The Relationship between Lead in Groundwater and Elementary School Students' Academic Performance in Missouri

Lynnette Xiangling Li

As with any intellectual project, the content and views expressed in this thesis may be considered objectionable by some readers. However, this student-scholar’s work has been judged to have academic value by the student’s thesis committee members trained in the discipline. The content and views expressed in this thesis are those of the student-scholar and are not endorsed by Missouri State University, its Graduate College, or its employees.

Follow this and additional works at: https://bearworks.missouristate.edu/theses

Part of the Hydrology Commons, Spatial Science Commons, and the Toxicology Commons

Recommended Citation
Li, Lynnette Xiangling, "The Relationship between Lead in Groundwater and Elementary School Students' Academic Performance in Missouri" (2016). MSU Graduate Theses. 3182.
https://bearworks.missouristate.edu/theses/3182

This article or document was made available through BearWorks, the institutional repository of Missouri State University. The work contained in it may be protected by copyright and require permission of the copyright holder for reuse or redistribution.

For more information, please contact BearWorks@library.missouristate.edu.
THE RELATIONSHIP BETWEEN LEAD IN GROUNDWATER AND
ELEMENTARY SCHOOL STUDENTS’ ACADEMIC
PERFORMANCE IN MISSOURI

A Masters Thesis
Presented to
The Graduate College of
Missouri State University

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science, Geospatial Sciences in Geography and Geology

By
Lynnette Xiangling Li
December 2016
THE RELATIONSHIP BETWEEN LEAD IN GROUNDWATER AND ELEMENTARY SCHOOL STUDENTS’ ACADEMIC PERFORMANCE IN MISSOURI

Geography, Geology and Planning

Missouri State University, December 2016

Master of Science

Lynnette Xiangling Li

ABSTRACT

Lead exposure can come from various sources, e.g., lead mines, industrial areas, and lead based paint. Missouri has more than 4,128 lead mines. Lead is a neurotoxin. The objectives of the study were to investigate i) the spatial relationship between lead mines and lead distribution in groundwater, ii) the statistical impact of lead in groundwater on students’ Missouri Assessment Program (MAP) scores, and iii) the geographic variations of lead’s impact on students’ MAP scores. Geographic Information Science spatial analysis tools were used to analyze the concentration of lead in groundwater within school districts in Missouri. Regression analysis was utilized to study the effect of neighborhood lead concentration in groundwater on students’ MAP scores. Geographically weighted regression (GWR) analysis results show the geographic variations of lead’s impact on students’ MAP scores across different areas in Missouri from significant to insignificant. The results of this study show no relationships between lead mines density and lead level distribution in groundwater. In localized areas, it shows the geographical variability of the spatial relationship between lead in groundwater and MAP scores. This study provides better understanding of the relationship between environmental factors and elementary students’ academic performance.

KEYWORDS: academic performance; cognitive development; geographic information systems; geographically weighted regression; lead; Missouri Assessment Program.

This abstract is approved as to form and content

Dr. Xiaomin Qiu
Chairperson, Advisory Committee
Missouri State University
THE RELATIONSHIP BETWEEN LEAD IN GROUNDWATER AND
ELEMENTARY SCHOOL STUDENTS’ ACADEMIC
PERFORMANCE IN MISSOURI

By

Lynnette Xiangling Li

A Master’s Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science, Geospatial Sciences in Geography and Geology

December 2016

Approved:

______________________________
Dr. Xiaomin Qiu

______________________________
Dr. Jun Luo

______________________________
Dr. Robert T. Pavlowsky

______________________________
Dr. Julie Masterson: Dean, Graduate College
ACKNOWLEDGMENTS

I would like to express my deepest and sincere gratitude to my thesis adviser, Dr. Xiaomin Qiu, for her unwavering support, wisdom, mentorship and selfless care throughout this thesis. I give thanks to call her a mentor in demonstrating excellent academic scholarship.

I would like to extend my thanks to faculty and staff of the Geography, Geology and Planning department who offered collegial guidance and support over the years: Dr. Jun Luo, Dr. Robert Pavlowsky, and Dr. Melida Gutierrez. Special thanks are due to Dr. Judith Meyer, for her mentorship during my early years as a graduate student.

My gratitude extends to a team of sojourners who have encouraged me on: Emily Rice-Austin, Kati Tomlin, and Carol Minton.

I would like to thank Dr. Nathan Meleen, Becky and Roger Bush for encouraging me to explore the world with great curiosity.

I want to thank my Mom, Alice, my sisters, Grace and Lynndy, who have been supportive in every way possible.

And to my companion and love in life, thank you for being on this journey with me.

I dedicate this thesis to my father, John Tong Seng Lee (1949-2007).
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter 1. Introduction</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose of Study</td>
<td>2</td>
</tr>
<tr>
<td>Research Objectives</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 2. Literature Review</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lead and its Effect on Children</td>
<td>11</td>
</tr>
<tr>
<td>Groundwater and its Uses</td>
<td>14</td>
</tr>
<tr>
<td>K-12 Student Assessment in Missouri</td>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 3. Methodology</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Framework</td>
<td>20</td>
</tr>
<tr>
<td>Academic Performance Factors</td>
<td>20</td>
</tr>
<tr>
<td>Methods</td>
<td>23</td>
</tr>
<tr>
<td>Kernel Density Analysis</td>
<td>25</td>
</tr>
<tr>
<td>Global (Traditional) Regression Analysis</td>
<td>26</td>
</tr>
<tr>
<td>Geographically Weighted Regression Analysis</td>
<td>27</td>
</tr>
<tr>
<td>Study Area</td>
<td>30</td>
</tr>
<tr>
<td>Data Processing and Editing</td>
<td>34</td>
</tr>
<tr>
<td>Lead in Groundwater Data</td>
<td>35</td>
</tr>
<tr>
<td>Socio-Economic and School Data</td>
<td>40</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 4. Results and Discussion</th>
<th>48</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kernel Density Analysis Results</td>
<td>48</td>
</tr>
<tr>
<td>Global (Traditional) Regression Analysis Results</td>
<td>51</td>
</tr>
<tr>
<td>Geographically Weighted Regression Analysis Results</td>
<td>58</td>
</tr>
<tr>
<td>Discussion</td>
<td>64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 5. Limitations of the Study and Conclusion</th>
<th>71</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limitations</td>
<td>71</td>
</tr>
<tr>
<td>Conclusion</td>
<td>72</td>
</tr>
</tbody>
</table>

| References                                        | 74 |
LIST OF TABLES

Table 1. Pearson Correlation Coefficients for All Paired Variables...............................54
Table 2. Regression Coefficients Statistics from OLS (Final Model)..........................56
Table 3. Regression Coefficients Statistics from OLS ....................................................57
LIST OF FIGURES

Figure 1. Drinkable Groundwater in Missouri ................................................................. 7
Figure 2. Lead Mines and School Districts in Missouri ..................................................... 10
Figure 3. A Statistical Model for School Performance .................................................... 21
Figure 4. Variables used in Traditional Regression Analysis ......................................... 24
Figure 5. Lead Mines in Missouri .................................................................................. 31
Figure 6. USGS Monitored Wells and NWIS Gauge Sites ............................................ 36
Figure 7. Location of Accepted Lead in Groundwater Samples .................................... 37
Figure 8. Example of a Raster Produced by Control Points .......................................... 40
Figure 9. Lead Zonal Mean for School Districts ............................................................ 41
Figure 10. Missouri Assessment Program Score: Communication Arts (Year 2007) .... 43

Figure 11. Maps Depicting: A) Percentage of Students Receiving Free Lunch (LuPer); B) Teacher Average Years of Experience (TeExp); C) Percentage of Teachers with Certification (TeCert); D) Percentage of Teachers with Master’s degree (TeMs); E) Teacher-student ratio (TeStuR); F) Teacher average total salary (TeSal) ...................... 46

Figure 12. Maps Depicting: A) Per Student Expenditure (PerStu_Exp); B) Percentage of College Degree Holders in School Districts (Edu_Bach); C) Married Couple Families with Children under 18 years (MCFC18) .................................................................... 47

Figure 13. Kernel Density of Lead Mines ........................................................................ 49
Figure 14. Lead Levels in Groundwater (County Level) .................................................. 50
Figure 15. Scatterplot of TeSal versus TeMs ................................................................ 55
Figure 16. Scatterplot of TeSal versus TeStuR ............................................................... 55
Figure 17. Local Residuals ($R^2$) from GWR (bandwidths of 30, 60 and 100) .......... 60

LIST OF FIGURES (CONTINUED)
Figure 18. Local $F$-value results from GWR (bandwidths of 30, 60 and 100) ..................62

Figure 19. Local Standardized Residuals (bandwidths of 30, 60 and 100) .....................63

Figure 20. Local Coefficient $t$-values of Lead in Groundwater from GWR (bandwidths of 30, 60 and 100) ..........................................................67
CHAPTER I: INTRODUCTION

Historically, some kingdoms, empires, and dynasties relied on naturally endowed resources to fuel growth and development. The dependence on mined resources ended the Stone Age and ushered in the Bronze Age. Similarly, the United States has a long history of mining resources such as aluminum, coal, copper, gold, and lead. Much of today’s economic growth can be attributed to the success in finding mineral resources and its extraction.

Missouri has a long mining heritage which can be traced to the Native Americans, followed by the French-Canadian explorers, and later by European settlers (Burford, 1978). In the early 1720s, lead played a fundamental role in the establishment of the state of Missouri. Lead mining hastened settlement growth. It influenced the construction of roads and railroads. From an economic standpoint, it drove the wheel of commerce and industry, and gave rise to an economic base for generations of Missourians. Apart from having the highest lead deposits globally, Missouri continues to be the largest lead producer in the lower 48 states (U.S. Geological Survey, 2014).

Missouri, with more than four thousand lead mines, has the most lead deposits in the world as stated in a report by the Missouri Department of Natural Resources (Missouri Department of Natural Resources, 2002). While lead mining in Missouri benefits the state’s economy and development of infrastructure, the shortcomings of environmental decay caused by lead mining are overwhelming and detrimental. This can be seen from the widespread of lead wastes, chat piles and tailings around lead mines. The implementation of the Environmental Protection Act in 1970, led to Federal agencies
attempting to contain waste products from mining, milling and smelting. Lead waste containment is one of the environmental recourse when it comes to the problems and residual effects of lead. Lead mining has physically altered the natural landscape of the environment.

**Purpose of Study**

Through the past decades, research on spatial distribution of environmental risks and hazards focused on the relationship between a specific hazard and its ramifications. Many of these studies have been instrumental in shaping governmental policies and strategies to create a more conducive and sustainable environment. A study conducted by Stack et al., states that an increase in lead concentration in domestic water contributes significantly to the total body lead burden to the children studied (Stack et al., 1975).

Lead is a neurotoxin. Lead contamination has affected the health of the population residing within the vicinity of a lead mine. In 1974, the Safe Drinking Water Act was passed. As a result, the Environmental Protection Agency (EPA) regulates the safe levels of different chemicals in drinking water. Their standard for lead in drinking water is held at 15 parts per billion (ppb), as the EPA believes that with the current technology, this is the lowest level attainable by water systems to control lead (EPA, 1999).

