</td>
<td>1.863</td>
<td>2.166</td>
<td>1.1628</td>
</tr>
<tr>
<td>Fe/Geology/Saprolite</td>
<td>-0.00004</td>
<td>1.827</td>
<td>2.169</td>
<td>1.1872</td>
</tr>
<tr>
<td>Fe/Saprolite</td>
<td>0.00126</td>
<td>1.825</td>
<td>2.171</td>
<td>1.1886</td>
</tr>
</tbody>
</table>

All models produced ME values centered around 0. Fe/geology produced the lowest RMS value (2.166), then Fe/geology/saprolite (2.169), Fe/saprolite (2.171) and finally, Fe/geology/elevation, Fe/saprolite/elevation and Fe/elevation produced the same, and highest, RMS value (2.173). Fe/well depth (1.0384), Fe/geology/elevation (1.0738), Fe/saprolite/elevation (1.0743) and Fe/elevation (1.0744) produced RMSE values lower than the kriged Fe surface, and the rest had higher RMSE values that the kriged surface.

The difference between RMS values and ASE values for each surface follow the same pattern as RMSE values. Fe/well depth had the least difference between RMS and ASE (0.081) followed by Fe/geology/elevation (0.149), Fe/saprolite/elevation (0.150), Fe/elevation (0.150), Fe/geology (0.303), Fe/geology/saprolite (0.343), and Fe/saprolite (0.346). Covariate combinations like Fe/saprolite/elevation, Fe/elevation/well depth, and Fe/elevation/well depth/geology did not change the model RMSE, ASE, and RMS, hence were not reported in the Table 2.4. All models produced from adding the covariates have RMSE > 1 and RMS > ASE indicate model underestimation in error variability as well as underestimation of the distance of the predicted value from the true value.
Figure 2.7: Kriging/cokriging surfaces for Fe and covariates: Fe (A), Fe/saprolite (B), Fe/well depth (C), Fe/geology (D), Fe/elevation (E), Fe/geology/elevation (F), Fe/saprolite/elevation (G), Fe/geology/saprolite (H) all shown in concentrations of mg/L.
5.2 Fe – Validation with 30% Test Dataset

Validation was performed on the kriging/cokriging models produced from the 70% training dataset using the 30% test dataset to ensure accuracy of the training dataset models. Table 2.5 lists the accuracy metrics from validation of 70% training data for Fe. All models had ME values close to 0. Fe/saprolite was closest to 0 (ME = 0.002) and Fe/geology was next closest to 0 (ME = -0.003) while Fe/geology/saprolite was farthest from 0 (ME = 0.014). Models producing surfaces using Fe (ME = 0.011), Fe/geology/saprolite (ME = 0.014), Fe/well depth (ME = 0.007) and Fe/saprolite (ME = 0.002) are overestimates of the errors in the model. Fe/geology/elevation (ME = -0.009), Fe/saprolite/elevation (ME = -0.009), Fe/elevation (ME = -0.009) and Fe/geology (ME = -0.003) underestimate the errors in the interpolated surface produced by the model. Fe/geology had the least difference between RMS and ASE (0.223) while Fe/saprolite retained the greatest difference (0.409).

<table>
<thead>
<tr>
<th>Kriging/Cokriging Surfaces</th>
<th>ME</th>
<th>RMS</th>
<th>ASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe/Geology</td>
<td>-0.003</td>
<td>1.832</td>
<td>2.055</td>
</tr>
<tr>
<td>Fe/Saprolite</td>
<td>0.002</td>
<td>1.833</td>
<td>2.242</td>
</tr>
<tr>
<td>Fe/Elevation</td>
<td>-0.009</td>
<td>1.834</td>
<td>2.098</td>
</tr>
<tr>
<td>Fe/Saprolite/Elevation</td>
<td>-0.009</td>
<td>1.834</td>
<td>2.099</td>
</tr>
<tr>
<td>Fe/Geology/Elevation</td>
<td>-0.009</td>
<td>1.834</td>
<td>2.095</td>
</tr>
<tr>
<td>Fe/Well Depth</td>
<td>0.007</td>
<td>1.846</td>
<td>2.173</td>
</tr>
<tr>
<td>Fe/Geology/Saprolite</td>
<td>0.014</td>
<td>1.859</td>
<td>2.136</td>
</tr>
<tr>
<td>Fe</td>
<td>0.011</td>
<td>1.859</td>
<td>1.934</td>
</tr>
</tbody>
</table>
5.3 Mn Kriging and Cokriging

