Estimating At-risk Population for Lead Service Lines Induced Lead Exposure and Their Correlation to Socioeconomically Disadvantaged Neighborhoods in Milwaukee, Wisconsin

by

Ariana Sul-ki Kim

A Thesis Presented to the
FACULTY OF THE USC DORNSIFE COLLEGE OF LETTERS, ARTS AND SCIENCES
University of Southern California
In Partial Fulfillment of the
Requirements for the Degree
MASTER OF SCIENCE
(GEOGRAPHIC INFORMATION SCIENCE AND TECHNOLOGY)

December 2021

To my Mom, Sam-chun, and my wife Blandine

Acknowledgements

I would like to thank my advisor, Dr. Leilei Duan, for her unending patience and guidance throughout this whole journey. Thank you for advocating on my behalf and believing in me when I struggled to believe in myself. Thank you to my committee members Dr. Robert Vos and Dr. An-Min Wu for their insightful feedback and being willing to meet at strange times due to wide geographical separation. A special thank you to Dr. Vanessa Osborne for helping me order the chaos of my thoughts into coherent text.

I am deeply grateful to all my professors and peers in the Spatial Sciences department for keeping me curious and open to exploring all the capabilities in the realm of GIS. I would like to acknowledge Kendrick Watson and Maureen Scott for their check-ins and providing me all the tools to succeed.

Huge shout out and thanks to Thomas Welcenbach at Lead Free MKE for being willing to speak with me about the Get the Lead Out cause in Milwaukee and for providing information and data resources. I hope my work will help in the fight to ensure clean drinking water for all.

Finally, I want to thank my family and friends for their encouragement and understanding through this hectic time. Thank you for being my cheerleaders and emotional support throughout this process. Words cannot express the depth of my gratitude. I love you all.

Table of Contents

List of Tables

List of Figures

Abbreviations

ACS	American Community Survey
ARA	Adjusted Residential Area
BA	Building Area
BLL(s)	Blood Lead Level(s)
CEDS	Cadastral-based Expert Dasymetric System
CDC	Centers for Disease Control
EJ	Environmental Justice
EJSEAT	Environmental Justice Strategic Enforcement Assessment Tool
EPA	Environmental Protection Agency
GIS	Geographic information system
GISci	Geographic information science
GWR	Geographically Weighted Regression
LCR	Lead and Copper Rule
LSL(s)	Lead Service Line(s)
MAUP	Modifiable areal unit problem
MHD	Milwaukee Health Department
MKE	Milwaukee
MOE	Margin of Error
MWW	Milwaukee Water Works
OLS	Ordinary Least Squares
PCA	Principal Component Analysis
PPB	Parts per billion

RA	Residential Area
RU	Residential Unit
SAR	Simultaneous Autoregressive (model)
SES	Socioeconomic status
SSI	Spatial Sciences Institute
USC	University of Southern California
WHO	World Health Organization

Abstract

The dangers of lead poisoning have posed a real threat to the population of the United States since the turn of the century. It has a cumulative effect in the human body and can therefore build up over time, even with low dose exposure. Children are especially susceptible to lead exposure due to their increased absorption rate of the metal and the lasting health issues that can persist throughout their lives. Minority communities with low socioeconomic status are especially vulnerable to exposure because they are more likely to live in close proximity to lead pollution sources, older homes, and have lower rates of toxicity screenings. Poisoning occurs primarily when lead is ingested through lead-based paint, lead contaminated water pipes, dust, and soil. Older cities across the United States are particularly prone to have populations with increased blood lead levels because lead was a common building material in the early 1900s. Milwaukee, Wisconsin is one such historical city where around 40% of the city's active residential water service lines are constructed of lead. This study quantifies how many people are at risk for lead poisoning based on the existence of lead service lines in their buildings by census tract. Given the deeply segregated history of Milwaukee, an issue that still plagues the city to this day, this study also examines the relationship between the number of at-risk people per census tract and a variety of socioeconomic indicators. Dasymetric mapping techniques as well as regression analysis were used to shed light on this environmental justice issue in Milwaukee. Results show that the number of at-risk people in a census tract has a positive linear relationship with the race, education level, and poverty status of neighborhoods. In the context of Milwaukee's demographics, the issue of lead exposure due to LSL disproportionately affects poorer communities of color.

Chapter 1 Introduction

The fight for environmental justice and equity is an ongoing battle across the nation. The Environmental Protection Agency of the United States defines environmental justice as the effort to provide environmental equity for all through implementation, enforcement, laws, regulations, and policies (US EPA 2014). Too often those in minority social groups – i.e., those of low socioeconomical status, racial minorities, or the unemployed – experience higher levels of environmental toxicity (Maantay 2002). The causes of these inequities, also known as environmental burdens, can include poor air quality, proximity to polluting industry, old lead water mains, or lack of funding to expand green spaces for greater access (Emer et al. 2020). The city of Milwaukee, Wisconsin is one of the many major urban centers where these environmental inequities are starkly apparent. The city has a long history of environmental justice issues such as increased asthma rates, childhood lead poisoning, and lack of access to green space within poor black and brown communities (Small 2019; Collins 2011).

In 2020, issues of environmental inequity were exacerbated with the advent of Covid-19 and the realization that mostly poor black and brown communities were disproportionately affected by the virus. The CDC has listed discrimination, access to healthcare, occupation, education/wealth/income gaps, and housing as social indicators for the increase in risk these minority groups face when infected by the virus (CDC 2020). The low-income population of Milwaukee often relies heavily on state run health programs such as BadgerCare and the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) for medical screenings, treatments, and education. Due to the current pandemic, many of these health centers have limited office visit availabilities. Not only have these closure contributed to the already poor access to health care facilities, but it also prevented many minority communities from receiving vital information about the dangers of Covid-19 and how to prevent the spread of infection.

A further consequence of lack of access to medical centers is a 34% decrease in lead screenings in the city (Dang et al. 2021). Given the worsening state of this health crisis, and Milwaukee's recent initiative to replace all lead service lines (LSL) (Jannene 2020), the examination of where these remaining lead pipes exist in the city, how many people are at risk for lead poisoning from these pipes, and their spatial correlation with neighborhoods of various socioeconomic statuses would ideally prove useful information to both city and health officials as well as environmental justice advocates. Therefore, the aims of this project will be two-fold. The first goal is to determine which census tracts in Milwaukee have the greatest number of people directly exposed to LSL, and therefore have higher risk for lead poisoning, by creating a dasymetric map of the city's population distribution and relating the population density with the corresponding number of lead service lines that need to be replaced. The second goal is to characterize regions of the city with high levels of lead exposure risk with racial and socioeconomic indicators by creating bivariate correlation matrices.

The targeted audience of this study would be policy or decision makers within the community, as well as grassroots organizations fighting for change. It would provide a tool of reference that could help inform them of the neighborhoods that would benefit from intervention to mitigate harmful environmental burdens. This project would also be accessible to the wider public as an educational tool that could potentially inspire them to become involved in community projects whose goals are to combat the impacts of environmental injustice or to contact their local government officials to incite change in harmful policies.

1.1. Study Area: Milwaukee, Wisconsin

The study area of this project is the city of Milwaukee in Wisconsin. Founded in 1846, Milwaukee has a population of roughly 590,000 and is located on the western banks of Lake Michigan (Figure 1) ("U.S. Census Bureau" 2019). The city is known for a plethora of breweries, Harley Davidson motorcycles, and being the home of the 2021 NBA Champions, the Milwaukee Bucks. Unfortunately, the city also has a reputation for being one of the most segregated cities in America. Due to systemically racist policies, such as red-lining, Milwaukee is one of the most racially and socioeconomically divided in America (Lynch and Meier 2020). Studies have been done about the increased exposure to lead toxicity, lower air quality, and inaccessibility to healthy foods in these neighborhoods. Several studies have been conducted highlighting the disparity between affluent neighborhoods and poorer neighborhoods in Milwaukee with respect to access to green space, childhood exposure to lead, and inequities with transport access (Emer et al. 2020; Heynen et al. 2006; Milwaukee Environmental Justice Lab n.d.).

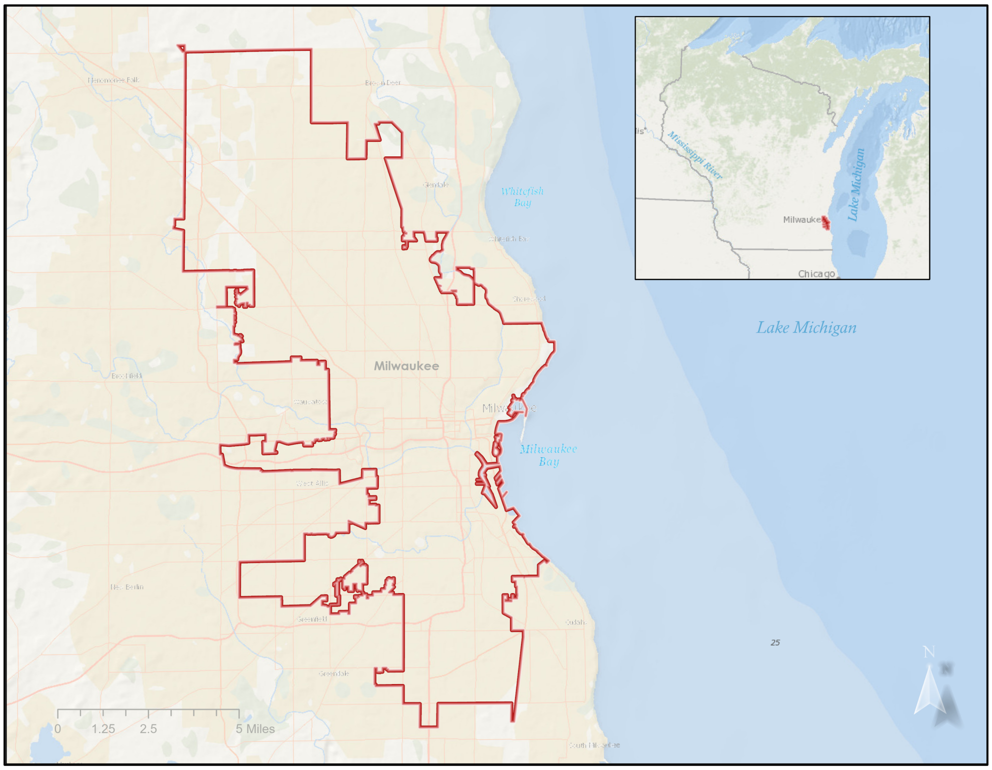

Figure 1: City boundary of Milwaukee, Wisconsin.

1.1.2 Milwaukee and Lead

The issue of lead poisoning in the city of Milwaukee is a long-standing problem that medical professionals, city leaders, and activist groups have been fighting for years. Being an older city, Milwaukee has a good number of homes that were built in the late 1800s to the early 1900s, a time where lead was a common construction material. The pipes that brought water to homes were made of lead, and lead-based paint was the go-to for decorating homes (City of Milwaukee 2016). Despite spreading awareness across the U.S. about the dangers of lead poisoning, especially among small children, Milwaukee continued to have a higher than the national average of children with elevated blood lead levels (BLL) based on data collected by the Milwaukee Health Department's (MPH) Childhood Lead Poisoning Prevention Program (CLPPP) (Public Health Foundation 2020). As a response to this alarming statistic, the Community Lead Outreach Project was started in 1995 by the Sixteenth Street Community

Health Center. Their goal was to decrease childhood lead poisoning by reaching out to the poorest communities in the city that live in the oldest homes. The project was successful in achieving this goal and health studies from the time showed a decline in children testing positive for elevated BLL from 46% to 23% over the years 1996-1999 (Schlenker et al. 2001). As of 2019, 9-10% of children in Milwaukee test positive for elevated BLL.

The Community Lead Outreach Project provided a strong case for the importance of primary intervention when it came to reducing childhood lead poisoning. They found that parental education as well as fostering close relationships with the communities they were working with helped contribute to the overall success of the project (Schlenker et al. 2001). Despite all the progress Milwaukee has made to reduce childhood lead poisoning, the fact remains that exposure to lead is a persistent threat to the overall health of the city's residents. To date, no such study has been conducted attempting to estimate the actual number of people exposed to lead via lead service lines, nor has research examining the correlation between lead exposure and socioeconomic indicator variables within neighborhoods. Considering there are still over 70,000 active residential lead service lines, this study could be useful for the decision makers of the city, lead-free activist organizations, and for the general public as a tool to educate themselves about an issue that impacts their daily lives.

1.2. Environmental Inequity

The causes of environmental exposure disparities are not as simple as residential segregation of minority communities. It is an issue that is deeply engrained into the fabric of society through years of discriminatory legislation and social conditioning. From the era of Jim Crow to the deeply problematic policy of redlining, racial discrimination in laws has lasting consequences today. In cities like Chicago and Milwaukee, neighborhoods can be distinguished based on their racial make-up as a result of redlining in the 1960s. Since black and brown people were denied mortgage loans to buy property is certain parts of the city, they were forced to specific areas where they could afford housing. Often the parts of the city that were available were undesirable for the affluent society because they were near polluting factories. Based on proximity to hazardous byproducts coming from these factories, the minority population that were forced to settle in the neighborhood become exposed to environmental toxins at higher rates (Shrader-Frechette 2002; Hillier 2003; Lester 2018).

From an economic standpoint, environmental inequities – access to urban green spaces for example – can be linked to the capitalistic commodification of these elements. Urban forests have become part of the process of production where those who can afford to "consume" more of the commodity are able to dictate where these spaces are introduced. Therefore, the distribution of the green spaces become uneven and neglected in the parts of a city that are of a lower socioeconomic status (Heynen et al. 2006). Commodification of urban spaces is likely at the root of many cases of environmental injustices minority communities face. The reality is people with low socioeconomic status in society lack the financial and social resources to combat their adverse environmental exposures (Kelly-Reif and Wing 2016). The task of mitigating these hazards is left to local and federal governments who, in the best-case scenario, can commission studies to be carried out to identify the exact issue at hand and the most effective strategy to combat the environmental injustice.

1.2.1 Quantifying and Analyzing Environmental Injustice

When studying communities to identify environmental injustice, there are three components that are regarded. First, there is an exposure assessment for locations within a geographic region. Second, there must be some way to quantify sociodemographic variables across the whole geographic region in question. And third, there is a presence of disease or other detriments to human health in the region (Waller et al. 1997). The methods behind studying the cause and effects of environmental injustice have become more sophisticated over the years. From collecting health data incorporating remote sensing and machine learning, the research on this subject has expanded significantly since its inception amongst the Civil Rights movement in the 1960s. The development of remote sensing technology increased data resources for environmental justice analysis by providing highly detailed views of urban landscapes. These data can be used to analyze multiple environmental burdens such as heat islands, air pollution, and access to green spaces (Weigand et al. 2019). Researchers have utilized principal component analysis (PCA) to create neighborhood deprivation indices based on social data taken from the US Census Bureau. Once these indices are created, regression analysis is used to examine correlations between the spatial phenomena being studies and neighborhoods that have low deprivation index scores. The results of the analysis can then be visualized on a map of various spatial scales and a hotspot analysis can be used to identify areas of potential intervention (Padilla et al. 2014).