While the threat of lead has diminished as a result of federal enforcements such as the Safe Drinking Act, a growing body of research suggests that there is no threshold for the adverse effects of lead (CDC, 2003; Tsekrekos and Buka, 2005). Rising evidences reflect the detrimental effects caused by lead that cannot be ignored (Bellinger et al., 1992; Canfield et al., 2003; Jusko et al., 2008; Lanphear et al., 2000; Moore et al., 1987;
Many studies report a significant negative correlation between children’s cognitive development and environmental lead exposure (Kim et al., 2008). For instance, an intelligent quotient (IQ) decline of 1-5 points is associated with an increase in lead in the blood (PbB) of 10µg/dL (Bellinger 2004; Koller et al. 2004; Lidsky and Schneirer 2003; Needleman 2004). On the other hand, a large number of experimental and human data support the claim that blood lead levels of levels less than 10µg/dL have injurious effects on brain (cognitive) function, including lowered intelligence, behavioral problems, and diminished school performance (Baghurst et al. 1992; Bellinger et al., 1992; Cory-Slechta, 1997; Dietrich et al, 2004; Ernhart et al., 1989; Wasserman et al., 1997; Yule et al., 1981).

According to the National Health and Nutrition and Nutrition Examination Survey (NHANES), more than 12.8 million children and adolescents born from 1972 and 1988 were adversely affected by environmental lead exposure with blood lead concentrations greater than 32.5µg/dL (CDC, 2013). Another study estimated that 1 in 20 children are affected by lead toxicity (Kaufmann et al., 1998). There are many studies on the effects of lead on cognitive development in children and adolescents (Hubbs-Tait et al., 2007). However, research on the geographical variability of lead in the environment influencing academic performance is not available.

When it comes to assessing the educational progress of Missouri, the National Assessment of Educational Progress (NAEP) ranks the state’s education as average (NAEP, 2009). Every state is required to conduct an end of grade assessment for its public school education. In Missouri, the Missouri Assessment Program (MAP) fulfills that requirement mandated by the United States Department of Education. MAP is a set
of standardized tests conducted on a yearly basis to assess the students’ comprehension of skills such as reading, science, and mathematics. The Department of Elementary and Secondary Education of the state of Missouri conducts MAP tests on an annual basis to evaluate students’ academic performance from grades 3 to 10. As MAP tests are compulsory for a school district to maintain its accreditation, most schools comply by participating in the assessment program. Data from the MAP scores allow for a comprehensive picture of the academic performance of students in Missouri’s public schools. In addition to providing useful data to educators and administrators about each school or school district’s academic progress, the MAP test results from these standardized tests are beneficial as they provide information about student’s learning progress in Missouri.

There are many factors that contribute to higher standardized test scores. A study conducted by Qiu and Wu, using American College Test (ACT) scores, global regression analysis shows that high schools with higher parent income and educational level, double parent family, larger class size and more experienced teachers tend to have significantly higher ACT scores (Qiu and Wu, 2010). Since lead is a neurotoxin and current literature gives evidence to its adverse effects on cognitive development particularly to young children, this study seeks to establish the statistical and geographical variability between lead in groundwater and the end of grade assessment scores for elementary students in Missouri’s public school. This study will investigate the relation between lead in groundwater and students’ MAP score. As such, the MAP results are used as one of the nine variables in spatial and traditional statistical analyses, and to show the geographic variations of lead’s impact on elementary students’ standardized test scores across
different areas in Missouri. This study sheds light on accessing how living in an environment richly endowed with lead can affect local children’s academic performance. It also provides a possible framework for further research on how environmental factors can contribute to elementary school students’ academic performance.

**Research Objectives**

The main goal of this study is to determine if there is a statistically significant relationship between the lead concentration in groundwater and the end of grade assessment scores for elementary students in Missouri through traditional statistical and spatial analyses. This study is intended to investigate i) the spatial relationship between lead mines and lead distribution in groundwater, ii) the overall statistical impact of lead concentration in groundwater on elementary students’ MAP scores in the state of Missouri, and iii) the geographically varied impact of lead concentration in groundwater on students’ MAP scores. The detailed research questions that this study investigates include the following:

1. Do lead mine locations and lead concentration levels in groundwater have a spatial relationship?

Groundwater and surface water are not isolated components of the hydrologic system, but instead interact in a variety of physiographic and climatic landscapes (Sophocleous, 2002). Hence, pollutants such as lead that contaminate surface water supplies can seep into groundwater sources. Studies have reported significant negative correlations between children’s mental development and environmental lead exposure (Pocock et al., 1994; and Rosen, 1995). Lead in the environment may result in deficits in IQ, verbal memory, attention, executive
function, visual-motor integration, fine motor coordination, as well as impairment in language development (Counter et al., 2008). The first law of geography states that, “everything is related to everything else, but near things are more related than distant things” (Tobler, 1970). Lead is treacherous with ill health effects on individuals, notably pre-school aged children (Boreland and Lyle, 2008). Many lead mines are found in the southern half of Missouri. The lead mines in Missouri are distinctly located in the south-east, central and south-west regions of the state (Figure 1). There are major aquifers and critical groundwater regions in those regions (refer to Figure 1). Does groundwater closer to lead mines contain more lead? With the use of kernel density, this study analyzes the concentration of lead in groundwater and its relationship to lead mines based on distance.
Figure 1: Drinkable Groundwater in Missouri
(2) Do lead concentrations in groundwater have an overall statistical impact on elementary students’ MAP scores?

For several decades, Missouri has been one of the main producers of lead in the country. Missouri’s lead mines supply the constant domestic demand for lead. Dewatering, the process of draining groundwater from the mine shaft by dewatering pumps is an integral part of lead ore extraction. Thus, in regions where lead mines are located, groundwater maybe contaminated with lead.

More than half the population in Missouri is dependent on groundwater for its drinking water supplies. Some of these groundwater sources are located near lead mines; many are within critical watersheds (Figure 1). Given the possibility of lead contamination in groundwater near lead mines, do lead concentrations in groundwater affect children’s learning? While other factors such as teachers’ class preparation, school demographics and educational resources can influence standardized test scores, this study questions the role that lead in groundwater plays on elementary school students’ MAP scores. Regression analysis is used to study the effect of neighborhood lead concentration in groundwater on MAP scores.

(3) Are there geographic variations of the impact of lead concentration in groundwater on elementary school students’ MAP scores?

As early as 1720, humans have extracted lead ore in Missouri (Lippincott, 1912). Currently, there are 4,128 lead mines. While early lead mines were relatively shallow, many recent mines have shafts as deep as 290 feet. Unfortunately, many of these deeper lead mines are located in critical watersheds (Figure 1).
The evidence of lead’s detrimental effects on children’s cognitive development is growing. Currently, there is no safe threshold for lead in groundwater. A significant amount of the nation’s population is dependent on groundwater for daily usage. Missouri is no exception. Figure 2 maps the presence of lead mines in the school districts in Missouri. In certain regions, such as the southwest and central portions of Missouri, there are comparatively more lead mines than in northern parts of Missouri. The lead concentration in groundwater is heterogeneous across the state of Missouri. Traditional regression analysis is informative in showing the relationship between lead in groundwater and public elementary school academic performances. However, it does not show the geographic variances of lead levels in groundwater in relation to school districts’ academic performances. This study uses geographically weighted regression to investigate the spatial variability of lead levels in groundwater and its influence on public school students’ standardized test performances.

The findings of this study will form a basis to inform policy makers of the need to regulate and manage the containment of lead tailings and its spread. It may heighten the awareness of educators, policy makers and city planners to work toward reducing lead contaminants in Missouri. The literature on the relationship between academic performances in Missouri using GIS analysis is largely dismissed. The results will help us understand how environmental factors can produce a significant effect on academic performance.
Figure 2: Lead Mines and School Districts in Missouri
CHAPTER 2: LITERATURE REVIEW

Lead and its Effect on Children

Lead is everywhere in the environment. It has molecular characteristics that enable it to float in the air or attach to objects. Lead is treacherous with ill health effects on individuals, notably pre-school aged children (Boreland and Lyle, 2008). Lead is a neurotoxin. It is especially toxic to young children and fetuses. According to Koller et al, (2004), children are more susceptible to lead exposure due to the following three reasons: (1) young children are at greater risk of ingesting environmental lead through normal mouthing behaviors; (2) absorption from the gastrointestinal tract is higher in children than adults; and (3) the developing nervous system is thought to be far more defenseless to the toxic effects of lead compared to the developed brain.

Environmental conditions, such as, lead in the environment can delay cognitive development, can be credible reasons for school districts inability to perform. Many of the effects of lead on the central nervous system functioning are difficult to assess in infants or toddlers with the degree of reliability and precision that is possible in older children (Laphear et al., 2000). Needleman et al., found that asymptomatic (showing no evidence of disease) children with somewhat high lead levels in the dentin of shed deciduous teeth\(^1\) (greater than 20\(\mu\)g/g) were at significantly higher risk of failing to complete their secondary school program than children with lead levels less than 10\(\mu\)g/g. They were also at greater risk of achieving reading scores two or more grades below the expected level.

\(^1\) The research tested lead content in the teeth of children.
Important sources of lead exposure affecting blood lead concentrations in the general population are drinking water distributed through lead pipes (Meyer et al., 2003; and Watt et al., 2000), house paint containing lead (Lanphear and Roghmann, 1997), automobile exhausts from leaded gasoline (Rodamilans et al., 1996; Cowie et al., 1997), lead-glazed household ceramics (Rojas-Lopez et al., 1994; Romieu et al., 1995), industrial emissions and mining activity (Trepka et al., 1997; Morales Bonilla and Mauss, 1998).

The United States Department of Health and Human Services (DHHS) has determined that lead and its compounds are human carcinogens based on the evidence from studies in both human and animal (DHHS, 2004). Studies carried out by the Centers for Disease Control and Prevention (CDC) show that the levels of lead in blood in U.S children have been getting significantly lower and lower (CDC, 2012). This is primarily due to the banning of lead from gasoline, residential paint, and solder used for food cans and water pipes. Overall, blood lead concentrations in children in the US have drastically decreased over time (Jones et al., 2009). In 1997, out of the children tested, 7.6% had blood lead levels above the elevated levels of 10µg/dL. By 2006, the percentage of children with elevated blood lead levels decreased to 1.21% of those tested (CDC, 2013). Missouri reflects a relatively similar downward trend. Although encouraging, this does not negate the fact that a sizeable number of children have blood lead concentrations that exceed 0.48µmol/L, the CDC’s current intervention level. As discussed earlier, the adverse effects of lead on cognitive development can occur at blood lead concentrations below 0.48µmol/L. This highlights that children with blood lead levels above this level will significantly underestimates the size of the affected population.
The World Health Organization (WHO) and the Center for Disease Control and Prevention (CDC) recognized that there is no discernable threshold for the adverse effects of lead exposure (Canfield et al., 2003), however too few studies had examined children with blood lead levels less than 10µg/dL to support any firm conclusions (Lanphear et al., 2005). There is emerging evidence that lead associated intellectual deficits occur at blood levels less than 10µg/dL. The impact of low-level environmental lead exposure on the health of the public is substantial. Environmental lead exposure has been linked with an increased risk for numerous conditions and diseases that are prevalent in industrialized society (Kaufman et al., 2001; Kordas et al., 2004). Studies have reported significant negative correlations between children’s mental development and environmental lead exposure (Chen, et al, 2005; Pocock et al., 1994; and Rosen, 1995). Examples of such are deficits in IQ, verbal memory, attention, executive function, visual-motor integration, fine motor coordination, as well as impairment in language development (Canfield et al., 2003; Counter et al., 2008). In addition, reading problems, school failure, delinquent behavior, hearing loss, tooth decay, spontaneous abortions, renal disease, and cardiovascular disease are effects associated to lead exposure (Borja-Aburto et al., 1999; Dietrich et al., 2001; Factor-Litvak et al., 1999; Lin et al., 2003; Moss et al., 1999; Nash et al., 2003; Needleman et al., 2002; Schwartz and Otto, 1991).