100% of the data were also used to create an interpolation model for Mn concentrations as well as the surfaces created from multivariate cokriging analyses (Figure 2.8). The most optimal Mn model used a circular semivariogram with no anisotropy (RMSE = 1.211). Mn concentrations data were cokriged with variables such as well depth, saprolite depth, geology, elevation, and combination of variables like geology/elevation, geology/saprolite, saprolite/elevation, and others. The accuracy metrics for cokriging results among the Mn/covariate combinations is listed in Table 2.6. All models produced identical RMS values (0.093). This negligible difference between RMS values led to differentiating the models by using RMSE values along with corresponding least differences between RMS and ASE. The original Mn kriged surface produced RMSE value of 1.211. Two cokriging surfaces had values less than Mn kriging surface; Mn/Geology (1.155) and Mn/Saprolite/Elevation (1.184). Surfaces with RMSE values the same as Mn were Mn/Saprolite (1.211), Mn/Elevation (1.211), Mn/Well Depth (1.211), Mn/Geology/Saprolite (1.211), and Mn/Geology/Elevation (1.211). The interpolated surface with the least difference between ASE and RMS was Mn/Geology which also had the lowest RMSE value. Mn/Saprolite/Elevation had the next least difference between ASE and RMS and also had the next lowest RMSE value. The remaining surfaces produced identical results to the Mn kriging surface. Figure 2.8 is a map of the interpolated surface produced from Mn/Geology.
Figure 2.8: Kriging/cokriging surfaces for Mn and covariates: Mn (A), Mn/saprolite (B), Mn/well depth (C), Mn/geology (D), Mn/elevation (E), Mn/geology/elevation (F), Mn/saprolite/elevation (G), Mn/geology/saprolite (H) all shown in concentrations of mg/L.
Table 2.6: Accuracy metrics produced from the kriging and cokriging of Mn and covariates

<table>
<thead>
<tr>
<th>Cokriging</th>
<th>ME</th>
<th>ASE</th>
<th>RMS</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mn</td>
<td>-0.00796</td>
<td>0.077</td>
<td>0.093</td>
<td>1.211</td>
</tr>
<tr>
<td>Mn/Geology</td>
<td>-0.00874</td>
<td>0.080</td>
<td>0.093</td>
<td>1.155</td>
</tr>
<tr>
<td>Mn/Saprolite/Elevation</td>
<td>-0.00762</td>
<td>0.078</td>
<td>0.093</td>
<td>1.184</td>
</tr>
<tr>
<td>Mn/Saprolite</td>
<td>-0.00796</td>
<td>0.077</td>
<td>0.093</td>
<td>1.211</td>
</tr>
<tr>
<td>Mn/Elevation</td>
<td>-0.00796</td>
<td>0.077</td>
<td>0.093</td>
<td>1.211</td>
</tr>
<tr>
<td>Mn/Well depth</td>
<td>-0.00797</td>
<td>0.077</td>
<td>0.093</td>
<td>1.211</td>
</tr>
<tr>
<td>Mn/Geology/Saprolite</td>
<td>-0.00795</td>
<td>0.077</td>
<td>0.093</td>
<td>1.211</td>
</tr>
<tr>
<td>Mn/Geology/Elevation</td>
<td>-0.00795</td>
<td>0.077</td>
<td>0.093</td>
<td>1.211</td>
</tr>
</tbody>
</table>