A major consideration when assessing a geographic area for exposure to an environmental toxin is the appropriate level of spatial scale. The modifiable areal unit problem (MAUP) can be an issue when choosing the appropriate scale at which to conduct the analysis (Mennis 2003). If the spatial resolution is too big, the nuances of the environmental phenomena being studied can be lost. If the resolution is too fine, particularly in the case of census data, one runs the risk of having data with high levels of error and thus skewing the results of the analysis. Specifically in health studies, it is understood that environmental justice health issues are often spatially autocorrelated. It is therefore critical to understand and consider MAUP and its roll in potentially skewing the results of aggregated data by introducing spatial bias (Swift, Liu, and Uber 2008). Researchers can ensure that their results are as accurate, valid, and transparent as possible by maintaining the integrity of their methods and disclosing limitations in existing data.

1.3. Lead Toxicity

In the late 1800s, lead became the industry standard for the construction of water distribution pipes in the United States due to its ability to corrode at a slower rate than iron, as well as its superior malleability. Despite the slower corrosion rate, lead is still prone to break down over time. These particulates eventually leach into the water supply, and the residents who consume the contaminated water see a rise in their blood lead levels (BLLs) (Brown and Margolis 2012). The dangers of lead service lines (LSL) have been widely known with articles being published about their toxic effects as early as 1859. However, lead was too convenient of a material source to prohibit from its use to be mandated until 1986 when Congress passed the Safe Drinking Water Act Amendments (Rabin 2008). Despite this ban, hundreds of thousands of Americans are still at risk for lead poisoning due to existing lead service lines. In Milwaukee specifically, upwards of 70,000 residential structures still have LSL in use (Lewis et al. 2017).

When Milwaukee's original water services lines were placed in the early 1900s, they were made out of lead. In 2017, Milwaukee launched an initiative to raise awareness about the dangers of lead poisoning and to replace the lead laterals, the pipes that bring water to homes from the water main (Milwaukee Water Works n.d.). Due to the Covid-19 pandemic, the city has fallen far behind in their goal to replace all the city's lead laterals and is now projected to complete the project in 70 years (Jannene 2020). This delay is ultimately the most harmful to those in the lowest socioeconomic classes, primarily racial minority communities, in Milwaukee (Emer et al. 2020).

1.3.1 Blood Lead Levels and Exposure

Lead poisoning has a cumulative effect on the human body. The Centers for Disease Control (CDC) have stated that the maximum threshold of BLL is five micrograms per deciliter (µg/dL), however no amount of lead is safe for the human body, especially in young children (CDC 2021; Miranda et al. 2002). The general symptoms of short-term lead poisoning include abdominal pain, constipation, exhaustion, headaches, irritability, loss of appetite, memory loss, tingling in hands and feet, and feeling weak (CDC 2020). The long-term effects of lead exposure are similar to the short-term effects but also include mood disorders, decreased fertility, and difficulties concentrating. Arguably the more alarming and damaging impacts of lead poisoning occur in young children. Exposure to lead has been linked to developmental delay, learning difficulties, weight loss, hearing loss, and seizures (Mayo Clinic 2019). These symptoms can severely impact a child's quality of life, long past the time of their initial exposure. Research has shown that even low-level exposure can have adverse effects on cognitive development in children (Hou et al. 2013). In other words, children with lead exposure show decreases in their IQ scores and academic achievement when compared to children not exposed (Sorensen et al. 2019).

Children and adults can be contaminated with lead poisoning by means of paint, soil, dust, air, and water (Lynch and Meier 2020). Children who are exposed to lead in the United States usually live in a structure that was built prior to 1940, an era where it was common to use lead as a construction material in both the water pipes as well as in the paint (Chisolm 1971). Despite legislation and efforts to minimize the risk of lead contamination through water consumption, 10-20% of children and 40-60% of infants' lead intake can be traced to potable sources (Rabin 2008). Studies have proven that childhood lead exposure is a spatially correlated issue. These exposures also tend to disproportionately impact communities of color, especially those communities with families living below the poverty line or with low socioeconomic status (Oyana and Margai 2007). Due to lack of resources, communities with a low socioeconomic status are also more likely to be unable to afford the cost of replacing their LSL even if they wanted to (Sampson and Winter 2016).

1.3.2 Lead Service Lines

Lead service lines (LSL) are water pipes that run from the water main to the house (Figure 2). As mentioned earlier, lead was the most common construction material in the late 1800s to early 1900s. Unless replacements have occurred, areas of the city that were built around this time will have LSL that supply the houses with water from the main line. According to Milwaukee's water department, the water that leaves the water treatment facilities contains no lead. However, lead can leach into the water simply transporting water using an old LSL (Milwaukee Water Works 2021). According to Milwaukee Water Works (MWW), the city's water supply has been in compliance since 1996 with the EPA's 1991 Lead and Copper Rule (LCR) – legislation that sought to regulate lead and copper levels in drinking water. To be compliant with the LCR, there must be less than 15 parts per billion (ppb) concentration of lead found in tap water (US EPA 2015). While 15 ppb of lead is an extremely low concentration, it is a consensus that no level of lead is safe to consume. The EPA has recently updated the LCR to include earlier intervention to detect lead in communities' drinking water, push for complete LSL replacements as opposed to partial replacements, require lead level testing at schools and childcare facilities, and to make the locations of existing LSL available to the public (US EPA 2020). The obvious and most logical solution to preventing lead exposure by LSL is to replace the whole service line.

Figure 2: Water service line graphic (Denver Water 2021).

1.4. Thematic Mapping

When studying a spatial phenomenon, researchers often aggregate their data into polygon areal units, such as census tracts. From there, a thematic choropleth map can be generated to determine any spatial patterns in the mapped phenomena. Choropleth maps assume that the data are spread homogenously across the chosen areal unit. When studying population related data, it is known that the data are more heterogeneously dispersed. Further, it can prove challenging to choose the optimal areal unit to aggregate the data to begin with. Government drawn boundaries can be arbitrary and are for administrative or political reasons. They often do not have any relation to any underlying spatial occurrence – i.e., crime rates, public health issues, and land use (Maantay et al. 2007). It is important to note that policies are put in place to protect the privacy of people when data can contain sensitive information. For this reason, it can be difficult to obtain accurate data at a detailed level (Kennedy and Kennedy 2004).

Dasymetric mapping, first developed in the early 1900s as an alternative thematic mapping technique to choropleth maps, involves the division of data into homogenous zones that represent the underlying statistical surface (Eicher and Brewer 2001). Population density within census tracts is a prime example of the potential issues with arbitrary unit partitioning.

Populations within a census tract are assumed to be homogenously distributed across the areal unit. Depending on the size of the areal unit, this can lead to fallacious analytical claims that do not accurately represent the population distribution within the unit (Mennis 2003). Administrative boundaries drawn for bureaucratic purposes can often mask the underlying issue being studied. Therefore, when studying environmental justice issues that emphasize the importance of accurate population distribution, it is critical that a reliably detailed spatial scale is used. Dasymetric mapping techniques can redistribute the population within an aerial unit like a census tract by disaggregating the population data and repopulating it with ancillary data (such as residential parcels). This provides a much finer spatial resolution and more precisely represent the actual population distribution within the study area (Eicher and Brewer 2001; Maantay et al. 2007).

1.5. Racial and Socioeconomic Correlations with Environmental Justice

When traveling through any major city in the United States, it is usually quite apparent which areas of town are wealthier than others. Just by means of observation, one can differentiate between these neighborhoods by their cleanliness, property values, access to healthy and a variety of grocery stores, etc. Drive-by observations aside, researchers have developed ways to quantify the socioeconomic disparities in order to perform meaningful analysis on either the reasons for such stark differences within one city or the impacts of these divides, both direct and indirect, on peoples' lives. Research has determined that certain social indicators, such as race, unemployment, poverty, and education level, tend to aggregate at the neighborhood level (Messer et al. 2006). With this understanding, there can be a multitude of variables that can indicate economic disadvantage. It is a complex, multi-faceted issue where one or many of these social indicators are linked to disparities in public health (Eicher and Brewer 2001). Studying the relationship between socioeconomic variables and the at-risk population for lead exposure in Milwaukee can provide greater context and insight into the underlying issues that are contributing to this environmental inequity.

Correlation analyses are used in environmental justice studies to examine relationships between a dependent variable and explanatory socioeconomic factors (Jerrett et al. 2001; Raddatz and Mennis 2013). According to Tobler's first law of geography, "everything is related to everything else, but near things are more related than distant things" (Tobler 1970, 236). The concept of spatial autocorrelation as outlined by Tobler's famous law can pose problems to spatial statisticians when attempting to run regression or correlation tests because their base assumption lies in the independence of the observations and errors (Chakraborty 2011). Therefore, it is highly important to take spatial autocorrelation and multicollinearity between variables into account when running correlation tests because of the bias they create in the data.

In the following chapter, studies related to the topics of environmental justice, lead poisoning, dasymetric mapping techniques, and regression analyses will be explored. Concepts will be reviewed, methodology outlined, and a case for the justification of this study will be addressed as well.

Chapter 2 Related Work

The damaging effects of ingesting lead are not as recent a discovery as one might think. Texts from the second century BC suggest that Hellenistic physicians were aware of the toxic effects of high lead exposure. It was not until the Industrial Revolution when the chronic effects of lead exposure were made known to the larger medical community (Riva et al. 2012). Within the last 80 years, research about the damages of lead poisoning, as well as about sources of lead exposure has exploded. The literature on the subject has expanded from a medical perspective to encompass environmental points of view. Legislation was introduced in the 20th century to restrict the use of lead-based materials in construction to help lower exposure rates among the population. Overall, the instances of lead poisoning and exposure have significantly decreased since the Industrial Revolution due to enhanced research and awareness (Brown and Margolis 2012; Chisolm 1971; Rabin 2008).

Issues of lead in the water and the links to environmental justice have been extensively researched for the past century. The literature has revealed that lead exposure disproportionately affects minority communities across America (Sampson and Winter 2016). Public attention to this injustice increased following the Flint, MI water crisis in 2016. The outcry of poor black and brown communities, the ones most impacted by lead exposure, was finally gaining national attention (Butler et al. 2016). Other cities, including Milwaukee, began to put renewed efforts into their clean water programs and removing remaining lead service lines (Jannene 2020; Public Health Foundation 2020; Wisconsin DNR 2021). In 2017, Milwaukee's water department announced an initiative to replace all remaining LSL in the city. Unfortunately, circumstances like Covid-19 have drastically slowed their progress toward their goal. The delays mean that

Milwaukee's residents, particularly historically black and brown neighborhoods, will continue to be at-risk of poisoning from exposure to lead in their water supply.

The following chapter explores the related literature researching lead exposure from a historical and environmental justice lens. It also explores the ways in which dasymetric mapping can be used to accurately estimate the number of people at-risk. Finally, examples of regression analysis will be explored to show how these techniques can help shed light onto the underlying patterns of toxicity exposure in relation to socioeconomic factors.

2.1 Environmental Justice: Milwaukee

Milwaukee has been the case study for environmental justice issues for many years. It is a city where the social divisions between neighborhoods are starkly apparent and have root in years of divisive public policy. Mary Collins published a study in the American Journal of Public Health in 2011 called "Risk-Based Targeting: Identifying Disproportionalities in the Sources and Effects of Industrial Pollution." The most notable point in this study is the fact that Collins's study area was the City of Milwaukee. The aim of her study was to prove that industrial pollution from a few key polluters in the region disproportionally affects low-income and minority communities. She used the risk screening environmental indicators model to conduct her methodology. She first assessed the efficacy of the current mode of monitoring environmental burdens in the community, the Environmental Justice Strategic Enforcement Assessment Tool (EJSEAT). Although it contains 18 variables in its assessment process, race is excluded from the method. The study sought to add race as a variable in the EJSEAT and include more specific data to effectively measure environmental justice concerns within Milwaukee. Her methodology for calculating race and socioeconomic class dissimilarity by using percentages will most likely be used in this project, as both will be examined as a social context indicator (Collins 2011).

A 2013 health report for Milwaukee conducted by the Center for Urban Population Health found significant health disparities between the socioeconomic classes within Milwaukee. One of the categories that disproportionately affects the lower socioeconomic class is childhood lead poisoning (Greer et al. 2013). Lead poisoning is a serious issue because it is a cumulative toxicant that can affect multiple systems in the body. Small children are especially susceptible because they absorb four to five times more lead than adults when ingested (WHO 2019).

2.2 Lead and Water

Studies in the United States examining the toxicity of lead in the drinking water supply can be traced back as early as 1845 (Brown and Margolis 2012). Despite the fact that many US cities were voting to move away from using lead for water pipes by the 1920s, the national plumbing codes continued to approve lead as a viable material source until well into the 1980s, around the time Congress passed the Safe Drinking Water Act Amendments (Rabin 2008). With all these efforts to ensure safe, toxin-free water for all, the Environmental Protection Agency still allows up to 15 parts per billion (ppb) of lead in drinking water (US EPA 2016). The EPA, in partnership with the Centers for Disease Control (CDC), recognizes that no amount of lead is safe in the human body, especially in children (US EPA 2016). There is a worldwide consensus on this view (WHO 2019). An added layer to the environmental issue of lead exposure is observing which communities are most at risk. It has been proven time and time again that lead poisoning disproportionately affects minority communities throughout the United States (Brown and Margolis 2012; Lewis et al. 2017; Butler et al. 2016; Rabin 2008; Sampson and Winter 2016).

2.2.1 Flint, Michigan Water Crisis

A landmark example that occurred within the past decade is the Flint, Michigan water crisis. The case in Flint was overwrought with regulation violations and false reporting. The House Committee of Oversight and Government Reform formed a bipartisan conclusion that the city officials had been negligent in their duties to ensure the safety and health of its residents (Chaffetz et al. 2016). It was revealed that after the city switched water sources, high levels of lead began to be reported in the city's drinking water.

When conducting lead and copper rule (LCR) sampling, reports found that the Flint's water treatment plant improperly collected samples from homes that were not at high risk for lead contamination. On top this sampling error, the Michigan Department of Environmental Quality gave faulty instructions to the residents of the town, telling them to pre-flush their taps before the sample collected (Butler et al. 2016). This is against the protocol for LCR sampling and skewed the city's reported lead levels. A major justification for the delayed and poor response of the local government to intervene is the direct result of the economic and racial makeup of the city. The population of Flint is 62.6% people of color and 41.6% of individuals live below the poverty line (Butler et al. 2016).

The BLLs of children under five were compared before and after the change in water source for the city, and significant increases in elevated BLLs were recorded. Overall, the results showed that 2.4% of children had elevated BLLs before the change, and 4.9% were recorded to have elevated levels after the fact (Hanna-Attisha et al. 2015). In areas identified with high water lead levels, the jump went from 4% to 10.6%. The same study compared these results against socioeconomic disadvantage scores and found the areas with elevated BLLS to be statistically positively correlated with areas with high levels of socioeconomic disadvantage (Hanna-Attisha et al. 2015).