Some studies demonstrate a decreased performance on standardized IQ tests for school-aged children due to lead found in their blood (Bellinger et al., 1992; Canfield et al., 2003; Chiodo et al. 2004; Dietrich et al., 1993; Schnaas et al., 2006; Tong et al., 1996). The damage to the nervous system and brain caused by lead is irreversible and in extreme cases, permanent nerve damage (Luo et al., 2012). Some studies indicate that
blood lead level (BBL) of less than 100mg/L or less, it can cause intellectual impairment (Canfield et al., 2003), some serious cognitive impairment (Bellinger et al., 2003) and neurobehavioral deficits (Chiodo et al., 2004; Chiodo et al., 2007).

With regard to school performance, a Taiwanese study of 934 grade three children (with a mean blood lead level concentration of 0.27µmol/L) demonstrated that children with higher blood lead levels had lower class rankings in Chinese, history, mathematics and science (Wang Cl et al., 2002). A similar finding was reported in a study of a cohort of about 200 Danish second graders (Lyngbye et al., 1990). The results indicate that lead has a greater influence on linguistic capabilities than mathematical abilities. Early childhood low-level lead exposure has been associated to high school failure and reading disabilities (Miranda et al., 2007; Needleman et al., 1990). The adverse effects of lead exposure on reading and other deficits in language-based abilities are of particular importance, primarily due to the fact that they are potent predictors of academic achievement and antisocial behavior (Moffitt and Henry, 1991; Ruff et al., 1993). In addition, delinquent behavior has been linked to asymptomatic lead exposure (Needleman et al., 2002).

**Groundwater and its Uses**

Fifty percent of the United States population depends on groundwater for daily drinking water. Groundwater is generally a safe source of drinking water. Unfortunately, groundwater is susceptible to pollutants. Groundwater and surface water are not isolated components of the hydrologic system, but instead interact in a variety of physiographic
and climatic landscapes (Sophocleous, 2002). Hence, pollutants such as lead that contaminate surface water supplies can make way to groundwater sources.

Groundwater contamination can be caused by human activity such as lead mining, soluble or insoluble substances are introduced into the hydrogeologic environment. In localities where groundwater contamination is severe, the continued use of water could lead to serious health problems. Groundwater contamination is difficult to contain or control. Governmental policies have been directed at its early detection and treatment.

Mining activities have a significant influence on the quality and quantity of water resources in their surrounding environment. The extraction of lead consists of four stages: crushing and grinding, flotation, filtering and dewatering, and tailings disposal. It is important to note that water is added at all stages of this process. Thus, due to the large amount of water used in processing lead ore, lead has the ability to enter groundwater resources directly from industrial sources (Rösner, 1998). Mine drainage from existing and abandoned lead mine sites may be considered as the most significant sources of environmental contamination association with mining activity (Aslibekian and Moles, 2003). Metals from mining areas may enter groundwater aquifers by seepage from tailings or mine-water impoundments, discharge from passive treatment systems, or erosion of tailings deposits during runoff events (Besser et al., 2007; Jennette et al., 1979). In addition, lead ore extraction requires the mine to be dewatered. The process of dewatering is to drain groundwater from the mine shaft; often done by dewatering pumps. This prevents groundwater from flooding the mines. Between the 1950’s and 1960’s, when pumpage declined, the abandoned mine drifts and shafts filled with water (USGS,
Hence, the interaction between groundwater and lead is greater in these regions with lead mines.

Nationally, much is known about the hazards of lead once in the body, but little research has been done to determine actual lead exposure from drinking water, and the information that does exist is dated (US Government Accountability Office, 2004). Surveys conducted in Canada and the United States indicate that drinking-water supplies leaving treatment plants contain 2-8µg/L lead (EPA, 1986; Dabeka et al., 2002). The EPA estimated that less than 1% of the public water systems in the USA have water entering the distribution system with lead concentrations above 5µg/L (EPA, 1991). While lead pollution from localized, obvious sources (point sources) may be easy to identify and control, resulting improvements in water quality may be disappointing (Levin et al., 2008). Consequently, attention has shifted in recent years to the more diffuse (nonpoint) sources of pollution, which are more difficult to control, and to those ubiquitous pollutants, such as lead, which originate from point sources but disperse before entering water bodies (Chesters and Schierow, 1985). Until recently, few people realized the extent of underground contamination or its adverse impact on groundwater quality.

Groundwater contamination may be due to infiltration of contaminated surface water. Lead is likely transported in groundwater by mobile particulate matter (McDowell-Boyer et al., 1986; Well et al., 1989). A study conducted by El Khalil et al. (2008), shows that metal toxicity was highest in monitoring stations closest to lead tailings deposits and decreased as the distance from the tailings increased. The study concludes that stream waters heavily contaminated with toxic metals contribute to the contamination of the soils.
located along the water flow and may seep through fissured and faulted zones, leading to groundwater contamination.

Surface water can be contaminated by lead in many ways such as runoff from tailing piles, airborne lead particles and lead in the environment. This study will not regard surface water as a variable that would affect cognitive development in Missouri students, Missouri households do have a strong dependence on it as they do with groundwater. Groundwater will be analyzed with geographic variance to show if there is a relationship between MAP scores and lead in groundwater.

K-12 Student Assessment in Missouri

For decades, the United States led the world as a successful and economic powerhouse. This is largely due to the quality and attainment in education of the American population. There is an indisputable correlation between the economic prosperity of a state, the quality and level of education, and skill of its citizens (Missouri Senate Educated Citizenry Commission 2020, 2010). In recent years, America has dwindled in rank to 10th in its industrialized peers of young adults with college degrees. The state of Missouri stands at 30th in the nation in 25-34 year olds with a degree past high school. This trails far behind neighboring states such as Kansas, Nebraska, Iowa and Illinois.

The United States’ Department of Education identifies the need for performance assessment systems in every state. The rationale is held together by the necessity to (a) monitor student progress toward desired outcomes, (b) hold schools and educators accountable for students’ academic performance, (c) certify students’ ability to
demonstrate academic robustness of skills and capabilities, (d) comparatively assess national educational reform efforts, and (e) influence and promote shifts in pedagogy (strategies of instruction) to emphasize higher order thinking and analytical skills (U.S. Department of Education, 1997). In 1993, Missouri’s Department of Elementary and Secondary education (DESE) developed its “Show-Me Standards.” The aim was to produce high school graduates who lead productive and successful lives as they continue their education, enter the work force and assume civic responsibilities.

As such, the Missouri Assessment Program (MAP) was developed to nurture young Missourians in that direction. The MAP assessment covers six subject areas: mathematics, communication arts, science, social studies, physical education, and fine arts. It evaluates students’ proficiency in the knowledge, skills and competencies. Its purpose is to provide precise, clear and relevant information about the effectiveness of Missouri’s schools. Such information is vital in monitoring districts’ progress, the effectiveness of educational programs, and a means to compare Missouri students to other schools and students within the state. As of spring of 2001, the MAP assessment was mandatory for schools as well as school districts to maintain their accreditation. The DESE is aware of a variety of sources for assessment error. While MAP scores are the most reliable yardstick to measure and comprehend Missouri students’ academic performances, standardized tests do not show the totality of what public school students know. With regards to this research, the MAP test scores provide this study a consistent and reliable method to measure public elementary school students’ academic proficiency in Missouri.
In this study, the relationship between lead in groundwater and students’ academic performance is investigated. Children are particularly sensitive to the effects of lead and are more vulnerable due to their rapidly developing nervous systems (Bellinger 2004; Koller et al. 2004; Lidsky and Schneirer 2003; Needleman 2004). Hence, it was more apt to use the standardized test scores of the Missouri Assessment Program (MAP) as the data reflects the demographics I am interested in.
CHAPTER 3: METHODOLOGY

Research Framework

**Academic performance factors.** The performance of a school district is often measured by its students’ standardized scores. There are many methods to measure academic performances. Standardized testing such as the Scholastic Aptitude Test (SAT) and American College Test (ACT). Past studies state that school characteristics are a significant factor in producing well-performing schools (Ehrenberg and Brewer, 1997; Fowler and Walberg, 1991; and Hanushek, 1986). These studies assume that inputs such as teachers’ qualification, classroom size and curricula positively affects student standardized test scores.

While research pertaining to students’ academic performances has mainly focused on school characteristics and individual-level variables, few studies investigate the environmental and geographic variations of students’ performances. A study conducted by Qiu and Wu, adopted a multi-leveled framework to combine nine variables such as American College Test (ACT) scores, school characteristics, teacher characteristics, and student characteristics to study their spatial relationship (Qiu and Wu, 2011). The nine school-level variables belong to three categories of student characteristics, teacher
characteristics and school characteristics respectively (Figure 3).

![Diagram of school performance factors]

**Figure 3: A Statistical Model for School Performance**
LuPer = Percentage of students receiving free lunch; Edu_Bach = college degree holders; MCFC18 = Married coupled families with children under 18 years; TeExp = years of teaching experience; TeCert = teachers certification percentage; TeMs = teachers master’s degree percentage; TeStuR = Teacher-student ratio; TeSal = teacher salary; PerStu_Exp = Per student expenditure

Firstly, studies show that the affluence and educational attainment of parents contribute to higher test scores (Berlin and Sum, 1988; Hanushek, 1986, Nobel et al. 1999). In addition, family background traits such as single parent households influence test score negatively (Ehrenberg and Brewer 1997; Fotheringham et al. 2002). Students
who are eligible for free or reduced lunch are from lower income households. This information is used as a socio-economic indicator of the level of affluence in the school districts. The factors for student characteristic are percentage of students receiving free lunch (LuPer), percentage of people age twenty-five or over in the school district with a college degree (Edu_Bach), and percentage of married-couple families with children under eighteen within the school district (MCFC18).

Secondly, based on literature in the field of education, effective teachers produce students with comparatively higher test scores (Goldhaber and Brewer, 2000; Tuckman, 1971). Students who have teachers with master degrees are found to have higher test scores. The more education and experience on the part of the teacher, the higher the test scores comparatively (Goldhaber and Brewer, 1997; Hanushek, 1986). Some studies indicate that qualities of an effective teacher include teachers who are well prepared on subject matters. The type of certification a teacher holds is reported to be a vital determinant of student outcomes (Shen, 1997; Strauss, and Sawyer, 1985). Therefore, the explanatory factors in teacher characteristic category are teacher average years of experience (TeExp), percentage of teachers with certification for their teaching assignment (TeCert), and percentage of teachers with master’s degree (TeMs).