5.4 Mn – Validation with 30% Test Dataset

Cross-validation was also performed with the Mn dataset. A separate dataset consisting of 70% of the Mn data were used to create interpolation models for all the covariates, which were verified using cross-validation procedures on the remaining 30% of the data. Table 2.7 shows accuracy metrics from cross-validation of Mn kriging/cokriging models.
Table 2.7: Accuracy metrics created from 30% testing Mn dataset validation

<table>
<thead>
<tr>
<th>Kriging/Cokriging Surfaces</th>
<th>ME</th>
<th>RMS</th>
<th>ASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mn/Saprolite</td>
<td>0.12766</td>
<td>0.06691</td>
<td>0.08347</td>
</tr>
<tr>
<td>Mn/Elevation</td>
<td>0.12767</td>
<td>0.06691</td>
<td>0.08347</td>
</tr>
<tr>
<td>Mn/Geology</td>
<td>0.12767</td>
<td>0.06691</td>
<td>0.08347</td>
</tr>
<tr>
<td>Mn/Well depth</td>
<td>0.12767</td>
<td>0.06691</td>
<td>0.08347</td>
</tr>
<tr>
<td>Mn/Geology/Elevation</td>
<td>0.12767</td>
<td>0.06691</td>
<td>0.08347</td>
</tr>
<tr>
<td>Mn/Geology/Saprolite</td>
<td>0.12767</td>
<td>0.06691</td>
<td>0.08347</td>
</tr>
<tr>
<td>Mn/Saprolite/Elevation</td>
<td>0.12766</td>
<td>0.06691</td>
<td>0.08347</td>
</tr>
<tr>
<td>Mn</td>
<td>0.12694</td>
<td>0.06772</td>
<td>0.08602</td>
</tr>
</tbody>
</table>

The difference among ME values for all the models were negligible. Mn had the smallest ME (0.127) but had the greatest difference between ASE and RMS. The remaining cokriging surfaces all had the same difference between ASE and RMS. All ME values for the models were greater than 0 indicating model overestimation of errors.

5.5 Percent Area Exceeding Acceptable Limit

For Fe, the percentages in which the total interpolated surface area that is occupied by concentrations > 0.30 mg/L for all surfaces (including cokriged surfaces) is approximately 84.0%. This portion of the surface is the largest area and is distributed throughout the county. Table 2.8 gives a summary of the percentage of each surface that contains areas with concentrations higher than 0.3 mg/L.
Table 2.8: Percent area exceeding acceptable Fe concentrations

<table>
<thead>
<tr>
<th>Surface</th>
<th>% of surface greater than 0.3 mg/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe/Saprolite</td>
<td>82.9</td>
</tr>
<tr>
<td>Fe/Geology/Saprolite</td>
<td>83.3</td>
</tr>
<tr>
<td>Fe</td>
<td>84.0</td>
</tr>
<tr>
<td>Fe/Geology/Elevation</td>
<td>84.1</td>
</tr>
<tr>
<td>Fe/Saprolite/Elevation</td>
<td>84.2</td>
</tr>
<tr>
<td>Fe/Elevation</td>
<td>84.3</td>
</tr>
<tr>
<td>Fe/Geology</td>
<td>84.6</td>
</tr>
<tr>
<td>Fe/Well depth</td>
<td>87.1</td>
</tr>
</tbody>
</table>

Approximately 32.6% of the Mn, Mn/saprolite, Mn/well depth, Mn/geology, Mn/elevation, and Mn/geology/elevation surfaces correspond to areas that are predicted to contain concentrations of Mn greater than 0.05 mg/L. Mn/saprolite/elevation predicts 30.9% of the interpolated surface will contain these concentrations, and Mn/geology/saprolite predicts 30.2% of the interpolated surface will contain areas that have concentrations of Mn greater than the threshold limit. Table 2.9 lists all interpolated surfaces along with the corresponding percentages of the surfaces that are predicted to contain concentrations greater than 0.05 mg/L.