2.3 Dasymetric Mapping

When analyzing spatial phenomena utilizing census data, it has become common practice to use dasymetric mapping techniques to gain an accurate visualization of the topic of study. The boundaries used in census data are meant to serve government purposes and are not always ideal when creating a map to study population distribution as an example (Mennis 2003). Traditional choropleth mapping assumes the data are spread evenly across the chosen areal unit. In the case of census boundaries, which are drawn with no consideration to any sort of spatial subject, thematic maps can be misrepresented, and the underlying pattern masked (Maantay et al 2007).

There are several different ways to approach dasymetric mapping depending on the subject matter being mapped and the research question. A common approach for disaggregating population is areal interpolation. Areal interpolation involves the transfer of data from a source dataset to a target dataset of overlapping areal units. It is assumed that population is distributed evenly across the source layer. When estimating the population based on the overlap between the source and target zones, the ratio of the overlap between the two datasets is applied to the population of the source zone, thus yielding the estimated population of the target zone (Maantay et al. 2007). The simplest version is a binary method of areal interpolation that can be used to estimate population density based on land-use data that filters out regions of uninhabitable land (Eicher and Brewer 2001).

There are some shortcomings with areal interpolation methods for population disaggregation. For one, areal interpolation operates under the assumption that all residential areas have homogenously distributed population density (Maantay et al. 2007). Residentially zoned areas in a municipality do not necessarily have even population density throughout, therefore a similar fallacy in the analysis could occur to if one used the census boundaries as the areal unit of disaggregation. The spatial resolution of land-use data may not be as fine-tuned as it needs to be to present a true representation of population distribution over a given area. When seeking to build as true of a representation as possible for population distribution within a city, a more refined approach must be taken that is able to disaggregate population data to the finest possible resolution possible.

2.3.1 The Cadastral-based Expert Dasymetric System

The Cadastral-based Expert Dasymetric System (CEDS) was first introduced by Maantay et al.'s (2007) study comparing its novel methodology against different, more established, dasymetric mapping techniques. Notably, the researchers compared their CEDS disaggregation method against simple and filtered areal weighting disaggregation methods – two of the most common ways to disaggregate population data. While the areal disaggregation methods use remotely sensed land-use data as their ancillary data, the CEDS method uses cadastral data to estimate population distribution. Maantay et al. (2007) developed a method that disaggregates data from a more general census block group level down to the tax lot level. The researcher used residential area (RA) and residential units (RU) as stand-ins for the population distribution in their calculations. The adjusted residential area (ARA) is the total livable area within a building, and it is derived by multiplying the tax lot's building area by the ratio of residential units and total units. The information about the RU and the calculated ARA were aggregated up from the tax lot level to the census block group and census track levels (Figure 3).

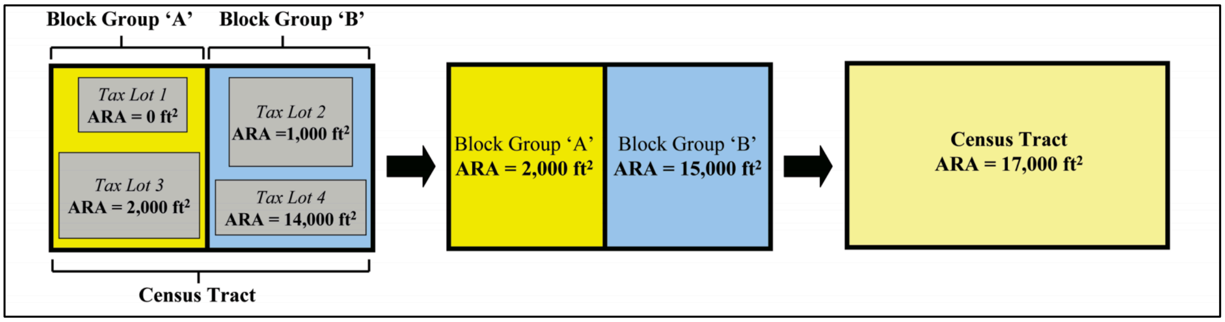

Figure 3: Visual representation of the ARA calculation

From there, the census population data can be multiplied by the proxy population unit ratios (RU or ARA) to derive the final dasymetrically calculated population density. Maantay et al. (2007) could then determine which proxy unit (RU or ARA) was most appropriate for tract population estimation by a tract-by-tract basis. To perform this check, population data was disaggregated back down to the tax lot level, which was then aggregated back up to the block group level. By doing this disaggregation and aggregation for both RU and ARA proxy units, the researchers were able to compare the calculated block group population estimate with the census block group population data. Whichever proxy unit with the smallest difference between the two values was chosen as the preferred proxy unit for the dasymetric calculations for that census tract. The argument that makes the CEDS method superior to the areal interpolation disaggregation methods is that CEDS uses tax-lot information which is a much finer spatial resolution than anything that the areal interpolation methods work with. Thus, the final dasymetric population density result can display the nuances in human population distribution better than areal interpolation methods (Maantay et al. 2007). A slightly modified version of the CEDS method of dasymetric mapping was used in this study when calculating the number of at-risk people for lead exposure due to the presence of LSL.

2.4 Quantifying Socioeconomic Disadvantage

Being able to provide context behind mapping spatial phenomena brings meaning to the analysis of the issue. The ability to answer the question “why?” in reference to the spatial event being visualized pushes the research in the direction of problem solving over simply pointing out a potentially problematic situation. Conducting correlation tests to characterize the regions where an environmental justice issue is occurring can provide valuable evidence to support the plight of the victims of such an injustice. Being able to identify areas within a city with shared attributes that experience environmental inequity can help decision makers and advocacy groups know where to target their outreach programs and solution actions.

The literature on quantifying socioeconomic disadvantage within neighborhoods contains a wide variety of different methodologies. One route relies on inductive and deductive reasoning to select the most appropriate social vulnerability indicators. The concept would be to initially select the vulnerability indicator variables using deductive reasoning by researching with background literature. From there, the variables can be verified and selected for analysis with inductive reasoning – looking at current data and statistical studies that utilize the variables in question (Hinkel 2011). Building on Hinkel’s methodology for indicator identification, Samuel Rufat et al. (2015) compiled worldwide case studies on social vulnerability to floods. The researchers were able to identify seven social indicators based on frequency of appearance in the 125 studied cases. These indicators were: demographic characteristics, socioeconomic status (SES), health, coping capacity, risk perception, neighborhood quality of life, and land tenure (Rufat et al. 2015). Hinkel’s (2011) method for social indicator selection is highly labor intensive and scrupulous research on the subject at hand. It can contextualize the indicators that are selected and seeks to provide a global standard for social vulnerability indicators.

To select the socioeconomic indicators for this study, a similar approach to Hinkel (2011) was taken in that extensive background research was conducted to select out appropriate SES indicator variables from the literature. Given the historical segregation of the city, Milwaukee has no shortage of socioeconomic studies as they relate to health and environmental justice issues. In a 2013 health report, the Center for Urban Population Health conducted a study of health disparities in Milwaukee by socioeconomic status. The researchers derived a SES index by using household median income and education data. Each zip code within the city was assigned a SES index score based on the calculation from the two indicator variables. The results of this report showed major health disparities within the city and upheld previous research conclusions that socioeconomic status is one of the most telling indicators of peoples' health outcomes (Greer et al. 2013). A more recent study conducted by Lynch and Meier (2020) examined the intersection of poverty, home ownership, and race on childhood blood lead levels in Milwaukee. All the SES data in this study were taken from a census tract level while the childhood blood lead levels were measured as a continuous mean across census tracts. The results of Lynch and Meier's (2020) research concluded that socioeconomic and racial minority neighborhoods have higher average childhood BLL. Additionally, high BLL risk greatly increases if the neighborhood's population has multiple risk variables (Lynch and Meier 2020). These studies were just two key examples from literature studying inequities in Milwaukee based on socioeconomic status. Based on the precedent these studies provided, the socioeconomic variables used in this study are median household income, poverty status, race, ownership of the lived-in residence, and education level.

2.5 Environmental Justice Correlation Studies

Correlation research between racial and social factors and environmental justice issues are found in abundance within the literature. Amongst the most studied relationships is that of pollution with communities of color and/or those experiencing poverty (Banzhaf et al. 2019). In 1992, researchers Mohai and Bryant found that racial and class biases are directly related to issues of proximity to environmental hazards. This conclusion was reached after studying sixteen environmental justice case studies and conducting a public perception study in the Detroit area. While their research showed class to be a significant factor in exposure to environmental toxins, race was the more strongly related factor with environmental hazard exposure (Mohai and Bryant 1992). Mohai and Bryant’s 1992 study was just one example of the increasing number of environmental justice correlation studies that have been conducted over the past two decades.

2.5.1 Understanding Ecological Fallacy

Significance of racial and socioeconomic inequalities with hazard exposure vary in the studies, some finding the two to be strongly related, while others finding little correlation (Anderton et al. 1994). A major source of oversight in studies is the failure to consider ecological fallacy. Ecological fallacy occurs when one attempts to draw conclusions about spatial relationships by comparing two sets of data with different aggregation scales. Specifically, error is introduced when assuming variations at a larger scale are the same at a smaller, more individual scale (de Munck 2005). A way to help account for ecological fallacy in correlation analyses is to choose units of analysis as small as possible without compromising the integrity of the data (Banzhaf et al. 2019). The issue remains that when comparing different variables in a correlation calculation, the aggregation methods of the data are not guaranteed to be the same. One could measure estimates on the individual level, while another could represent the data collectively across the chosen spatial unit. Therefore, a good practice is to describe the overall characteristics of the population in the region of study, especially when incorporating multiple social explanatory variables in the study (de Munck 2005).

2.5.2 Hazard Proximity Methods

Another possible reason for the wide variation in statistical uncertainties among environmental justice correlation studies is a lack of standardized methodology for calculating environmental toxin exposure proximities. The classical approach is to define geographical units, such as counties or zip codes, identify those with and without the environmental burden present, then compare the demographic data between the two sets of units. This method of determining hazard exposure is referred to as the spatial coincidence approach and it assumes the population living in the geographic units that contain the environmental toxin are automatically closer to the source than those who live in units that do not contain the hazard. The logic behind the thought of the spatial coincidence methodology is flawed because one, it assumes a spatial unit's population is evenly distributed in the distance to the hazard source and two, it does not account for edge effects among the neighboring units.

Distance-based methods of proximity analysis can account for the issues found in the classical, coincidence approach. With distance-based approaches, the precise locations of the hazards are included, and demographic variables within a set distance of the environmental burden from any geographic unit can be compared against those that occur farther away. In other words, the data are aggregated based on their closeness to the hazard locations instead of by more arbitrary geographic boundaries (Mohai and Saha 2006).

While distance-based methods of examining correlations between social and racial variables and an environmental injustice, it relies heavily on very fine spatial resolutions of data.

The level of detail that would be required to conduct an accurate study with distance-based methods can be difficult to obtain from public access databases. For this reason, this study will attempt to quantify the characteristics of the regions of the city with higher levels of lead exposure risk based on bivariate correlation analysis on demographic indicators obtained from the Census Bureau. The following chapter explains the methodology used to derive more accurate population distributions in the city using a dasymetric mapping technique and the steps of the correlation analysis using the results of the dasymetric map as the dependent variable and demographic data as the explanatory variables.

Chapter 3 Methods

The following chapter contains the details for the methods of this project. It begins with an overview of the project's design followed by a section containing a description of the data used in the analysis. The last two sections go through the creation of the dasymetric map using the Cadastral-Based Expert Dasymetric System methodology and regression analysis using ordinary least squares.

A major consideration with data selection in this study was selecting the most appropriate spatial resolution for the data analysis. Utilizing spatial scale that was at the finest resolution possible without compromising the accuracy of the data was critical to ensure the integrity of the analyses being conducted. Datasets that are found in the Census Bureau's website all contain the variable estimates as well as a calculated margin of error (MOE) to account for variation within the data. The MOE from the American Community Survey data were calculated at a 90% confidence interval (Berkley 2017). The smallest spatial scale that population estimates, as well as the social indicators from the ACS could be obtained were at the census block group level. Upon further inspection of these data, MOE for each block group was far too high to perform reliable statistical analysis. In some cases, the MOE exceeded that of the data estimate itself. The high error found in the census block group data also disqualified it from being used to validate the results of the CEDS map. The next spatial scale up was the census tract data. These tables were much more accurate in terms of a lower MOE for each variable estimate. Therefore, the spatial scale for all tabular data in this study is at the census tract level.

When it came to the explanatory variable selection for the bivariate correlation analysis, an inductive approach was taken. Results from previous environmental justice and health equity studies in Milwaukee were examined and the selection of socioeconomic indicator variables were compiled as results of the research. There are statistical methods that can be used to select socioeconomic variables for a regression analysis, but these are beyond the scope of this study and are addressed in Chapter 5 of this paper.

3.1 Project Design

Given the historical and modern relevance of the damages lead pipes have caused in Milwaukee, as well as the abundance of data related to the subject, the environmental burden variable chosen for this study is the presence of lead water mains within the City of Milwaukee. As mentioned above, it is difficult to conduct an accurate measurement of exposure to a toxin if the population estimates that are used in the analysis are assumed to be distributed homogenously across a census tract. Nuances in the living circumstances of the population can be lost when using such a large spatial unit. Further, it is difficult to tell a precise number of people who are directly at-risk for lead exposure based on just the total number of LSLs still active in the city. This is due to the reality that one LSL does not equal one person exposed. A single LSL can feed water into a house or apartment complex where multiple people reside. If one wants to see how many people are at-risk by census tract, a simple proportion of number of LSL addresses to the tract population will not suffice. By using the census tract as the aerial unit of measure, it assumes that the populous is spread homogenously across the tract. This rationale is an obvious fallacy as humans disperse heterogeneously across a given space. A dasymetric population distribution map is a way to solve the issue of spatial scale by disaggregating the data into smaller units across the study area. This involved a series of joins and field calculations to disaggregate and reaggregate census tract data down to city parcel resolution, then back up to tract data. Once the final at-risk population totals were obtained per census tract, a hotspot analysis was run to find the statistically significant hot and cold spots by census/ tract across the city.

The second step in this project will be to determine the relationship of certain social indicators and the census tracts with the highest number of people exposed to LSLs. The social context indicators most widely researched in environmental justice issues are low-income and racial minority groups (Alexeeff et al. 2012). Based on a compilation of socioeconomic indicators used in other environmental justice studies in Milwaukee, the final list of explanatory variables was as follows: median household income, poverty status within the last twelve months, tenure status of the property (is it being rented or is it owned by its occupants), race, and level of education. An exploratory ordinary least squares regression analysis was then conducted on the dependent variable (the number of people at-risk for lead poisoning in each census tract) with the above listed socioeconomic indicators as the explanatory variables. The variables exhibiting high multicollinearity were thrown out of the model and the remaining explanatory variables were run through a local bivariate relationship analysis to visualize each independent variable's relationship with the dependent variable.

3.2 Data Description and Data Processing

The city lead pipe data was obtained from the water department's website, which keeps up-to-date records of the existing pipes. These addresses were geocoded using ArcGIS Pro's geocoding service and mapped as a dot density layer (Figure 4). Of the 74,225 addresses that were converted to points, two landed outside of the city boundary and were thus excluded from the study. ArcGIS Pro's geocoding service placed all points on the street facing edge of the parcel boundaries. On occasion, the point was not quite inside the parcel boundary. For this reason, a five-foot buffer was applied to the parcels when spatially joining the LSL points to the parcel polygons where if the point fell inside of the five-foot buffer, it was joined to the parcel.