Thirdly, various school characteristics produce a more conducive environment for student learning. Most teacher salaries are directly related to years of teaching experience and educational levels. Thus, the higher teacher salaries positively affect student performance (Hanushek, 1986). Studies show that variables such as expenditure per student and class room size may explain the vast variation in student scores (Hanushek, 1986, Darling-Hammond, 1999; Goldhaber and Brewer, 2000). As such, independent
variables for school characteristics are teacher-student ratio (TeStuR), teacher average total salary (TeSal), and school district per student expenditure (PerStu_Exp).

As this study seeks to explore statistical and geographic variations of the impact of lead concentration in groundwater on elementary students’ score, lead in groundwater will serve as the tenth predictor variable (Pb_ZnMean). For the purpose of this study, the dependent variable is the 2007 MAP scores for grade four communication arts (COM_G4). Based on the literature of previous studies, results indicate that lead has a greater influence on linguistic capabilities than mathematical abilities (Lyngbye et al., 1990). In addition, early childhood low-level lead exposure has been associated to reading disabilities (Needleman et al., 1990). The adverse effects of lead exposure on reading and other deficits in language-based abilities are of particular importance, primarily due to the fact that they are potent predictors of academic achievement (Moffitt and Henry, 1997). Thus, the MAP scores for grade 4 public school students’ communication arts is a good fit for the use of this study. With that, there are a total of 11 variables used for the traditional regression analysis (Figure 4).

**Methods.** This study is intended to answer the following three questions: (1) Do lead mine locations and lead concentration levels in groundwater have a spatial relationship? (2) Do lead concentrations in groundwater have an overall statistical impact on elementary students’ MAP scores? (3) Are there geographic variations of the impact of lead concentration in groundwater on elementary school students’ MAP scores? The first question seeks to determine a spatial relationship between the location of lead mines and lead concentration levels in groundwater. For the purpose of the study, kernel density
Figure 4: Variables used in Traditional Regression Analysis

analysis is used to show the relationship between lead mines and lead concentration in groundwater. The second question seeks to determine if lead concentration levels in groundwater have a statistical impact on elementary students’ MAP scores. For the purpose of the study, global regression analysis is employed to study the global effect of neighborhood lead concentration in groundwater on students’ MAP scores. Lastly, the third question seeks to determine if the impact of lead in groundwater on MAP test scores produce spatial variability that can be observed. Geographically weighted regression
(GWR) analysis results is used to demonstrate the local impact of lead concentration in groundwater on elementary students’ MAP scores across different areas in Missouri.

**Kernel Density Analysis.** The first objective of this study is to see the relationship between lead mines and the concentration of lead in groundwater. Lead mines are not homogenously distributed (Figure 1). There are major aquifers and critical groundwater regions in those regions (Figure 2). Groundwater and surface water are not isolated components of the hydrologic system, thus this portion of the study uses kernel density to analyze the geographic relationship between lead mines and the levels of lead in groundwater.

Density analysis distributes the measured quantities of an input feature layer across a landscape to create a continuous raster surface. It is based on the quantity that is measured at each location and the spatial relationship of the locations of the measured quantities. Kernel density estimation calculates the density of features in a neighborhood around those features. Theoretically, it uses a kernel function to fit a smoothly curved surface over each point. In the raster output of kernel density, the location of the point feature has the highest surface value. There is an inverse relationship between the surface value and the location of the point feature. As the distance from the point increases, the surface value decreases. The density at each output raster cell is calculated by adding the values of all the kernel surfaces where they overlay the raster cell center (Silverman, 1986).

In this study, the input feature is represented by the lead mines. The kernel function calculates the density of the lead mines around each output raster cell. The surface value of the raster is highest at the location of the point and diminishes with
increasing distance from the point, reaching zero at the user-defined search radius
distance from the point. Validating the search radius has always been a problematic issue
for geographic information scientist. For the purpose of the study, search radiiuses
between 20,000 to 150,000 meters were experimented at increments of 5,000 meters. The
search radius of 50,000 meters is the optimum distance as it had the best fit. The cell
sized used for the interpolation of the level of lead in groundwater in Figure 13 was
50meters, thus the cell size for kernel density estimation analysis would be similar.

Global (Traditional) Regression Analysis. The second objective of this study is to
investigate the overall statistical impact of lead in groundwater and public school
elementary students’ MAP scores. Regression analysis used to explore the relationship
between communication arts MAP scores (the dependent variable) and the set of ten
predictor variables (Figure 4). It has been commonly used to test the usefulness and
functional relationship between predictor variables. The strengths of using regression
analysis are its simplified perspective of the relationship between variables, use of fitting
the model with given data, and its method for demonstrating the importance of variables
and the correctness of the model (Rogerson, 2006). For a single linear regression,
whereby there is an independent and one explanatory variable, the following equation is
used to find the best fit among the set of given points:

\[ \hat{y} = \beta_0 + \beta_1 x \]

where \( \hat{y} \) is the predicted value of the dependent variable, the observed value of the
independent variable is \( x \), the point where the line crosses the vertical axis is \( \beta_0 \), and the
slope of the line is represented by \( \beta_1 \). A regression that has one independent variable is
called a bivariate regression. For this study, a multiple regression analysis is used as there
are ten predictor variables that contribute to students’ academic performances. The following is its equation,

$$\hat{y} = a + b_1 + b_2 + \cdots + b_p$$

where \( p \) is the independent, and \( \hat{y} \) is the predicted value of the dependent variable. From given observations on the dependent (\( y \)) and independent (\( x \)) variables, the parameters of \( a \) and \( b_1, b_2, \ldots, b_p \) is sought. This is done by minimizing the sum of the squared residuals,

$$\{ \hat{y}_i - y_i \}^2 = \left( \hat{y} - \right)^2$$

The global regression weakness is that when the scale of the study area is comparatively large (i.e., for a state), the results from multiple regression may hide some noteworthy and vital local differences. The assumption that the phenomena or trend is homogenous across a large study area (i.e., entire states or countries) does not allow for the statistical evaluation of geographic differences (Rogerson, 2008).

**Geographically Weighted Regression Analysis.** Global statistics are normally single-valued and a mere description of the spatial autocorrelation in a data set. In the global regression model, depending on magnitude of the study area, statistical evaluation comparing geographic differences cannot be accounted for (King, 1969). Thus, a global model may be inappropriate when variable relationship is not consistent across the study area (Qiu and Wu, 2010).

Local statistics, such as GWR, allows for any descriptive statistic associated with a spatial data set whose value varies from place to place (O’Sullivan and Unwin, 2010). GWR technique is primarily based on views of regression from each observed data location (Rogerson, 2006). Observations that are near to the location are given larger weights compared to observations that are further. The reason is that observed data near
to point $i$ is assumed to have a greater influence in the estimation of regression coefficients compared to data located further from point $i$ (Fotheringham et al., 2000).

At location $i$, the dependent variable is modeled as follows:

$$ y_i = \beta_0 + \sum_{j=1}^{n} \beta_j x_{ij} + \epsilon_i $$

where $y_i$ is the dependent variable, $x_{ij}$ is the independent variable, and $\epsilon_i$ is the observation on variable $j$ at location $i$. In the model's equation, the $\beta$ coefficients have $i$ subscripts, it implies that they are specific to the location of observation $i$. The parameters change as the values of the location $i$ changes over the study area, thereby resulting in a surface of parameter estimates that is used to identify spatial variability.

In geographically weighted regression, every observation is weighted according to its proximity of location $i$. The weights of the weighted observations are assigned using the following formulas whereby:

$$ \ast = \sqrt{\ldots} $$

$$ \ast = \sqrt{\ldots} \quad j = 1 \ldots , n $$

At the given location of $i$, linear regression of the * on * is processed. From the results, the coefficients are used to predict the value of $y$ at location $i$. Next, the squared difference between the observed value of $y$ and the predicted value

$$ \{ y - \hat{y}(\ldots) \}^2 $$

At location $i$, the predicted value of the dependent variable at location $i$ is $\hat{y}(\ldots)$. This is done when observation $i$ is not used in the estimation and the $\beta$ indicates that the above prediction was made using a specific value of $\beta$. After this is
repeated for each location $i$, the computation of the total sum of squared deviations between observed and predicted values as:

$$
(\ ) = \sum_{1}^{n} ( - \hat{y}( \ ))^2
$$

After which, this process is repeated for many values of $\beta$ to choose the optimal value of $\beta$ to minimize the score of $s(\beta)$. At the end of the day, the final value of $\beta$ produces the best sets of weights. From there, the final regression coefficients at each location uses weighted based on the optimal value of $\beta$ to set the parameters of the weights, and then regress $y^*$ and $x^*$ using all the observations.

The choice of bandwidth has a tremendous impact on the results obtained from Geographically Weighted Regression (GWR). The bandwidth can be understood as a smoothing parameter, with larger bandwidths causing greater smoothing. The goal of a GWR user is to find an optimal bandwidth that provides a good equilibrium between the extremes of an oversmooth and undersmoothed model.

There are two methods used to find the optimal bandwidths for GWR. They are user-supplied bandwidth, and the Golden selection search. The user-supplied bandwidth is when the GWR analyst has a suitable value for the bandwidth. Most often this is from previous experience and computation of GWR with a similar or particular model. On the flip side, the user may have a theoretical rational for employing a certain bandwidth. The next method estimation by cross-validation is there is no prior justification for supplying a particular bandwidth; the GWR user may let the software choose an appropriate bandwidth (Fotheringham et. al, 2002). For the purpose of this study, the user-supplied bandwidth is used. The bandwidths are 30, 60 and 100.
Study Area

Missouri has 8,371 mines that are spread over 69,677 square miles. These mines produce mineral ore such as aluminum, barite, barium, copper, iron, lead, silver, and zinc. Of the 8,371 mines, 4,128 are lead mines. According to historic records, in as early as the 1700, Missouri had small lead mines that produced approximately 1,500 pounds of lead ore per day. This commodity was exported to France via the Mississippi River (Burford, 1978). In 1967, the General Assembly of Missouri designated Galena as the official state mineral of Missouri. Galena (lead sulfide), has propelled Missouri as the premier lead producer of the world.

There are 115 counties in Missouri. List 1 accounts for the number of lead mines in each county. Lead mines are heavily concentrated in the south-west, central and south-east regions of Missouri (Figure 5). The distribution of lead mines is predominantly in the central mining district, new lead belt, old lead belt and Tri-State Lead-Zinc district. Twenty-four counties have no lead mines. Seventy-four counties have less than ten lead mines within their county boundaries. There are thirteen counties with more than a hundred lead mines (List 1). Figure 5 shows the number of lead mines in Missouri counties. Jasper and Newton counties have the most number of lead mines of 2,482 and 1,716 respectively.