Table 2.9: Percent area exceeding acceptable Mn concentrations.

<table>
<thead>
<tr>
<th>Surface</th>
<th>% of surface greater than 0.05 mg/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mn/Geology/Saprolite</td>
<td>30.2</td>
</tr>
<tr>
<td>Mn/Saprolite/Elevation</td>
<td>30.9</td>
</tr>
<tr>
<td>Mn/Saprolite</td>
<td>32.6</td>
</tr>
<tr>
<td>Mn</td>
<td>32.6</td>
</tr>
<tr>
<td>Mn/Geology/Elevation</td>
<td>32.6</td>
</tr>
<tr>
<td>Mn/Elevation</td>
<td>32.6</td>
</tr>
<tr>
<td>Mn/Geology</td>
<td>32.6</td>
</tr>
<tr>
<td>Mn/Well depth</td>
<td>32.6</td>
</tr>
</tbody>
</table>
6.0 Discussion

Due to economic and health damages and costs, determining the locations of areas in which soluble Mn and Fe exist in high concentrations in the groundwater is a benefit for residents in Buncombe County. The primary goals of this research were to determine the areas that are affected by Mn and Fe, interpolate these concentrations to areas that do not have well test data, and determine environmental variables associated with high concentrations. In this study, the entire area of Buncombe County was examined; however, due to the unavailability of well data in urban areas which predominately rely on city water as well as national forest land areas, which do not normally contain private properties, were not included in the investigation.

6.1 Exploratory Non-Spatial and Spatial Statistics

Non-spatial exploration of Fe and Mn data identified a positive correlation between these two variables, which may be due to inherent bedrock geology patterns which have also been noted in other studies (Hem 1972; Jaudon et al. 1989; Rönkä 1983). The analysis also identified a negative correlation between Mn and saprolite depth (thickness) indicating that as the saprolite thickness increases, Mn ion concentration in the groundwater decreases. This might be due to the binding potential of Mn ions in the saprolite layer. However, this trend was not seen in the kriging surfaces.

The results of the analysis using GeoDa predicted spatial autocorrelation existed among variables Fe and Mn. The presence of spatial autocorrelation among the variables satisfied the requirements and provided the framework for kriging and cokriging analysis.
6.2 Final Spatial Statistics and Prediction Maps

Non-spatial and spatial statistical results supported the use of kriging and cokriging as the next analysis to predict Fe and Mn locations in unmeasured areas. To determine the environmental factors responsible for variable concentrations of Fe in Buncombe Co., cokriging with well depth, geology, saprolite depth and elevation was performed. The intermodel cross validation test for Fe concentrations indicated a ME value close to 0 and a low RMSE value. Based on the ME and RMSE results, there were only slight improvement in the kriging model after adding the covariates. That is, the RMSE value was low (1.155) in the kriging model, and was reduced after adding the covariates (i) well depth (1.0384) and (ii) elevation and bedrock geology (1.0738). Addition of other covariates did not improve the model performance. The validation test using training and test data for Fe concentration prediction produced similar results when compared with cross validation results. It should be noted that when all three covariates were used in cokriging, the model did not improve. Thus, the Fe and well depth model was chosen to be the optimal prediction model, but the contribution of elevation and geology was taken into account. All models produced from adding the covariates have RMSE > 1 and RMS > ASE indicating model underestimation in error variability as well as underestimation of the distance of the predicted value from the true value.