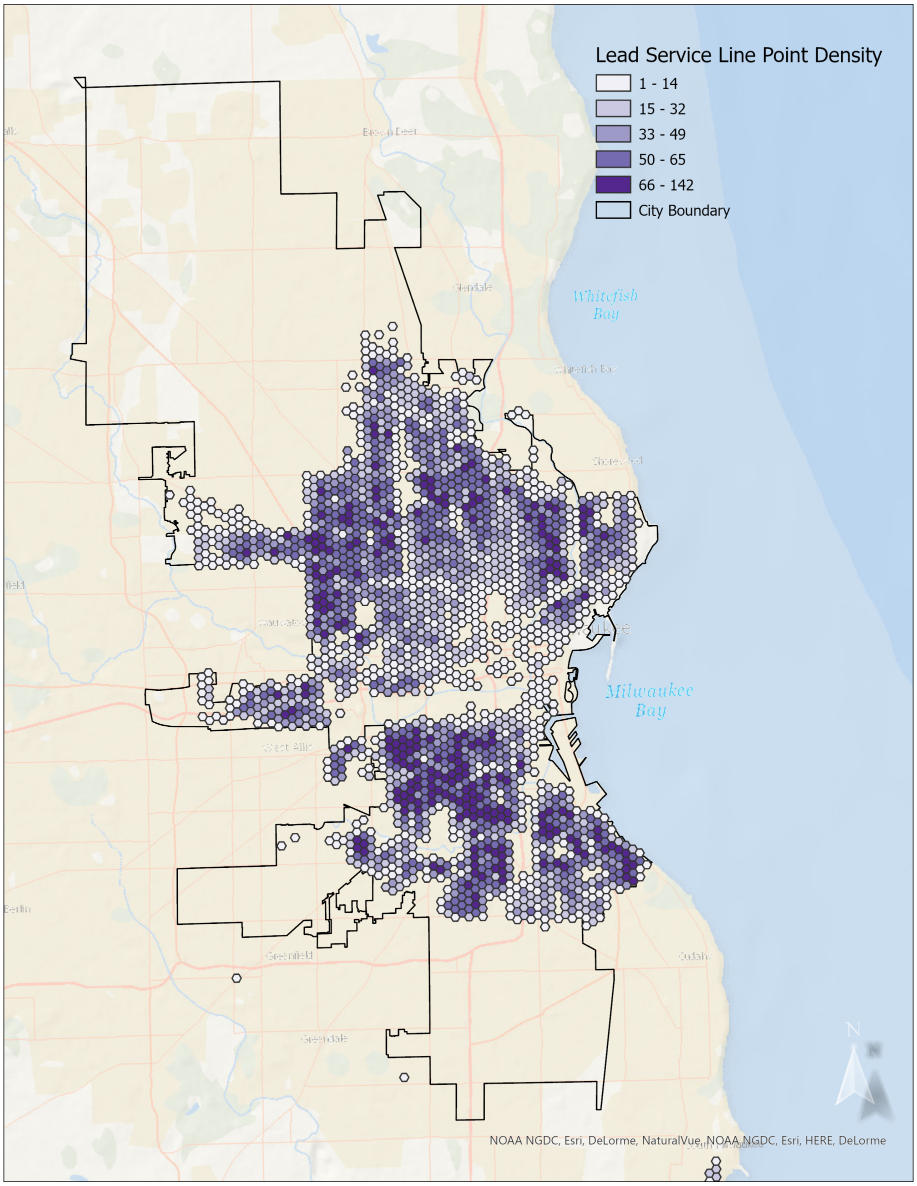

Figure 4: Dot density of residential addresses with lead service lines

The census tract boundaries were obtained from the Milwaukee County's open data portal in the form of a polygon shapefile. The tracts that were within Milwaukee's city boundaries were selected out from the county shapefile. It was found that two of the 211 census tracts contained no residents so those two were excluded from the study, leaving a total of 209 tracts used in this study. City boundary and parcel data were obtained from the city's open data portal. The parcel layer contained data about the parcel, including a field with the total building area of the property. The "Building_Area" field was used for the dasymetric mapping portion of the methods. The "Building_Area" is an attribute found in the parcel layer file and it is defined as "the total usable floor area of the structure in square feet" (City of Milwaukee 1999, 10). The residential parcels were selected out of the parcel layer using the select by attribute tool in ArcGIS Pro.

The census tract population estimates and social indicator data – median household income, poverty status, education level, race, and tenure – were obtained from the Census bureau's website using the American Community Survey (ACS) formatted as comma-separated values (csv) files. The race and education level indicators had data tables that contained several different categories within the files. The race domain was broken up into individual races that people identify as, and those of Hispanic origin were recorded in a separate dataset all together. In an effort to reduce the number of redundant variables within the study, the data for all non-white races and Hispanic origin were combined into a people of color (POC), or non-white, category. So, the race indicator in this study was comprised of two variables, white and non-white. The education level dataset also contained a multitude of categories where the ACS created separate columns of data for each level of education attained – from kindergarten to doctorate degrees. The indicator variable used in this study was the number of people with a high school degree or less. To create this high school or less education category, the data were combined in excel from kindergarten to high school or high school equivalent. All the variable data used in the analysis were extracted from each table and joined to a master csv file with all the social indicators. This master csv file was then combined with the map of the city's census tracts via a table join. Table 1 lists all the data and their sources that were used in this project.

Table 1: Datasets and Sources

Dataset	Description	Format	Data Type	Spatial Scale	Reporting Period	Source
City Boundary	Administrative boundary of the city of Milwaukee	.shp	Vector data (polygon)	City limits	2021	Milwaukee Open Data
Parcel Polygons	Shapefile containing the city's parcel polygons with master property file data	.shp	Vector Data (polygon)	City Parcels	2021	Milwaukee Open Data
Census Tracts	Boundaries of Milwaukee County's census tracts	.shp	Vector Data (polygon)	Census tracts	2020	Milwaukee County Land Information Office
Population Estimate	Dataset reporting estimated population size	.csv	Aggregated census tract population	Census Tract	2019	U.S. Census Bureau ACS data
Education Level over 25	Dataset reporting estimated education level attainment for individuals over the age of 25	.csv	Aggregated census tract population	Census Tract	2019	U.S. Census Bureau ACS data
Median Income	Dataset reporting estimated median income level per household	.csv	Aggregated census tract population	Census Tract	2019	U.S. Census Bureau ACS data
Poverty Level	Dataset reporting estimated household poverty status within the last 12 months	.csv	Aggregated census tract population	Census Tract	2019	U.S. Census Bureau ACS data
Race	Dataset reporting estimated numbers of individuals identifying within racial categories	.csv	Aggregated census tract population	Census Tract	2019	U.S. Census Bureau ACS data
Hispanic/Latino Origin	Dataset reporting estimated individuals that identify as Hispanic or Latino origin	.csv	Aggregated census tract population	Census Tract	2019	U.S. Census Bureau ACS data
Lead Service Lines	Master list of residential addresses within the city with LSL	.csv	Point data with text field addresses	City Parcels	2021	Milwaukee Water Works

3.2.1 American Community Survey Margin of Error

To address the margin of error in the ACS data, a coefficient of variation (CV) was calculated for each variable dataset at a 90% confidence level using methodology outlined by Tufts GIS Center (Parmenter and Lau 2013).The coefficient of variation is a metric used to establish the reliability of the ACS data estimate and was calculated for all 209 census tracts in Milwaukee. First, the standard of error (SE) for each tract estimate was calculated using Equation 1.

$$\mathrm{SE} = \mathrm{MOE} \div 1.645 \tag{1}$$

Once the SE was calculated for each census tract variable's estimate, the CV was derived using Equation 2 (Parmenter and Lau 2013).

$$\mathrm{CV} = (SE \div Estimate) \times 100 \tag{2}$$

In the cases of the number of non-white people and those with or less than a high school education, the data for these two variables were aggregated from smaller categories within each of their respective datasets. The MOEs for the aggregated data variables were calculated using the formula in Equation 3, where the letter 'c' represents each individual data estimate to be included in the aggregation (CCRPC 2015).

$$\mathrm{MOE_{agg}} = \pm\sqrt{\textstyle\sum_c MOE_c^2} \tag{3}$$

The CVs were then categorized into high, medium, and low reliability. The thresholds used for each were taken from ESRI's guidelines where high reliability CV scores are anything less than or equal to 12, medium CV scores are between 12-40, and low reliability scores are anything above 40 (Herries 2021). The totals for each reliability category were then added together and recorded as percenta0ges. Table 2 shows all the explanatory variables taken from the ACS and their respective CV reliability percentages.

Table 2: Coefficient of Variation Reliabilities for ACS Data

CV Categories	Population	Rents	Median Income	Non-White	Owns	White	HS Education	Poverty Status
High	93.78%	73.68%	54.55%	50.24%	45.19%	43.27%	23.92%	1.91%
Medium	6.22%	25.84%	44.02%	49.76%	51.92%	42.79%	74.16%	91.39%
Low	0.00%	0.48%	1.44%	0.00%	2.88%	13.94%	1.91%	6.70%

3.3 Analysis Tasks Details

The methodology for the analysis tasks of this project were broken into two main sections. The first being the CEDS mapping methodology to estimate the number of Milwaukeeans at-risk more accurately for lead poisoning based on the presence of LSL at their place of residence. The second part of the study examined the relationship between the number of people at-risk for lead exposure per census tract with social indicators obtained from the Census Bureau's American Community Survey data. The idea is to quantify the characteristics of census tracts that have a higher number of individuals at-risk for lead exposure. A bivariate correlation matrix was constructed using Microsoft Excel to determine the nature of the relationships between the independent, social indicator variables and the dependent, number of people at-risk variable.

3.3.1 Cadastral-based Expert Dasymetric System

The Cadastral-Based Expert Dasymetric System (CEDS) methodology was used in this study to create a dasymetric population distribution estimate in the City of Milwaukee. This method disaggregates the census tract data and redistributes the population estimates by assuming the number of individuals residing in a parcel is proportionate to the size of the building area within that parcel. It should be noted that in the original methodology put forth by Maantay et al. (2007), the spatial unit for disaggregation was tax lot data. In this study, parcel information is used as a synonym for tax lot because the data obtained from city websites were named “parcels” rather than “tax lot.”

In ArcGIS Pro, the ACS population estimates per census tract layer were joined by table to the county census tract shapefile. Census tracts within the city boundaries were selected out and created into a separate layer. In the city parcels shapefile, the parcels coded “residential” were selected out and spatially joined to the city census tracts layer. In the spatial join, the merge rule for the parcels layer was set to “sum” for the “Building_Area” (BA) field.

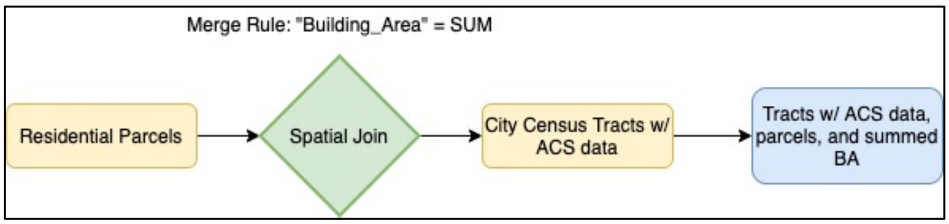

Figure 5: First spatial join between parcels and tracts with BA summed

The city census tracts with the BA Summed field were then spatially joined back down to the parcels layer.

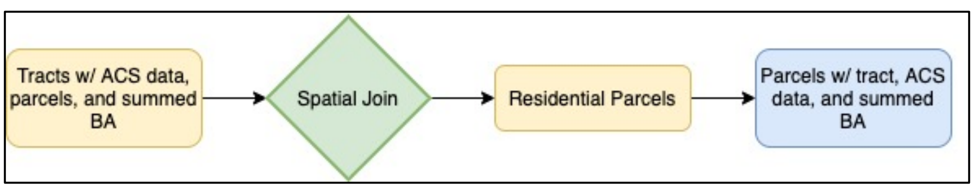

Figure 6: Second spatial join between new tracts and parcels with “BA_Sum”

A new field called “Building_Area_Proportion”(BA_Prop) was created and the calculate field tool was used to divide each individual parcel’s “Building_Area” by the tract’s “BA_Sum.”

$$BA\_Prop = BA / BA\_Sum \quad (1)$$

A new field, “Parcel_Population”(Parcel_Pop) was also created and calculated by multiplying the BA_Prop value with the tract’s population (Tract_Pop) field to yield the estimated population per parcel.

$$Parcel\_Pop = BA\_Prop * Tract\_Pop \quad (2)$$

The lead pipe address point layer was then joined to the parcel layer containing the "Parcel_Population" field with another spatial join and the parcels containing lead pipes were selected into a new layer.

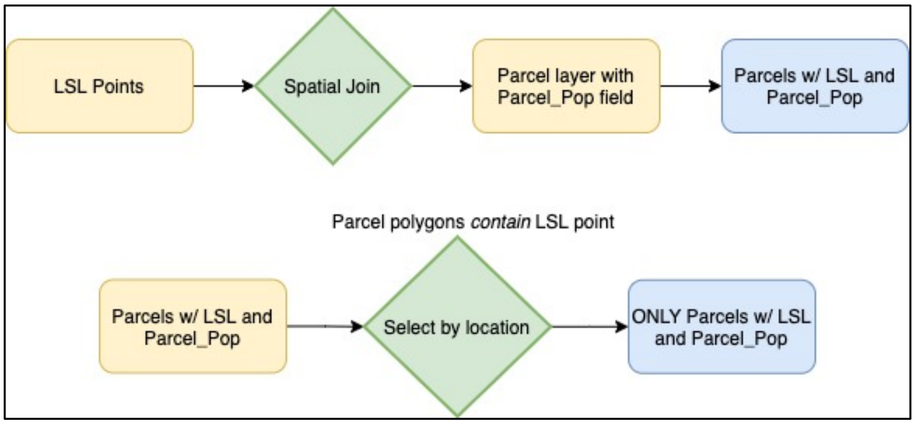

Figure 7: Process for layer with only parcels with LSL

A "sum" summarize field calculation with run for the Parcel_Pop field in the "Only_Parcels_with_LSL" layer to get the total estimated number of people at-risk for lead exposure. Parcels containing the parcel population and with a lead pipe associated address were spatially joined back up to the census tracts with ACS data layer with the merge rule for Parcel_Pop being "sum."

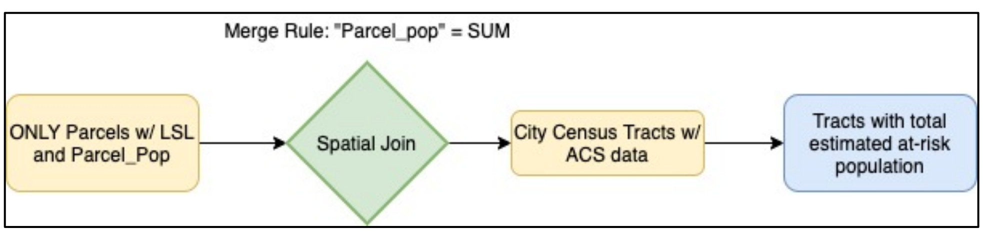

Figure 8: Final spatial join to yield total population at-risk for lead exposure in each census tract

The last join yielded the final result of the dasymetric map, visualizing the estimated number of people at-risk for lead poisoning due to the presence of lead pipes per census tract.