The total all-time production estimate of lead is more than 17 million tons and valued at approximately 5 billion US dollars (Missouri Department of Natural Resources, 2006). The lead mines in Missouri are distinctly located in the south-east, central and south-west regions of the state (refer to Figure 5). Today, all of Missouri’s lead
Figure 5. Lead Mines in Missouri

production comes from the Viburnum Trend district over a very narrow, 56.327 km (35 miles) long ore district. In 1992, nine Missouri mines produced a total of 330,000 tons of lead (75 percent of the U.S total) valued at more than $230 million (Missouri DNR, 2002). In 2000, Missouri produced 313,105 tons of lead that was estimated at $128,838,880 (Missouri DNR, 2002).
<table>
<thead>
<tr>
<th>County</th>
<th>Lead Mines</th>
<th>County</th>
<th>Lead Mines</th>
<th>County</th>
<th>Lead Mines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adair</td>
<td>3</td>
<td>Greene</td>
<td>154</td>
<td>Pemiscot</td>
<td>0</td>
</tr>
<tr>
<td>Andrew</td>
<td>9</td>
<td>Grundy</td>
<td>2</td>
<td>Perry</td>
<td>16</td>
</tr>
<tr>
<td>Atchison</td>
<td>0</td>
<td>Harrison</td>
<td>1</td>
<td>Pettis</td>
<td>7</td>
</tr>
<tr>
<td>Andrain</td>
<td>2</td>
<td>Henry</td>
<td>13</td>
<td>Phelps</td>
<td>30</td>
</tr>
<tr>
<td>Barry</td>
<td>71</td>
<td>Hickory</td>
<td>49</td>
<td>Pike</td>
<td>4</td>
</tr>
<tr>
<td>Barton</td>
<td>6</td>
<td>Holt</td>
<td>2</td>
<td>Platte</td>
<td>0</td>
</tr>
<tr>
<td>Bates</td>
<td>6</td>
<td>Howard</td>
<td>0</td>
<td>Polk</td>
<td>23</td>
</tr>
<tr>
<td>Benton</td>
<td>29</td>
<td>Howell</td>
<td>23</td>
<td>Pulaski</td>
<td>8</td>
</tr>
<tr>
<td>Bollinger</td>
<td>2</td>
<td>Iron</td>
<td>35</td>
<td>Putnam</td>
<td>2</td>
</tr>
<tr>
<td>Boone</td>
<td>8</td>
<td>Jackson</td>
<td>7</td>
<td>Ralls</td>
<td>4</td>
</tr>
<tr>
<td>Buchanan</td>
<td>4</td>
<td>Jasper</td>
<td>2484</td>
<td>Randolph</td>
<td>1</td>
</tr>
<tr>
<td>Butler</td>
<td>2</td>
<td>Jefferson</td>
<td>204</td>
<td>Ray</td>
<td>20</td>
</tr>
<tr>
<td>Caldwell</td>
<td>1</td>
<td>Johnson</td>
<td>9</td>
<td>Reynolds</td>
<td>18</td>
</tr>
<tr>
<td>Callaway</td>
<td>5</td>
<td>Knox</td>
<td>1</td>
<td>Ripley</td>
<td>10</td>
</tr>
<tr>
<td>Camden</td>
<td>54</td>
<td>Ladede</td>
<td>7</td>
<td>Saline</td>
<td>11</td>
</tr>
<tr>
<td>Cape</td>
<td>2</td>
<td>Lafayette</td>
<td>0</td>
<td>Schuyler</td>
<td>0</td>
</tr>
<tr>
<td>Girardeau</td>
<td></td>
<td>Lawrence</td>
<td>574</td>
<td>Scotland</td>
<td>0</td>
</tr>
<tr>
<td>Carroll</td>
<td>0</td>
<td>Lewis</td>
<td>3</td>
<td>Scott</td>
<td>1</td>
</tr>
<tr>
<td>Carter</td>
<td>6</td>
<td>Lincoln</td>
<td>1</td>
<td>Shannon</td>
<td>8</td>
</tr>
<tr>
<td>Cass</td>
<td>0</td>
<td>Linn</td>
<td>0</td>
<td>Shelby</td>
<td>0</td>
</tr>
<tr>
<td>Cedar</td>
<td>1</td>
<td>Livingston</td>
<td>0</td>
<td>St. Charles</td>
<td>0</td>
</tr>
<tr>
<td>Chariton</td>
<td>9</td>
<td>Macon</td>
<td>0</td>
<td>St. Clair</td>
<td>16</td>
</tr>
<tr>
<td>Christian</td>
<td>174</td>
<td>Madison</td>
<td>120</td>
<td>St. Francois</td>
<td>247</td>
</tr>
<tr>
<td>Clark</td>
<td>3</td>
<td>Maries</td>
<td>9</td>
<td>St. Louis</td>
<td>4</td>
</tr>
<tr>
<td>Clay</td>
<td>1</td>
<td>Marion</td>
<td>0</td>
<td>St. Louis</td>
<td>1</td>
</tr>
<tr>
<td>Clinton</td>
<td>0</td>
<td>McDonald</td>
<td>5</td>
<td>City</td>
<td></td>
</tr>
<tr>
<td>Cole</td>
<td>77</td>
<td>Mercer</td>
<td>1</td>
<td>Ste.</td>
<td>35</td>
</tr>
<tr>
<td>Cooper</td>
<td>15</td>
<td>Miller</td>
<td>352</td>
<td>Genevieve</td>
<td></td>
</tr>
<tr>
<td>Crawford</td>
<td>94</td>
<td>Mississippi</td>
<td>0</td>
<td>Stoddard</td>
<td>4</td>
</tr>
<tr>
<td>Dade</td>
<td>39</td>
<td>Moniteau</td>
<td>143</td>
<td>Stone</td>
<td>15</td>
</tr>
<tr>
<td>Dallas</td>
<td>25</td>
<td>Monroe</td>
<td>0</td>
<td>Sullivan</td>
<td>6</td>
</tr>
<tr>
<td>Daviess</td>
<td>0</td>
<td>Montgomery</td>
<td>0</td>
<td>Taney</td>
<td>18</td>
</tr>
<tr>
<td>DeKalb</td>
<td>0</td>
<td>Morgan</td>
<td>334</td>
<td>Texas</td>
<td>11</td>
</tr>
<tr>
<td>Dent</td>
<td>9</td>
<td>New Madrid</td>
<td>0</td>
<td>Vernon</td>
<td>3</td>
</tr>
<tr>
<td>Douglas</td>
<td>15</td>
<td>Newton</td>
<td>1716</td>
<td>Wayne</td>
<td>11</td>
</tr>
<tr>
<td>Dunklin</td>
<td>0</td>
<td>Nodaway</td>
<td>15</td>
<td>Webster</td>
<td>50</td>
</tr>
<tr>
<td>Franklin</td>
<td>207</td>
<td>Oregon</td>
<td>6</td>
<td>Worth</td>
<td>0</td>
</tr>
<tr>
<td>Gasconade</td>
<td>5</td>
<td>Osage</td>
<td>2</td>
<td>Wright</td>
<td>70</td>
</tr>
<tr>
<td>Gentry</td>
<td>1</td>
<td>Ozark</td>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
An aquifer is a geological unit with two fundamental functions. Firstly, they act as a water-bearing storage formation. And secondly, they function as a network of conduits (Driscoll, 1986). In the conduits, groundwater is constantly moving under a local hydraulic gradient. The rates of movement may vary from feet per year to feet per day. The groundwater resources in Missouri vary greatly across the state. They are categorized into groundwater provinces based on factors such as aquifer characteristics, changes in groundwater quality, and aquifer boundaries. Many lead mines are found in the southern half of Missouri. There are near major aquifers and groundwater sources in those regions (Figure 1).

According to the 2000 census, there were 5,595,211 people living in Missouri. The U.S Census Bureau estimates that in 2009, there were 5,987,580 residents in the state (U.S. Census Bureau, 2010). In the school year of 2009, a total of 892,279 students were enrolled in public elementary, middle and high schools (Missouri Department of Elementary and Secondary Education, 2010). There are 522 school districts with a total of 2,305 public schools to cater to the educational needs of Missourians. The lead mines are predominantly in the south-west, central and south-east regions of Missouri. Figure 2 shows the public schools that lie within regions that have lead in groundwater.

Data Collection, Editing and Processing

A major source of the spatial data used in this study was the Center for Applied Research and Environmental Systems. Block group level demographic data from the Census 2000 were obtained here. Data for school district statistics, for the 2007-2008 school year was obtained from Missouri Department of Elementary and Secondary Education.
Education (DESE, 2010) website. In addition, groundwater data and the locations of monitored wells were obtained from the National Water Information System (NWIS) and U.S Geological Survey (USGS) respectively.

**Lead in Groundwater Data.** For the purpose of the study, lead in groundwater was used as one of the nine variables. The rationale for this was the availability of data provided by geoscientists from the United States Geological Survey (Rolla, Missouri). Groundwater was sampled from various types of existing wells such as observation wells, monitor wells and water-supply wells (USGS, 2005). Observation wells are wells installed to observe and determine hydrologic characteristics of an aquifer from collected hydrologic data. Monitor wells are observation wells installed specifically for assessment of physical, chemical, and biological characteristics of the aquifer (USGS, 2005). Most of the groundwater data used in this study are from observation wells and monitor wells. The geographical coordinates of the locations of the USGS monitored wells and NWIS gauge sites are shown in Figure 6.

During site visits, a team of USGS hydrologic technicians compiled an inventory of the groundwater site. The individual well was purged to replace stagnant water with fresh groundwater. Immediately after the completion of the purging process, the well is sampled. An inventory of site and well information was entered into the NWIS water-quality and groundwater site inventory databases. Data of 2419 samples of detected dissolved lead in groundwater were obtained from NWIS. The samples were taken from wells with well depths between 30 to 1,600 feet. These samples were validated and accepted by a scientist from USGS. A geoscientist from the USGS validated the process.
For the purpose of this study, lead in groundwater samples with the term “accepted” in the database is used.

The EPA regulates the limit of lead in drinking water to 0.015 milligrams per liter (mg/L). Samples that exceed EPA’s regulation were used for the purpose of this study. A

Figure 6. USGS Monitored Wells and NWIS Gauge Sites
total of 2295 samples with dissolved lead from the original 2419 samples were used for interpolation purposes. The locations of wells where these samples were obtained are shown in Figure 7.

Figure 7. Location of Accepted Lead in Groundwater Samples

Spatial interpolation is the estimation of values of attributes at unsampled locations from data made at control points in the same area (O’ Sullivan and Unwin,
It attempts to create a reasonable estimation based on the values of a continuous field at places where the field has not actually been measured. There are three methods of spatial interpolation—Thiessen polygons, inverse-distance weighting, and Kriging.