Fe concentrations were found to decrease as the depth of the well increased. Deeper wells in the county were primarily located on the ridges or highlands because in the higher elevations or ridge areas the groundwater table is lower and deeper wells are required to access the groundwater (Fetter 2001). However, due to the presence of crystalline bedrock of Blue Ridge physiographic provinces deep wells were also present in valleys. Generally, the crystalline bedrock ranges from not-fractured to highly-fractured type of rock in various places and if the
drinking water wells do not intercept any fractures, the water yield becomes negligible; hence deeper wells are needed for sustainable water yield (Chapman 2001). The Fe concentration was reduced in groundwater under the ridges. A possible explanation for this phenomenon is related to the ion exchange process that occurs in deeper groundwater (Rosenshein and Hunn 1968). Certain ions present in the groundwater, such as calcium, magnesium, and Fe, are adsorbed into the sediments and rocks as the groundwater percolates to the aquifer. In exchange for these ions, sodium is released into the groundwater. As a result, lower concentrations of Fe and higher concentrations of sodium are found at the deeper wells. Similar results were found in other studies where low calcium and high sodium concentrations, i.e., softer water, was present in deeper groundwater (Giese et al. 1986; Lee 1981). In addition, a lower concentration of Fe in the groundwater was found in the ridges and a higher Fe concentration was found in the valleys. The ridges are groundwater recharge areas and low concentration of Fe may be due to supply of fresh water through infiltration which might dilute the metal concentration. The valleys are groundwater discharge areas and high concentrations of Fe in discharge area is attributed to slow movement of ground water in crystalline fractured bedrock, which increases the residence time of ground water along the flow path from recharge to discharge zones thereby providing more time for dissolution of bedrock (Campbell 2011; Nelson 2002).

Bedrock geology is a factor that contributes to Fe concentration especially in the crystalline aquifers. The longer the groundwater residence time, the more opportunity the water has to dissolve the metal ions into solution. Cokriging results from the comparison of geology and Fe support this finding. High concentrations of Fe were found in regions of the county that contain rocks including sulfidic rocks, muscovite-biotite gneiss, amphibolite, biotite granitic gneiss and migmatitic biotite-hornblende gneiss.
The final Fe concentration map from the cokriging surface determined high Fe concentrations beginning in the county’s centroid and trending northeastward to the county border (Figure 2.9). Concentrations less than 0.3 mg/L were found in the remaining area excluding a high concentration area in the very southwestern tip. Figure 2.10 shows the associated standard errors involved in predicting the Fe concentration in groundwater. The map indicated that Fe concentration was predicted fairly well throughout the county except in the national forest areas and in the urban areas, which was due to the lack of public drinking water well data in these areas. Therefore the cokriging interpolation surface indicates higher prediction error in the national forest areas and in the urban areas.
Figure 2.9: Final prediction map of Fe concentrations cokriged with well depth
This analysis was also found to be on par with a previous local study conducted in 1961 by the North Carolina Department of Water and Air Resource, Division of Groundwater (Trapp 1970). The study analyzed water samples from 7 wells installed in areas around the county. The Fe concentrations from the samples ranged from 0.30 – 12 mg/L. These samples’ Fe concentrations correlated with Fe concentrations found from the DHHS samples in this research (Figure 2.11), except in two areas (one in northeastern and the other in western part), where the previous study found Fe in low concentration.
Non-spatial exploration of the relationship between Mn and the other variables revealed a negative correlation between Mn and saprolite depth. However, cokriging estimates did not support this finding. Using cokriging, other covariates were analyzed to determine associations between potential non-point sources and high Mn concentrations. The cross validation test for Mn concentrations indicated that the RMSE value (1.211) produced from kriging was slightly improved after adding the covariate bedrock geology (1.155), and addition of other covariates did not change or improve the model performance. The validation test using training and test data for Mn concentration prediction produced somewhat similar results when compared with
cross validation result, but the overall accuracy metrics indicated that the results were not very sensitive to the covariates. There might be two potential reasons for the finding (i) this may be a result of environmental factors not contributing much to Mn concentrations, but rather, anthropogenic factors such as deteriorating infrastructure might play an important role, or (ii) the geochemistry of groundwater, i.e., the Eh and pH may have a prevailing effect on Mn concentration in the study area. The RMSE of all Mn prediction models were greater than 1 indicating the model is an underestimate of the variability in the errors. Furthermore, ASE (0.077) is less than RMS (0.093) indicating model underestimation in the distance between the predicted value from the true value. The Mn/geology model was chosen to be the optimal prediction model, and bedrock geology was found to be a potential source of this contaminant. Due to similar chemical properties of Fe and Mn, Mn frequently substitutes Fe in the chemical composition of rock, and as a result, Mn can be found frequently in Fe bearing rocks (Drever 2002; Merritts et al. 1998). Similar to Fe, muscovite biotite and granitic gneisses are found to release Mn in the groundwater. Concentrations greater than the reporting limit (0.05 mg/L) occupy 32.6% of the total area in the county with the highest concentrations accumulating to 4.0% of the county which are located in the southwest corner, south-central border and north-central area of the county (Figure 2.12). Several muscovite-biotite and granitic gneisses are also found in the bedrock of the same areas of high Mn concentrations. Figure 2.13 shows the associated standard errors involved in predicting the Mn concentration in groundwater. The map indicates Mn concentrations were predicted fairly well throughout the county. Similar to Fe concentrations, Mn concentrations were excluded in the national forest areas and in the urban areas. The lack of public drinking water well data in these areas generated higher prediction errors in the cokriging model.
Figure 2.12: Final prediction map of Mn concentrations cokriged with geology
Figure 2.13: Standard error prediction map of Mn/geology cokriged surface