The field containing the estimated people at-risk was summed to yield the overall total of people at-risk for lead exposure for the whole city.

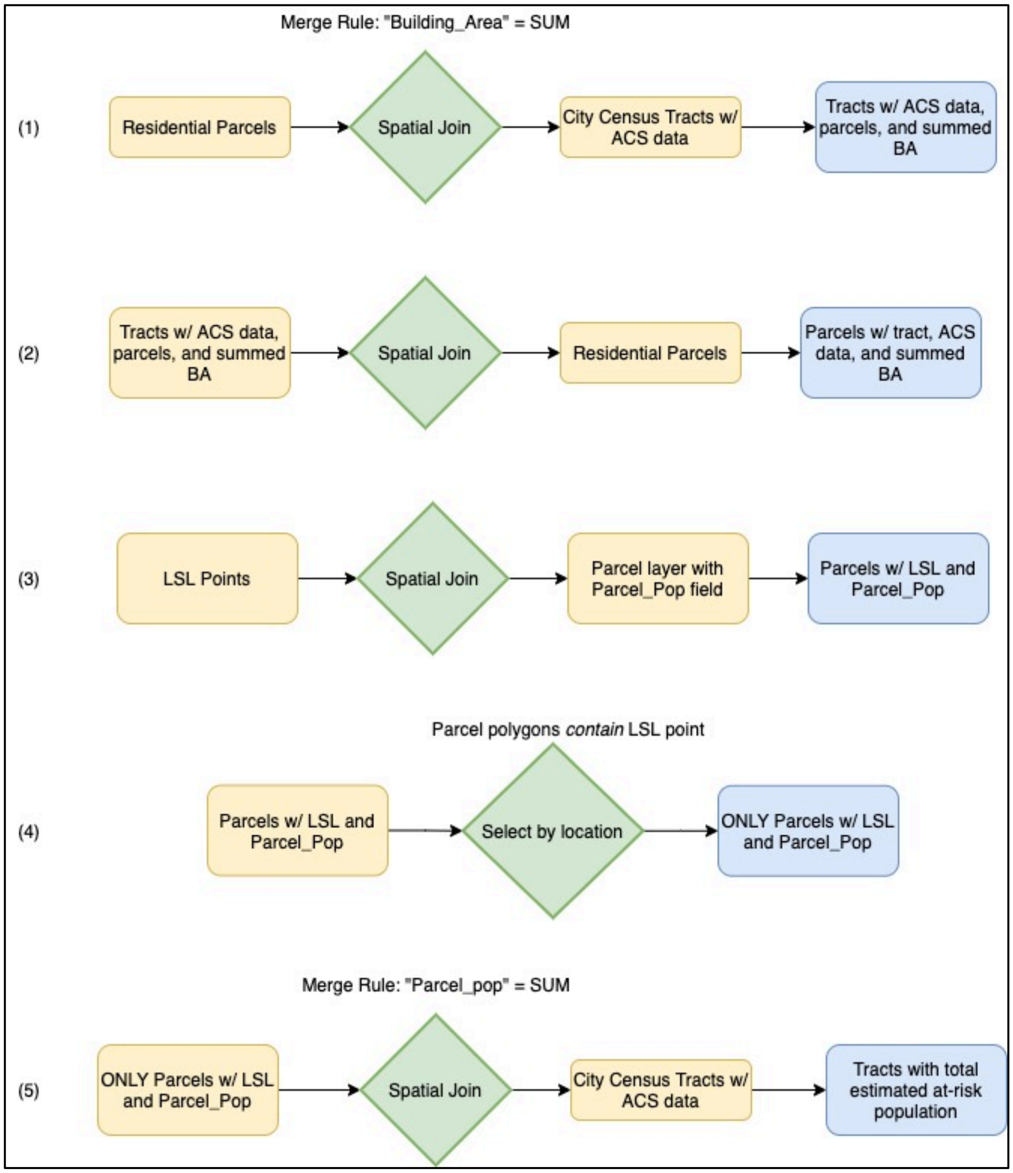

Figure 9: Final workflow of the CEDS method

A Getis-Ord Gi* statistic, or hotspot analysis, was then run to highlight the statistically significant hot spots and cold spots of the at-risk census tracts. The tool examines each feature within a neighborhood of features to determine statistically significant hot and cold spots. A calculation of the local sum of the feature values is compared with the total sum of all the features in the study area. If the difference between the local and total sums is great enough to not be considered a random occurrence, a statistically significant z-score is calculated. ArcGIS Pro has an Optimized Hotspot Analysis geoprocessing tool which automatically calculates the best settings, such as an appropriate distance band value, to produce the most accurate result. The positive significant z-scores represent hot spot clusters amongst the features, and the negative significant z-scores represent cold spots. For the purposes of this study, hot spots would represent areas of high numbers of people at-risk for lead exposure, while cold spots would represent areas of low numbers of people at-risk.

3.3.2 Bivariate Correlation Analysis

To establish the characteristics of the census tracts with the highest number of people at-risk for lead exposure, a bivariate regression analysis was run for seven social indicator explanatory variables obtained from the Census Bureau's ACS data.

An exploratory ordinary least squares analysis was conducted to assess multicollinearity amongst the explanatory variables, as well as to see if the data were spatially autocorrelated. Because OLS assumes normal distribution, the data of the dependent and all seven independent variables were checked for a normal distribution curve. The at-risk population dependent variable, median household income, poverty percentage, percentage of white residents, and percentage of owner residents' data were found to be not normally distributed. The square-root transformation was applied to all five of the not normally distributed datasets because it more accurately corrected for the skew in the data than a logarithmic transformation. While OLS is used to model explanatory variable relationships with a dependent variable, the OLS tool in ArcGIS Pro's spatial statistics toolbox also includes VIF calculations for each individual explanatory variable. Redundancy amongst the explanatory variables can lead to difficult to interpret results as well as the risk of overfitting the model. The OLS model calculates variance inflation factor (VIF) values for each of the explanatory variables to measure any redundancy (ESRI 2021). If the VIF value is too high, it would suggest said variable has multicollinearity with another variable. The suggested cutoff for this value is anywhere from 5-10 (Craney and Surles 2002). The OLS analysis revealed minimal multicollinearity amongst the explanatory variables as all had VIF values within the acceptable less than 10 range.

A Global Moran's I test was run on the residuals of the OLS model to check if the variables were spatially autocorrelated. K nearest neighbors was selected as the conceptualization of spatial relationships parameter with 25 being set as the number of neighbors. The number of neighbors was derived by running the average nearest neighbor tool in ArcGIS Pro on the studied census tracts. The results of the Moran's I showed the residuals were clustered and therefore the variables were spatially autocorrelated.

After the exploratory OLS analysis, a bivariate correlation matrix using Pearson's r was constructed for the dependent and seven independent variables in Excel. Two matrices were constructed, one with the explanatory variables inputted as proportions per census tract, and another with the actual numbers for each explanatory variable. The only explanatory variable not calculated as a proportion was median household income. While the at-risk population data were obtained through dasymetric mapping, there is no way of knowing the same distribution information with the explanatory variable data. Therefore, the proportions are used to characterize the nature of each census tract and how these characterizations relate to the number of people at-risk. Equation 4 is the Pearson's r formula where r is the correlation coefficient, x_i represents the values of the x-variable, $\bar{x}$ is the mean of the x-variable values, y_i is the y-variable values, and $\bar{y}$ is the mean of the y-variable values.

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \tag{4}$$

The correlation matrix computes the Pearson's r coefficient for each variable, dependent and independent, against each other thus calculating the type of linear relationship between each variable (Lane et al. 2013).

Chapter 4 Results

The threat of lead poisoning due to exposure from lead service lines is an issue that thousands of Milwaukeeans face daily. Dasymetric mapping methods help to more accurately display population distribution over a given aerial unit by disaggregating the population data and transferring it into a finer unit of analysis. The methodology explored in this study is the cadastral-based expert dasymetric system which uses breaks down census population estimates into tax lot data. The results of the dasymetric map highlighting which census tracts have the greatest number of people at-risk for lead poisoning are shown in this chapter.

Cases of environmental injustices have been seen time and time again in the historically segregated city of Milwaukee. As shown in Chapter 2, black and brown communities in Milwaukee exhibit higher rates of poverty, lack of access to healthcare, and are more likely to be exposed to environmental toxins. To determine if the topic of lead exposure is indeed another environmental justice issue, bivariate correlation analyses were conducted on seven socioeconomic variables to examine their relationship with the number of at-risk people for lead exposure in each census tract. The results of the analyses are explained in this chapter.

4.1. Dasymetric Mapping Results

The results of the dasymetric map of number of people at risk for lead poisoning per census tract are based on the estimated number of residents living in a building that contains a lead service line. The rationale behind a dasymetric approach being often the generalized results of census tract data can mask an underlying environmental injustice. Figure 10 shows the estimates of people at-risk for lead exposure per census tract, while Figure 11 is a visualization of the at-risk population percentages across census tracts. A graduated color scheme was used with natural breaks to symbolize both maps. The results are nearly identical between the two maps. The final total for the estimated at-risk population for the entire city was 296,327.13 residents. The estimated total population of the city is 594,772 people. According to these findings, 49.8% of Milwaukee's residents are getting their water from LSL and therefore are at higher risk for lead poisoning.

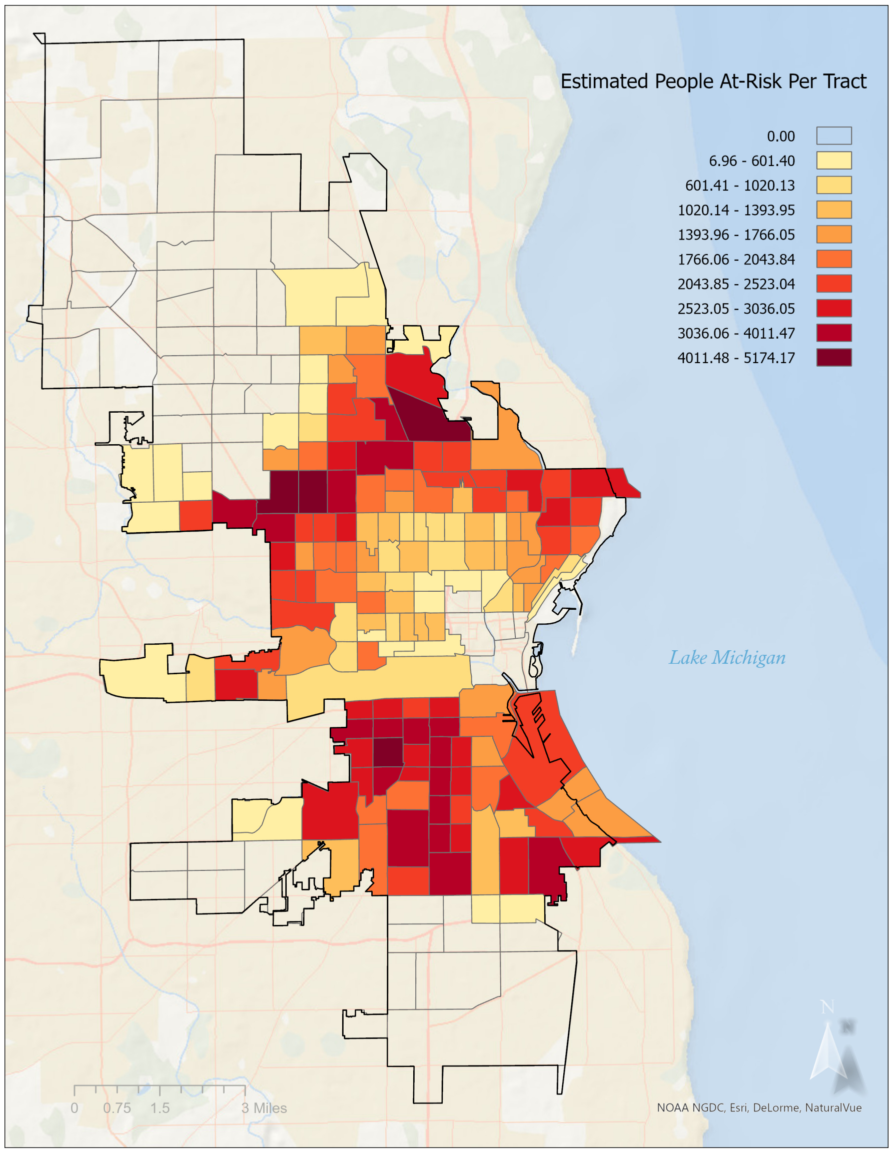

Figure 10: Dasymetric map of the estimated people at-risk for lead exposure by census tract.

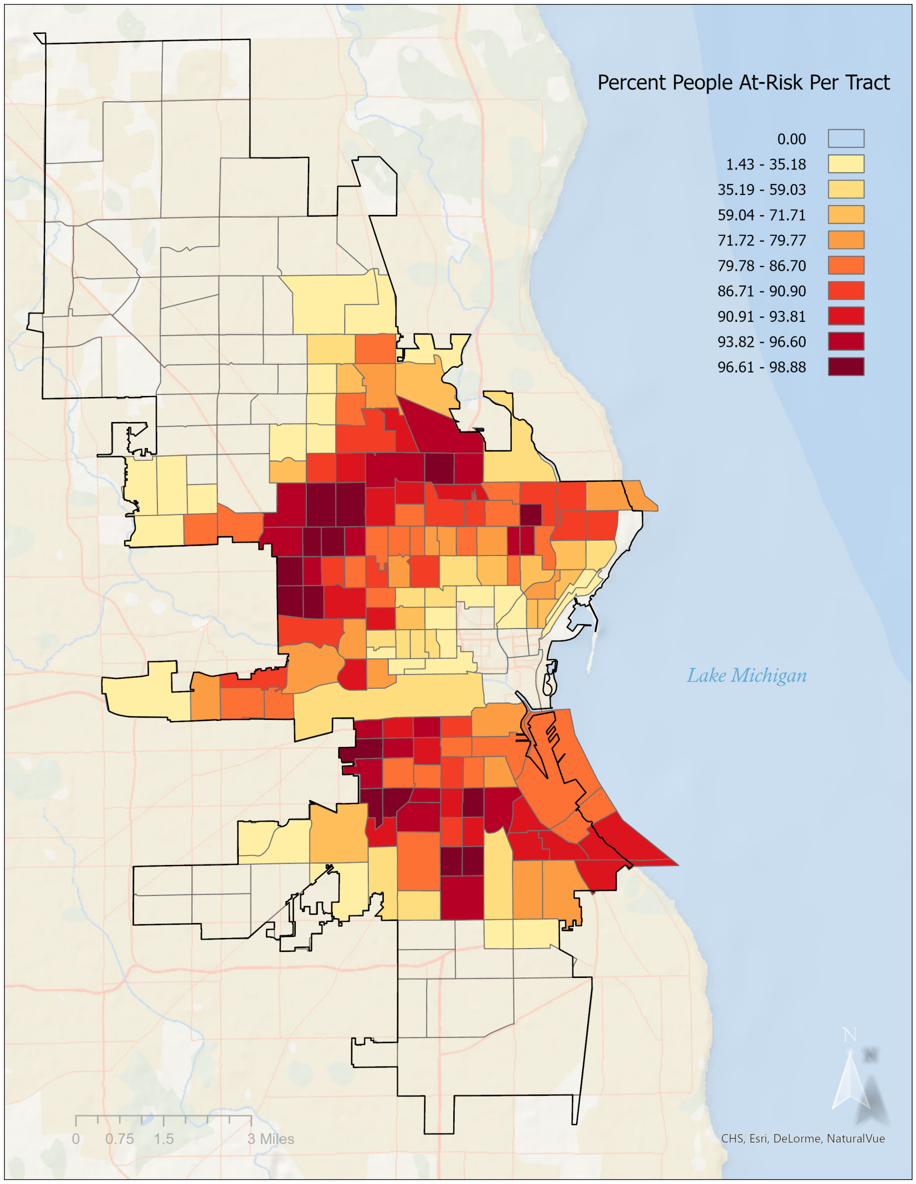

Figure 11: Dasymetric map of the percentage of at-risk people per census tract.