Each method of spatial interpolation has their strengths and weaknesses. The Thiessen polygon method is relatively simplistic. This method builds a polygon around a nearest control point and assumes that each polygon has a uniform value as the control point. The benefit of using the Thiessen polygons for spatial interpolation is its simplicity. However, this method’s weakness is that there are abrupt spikes in values between polygons. Hence, this study will not use Thiessen polygons for spatial interpolation. The inverse-distance weighting (IDW) method of spatial interpolation is helpful in calculating the local mean by its capacity to give weights. It gives a spatial weight to locations that are nearer to control points. This method works best when control points are laid out in a grid of points over an area. While IDW produces a continuous and smooth interpolated surface, the spatial distribution of the groundwater sampling sites is in certain places sparse. Hence, this method will not be used. Of the three methods, Kriging is used for this study as the control point data are used to estimate the spatial structure in the underlying surface and this information is thereby used to determine appropriate spatial weights (O’Sullivan and Unwin, 2010). Spatial interpolation using Kriging allows for the smoothness of the output to be used in a statistically meaningful way (Longley, et al., 2005).

Data of lead in groundwater from monitored wells obtained from the NWIS database were used as control points. The map below shows the location of these points.
within the 522 school districts of Missouri. These points were interpolated using Kriging method with a cell size of 50m to create a raster.

Figure 8, shows an example of the raster produced by this process. However, this raster is slightly smaller than the geographic boundaries of the school districts. As such, it was necessary to create a new interpolated raster. In ArcMap, under the environmental settings for spatial interpolation, the perimeter for the processing extent for interpolation was changed to the entire extent of the school district polygon layer. The raster produced was used to calculate the zonal mean for the amount of lead in groundwater in a given school district.

The lead level in groundwater for each school district is calculated using zonal analysis to produce a zonal mean for each school district (thereby known as Pb_ZnMean). The Pb_ZnMean is the spatial measure of central tendency of the dispersion of lead in groundwater within the given school district. In order to obtain the zonal mean of lead levels in groundwater for school districts, ArcToolbox’s conversion tool is used to convert the school district polygon feature layer to a raster. The output is a school district raster layer which is used to compute the zonal statistics.

The initial lead zonal mean for five school districts showed negative readings of lead in groundwater. In order to have no negative readings, for these five school districts, the following equation was applied to transform the negative zonal mean scores to between zero and the next positive zonal mean:

\[
Y = \left( \frac{+7.06646 \times 0.0424955}{7.06646} \right)
\]
At this point, the data for the lead zonal means for all school districts are in a spreadsheet. The lead zonal mean statistics is joined to the school district polygon feature layer. The result is shown in Figure 9, a map showing the lead zonal mean for school districts.

**Socio-Economic and School Data.** Block group level demographic data to obtain the following variables were taken from the 2000 Census of the Center for Applied Research and Environmental Systems (CARES) database: percentage of people age
twenty-five or over in the school district with a college degree (Edu_Bach), and percentage of married-couple families with children under eighteen within the school district (MCFC18). These variables are under the student characteristics that affect academic performance.

![Lead Zonal Mean in Groundwater](image-url)

Figure 9. Lead Zonal Mean for School Districts

Demographic data from the 2000 Census block is used to derive the percentage of college degree holders in the school district over 25 years (Edu_Bach) and the percentage of married-couple families with children under eighteen within the school district (MCFC18).

---

Pb_Zn refers to the Lead Zonal mean (mg/L). It is the spatial measure of central tendency of the dispersion of lead in groundwater within the given school district.
Data for the following variables were taken from DESE’s database: percentage of students receiving free lunch (LuPer), teacher average years of experience (TeExp), percentage of teachers with certification for their teaching assignment (TeCert), percentage of teachers with master’s degree (TeMs), teacher-student ratio (TeStuR), and teacher average total salary (TeSal). As each school district has a unique identifying code, the data of LuPer, TeExp, TeCert, TeMs, TeStuR, and TeSal obtained from DESE’s database are used. The per-student expenditure for each school district (PerStuExp) is calculated by using total expenditure and student enrollment.

Data for the MAP test scores were obtained from the Missouri Department of Elementary and Secondary Education’s website. It has a comprehensive inventory of the data of the demographic, performance and resource reports of all the school districts in Missouri. Based on the literature of previous studies, from a theoretical standpoint the MAP scores for grade 4 public school students’ communication arts is selected a good fit for the use of this study. Using the identifying school district code, COM_G4 is added to its corresponding school district the Microsoft Excel spreadsheet. As such, the spreadsheet contains all 11 variables as shown in Figure 4.

There are 522 school districts in Missouri. Lead mines are heavily concentrated in the south-west, central and south-east regions of Missouri (Figure 2). The ground water in these regions has relatively high levels of lead (Figure 9). The school districts that are within those regions seem to perform between average and below average comparatively in MAP communication arts tests (Figure 10). The reason for using 2007 MAP communication arts scores is discussed in chapter 4.
Data for the following variables were taken from DESE’s database: percentage of students receiving free lunch (LuPer), teacher average years of experience (TeExp), percentage of teachers with certification for their teaching assignment (TeCert), percentage of teachers with master’s degree (TeMs), teacher-student ratio (TeStuR), and teacher average total salary (TeSal).

The percentage of students receiving free lunch (LuPer) is comparatively higher in the southeastern school districts of Missouri (seen from Figure 111a). On the other side, school districts that are near Kansas City and St Louis, have a comparatively lower
percentage of students requiring lunch assistance. The correspondence between communication arts MAP score (Figure 010) and LuPer is evident, more so toward the regions in blue and navy blue. This suggests that school districts with higher percentages of students receiving free lunch tend to have lower MAP scores.

The average years of teaching experience (TeExp) is shown in Figure 111b. Out of 522 school districts, thirty-four school districts (mostly located in the upper half of the state) have teachers with average teaching experience of greater than 15.1 years. Next, Figure 111c, illustrates the percentage of teachers with certification for their teaching assignment (TeCert). It shows a relatively random spatial distribution.

Based on Figure 111d, eighteen school districts have a comparatively higher percentage of teachers with master’s degrees (greater than 65 percent). Majority of school districts do not have greater than 30 percent of teaching faculty with masters’ degrees. Studies show that the more education and more experience on the part of the teacher, the better the standardized test performances (Battistich, et, al., 1995, Summers and Wolfe, 1977 and Tuckman, 1971). This may be true to a certain extend as school districts that have higher percentages of teachers with masters’ degree do have comparatively higher MAP test score in communication arts. This does not imply that the educational attainment of teaching faculty is the sole cause for better standardized test scores.

Under school characteristics, teacher-student ratio in Missouri is relatively high (seen in Figure 111e). This is more evident in school districts near population dense regions such as Jefferson City, Kansas City, St Charles, Springfield and St Louis. School districts in the north and northwest Missouri have a comparatively lower teacher-student ratio. This may be due to the smaller class size. In Figure 111f, teachers’ salaries are
comparatively higher in Kansas City and St Louis. In addition, there seem to be a correspondence between the school districts with higher teachers’ salary and per-student expenditure (PerStu_Exp) as seen in Figure 111g.

The process to derive the percentage of college degree holders in the school district over 25 years, the demographic data, Bachelor degree education attainment, was applied for the 2000 Census Block. A spatial join for the number of Bachelor degree holders at the Block group level was joined with the school districts layer. Next, the number of Bachelor degrees attained within the school district is divided by the number of population older than 25 years of age. The result of this computation, seen in Figure 111h, is the percentage of college degree holders in the school district over 25 years of age (Edu_Bach). Figure 111h., shows a comparatively higher percentage of college degree holders in areas such as Kansas City, Springfield, and St Louis. This could be attributed to the number of universities within the mentioned regions.

Figure 111i., shows the distribution of the percentage of married couple families with children under 18 years of age (MCFC18). The correspondence between communication arts MAP score (Figure 010) and MCFC18 is evident, more so toward the regions in blue and navy blue. This suggests that school districts with higher double parent families tend to have higher MAP scores, vice versa.
Figure 11. Maps Depicting: (A) Percentage of students receiving free lunch (LuPer); (B) Teacher average years of experience (TeExp); (C) Percentage of teachers with certification (TeCert); (D) Percentage of teachers with Master’s degree (TeMs); (E) Teacher-student ratio (TeStuR); (F) Teacher average total salary (TeSal)
Figure 12. Maps Depicting: (A) Per student expenditure (PerStu_Exp); (B) Percentage of college degree holders in school districts (Edu_Bach); and (C) Married couple families with children under 18 years (MCFC18).
CHAPTER 4: RESULTS AND DISCUSSION

Kernel Density Analysis Results

This study is interested in calculating the density of a phenomenon across a continuous space. As such, kernel density estimation (KDE), an extension of ArcGIS Spatial Analyst is used to show the presence or lack of relationship between a point feature (lead mine) and the phenomena (the varied levels of lead in groundwater in Missouri). The raster shown in Figure 212 is its product. The strongest relationship between lead mines and lead levels in groundwater are found in Jasper, McDonald and Newton counties. Between the counties of Newton and Jasper, there are 4,327 lead mines. These counties along with Lawrence County have relatively high levels groundwater lead as seen in Figure 22. Two clusters, as highlighted by the red outlines in Figure 212, show a weak relationship between lead mines and groundwater lead levels. The cluster, indicated by the solid red outline, with the counties of Coller, Camdem, Miller, Moniteau and Morgan, are in an area with comparatively lower level of lead in groundwater. There are more than 960 mines within this cluster. The cluster highlighted by a dashed red outline, with the counties of Crawford, Franklin, Jefferson, Ste. Genevieve, St. Francois, and Washington, is in an area with comparatively higher levels of lead in groundwater. It has more than 1450 lead mines. In Figure 312, Jefferson County has a sharp peak in lead levels in groundwater. This along with the fact that there is a larger presence of lead mines within the farther east cluster could explain for the comparatively stronger relationship between lead mines and lead levels in groundwater.
Figure 13. Kernel Density of Lead mines
Figure 14. Lead levels in Groundwater (County Level)
Regions with significantly high levels of lead in groundwater show no relationship between the density of lead mines and lead level distribution in groundwater. The northern counties of Clay, Caldwell, Carroll, Charlton, Jackson, Lafayette, Macon, Randolph and Ray (circled black in Figure 313) have a comparatively high lead level in groundwater. Between these counties, there are thirty-seven lead mines. The southeastern counties of Carter, Dent, Howell, Reynolds, Shannon and Texas (circled with black dashes in Figure 313) have a relatively high level of lead in groundwater. There are thirty-five lead mines within the mentioned counties. Based on the kernel density output, in Figure 212, there is no relationship between the location of lead mines and the distribution of lead in groundwater. This is evident as regions that do not have a large presence of lead mines, circled in black in Figure 313, have higher levels of lead in groundwater. Hence, from this standpoint, there is no relationship between the geographic positions of lead mines and lead levels in groundwater.