6.3 Limitations

As part of the UNC Superfund Research Program (2014), Fe and Mn mapping was performed on a statewide scale; providing a general overview of the contaminant problem in North Carolina. The research presented here has produced the first county-level analysis of Mn and Fe prediction map in groundwater using kriging and cokriging. The research did have several limitations. Since water quality testing of well water is not a requirement by the USEPA, many residents have not tested their well water, and the well dataset is therefore incomplete. In addition, private water quality testing companies in Buncombe County that offer water quality testing for the
public do not make their results publicly available. Although this study used 1167 wells across the county, additional data may provide a better prediction model. Another issue with the drinking water wells is the absence of data points in the southwest and northeast areas outside of the Asheville jurisdiction. These areas belong to private communities/municipalities and may have municipal water supply systems or private communities on their own water systems, thus well test data were limited in these areas. Another problem with the data collected from DHHS was the absence of dissolved oxygen (DO) and redox potential (Eh) data for each sample at the sampling point. USEPA does not require the sample be tested for Eh or DO at the time of collection. Fe and Mn concentration is directly related to the geochemistry of the groundwater, and lack of Eh and DO data adds a constraint in the study. This information is useful in determining speciation states of Fe and Mn (Hem 1972; Jaudon et al. 1989). Water chemistry, in particular certain pH, redox potential (Eh), and dissolved oxygen (DO) conditions, helps to dissolve and mobilize Fe and Mn ions in groundwater and control the concentration in the water environment. In lower pH (acidic water) and Eh conditions both Fe and Mn occur mainly as the reduced soluble ions and can move in water, but are oxidized to form precipitates in the presence of oxygen and at neutral pH conditions (Hem, 1972).