Just by looking at the two maps, it is clear that there are certain areas in the city with higher numbers of people at-risk for lead exposure. However, a statistical significance test of these hot and cold spots is useful to support the initial observation. The result of the optimized hotspot analysis run on the at-risk population estimate per census tract map can be seen in Figure 12. The statistically significant cold spots symbolized in shades blue, hot spots in shades of red, and not significant in white.

Most of the cold spots can be found where there are no lead service lines present. The cold spots seen in the middle of the city are in the Lower East Side, Haymarket, Yankee Hill, and Juneau Town neighborhoods (Figure 12). When compared with the dasymetric at-risk population map, these neighborhoods also have a lower number of people exposed to LSL. The hotspots in the city can be seen on the northern and southern sides of the city center. The hotspots can also be seen to have higher numbers of people exposed to LSL when compared with the dasymetric maps. Figure 12 also labels the neighborhoods that intersect with the hot and cold spot regions of the city.

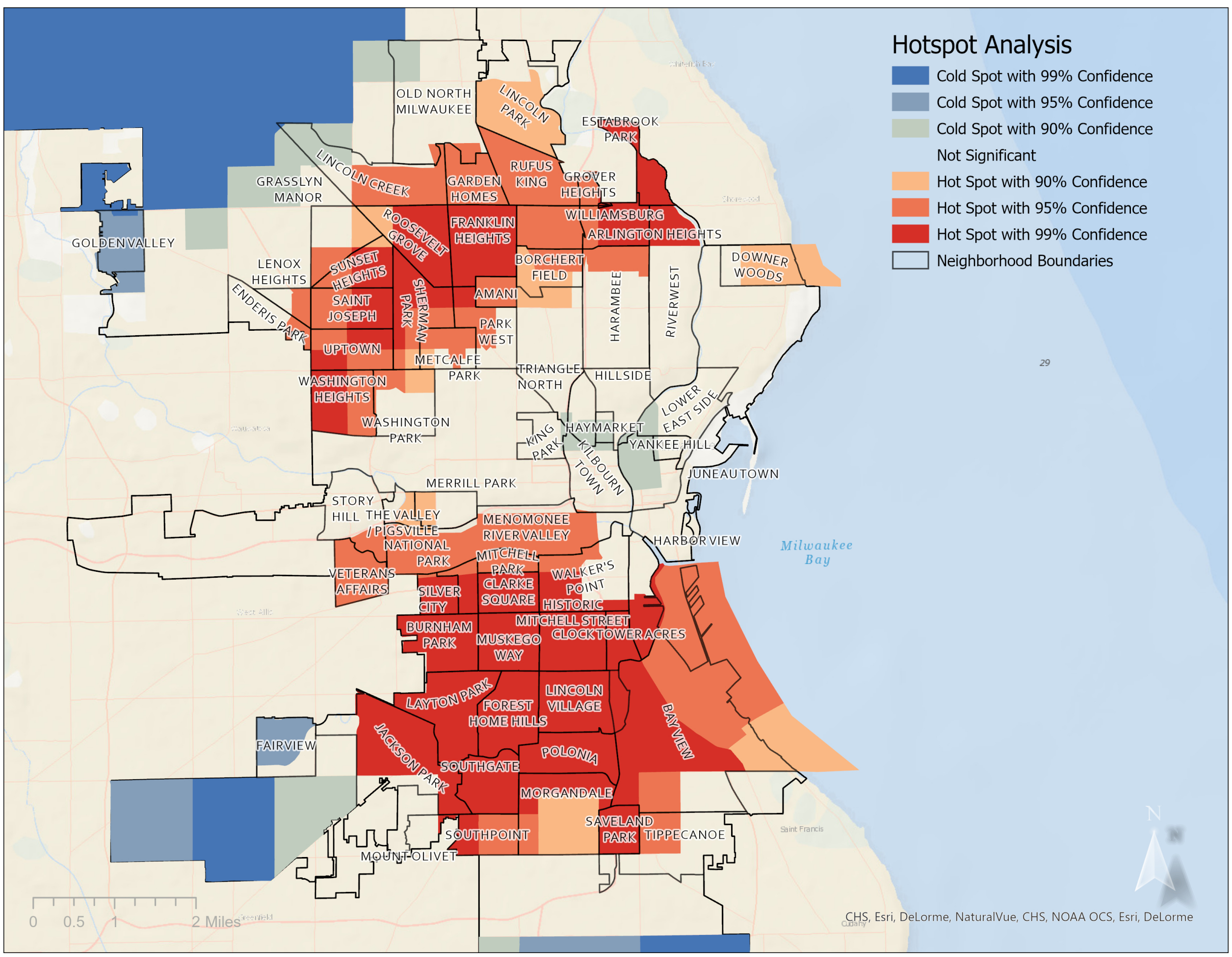

Figure 12: Neighborhoods in the city that intersect with the hot and cold spots

4.2. Bivariate Correlation Analysis Results

In order to characterize the census tracts with high numbers of people at-risk, seven racial and socioeconomic variables were assessed to determine their relationship with the dependent variable. As explained in the methods, an OLS analysis was conducted as the initial exploratory step to quantify the nature of the relationship between the dependent and independent variables as well as assess any bias between the explanatory variables themselves. The results of the analysis revealed strong multicollinearity between some of the explanatory variables and an overall model bias as evidenced by a $p < 0.01$ Jarque-Bera statistic. When the residuals were run through a Moran's I test, the p-value was found to be statistically significant, and the z-value was positive indicating the residuals are not dispersed at random but rather are found to be spatially clustered. Therefore, the residuals from the OLS were found to be spatially autocorrelated rendering the model a bias one.

Before the bivariate correlations were calculated, descriptive statistics of the minimum, maximum, mean, and standard deviation for each of the variables were determined and can be viewed in Table 3.

Table 3: Descriptive statistics for all variables

Variable	*Minimum*	*Maximum*	*Mean*	*Standard Deviation*
At-Risk Pop	0.00	5174.17	1417.83	1251.22
Non-White	6.76	100.00	64.83	27.06
HS Education	0.59	47.67	28.61	10.45
Owns	0.00	96.47	40.19	19.47
Rents	3.53	100.00	59.81	19.47
Poverty Status	2.20	59.92	26.40	14.03
White	0.00	93.24	35.17	27.06
Median Income	7917.00	113375.00	42320.37	17714.77

From there, the explanatory variables were thematically mapped against the dependent variable to provide an initial visualization of geographic patterns between each individual independent and the dependent variable. In these bivariate maps, the darker red regions show where the dependent and independent variable values were both high, and the lighter areas where both values were low. An additional element of transparency was added in the symbology where census tracts with higher populations were opaquer and those with lower populations more transparent (Figure 13). Each of the variables were symbolized in tertiles with cutoff values one standard deviation more and less from the mean of the dataset.

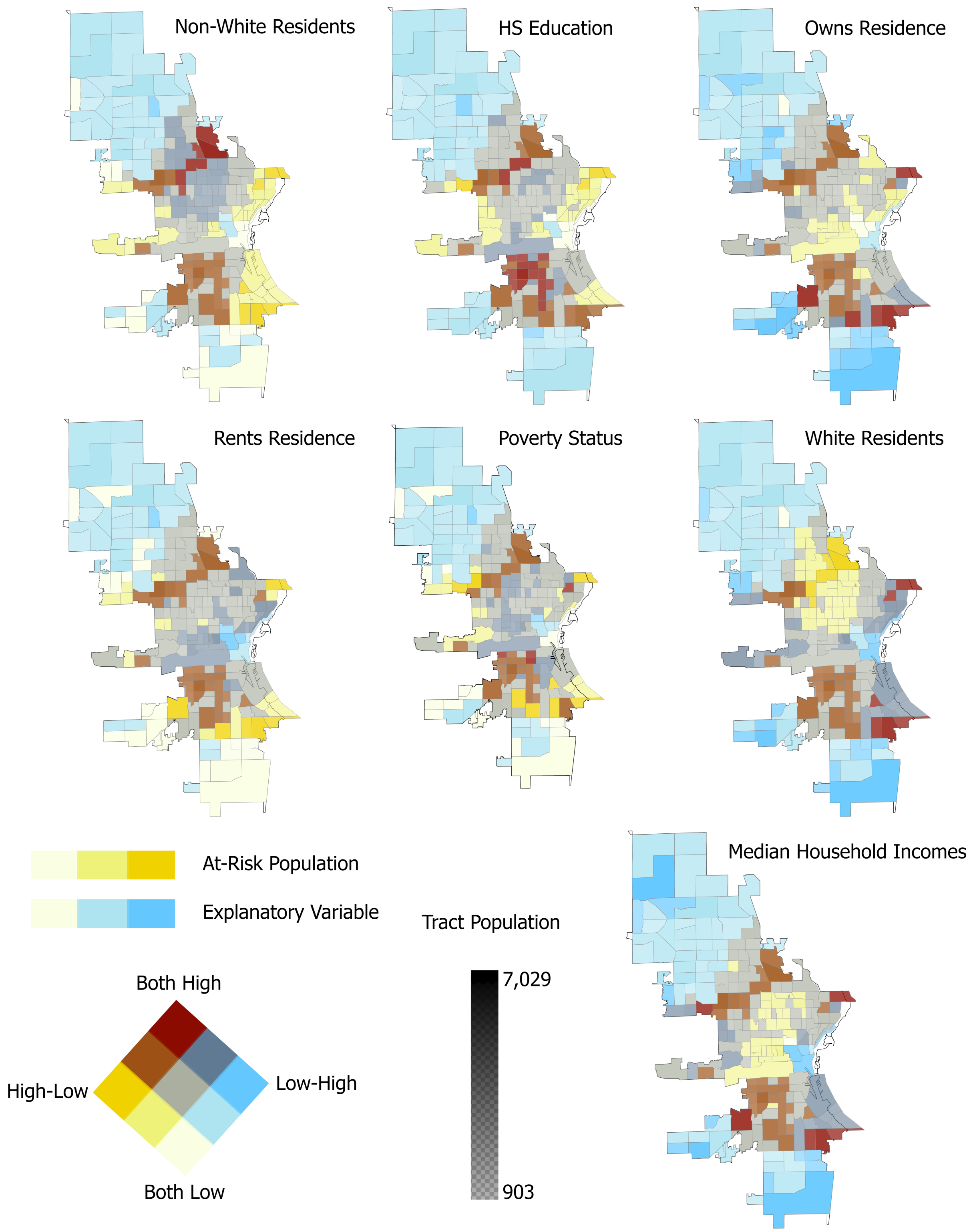

Figure 13: Bivariate maps displaying each independent variable with the dependent variable

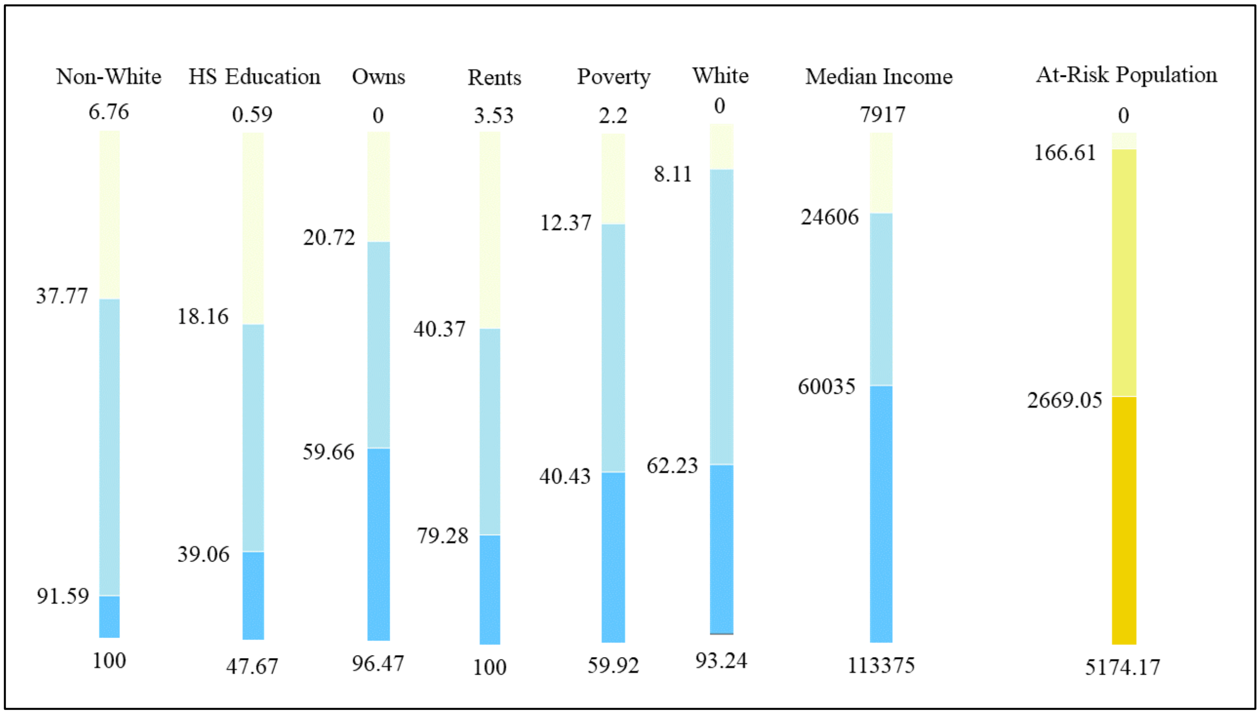

Figure 14: Legend showing the standard deviation cut-off values for each variable

The results of the bivariate correlation matrices can be seen in Tables 4 and 5. The first matrix (Table 4) was calculated using actual population estimates while the second matrix (Table 5) was calculated using the proportions of the population estimates over the total population in census tract for the explanatory variables. The values with a color background are the Pearson's r coefficients, while the values with a white background are the corresponding p-values. Values where $p \leq 0.05$ are denoted with a single asterisk and those where $p \leq 0.001$ are denoted with a double asterisk. The two correlation matrices are nearly identical with which variables have positive and negative relationships with the at-risk population according to their Pearson's r coefficient value. When comparing the at-risk population dependent variable with the seven explanatory variables, the only variable that differed between the two matrices was the white population data. When determining the relationship between the number of at-risk people with the number of white people, there is a weak positive correlation.