**Global (Traditional) Regression Analysis Results**

In addition to the 2007 communication arts MAP scores, the regression variables used are percentage of students receiving free lunch (LuPer), college degree holders (Edu_Bach), married coupled families with children under 18 years of age (MCFC18), years of teaching experience (TeExp), teachers certification percentage (TeCert), teachers master’s degree percentage (TeMs), teacher-student ratio (TeStuR), teacher salary (TeSal), per student expenditure (PerStu_Exp), level of lead in groundwater (Pb_ZnMean) and MAP communication arts scores of fourth grade public school students for 2007 (COM_G4).
A pairwise Pearson’s regression analysis is conducted to explore the relationships between regression variables to determine the global statistical impact of lead in groundwater to elementary students’ MAP scores. The output of the regression analysis, as seen in Table 1, shows negative relationships between six variables (COM_G4, TeMs, TeSal, TeStuR, LuPer, and Edu_Bach). The significant negative relationship between lead in groundwater (Pb_ZnMean) and the MAP score is of importance. It is an indication there is an inverse relationship between lead in groundwater and the MAP score. It implies that on a global scale, as the levels of lead in groundwater increases, the MAP score for the subject in the given year would decrease. At this juncture, the MAP scores for 2007 communication arts (COM_G4_07) had a strong significant negative relationship with lead in groundwater comparatively. Thus, the MAP scores for grade 4 public school students’ communication arts is a good fit for the use of this study.

In order to explore the relationships between regression variables, pairwise Pearson correlation test was conducted in Statistical Package for the Social Sciences (SPSS). This test also checks for the possibility of multicollinearity between variables. The degrees of correlation between TeSal and TeMs; as well as TeSal and TeStuR are relatively high with a coefficient of 0.700 and 0.628 respectively (Table 1); however, their bivariate pairwise scatterplots do not show a strong linear relationship (Figure 4 14 and 515). In real-world relationships, variables can be related in curvi-linear ways (Rogerson, 2006). While there can a chance for multicollinearity problem, all the variables are used for regression analysis are chosen for theoretical reasons. As such, none of the correlated variables are removed. It is important to note that Pb_ZnMean has negative relationships with COM_G4, TeMs, TeSal, TeStuR, LuPer and Edu_Bach. More
importantly, it implies that as Pb_ZnMean value increases, the communication MAP scores for students in grade four decreases.
Table 1. Pearson Correlation Coefficients for All Paired Variables

<table>
<thead>
<tr>
<th></th>
<th>COM G4</th>
<th>TeMs</th>
<th>TeSal</th>
<th>TeExp</th>
<th>TeStuR</th>
<th>TeCert</th>
<th>LuPer</th>
<th>PerStu_Exp</th>
<th>MCFC18</th>
<th>Edu_Bach</th>
<th>Pb_ZnMean</th>
</tr>
</thead>
<tbody>
<tr>
<td>COM G4</td>
<td>1</td>
<td>0.093</td>
<td>0.102</td>
<td>0.043</td>
<td>0.013</td>
<td>-0.059</td>
<td>-0.321</td>
<td>0.011</td>
<td>0.078</td>
<td>0.266</td>
<td>-0.051</td>
</tr>
<tr>
<td>TeMs</td>
<td>0.093</td>
<td>1</td>
<td>0.700</td>
<td>0.401</td>
<td>0.222</td>
<td>0.202</td>
<td>-0.265</td>
<td>-0.034</td>
<td>-0.089</td>
<td>0.431</td>
<td>-0.035</td>
</tr>
<tr>
<td>TeSal</td>
<td>0.102</td>
<td>0.700</td>
<td>1</td>
<td>0.401</td>
<td>0.628</td>
<td>0.348</td>
<td>-0.235</td>
<td>-0.064</td>
<td>-0.144</td>
<td>0.513</td>
<td>-0.034</td>
</tr>
<tr>
<td>TeExp</td>
<td>0.043</td>
<td>0.401</td>
<td>0.401</td>
<td>1</td>
<td>0.238</td>
<td>0.360</td>
<td>-0.026</td>
<td>-0.098</td>
<td>-0.144</td>
<td>0.091</td>
<td>0.024</td>
</tr>
<tr>
<td>TeStuR</td>
<td>0.013</td>
<td>0.422</td>
<td>0.628</td>
<td>0.238</td>
<td>1</td>
<td>0.240</td>
<td>-0.152</td>
<td>-0.576</td>
<td>-0.006</td>
<td>0.085</td>
<td>-0.152</td>
</tr>
<tr>
<td>TeCert</td>
<td>-0.059</td>
<td>0.202</td>
<td>0.348</td>
<td>0.360</td>
<td>0.240</td>
<td>1</td>
<td>0.114</td>
<td>-0.045</td>
<td>-0.004</td>
<td>-0.006</td>
<td>0.062</td>
</tr>
<tr>
<td>LuPer</td>
<td>-0.321</td>
<td>-0.265</td>
<td>-0.235</td>
<td>-0.026</td>
<td>-0.152</td>
<td>0.114</td>
<td>1</td>
<td>0.057</td>
<td>-0.410</td>
<td>-0.523</td>
<td>-0.009</td>
</tr>
<tr>
<td>PerStu_Exp</td>
<td>0.011</td>
<td>-0.034</td>
<td>-0.064</td>
<td>-0.098</td>
<td>-0.576</td>
<td>0.045</td>
<td>0.057</td>
<td>1</td>
<td>-0.122</td>
<td>0.233</td>
<td>0.117</td>
</tr>
<tr>
<td>MCFC18</td>
<td>0.078</td>
<td>-0.089</td>
<td>-0.144</td>
<td>-0.006</td>
<td>-0.004</td>
<td>-0.410</td>
<td>-0.122</td>
<td>1</td>
<td>0.069</td>
<td>0.033</td>
<td>0.033</td>
</tr>
<tr>
<td>Edu_Bach</td>
<td>0.266</td>
<td>0.431</td>
<td>0.513</td>
<td>0.091</td>
<td>0.085</td>
<td>-0.006</td>
<td>-0.523</td>
<td>0.233</td>
<td>0.069</td>
<td>1</td>
<td>-0.010</td>
</tr>
<tr>
<td>Pb_ZnMean</td>
<td>-0.051</td>
<td>-0.035</td>
<td>-0.034</td>
<td>0.024</td>
<td>-0.152</td>
<td>0.062</td>
<td>-0.009</td>
<td>0.117</td>
<td>0.033</td>
<td>-0.010</td>
<td>1</td>
</tr>
</tbody>
</table>

TeMs = teachers master’s degree percentage; TeSal = teacher salary; TeCert = teachers certification percentage; TeExp = years of teaching experience; TeStuR = Teacher-student ratio; LuPer = Percentage of students receiving free lunch; PerStu_Exp = Per student expenditure; MCFC18 = Married coupled families with children under 18 years; Edu_Bach = college degree holders; Pb_ZnMean = Zonal mean of lead in groundwater.
Figure 15. Scatterplot of TeSal versus TeMs

Figure 16. Scatterplot of TeSal versus TeStuR
Next, to test the functional relationship between variables, Ordinary Least Squares (OLS) regression is employed to estimate the determinants of MAP scores through SPSS. This procedure uses the combination of the variables to predict an outcome. The output of the OLS regression analysis is seen in Table 2. Backward elimination (BE) method is used to remove variables that are insignificant within the model, one variable at a time, until all the remaining variables in the model do not meet the removal criteria.

Table 2. Regression Coefficients Statistics from OLS (Final Model)

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Standardized coefficient</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>762.683</td>
<td>5.671</td>
<td>134.478</td>
<td>.000</td>
</tr>
<tr>
<td>LuPer_07</td>
<td>-.413</td>
<td>.080</td>
<td>-5.195</td>
<td>.000</td>
</tr>
<tr>
<td>Edu_Bach</td>
<td>.769</td>
<td>.278</td>
<td>2.763</td>
<td>.006</td>
</tr>
</tbody>
</table>

Dependent variable: COM_G4= 2007 grade four communication arts MAP scores; LuPer = Percentage of students receiving free lunch; Edu_Bach = college degree holders.

Overall, most school districts have well certified teachers (Figure 11c). In addition, the percentage of teachers with certification is relatively high (greater than 90 percent). As such, in the OLS process, TeCert was one of the first variables to be eliminated in the initial OLS models.

The final OLS regression model retains student characteristic variables of LuPer, and Edu_Bach, of which, a statistically negative relationship exists between LuPer and COM_G4. This indicates that as the percentage of student receiving free lunch increases, the MAP score for COM_G4 decreases. On the other hand, Edu_Bach has a statistically
positive relationship with COM_G4. It indicates that school districts with more college
degree holders tend to have relatively higher MAP test scores.

The OLS final model did not include Pb_ZnMean as a significant variable. However, Pb_ZnMean was the last variable removed before the final model. It had a negative coefficient of -0.056 (Table 3). While OLS calculates a single equation to represent the variable relationship for all school districts in Missouri, a global model may be inappropriate when variable relationship is not consistent across the study area (Qiu and Wu, 2010). In this case, a global regression model is more of an average of the mix of complex relationship between lead level in groundwater and students’ MAP scores across the state of Missouri, instead of a representative of the relationship. As such, Pb_ZnMean along with LuPer and Edu_Bach are used to conduct geographically weighted regression analysis.

Table 3. Regression Coefficients Statistics from OLS

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Standardized coefficient</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>764.297</td>
<td>5.817</td>
<td>131.388</td>
<td>.000</td>
</tr>
<tr>
<td>LuPer_07</td>
<td>-.415</td>
<td>.079</td>
<td>-5.218</td>
<td>.000</td>
</tr>
<tr>
<td>Edu_Bach</td>
<td>.763</td>
<td>.278</td>
<td>2.742</td>
<td>.006</td>
</tr>
<tr>
<td>Pb_ZnMean</td>
<td>-.056</td>
<td>.045</td>
<td>-1.235</td>
<td>.217</td>
</tr>
</tbody>
</table>

Dependent variable: COM_G4 = 2007 grade four communication arts MAP scores; LuPer = Percentage of students receiving free lunch; Edu_Bach = college degree holders; Pb_ZnMean = Zonal mean of lead in groundwater.
**Geographically Weighted Regression Analysis Results**

Local statistics, such as Geographically Weighted Regression (GWR) allows for any descriptive statistic associated with a spatial data set whose values varies from place to place (O’ Sullivan and Unwin, 2010). In addition, local statistics are multi-valued, whereby different values of the statistic can occur in different locations within the study region. Each local statistic is a measure of the relationship being examined in the vicinity of a location within the study region. As such, when the location changes, its local statistics takes on different values. The goodness fit of the GWR model determines and validates the usage of local statistics. This can be observed by examining the model’s spatial patterns of high and low residuals (underestimation and overestimation respectively). When the residuals are relatively small and evenly distributed across the entire study region, it indicates the validity and goodness of fit of the GWR model.

GWR is employed to demonstrate the geographical variations between students’ MAP scores and groundwater lead concentration at the local scale. The predictor variables significant in the global regression model (Table 3) are used in GWR analysis. In this study, three different GWR neighborhood distance parameters (or termed bandwidth) are applied to determine the strength of the relationships between lead in groundwater and MAP scores in different areas. Results from GWR analysis is mapped at bandwidths of 30, 60 and 100. These three bandwidths range from the reasonable number of 30 observations (10 times the number of predictor variables), 60 and 100 (as it was an optimal bandwidth in Qiu and Wu’s study).