7.0 Conclusion

This study found a significant correlation between Fe and Mn in groundwater wells. The Fe and Mn concentration prediction maps for Buncombe County indicate the environmental factors elevation, bedrock geology, and well depth have slightly attributed to the elevated level of the ions in the groundwater. The degree of concentration of Fe is related to well depth, bedrock geology and elevation of the area. Mn was found to also be dependent on the rock type of the area. Geologic formations that contain rocks with Fe and Mn in their chemical makeup
contribute to elevated level of these metals particularly if the length of time the water spent in
contact with the bedrock is substantial. That is, the groundwater in the recharge area contains
less Fe and Mn, whereas groundwater in the discharge area in the valleys contains higher Fe and
Mn. In addition, deeper wells provide the ideal environment for ion-exchange of Fe for sodium.
The study clearly indicates the environmental factors solely could not explain the variability of
Fe and Mn concentration in groundwater. The geochemistry of the groundwater like Eh and pH
are important factors and might play significant role in the metal ion concentrations. Also, more
complete well log data from the overall county would realistically help model the saprolite layer
and well depth layer. The results of this study will be useful for residents or companies installing
new water wells to avoid areas that are known or projected to contain concentrations of Fe and
Mn greater than the established limit. This study may also be expanded by analyzing the
complete chemical analysis data of the water samples of the wells in the study area.

This study provided the first county level Fe and Mn concentration prediction map in
Buncombe County, NC. Similar analytical procedures can be used to predict other contaminants
in Buncombe County, or in other areas with similar geology and environmental settings.

8.0 References

Agbaire, P. O., Obi, C. G., 2009. Seasonal variation of some physico-chemical properties of
River Ethiope water in Abraka, Nigeria. Journal of Applied Science and Environmental


Anselin, L., Syabri, I., Kho, Y., 2006. An Introduction to spatial data analysis. Geographical
Analysis 38:5–22.


CHAPTER 3

DISCUSSION AND MAJOR FINDINGS

This study evaluated the spatial distribution of Fe and Mn in the crystalline aquifer in the western mountainous region of North Carolina in Buncombe County using both non-spatial and spatial geostatistical methods such as SPSS, GeoDa and kriging/cokriging.

The following are the major findings of this study:

1. The non-spatial statistics revealed a relationship between Fe and Mn, which may be due to the inherent bedrock geology of the area. This type of relationship has also been documented in other studies. Fe and Mn contain similar chemical properties and are often found coexisting in bedrock. Soluble forms of the ions in groundwater are due to the chemical weathering of geologic materials such as biotite mica, pyrite, amphibole, and clay minerals that contain Fe and Mn as part of their chemical makeup. Fe and Mn may often occur in groundwater due to redox conditions of the bedrock.

2. Mn concentrations were correlated with saprolite depth. Non-spatial statistical analysis of these two variables indicates a negative relationship. In other words, at greater saprolite depths, lower concentrations of Mn were found and at lower saprolite depths, higher concentrations of Mn were found. This relationship might be due to absorption of Mn in the saprolite material horizon (Nealson and Saffarini 1994). However, this same relationship was not significant in the cokriging analysis.

3. Spatial analysis using GeoDa confirmed the existence of spatial autocorrelation in the distribution of Fe and Mn in the groundwater. This finding suggested the presence of non-point source factors that might contribute to the elevated concentration of Fe and
Mn in the groundwater. The autocorrelations determined the need for further exploration using interpolation analysis.

4. The kriging interpolation method was used to produce a predicted spatial distribution map of Fe and Mn in groundwater.

5. Cokriging was used to determine potential contribution of non-point environmental factors responsible for Fe and Mn concentration in groundwater. The environmental factors evaluated for the study were well depth, geology, elevation, and saprolite depth.

6. The environmental factors slightly improved the interpolation surface. Fe concentration was found to be related to the well depth, elevation and geology. Ion-exchange, dilution and residence time in the aquifer were found to attribute to the general trend of Fe concentration. Deeper wells tended to present lower concentrations of Fe due to the ionic exchange of Fe for Na as the groundwater travels deeper vertically in the bedrock. This also provides the opportunity for dilution of Fe ions by the time the groundwater reaches the bottom of the well. The opposite is true for Fe ion concentrations found in shallow wells. These wells were shorter depth than the deep wells which provided no opportunity for ion exchange. In addition, shallower wells are found in the lower elevations. Finally, cokriging results indicated the inclusion of bedrock geology somewhat improved the model prediction; this result indicates the type of minerals present in the bedrock in areas containing high concentrations of Fe in wells correlate to Fe-bearing minerals.