Table 4: Correlation matrix between explanatory and dependent variables

	HS	MedianIncome	Poverty	White	Non-White	Owns	Rents	At-Risk_Pop
HS	1.0000	0.151	0.000**	0.001**	0.0000**	0.0000**	0.0002**	0.005*
MedianIncome	-0.0997	1.0000	0.0000**	0.0000**	0.0000**	0.0000**	0.108	0.138
Poverty	0.5428	-0.5685	1.0000	0.010	0.000**	0.051*	0.000**	0.001**
White	0.2376	0.6298	-0.1774	1.0000	0.150	0.000**	0.002	0.3562
Non-White	0.8181	-0.2995	0.7321	-0.0999	1.0000	0.003*	0.000**	0.000**
Owns	0.4996	0.5954	-0.1352	0.6452	0.2066	1.0000	0.423	0.611
Rents	0.2532	-0.1116	0.4512	0.2121	0.3399	0.0557	1.0000	0.393
At-Risk_Pop	0.1932	-0.1030	0.2370	0.0641	0.3061	-0.0353	0.0594	1.0000

Table 5: Correlation matrix between explanatory variable percentages and the dependent variable

	HS	MedianIncome	Poverty	White	Non-White	Owner	Rent	At-Risk_Pop
HS	1.0000	0.000**	0.000**	0.000**	0.000**	0.793	0.793	0.043
MedianIncome	-0.4745	1.0000	0.000**	0.000**	0.000**	0.000**	0.000**	0.138
Poverty	0.3942	-0.8331	1.0000	0.000**	0.000**	0.000**	0.000**	0.122
White	-0.5829	0.7250	-0.6494	1.0000	NA	0.000**	0.000**	0.342
Non-White	0.5829	-0.7250	0.6494	-1.0000	1.0000	0.000**	0.000**	0.342
Owner	-0.0182	0.6839	-0.6603	0.3458	-0.3458	1.0000	NA	0.431
Rent	0.0182	-0.6839	0.6603	-0.3458	0.3458	-1.0000	1.0000	0.431
At-Risk_Pop	0.1403	-0.1030	0.1074	-0.0661	0.0661	-0.0548	0.0548	1.0000

Meanwhile, the relationship with the variable when it is represented as a percentage of the population, it is shown to have a weak negative correlation with the dependent variable.
This could be because there are high numbers of white residents in census tracts across the city, but proportionally, there are certain census tracts where white people are the overwhelming minority. Using percentages to represent the explanatory demographic data can standardize the difference in estimates across the categories. None of the p-values for the coefficients calculated for the explanatory variable percentages and the dependent variable returned a significant value. However, in the real number correlation results, high school education, poverty, and the non-white population variables were all shown to have statistically significant relationships with the dependent variable.

When comparing the explanatory variables against each other, there were a few more differences between the two matrices in respect to positive and negative linear relationships. Further, the strength of the linear relationships between certain variables shifted between the real numbers and percentages. One notable example is the relationship between those with a high school education or less and the number of people who own their place of residence. In the real numbers, the two variables have a positive linear relationship with a Pearson's r coefficient of 0.4996. However, when the numbers are converted to percentages, the relationship between the two variables have a slightly negative correlation (Pearson's r = -0.0182). The shift not only in the direction of the relationship from positive to negative, but also in the strength of correlation is a testament to the difference calculating Pearson's r with real numbers versus percentages.

The results of the dasymetric mapping and analysis revealed patterns similar to other environmental justice research with toxin exposure and social indicators. In depth explanations and implication of the results will be discussed in Chapter 5 of this paper.

Chapter 5 Conclusion

The two main questions that this project sought to answer were one, how many people run the risk of being exposed to lead due to the presence of LSL at their place of residence, and two, what is the relationship between SES indicator variables and the number of people at-risk per census tract. In response to the first question, the results of the dasymetric map of lead exposure risk throughout the city based on presence of LSL reveal drastic differences in number of people at-risk per census tract. There are significant hot and cold spots of exposure by census tract. The results of the regression analyses helped shed light on the answers to the second question. Non-white individuals, those with a high school education, and those who have experienced poverty have a higher chance of being exposed to lead contaminated water.

In the following chapter, the implications of the dasymetric map are discussed as well as the results from the bivariate correlation analyses. Understanding how one's race and socioeconomic status in the community can raise the likelihood of lead exposure is a crucial step in the fight for environmental equity in Milwaukee. The results of this project represent just the beginning of understanding lead exposure due to LSL as an EJ issue in Milwaukee. The limitations, and there were many, are reflected on and recommendations for future research are presented at the end of the chapter.

5.1. Dasymetric Map of At-Risk Populations

The results of the dasymetric map paint a more precise picture of how many people are at risk for lead exposure due to the presence of lead pipes on their property. Upon visual inspection, there are certain census tracts that have a much higher number of people exposed to LSL, and these high-risk census tracts are geographically close together. Upon conducting a Hot Spot analysis, the hot spots and cold spots for LSL exposure are confirmed. As discussed in chapter 4, the estimated at-risk population for the entire city is 296,327.13 residents out of a total population of 594,772. Based on these population estimates, 49.8% of Milwaukee's residents are at-risk for lead exposure due to living in a residence linked to a lead service line. Nearly 50% of the city's population being at-risk for lead exposure seems like a large proportion considering MWW's claims that 40% of active service lines are made of lead. The final total of at-risk residents of Milwaukee being greater than the percentage of LSL in the city is confirmation that population density greatly varies in certain census tracts over others.

The method of disaggregating the population data and redistributing the population estimates (as a function of proportion of building area within each parcel) yielded a more precise estimate of at-risk people due to the fine spatial resolution of each parcel within a census tract. The methods used in this study were a slight variation on the CEDS method outlined by Maantay et al. (2007) due to differences in data availability. The residential unit and adjusted residential area values put forth in the original methodology were swapped for building area data found in the Milwaukee parcels schema. The building area per parcel served as the proxy unit for population dispersion and the ratio of each parcel's building area to the census tract's total building area yielded the estimated number of people living in each individual parcel. Because the tract data could be disaggregated down to such a fine resolution, the nuances of population variance within each census tract were accounted for when totaling the number of individuals living in buildings with LSLs. Despite the slight variance in the methodology, the CEDS technique of dasymetric mapping is a highly effective way to redistribute population to reflect real-world situations more accurately in Milwaukee (Maantay et al. 2007).

When examining the hot spot map, it is evident that the hot spots of the most people at-risk for lead poisoning can be found in the northern and southern parts of the city center.

Focusing in on the northern neighborhoods such as Harambee, Park West, Sherman Park, and Franklin Heights, it can be seen that the these historically black neighborhoods are at the center of the hot spot area. Franklin Heights for example, saw a major rise in its African American population in the 1980s after the construction of the freeway led to more affluent, white residents to move away from the area. Today, it is one of the poorer neighborhoods in the city with a primarily black population (Gurda 1999; Nelsen 2016). The southern hot spots are found primarily in the Hispanic neighborhoods of the city. Specifically, Walker's Point, Clark Square, Silver City, Burnham Park, Muskego Way, Historic Mitchell Street, Lincoln Village, Polonia, Southgate, and Layton Park all have majority Latinx populations (Johnson 2020). Many of these neighborhoods also have lower than average median household income when compared to Milwaukee as a whole and have less than a high school education (City-Data 2019). There is an isolated cold spot found in the middle of the city that intersects the Lower East Side, Haymarket, Yankee Hill, and Juneau Town neighborhoods. This is a far more affluent part of Milwaukee, with a majority white population, people with college degree educations, and median household incomes more than double that of the city overall (City-Data 2019). The demographic statistics of the neighborhoods that intersect with the hot and cold spots in Milwaukee closely reflect those obtained in this study and will be explored further in section 5.2 of this chapter.

5.1.1. Considerations and Limitations

The dasymetric methods used in this study were a solid starting point for assessing at-risk populations within the city. However, there were a few considerations and limitations that should be acknowledged. A major limitation was access to cadastral metadata. While the City of Milwaukee makes their data easily accessible for anyone to download, deciphering the coded schema often proved a challenge. Metadata documents were found in separate locations than the data portal and were unclear in referencing field codes between the documentation and the actual attribute table. An example is the metadata document for zoning codes for all the parcel data in the city. Beyond the general “residential” designation, according to the zoning ordinances volume two subchapter five, the residential zones within the city were broken down into 22 subcategories. These subcategories would distinguish between single family, two family, multi-family, etc. places of residence, but these coded categories would not always be present in the parcel attribute data (Legislative Reference Bureau 2020). Given the limited timeframe of this study, it was not possible to relate the zoning and parcel tables to then cross-reference the specific zoning code to estimate population distribution. Further, just because a residence is designated as a single-family home does not mean a single family is occupying the property. One of the indicators of a lower socioeconomic status is overcrowding in one’s place of residence (Galobardes 2006).

5.2. Racial and Socioeconomic Correlations

One must be vigilant when conducting research on potential correlations between dependent and explanatory variables. As the saying goes, correlation does not necessarily mean causation. The purpose of performing the bivariate correlation analyses in this study was to illuminate some of the social indicators that could explain the number of at-risk people exposed to lead in each census tract. To perform this analysis, a set of explanatory variables were selected from existing literature studying socioeconomic disparities within Milwaukee. Based on the background research, the explanatory variables selected for this study were median household income, poverty status, educational attainment, tenure status, and race. According to the resulting correlation matrices, those with a high school education, experiencing poverty, and non-white residents all have statistically significant positive linear relationship with the number of at-risk people (Tables 4 and 5). Meanwhile, the number of white residents and wealthier census tracts where people are property owners are shown to have negative linear relationships (Tables 4 and 5). The visualization of these relationships can be seen in the bivariate map series where each explanatory variable was mapped with the dependent variable. The overlapping variables with positive linear correlations follow a similar pattern to that of the hotspot analysis map. The results are further confirmed when examining the demographic characteristics of the neighborhoods that intersect the hot and cold spots described in section 5.1. Therefore, it can be concluded that the census tracts with higher numbers and percentages of non-white individuals, lower education, and low socioeconomic status are more likely to have high numbers of people at-risk for lead exposure due to an LSL at their place of residence.

It tracks that the number of non-white people, high school education level, and poverty status would be the variables with the most impact on the number of at-risk individuals per census tract. The three variables are shown to have strong positive linear relationships amongst each other. The overlap between race and socioeconomic status is a well recorded phenomenon in the United States (Raddatz and Mennis 2013). In the case of Black Americans, it has been found that indicators such as income level, education, and occupation are not as strong at predicting a Black American's socioeconomic status as it would a White American's. This is most likely due to a general lack of intergenerational wealth and not seeing the same financial returns for higher education than White Americans (Wolff et al. 2010). In other words, the strongest predictor of a Black American's socioeconomic status would be the fact that they are Black and navigating a world that is systemically not made for their benefit. In Milwaukee's case, historically racist policies such as redlining shaped the urban geographic landscape that is seen today where neighborhoods are for the most part still segregated by race. Discriminatory housing markets drove people of color, those of whom were primarily black, to settle in regions of the city where the buildings were older and therefore more likely to have LSL (Trotter 1985). The ripple effect of discriminatory legislation from years ago can be seen in health disparities, lack of social mobility, and increased exposure to environmental toxins within Milwaukee today. The results of this study confirm this pattern of inequity amongst poorer, less education communities of color within the city.

5.2.1. Considerations and Limitations

A limitation in this study was the issue of the margin of error in the ACS census tract data. It was found that the more people surveyed in the census tract, the lower the margin of error, while the inverse was true when fewer people were surveyed. While most of the tracts had MOE within an acceptable range, there were a few outliers with low population counts that caused the MOE to become very large. The coefficient of variation was calculated for each ACS demographic category to disclose which had higher levels of error (Table 2). The variable with the highest percentage of low reliability CV scores was the estimates for the number of white people per census tract. In future iterations, steps could be taken to further reduce the MOE by aggregating contiguous census tracts in areas with hot and cold spots.

5.3 Future Study Recommendations

The results of the current study were sufficient as an initial step when examining the environmental justice issue of populations within Milwaukee exposed to lead pipes. The questions that this study uncovered could lead to a plethora of future research in this field. For example, it would be interesting to construct a socioeconomic index using principal component analysis (PCA) to select the indicator variables from a wider range of domains. The advantage of this method is it keeps the decision making to a minimum and uses statistics to select out the best fit variables for an SES index. Once an index is created, values can be assigned to each census tract based on how their SES index scores. The values can then be used to compare SES with the at-risk for lead exposure population data using a generalized additive model. Such methodology is outlined in Padilla et al. and Lalloué et al., who both researched environmental justice issues in relation to air pollutants in France (Lalloué et al. 2013; Padilla et al. 2014). It would also be interesting to compare lead exposure between lead paint and LSL within Milwaukee. According to a grassroots coalition called Lead Free Milwaukee, city officials have been pressing the narrative that lead paint is the primary source for lead exposure as opposed to lead service lines leaching the metal into residents' drinking water. The coalition asserts that the city is downplaying the severity of the issue of lead in the water and is trying to redirect attentions and efforts to mitigating lead paint (Washington and Welcenbach 2019). Finally, given the recent revisions to the Lead and Copper Rule by the EPA, a study on the efficacy of partial versus full lead service line replacements on reducing lead concentrations in the water supply could be useful to residents of the city. When the city replaces an LSL, usually a partial replacement approach is carried out where the city will replace only the public portion of the LSL, and leave the private portion as is. Research on the effects of partial service line replacements has shown that in only replacing part of the LSL, more lead is found in the water supply due to the disturbance of the service line (Lewis et al. 2017).

The fight for environmental equity is a long, uphill battle. In characterizing the neighborhoods that experience the highest levels of exposure, decision makers in Milwaukee will know where to focus their mitigation efforts and activist groups can know where to go to educate people about the dangers of lead exposure. GIS is a powerful tool that can contribute to the growing research on EJ through techniques such as dasymetric mapping and correlation analysis.

By conducting these studies, geospatial scholars can aid in giving a voice to the voiceless. The fruits of their labor can empower those who want to see justice served and insure healthy communities worldwide.

References

Alexeeff, George V., John B. Faust, Laura Meehan August, Carmen Milanes, Karen Randles, Lauren Zeise, and Joan Denton. 2012. "A Screening Method for Assessing Cumulative Impacts." International Journal of Environmental Research and Public Health 9 (2): 648–59. https://doi.org/10.3390/ijerph9020648.

Anderton, Douglas L., Andy B. Anderson, John Michael Oakes, and Michael R. Fraser. 1994. "Environmental Equity: The Demographics of Dumping." *Demography* 31 (2): 229–48. https://doi.org/10.2307/2061884.