The local $R$-squared ($R^2$) of bandwidths 30, 60 and 100 are mapped to show their geographic relevance (Figure 1616). $R^2$ is the squared correlation between the observed and expect values of the dependent variable. As the bandwidth decreases, the
relationships of the local variables become more prominent. This is evident from Figure 16. The local R-squared with the neighborhood of 30 shows a comparatively even distribution. As the bandwidth increases to a neighborhood of 60, relationships between variables are more defined. Lastly, at the neighborhood bandwidth of 100, school districts in the mid-section of Missouri become more prominent. This indicates that the relationship between the three variables is more obvious comparatively.
The $F$-test in one-way analysis of variance is used to assess whether the expected values of a quantitative variable within several pre-defined groups differ from each other.
other. Local F-value is mapped to show their geographic relevance (Figure 717). In the local models, regions with high $R$-squares correspond with areas that have high $F$-values. This indicates that the local statistics of the local models are positively correlated. In addition, it implies that the prediction for the regions are relatively reliable.

Local standardized residuals are mapped to show the spatial patterns of underestimation (high residuals) and overestimation (low residuals) for local models (Figure 1818). When the residuals are relatively small and randomly distributed across the entire study region, it is an indicator of the goodness of fit of the GWR model. Comparing different GWR bandwidths, it can be observed that the residuals are relatively small and evenly distributed across the entire study area, though more outliers of high local standardized residuals can be found at the bandwidth of 100. This indicates the validity of using the three variables of Pb_ZnMean, LuPer and Edu_Bach as input data in GWR.
Figure 18. Local $F$-value results from GWR (bandwidths of 30, 60 and 100).
Figure 19. Local Standardized Residuals (bandwidths of 30, 60 and 100).
Discussion

In this study, the relationship between the locations of lead mines and the varied levels of lead in groundwater was examined. The results show that there is no distance relationship between lead mines and the level of lead in groundwater. Lead is most likely transported in groundwater by mobile particulate matter (McDowell-Boyer et al., 1986; Wells et al., 1989). Groundwater and surface water interact in a variety of physiographic and climatic landscapes (Sophocleous, 2002). Lead concentration in groundwater could originate from several anthropogenic and natural sources. For instance, atmospheric deposition and the casing materials used in constructing wells.

There are eight groundwater provinces in Missouri. In each province, lead in groundwater is going to behave differently in relation to variations in localized hydraulic gradients (Missouri Department of Natural Resources, 2016). The Ozark aquifer, found in the Salem Plateau groundwater province, is mostly unconfined. Within the Salem Plateau groundwater province, is the Old Lead Belt and the Viburnum trend where lead mines are prevalent (Seeger, 2008). Many of Missouri’s active mines are found within this region. Due to the fact that aquifers in the Salem Plateau groundwater province are unconfined, and the prevalence of active lead mines, the lead in groundwater is comparatively higher. In the Springfield Plateau groundwater province, the Ozark aquifer is confined through most of the province. This aquifer is mainly recharged by precipitation. Some of the recharge is the gradual downward infiltration of water from precipitation through the soil materials, into the shallow bedrock, until it reaches the water table (Brookshire, 2012).

Much of the Tri-State lead mines are located in the Springfield Plateau groundwater province. This groundwater province, like the Salem Plateau groundwater
province, has a higher lead level in groundwater comparatively. While there are many possible reasons for the variations in lead levels in groundwater, it is important to note that kernel density estimation conducted in this study examined the relationship between the proximity of lead mines and presence of lead in groundwater. It does not attempt to show the causation of lead in groundwater.

As the study area is statewide, larger samples will allow for a more accurate estimate of lead in groundwater zonal means per school district. Ideally, having a stratified sampling of lead in groundwater would be the best sampling scheme. However, in reality, this is not feasible due to limitation in manpower and financial resources. Hence, using the given USGS groundwater data seems best for this study.

Of the 8371 lead mines in Missouri, there were 578 lead mines that had data of the estimated weight of lead ore extracted. The availability of the volume of lead ore extracted for all lead mines would be helpful to conduct a regionalized study the using the weightage of extracted lead ore for kernel density estimation. However, as the scale of this study is for the entire state of Missouri, a regional relationship between the amount of lead ore extracted and lead groundwater concentration will not suffice.

Lead in groundwater showed a globally insignificant negative effect on grade four communication arts MAP scores for the 522 school districts. However, the level of lead in groundwater is not distributed evenly throughout Missouri. As such, the intensity of the relationship between lead in groundwater and MAP communication arts scores is not homogeneous in the state of Missouri. In addition, lead level in groundwater should be regarded as an anomaly, and its effect on elementary school students’ MAP score is more likely to appear significant in local areas. The local coefficient $t$-values of $Pb\_ZnMean$
allows for the examination of how Pb_ZnMean influences MAP scores spatially on the local regression scale (Figure 1919). In the local regression, local influences become more prominent as bandwidth decreases (Qiu and Wu; 2010).
Figure 20. Local coefficient t-values of Lead in Groundwater from GWR (bandwidths of 30, 60 and 100).
In the local regression analysis, based on a neighborhood of 30 observations, the influence of lead in groundwater varies over the state of Missouri. 13 school districts in the upper north-east of Missouri (shaded in dark blue with $t$-value $\leq -1.68$, Figure 19, $BW = 30$) have a negative relationship between groundwater lead level and communication arts MAP scores. The localized statistics of these 13 school districts showed a significantly negative effect of lead in groundwater on MAP score. Based on Figure 9, the lead in groundwater in the upper northeast region is between average and high. The MAP communication arts scores for the upper north-east region range between low, relatively high and very high (shaded in orange and deep red respectively, Figure 010). The inverse relationship between lead in groundwater and communication arts MAP scores infers that as levels of lead in groundwater decreases, the MAP score for communication arts will increase. While correlation does not imply causation, local regression statistics show the negative effect of lead in groundwater on the MAP test scores of 13 school districts.

The local regression of 30-school district observation showed 17 school districts in central Missouri and 17 school districts in Southeastern Missouri (shaded in dark red $t$-value $\geq 1.68$, Figure 1919, $BW = 30$) with a positively relationship between lead in groundwater and communication arts MAP score. These school districts are in regions with relatively low levels of lead in groundwater (blue and dark blue in Figure 13). In addition, they are school districts that have performed relatively well (many have scores above 747, Figure 010). This may infer that within a very low range of lead levels in groundwater, school districts can perform comparatively well.
At the local regression neighborhood of 60 observations, there are four clusters with 13 school districts that show a negative relationship between lead level and communication arts MAP scores (shaded in dark blue with t-value <= -1.68 in Figure 19, BW = 60). These clusters are located in the upper northeast, central and southeast of Missouri. The cluster of school districts outlined in green (Figure 19, BW = 60) correspond to very low levels of lead in groundwater (Figure 13). These school districts performed relatively well communication arts MAP scores (shaded orange MAP score 767 to 796, Figure 010). On the flipside, 20 school districts show a positive relationship between the lead in groundwater and communication arts MAP score (shaded in light red with t-score >= 1.68, Figure 19, BW = 60). The 20 school districts are found in a region that has significantly low levels of lead in groundwater (Figure 13). Due to the relatively low levels of lead in groundwater, the effect of lead’s influence that can be tested here is weak. These results show a pattern whereby school districts with negative relationships have higher communication arts MAP scores when the level of lead in groundwater is low, and vice versa.

The local regression analysis based on 100-school district observations finds a large portion of school districts exhibit no significant influence of lead in groundwater on MAP scores (shaded light blue and light pink with t-scores of -1.67 to 0 and 0.01 to 1.67 respectively, Figure 19, BW=100). These school districts correspond to regions with low levels of lead in groundwater (Figure 13). There are nine school districts, located in the upper northwest of Missouri, showing a significant negative relationship between lead level and communication arts MAP scores (shaded in dark blue with t-value <= -1.68 in Figure 19, BW = 100). Their inverse relationship with the level of lead in groundwater is
evident as these school districts performed relatively well in their communication arts MAP scores (Figure 010) and have relatively low levels of lead in groundwater. This ascertains that in significant relationships found in local regression statistics, the lower the level of lead in groundwater, the better communication arts MAP scores are.
CHAPTER 5: LIMITATIONS OF THE STUDY AND CONCLUSION

Lead in environment and its effects are important to policy makers, educators, health officials and environmental organizations. Environmental lead exposure has been linked with an increased risk for numerous conditions and diseases that are prevalent in industrialized society. The impact of low-level environmental lead exposure on the health of the public is substantial. Missouri has the most lead deposits in the world. The results from this study show that there is no relationship between the distance of lead mines and lead groundwater concentration. This further shows the complex nature of lead and its means of transportation to groundwater sources.

Limitations

This is a pilot study undertaken to examine the effects of lead in groundwater on students’ academic performance. Higher levels of lead in groundwater are consistently associated to lower MAP communication scores. In this study, the lower MAP scores here might have a closer relationship to high levels of lead level in groundwater. Due to the processes involved in producing the localized mean for lead concentration in groundwater, the smoothing filter applied may dilute or dull the effects of lead in groundwater on elementary school students’ communication arts MAP scores. Although it is not clear if poor performing school districts are directly due to the effects of lead in groundwater, this study helps to explore the variations in academic performance in Missouri school districts with regard to lead concentration in environment.
This study sought to determine the geographic variability between lead in groundwater and elementary school students’ performance. One of the limitations for this study was the use of the administrative state boundaries for Missouri. This placed a limitation for spatial analysis on lead in groundwater within the geographic boundaries of the Missouri Stateline. The effects of lead in its natural environment are not isolate within state administrative boundaries. However, due to the absence of a nationwide standardized test for elementary students in the U.S, comparative studies for academic performance across different states makes it problematic for this study. Hence, this study only used the standardized test scores for Missouri.

Lead can be found in soil, surface water and groundwater. The availability of the data along with its use for human consumption led to the decision to use groundwater data in this study. Future work can be conducted on smaller-scaled analyses for interested local areas using lead in surface water as well as soil. In addition, it would be beneficial to include neighboring area outside Missouri to have a more comprehensive analysis of lead’s effect on students’ academic performance.

**Conclusion**

In conclusion, the literature that surrounds the relationship between academic performances and environmental factors (e.g. lead) using GIS analysis and spatial statistics are in lack in the past. This study investigated geographic variations of lead concentration in groundwater on elementary school students’ MAP scores across 522 school districts in the state of Missouri. This study will provide a better understanding of how environmental factors can contribute to elementary students’ academic performance.
It may heighten the awareness of educators, policy makers and city planners about lead contaminants in Missouri.
REFERENCES


Canfield, Richard L., Charles R. Henderson Jr, Deborah A. Cory-Slechta, Christopher
Cox, Todd A. Jusko, and Bruce P. Lanphear. "Intellectual impairment in children with blood lead concentrations below 10 \text{g per deciliter.}" New England journal of medicine 348, no. 16 (2003): 1517-1526.


Cowie, Christine, Deborah Black, and Ian Fraser. "Blood lead levels in preschool


Kaufman, Alan S. “How dangerous are low (not moderate or high) doses of lead for


Lin, Ja-Liang, Dan-Tzu Lin-Tan, Kuang-Hung Hsu, and Chun-Chen Yu. "Environmental


Schnaas, Lourdes, Stephen J. Rothenberg, Maria-Fernanda Flores, Sandra Martinez,