7. Mn concentration was related to bedrock geology for the same reason as the contributing factors controlling Fe. Mn and Fe often coexist and in addition, express very similar chemical behaviors. These transition metals can often exchange and substitute for each other within a Fe- or Mn-bearing mineral. As a result, Mn will often substitute for Fe in
the chemical makeup of the Fe-bearing mineral. Furthermore, muscovite, biotite and granitic gneisses readily contribute Mn ions to the groundwater.

**Study Limitations**

A few limitations were present in this study.

1. Not every resident in Buncombe County relies on groundwater for drinking supplies; some residents are located in the city or township jurisdictions that provide municipal drinking water for these residents. These areas exist in the center as well on the western flank of the county in the 6 city/township jurisdictions. In addition, federally-owned national forest land lies to the northwest of Asheville city limits. There is no private drinking water well data available for these areas.

2. Another limitation in this study was the lack of geochemical data associated with each well record. In order to analyze more completely the effect of environmental factors on the level of concentrations of Fe and Mn, other important geochemical information such as pH, redox potential parameters, and dissolved oxygen need to be considered, which were mostly unavailable for this study. Among them, only pH was recorded in the DHHS geochemical data, but it was not clarified at which point the pH measurement was taken and recorded. Temperature has a substantial effect on pH and samples should be in room temperature in order for the pH value to be valid (Barron et al. 2014).

3. Well permit data, i.e., GW1 data (well driller’s logs) were only available for 324 wells, whereas Fe and Mn data were extracted from 1099 and 1102 wells, respectively. Additionally, GW1 data were not complete and often did not include the groundwater table data, casing and pipe materials. This information could help in the model prediction.
4. Due to the presence of spatial autocorrelation, confirmation of the contribution of Fe and Mn from bedrock geology does exist; however, the presence of these ions in drinking water from leaching of distribution systems can neither be confirmed nor rebutted. GW1 data was collected sporadically for each well; not every well had documentation from the well drilling companies. With more of this information, point-source contribution from distribution systems can be more closely studied.

5. The prediction error associated with Fe and Mn concentrations in groundwater was related to lack of public drinking water well data in the national forest areas and in the urban areas.

**Future Research**

1. Although this study focused mainly on the existence of Fe and Mn concentrations in Buncombe County, other geochemical information recorded from DHHS drinking water analysis results, i.e., arsenic, nitrates/nitrites, sulfates, total alkalinity, and total hardness as well as other cation and anion information. This information can be used to determine an overall baseline for water quality in Buncombe County and can be used in further geochemical trend analysis studies.

2. Additional studies which include Piper and Stiff diagrams can further provide evidence of groundwater sources and the geographical location of different ground water types.

3. The higher prediction error, due to lack of well records in the national forests and in urban areas in Buncombe County, indicate a future study should be performed by excluding these areas.
REFERENCES


VITA

CRYSTAL DEANNE JOHNSON

Education: M.S. Geosciences, East Tennessee State University, Johnson City, Tennessee 2015
B.S. Environmental Science, University of North Carolina – Asheville, Asheville, NC 2010

Professional Societies: Association of American Geographers
Association of Environmental and Engineering Geologists

Employment: GIS Help Desk Manager, ETSU, Spring 2015
GIS Technician, NCDOT, January 2014
Graduate Assistant, ETSU, Fall 2013

Posters: 2014 AAG Poster Presentation
What’s in your water? Using GIS to study groundwater quality in Rural Buncombe County, NC

2014 Appalachian Student Research Forum – ETSU
Using GIS to Visualize Water Quality Contaminants Present in Private Water Wells in Rural Communities in Western North Carolina

2015 Appalachian Student Research Forum – ETSU
Iron and Manganese in Groundwater: Using Kriging, Cokriging and GIS to Visualize Water Quality in Buncombe County, NC