Banzhaf, Spencer, Lala Ma, and Christopher Timmins. 2019. "Environmental Justice: The Economics of Race, Place, and Pollution." *Journal of Economic Perspectives* 33 (1): 185–208. https://doi.org/10.1257/jep.33.1.185.

Berkley, Jennifer. 2017. "Using American Community Survey Estimates and Margins of Error Webinar - Transcript." Decennial Statistical Studies Division, April 19. https://census.webex.com/census/ldr.php?RCID=714c57425cd24a5b15fb4eea3da438ec.

Brown, Mary Jean, and Stephen Margolis. 2012. "Lead in Drinking Water and Human Blood Lead Levels in the United States." Supplement 61. Morbidity and Mortality Weekly Report. Centers for Disease Control and Prevention.

Butler, Lindsey J., Madeleine K. Scammell, and Eugene B. Benson. 2016. "The Flint, Michigan, Water Crisis: A Case Study in Regulatory Failure and Environmental Injustice." Environmental Justice 9 (4): 93–97. https://doi.org/10.1089/env.2016.0014.

Cameron, Ryan. 2020. "Mapping Future Population Impacts Caused by Sea Level Rise in Huntington Beach and Newport Beach: Comparing the Cadastral-Based Dasymetric System to Past Dasymetric Mapping Methods." University of Southern California.

CCRPC. 2015. "Understanding and Using the U.S. Census Bureau's American Community Survey." Champaign County Regional Planning Commission.

CDC. 2020a. "Community, Work, and School." Centers for Disease Control and Prevention. February 11, 2020. https://www.cdc.gov/coronavirus/2019-ncov/community/health-equity/race-ethnicity.html.

CDC. 2020b. "Health Problems Caused by Lead | NIOSH | CDC." September 10, 2020. https://www.cdc.gov/niosh/topics/lead/health.html.

CDC. 2021. "Lead (Pb) Toxicity: What Are U.S. Standards for Lead Levels? | Environmental Medicine | ATSDR." February 9, 2021. https://www.atsdr.cdc.gov/csem/leadtoxicity/safety_standards.html.

Census Bureau. 2020. "U.S. Census Bureau QuickFacts: Milwaukee City, Wisconsin." 2020. https://www.census.gov/quickfacts/milwaukeecitywisconsin.

Chaffetz, Jason, John L Mica, Michael R Turner, John J Duncan, Jim Jordan, Tim Walberg, Justin Amash, et al. 2016. *Examining Federal Administration of the Safe Drinking Water Act in Flint, Michigan*.

Chakraborty, Jayajit. 2011. "Revisiting Tobler's First Law of Geography: Spatial Regression Models for Assessing Environmental Justice and Health Risk Disparities." In Geospatial Analysis of Environmental Health, edited by Juliana A. Maantay and Sara McLafferty, 337–56. Geotechnologies and the Environment. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-0329-2_17.

Chakraborty, Jayajit, Timothy W. Collins, Marilyn C. Montgomery, and Sara E. Grineski. 2014. "Social and Spatial Inequities in Exposure to Flood Risk in Miami, Florida." Natural Hazards Review 15 (3): 04014006. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000140.

Chisolm, J. Julian. 1971. "Lead Poisoning." Scientific American 224 (2): 15–23.

City-Data. 2019. "Lower East Side Neighborhood in Milwaukee, Wisconsin (WI), 53202 Subdivision Profile - Real Estate, Apartments, Condos, Homes, Community, Population, Jobs, Income, Streets." 2019. https://www.city-data.com/neighborhood/Lower-East-Side-Milwaukee-WI.html.

City-Data. 2019. "Muskego Way Neighborhood in Milwaukee, Wisconsin (WI), 53204 Subdivision Profile - Real Estate, Apartments, Condos, Homes, Community, Population, Jobs, Income, Streets." 2019. https://www.city-data.com/neighborhood/Muskego-Way-Milwaukee-WI.html.

City of Milwaukee. 1999. "City of Milwaukee Master Property File Documentation."

City of Milwaukee. 2016. "Lead-Safe Milwaukee." 2016. https://city.milwaukee.gov/LeadSafeMKE.

Collins, Mary B. 2011. "Risk-Based Targeting: Identifying Disproportionalities in the Sources and Effects of Industrial Pollution." American Journal of Public Health 101 (Suppl 1): S231–37. https://doi.org/10.2105/AJPH.2011.300120.

Dang, Do, Morgan Lively, and Alonzo Jalan. 2021. "Lead Poisoning and Racism in the Time of COVID-19." Wisconsin Medical Journal, March, 3.

Denver Water. 2021. "Lead Service Lines." Denver Water. 2021. https://www.denverwater.org/your-water/water-quality/lead/lead-service-lines.

Eicher, Cory L., and Cynthia A. Brewer. 2001. "Dasymetric Mapping and Areal Interpolation: Implementation and Evaluation." Cartography and Geographic Information Science 28 (2): 125–38. https://doi.org/10.1559/152304001782173727.

Emer, Lindsay R., Amy E. Kalkbrenner, Mallory O'Brien, Alice Yan, Ron A. Cisler, and Lance Weinhardt. 2020. "Association of Childhood Blood Lead Levels with Firearm Violence

Perpetration and Victimization in Milwaukee." Environmental Research 180 (January): 108822. https://doi.org/10.1016/j.envres.2019.108822.

ESRI. 2021. "How Hot Spot Analysis (Getis-Ord Gi*) Works." 2021. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-hot-spot-analysis-getis-ord-gi-spatial-stati.htm.

ESRI. 2021. "Ordinary Least Squares (OLS) (Spatial Statistics)—ArcGIS Pro | Documentation." April 2021. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/ordinary-least-squares.htm.

Galobardes, B. 2006. "Indicators of Socioeconomic Position (Part 1)." Journal of Epidemiology & Community Health 60 (1): 7–12. https://doi.org/10.1136/jech.2004.023531.

Greer, Danielle M., Dennis J. Baumgardner, Farrin D. Bridgewater, David A. Frazer, Courenay L. Kessler, Erica S. LeCounte, Geoffrey R. Swain, and Ron A. Cisler. 2013. "Milwaukee Health Report 2013: Health Disparities in Milwaukee by Socioeconomic Status." Center for Urban Population Health: Milwaukee, WI.

Gulledge, Kelly W. 2020. "An Analysis of Racial Disparity in the Distribution of Alcohol Licenses and Retailers in Orange County, California." University of Southern California.

Gurda, John. 1999. *The Making of Milwaukee*. Milwaukee, Wis: Milwaukee County Historical Society.

Hanna-Attisha, Mona, Jenny LaChance, Richard Casey Sadler, and Allison Champney Schnepp. 2015. "Elevated Blood Lead Levels in Children Associated With the Flint Drinking Water Crisis: A Spatial Analysis of Risk and Public Health Response." American Journal of Public Health 106 (2): 283–90. https://doi.org/10.2105/AJPH.2015.303003.

Herries, Jim. 2021. "A Straightforward Approach to Mapping Margins of Error." *ArcGIS Blog* (blog). June 7, 2021. https://www.esri.com/arcgis-blog/products/arcgis-living-atlas/mapping/a-straightforward-approach-to-mapping-margins-of-error/.

Heynen, Nik, Harold A. Perkins, and Parama Roy. 2006. "The Political Ecology of Uneven Urban Green Space: The Impact of Political Economy on Race and Ethnicity in Producing Environmental Inequality in Milwaukee." Urban Affairs Review 42 (1): 3–25. https://doi.org/10.1177/1078087406290729.

Hillier, Amy E. 2003. "Spatial Analysis of Historical Redlining: A Methodological Exploration - ProQuest." Journal of Housing Research 14 (1): 31.

Hinkel, Jochen. 2011. "'Indicators of Vulnerability and Adaptive Capacity': Towards a Clarification of the Science–Policy Interface." Global Environmental Change 21 (1): 198–208. https://doi.org/10.1016/j.gloenvcha.2010.08.002.

Hou, Shuangxing, Lianfang Yuan, Pengpeng Jin, Bojun Ding, Na Qin, Li Li, Xuedong Liu, Zhongliang Wu, Gang Zhao, and Yanchun Deng. 2013. "A Clinical Study of the Effects of Lead Poisoning on the Intelligence and Neurobehavioral Abilities of Children." Theoretical Biology and Medical Modelling 10 (1): 13. https://doi.org/10.1186/1742-4682-10-13.

Jannene, Jeramey. 2020. "City Hall: Milwaukee on 70-Year Pace To Replace Lead Laterals." Urban Milwaukee. March 13, 2020. https://urbanmilwaukee.com/2020/03/13/city-hall-milwaukee-on-70-year-pace-to-replace-lead-laterals/.

Jerrett, Michael, Richard T Burnett, Pavlos Kanaroglou, John Eyles, Norm Finkelstein, Chris Giovis, and Jeffrey R Brook. 2001. "A GIS–Environmental Justice Analysis of Particulate Air Pollution in Hamilton, Canada." Environment and Planning A: Economy and Space 33 (6): 955–73. https://doi.org/10.1068/a33137.

Johnson, John D. 2020. "John D. Johnson | Neighborhoods Where Milwaukee Isn't Segregated." June 17, 2020. https://johndjohnson.info/post/milwaukee-neighborhoods-race-2018/.

Kelly-Reif, Kaitlin, and Steve Wing. 2016. "Urban-Rural Exploitation: An Underappreciated Dimension of Environmental Injustice." Journal of Rural Studies 47 (October): 350–58. https://doi.org/10.1016/j.jrurstud.2016.03.010.

Kennedy, Erika, and Leslie W. Kennedy. 2004. "Using Dasymetric Mapping for Spatially Aggregated Crime Data." Journal of Quantitative Criminology 20 (3): 243–62. https://doi.org/10.1023/B:JOQC.0000037733.74321.14.

Lalloué, Benoît, Jean-Marie Monnez, Cindy Padilla, Wahida Kihal, Nolwenn Le Meur, Denis Zmirou-Navier, and Séverine Deguen. 2013. "A Statistical Procedure to Create a Neighborhood Socioeconomic Index for Health Inequalities Analysis." International Journal for Equity in Health 12 (1): 21. https://doi.org/10.1186/1475-9276-12-21.

Lane, David M., David Scott, Mikki Hebl, Rudy Guerra, Dan Osherson, and Heidi Zimmer. 2013. *Introduction to Statistics*. Online. Rice University and University of Houston.

Legislative Reference Bureau. 2020. "Zoning 295-501 Subchapter 5 Residential Districts." In Milwaukee Code of Ordinances: Volume 2 Building and Zoning Code, 2:598. Common Council.

Lester, James. 2018. Environmental Injustice In The U.S.: Myths And Realities. Routledge.

Lewis, Carrie M., Lon A. Couillard, Patricia J. Klappa, and Terrence D. Vandenbush. 2017. "Lead Water Service Lines: Extensive Sampling and Field Protocol Protect Public Health." Journal - American Water Works Association 109 (January): 34–41. https://doi.org/10.5942/jawwa.2017.109.0016.

Lynch, Emily E., and Helen C. S. Meier. 2020. "The Intersectional Effect of Poverty, Home Ownership, and Racial/Ethnic Composition on Mean Childhood Blood Lead Levels in Milwaukee County Neighborhoods." PLoS ONE 15 (6). https://doi.org/10.1371/journal.pone.0234995.

Maantay, Juliana. 2002. "Mapping Environmental Injustices: Pitfalls and Potential of Geographic Information Systems in Assessing Environmental Health and Equity." Environmental Health Perspectives 110 (suppl 2): 161–71. https://doi.org/10.1289/ehp.02110s2161.

Maantay, Juliana Astrud, Andrew R. Maroko, and Christopher Herrmann. 2007. "Mapping Population Distribution in the Urban Environment: The Cadastral-Based Expert Dasymetric System (CEDS)." Cartography and Geographic Information Science 34 (2): 77–102. https://doi.org/10.1559/152304007781002190.

Maantay, Juliana, and Andrew Maroko. 2009. "Mapping Urban Risk: Flood Hazards, Race, & Environmental Justice in New York." Applied Geography 29 (1): 111–24. https://doi.org/10.1016/j.apgeog.2008.08.002.

Mayo Clinic Staff. 2019. "Lead Poisoning - Symptoms and Causes." Mayo Clinic. December 2019. https://www.mayoclinic.org/diseases-conditions/lead-poisoning/symptoms-causes/syc-20354717.

Mennis, Jeremy. 2003. "Generating Surface Models of Population Using Dasymetric Mapping*." The Professional Geographer 55 (1): 31–42. https://doi.org/10.1111/0033-0124.10042.

Messer, Lynne C., Barbara A. Laraia, Jay S. Kaufman, Janet Eyster, Claudia Holzman, Jennifer Culhane, Irma Elo, Jessica G. Burke, and Patricia O'Campo. 2006. "The Development of a Standardized Neighborhood Deprivation Index." Journal of Urban Health : Bulletin of the New York Academy of Medicine 83 (6): 1041–62. https://doi.org/10.1007/s11524-006-9094-x.

Milwaukee Environmental Justice Lab. n.d. "A Community / University Partnership Around Emancipatory Storytelling and Transformative Actions." MKEEJLAB. Accessed February 15, 2021a. https://mkeejlab.weebly.com/.

Milwaukee Environmental Justice Lab. n.d. "Transit History." Climates of Inequality. Accessed February 13, 2021b. https://mkeejlab.weebly.com/transit-history.html.

Milwaukee Water Works. 2021. "Lead and Water." 2021. https://city.milwaukee.gov/water/WaterQuality/LeadandWater.

Miranda, Marie Lynn, Dana C Dolinoy, and M Alicia Overstreet. 2002. "Mapping for Prevention: GIS Models for Directing Childhood Lead Poisoning Prevention Programs." Environmental Health Perspectives 110 (9): 947–53. https://doi.org/10.1289/ehp.02110947.

Mohai, Paul, and Bunyan Bryant. 1992. "Environmental Injustice: Weighing Race and Class as Factors in the Distribution of Environmental Hazards." *University of Colorado Law Review* 63 (4): 13.

Mohai, Paul, and Robin Saha. 2006. "Reassessing Racial and Socioeconomic Disparities in Environmental Justice Research." *Demography* 43 (2): 383–99. https://doi.org/10.1353/dem.2006.0017.

Munck, Victor C. de. 2005. "Units of Analysis." In *Encyclopedia of Social Measurement*, edited by Kimberly Kempf-Leonard, 899–907. New York: Elsevier. https://doi.org/10.1016/B0-12-369398-5/00043-8.

Nelsen, James. 2016. "Franklin Heights." *Encyclopedia of Milwaukee*. Retrieved September 2, 2021 (https://emke.uwm.edu/entry/franklin-heights/).

Oyana, Tonny J., and Florence M. Margai. 2007. "Geographic Analysis of Health Risks of Pediatric Lead Exposure: A Golden Opportunity to Promote Healthy Neighborhoods." Archives of Environmental & Occupational Health 62 (2): 93–104. https://doi.org/10.3200/AEOH.62.2.93-104.

Padilla, Cindy M., Wahida Kihal-Talantikite, Verónica M. Vieira, Philippe Rossello, Geraldine Le Nir, Denis Zmirou-Navier, and Severine Deguen. 2014. "Air Quality and Social Deprivation in Four French Metropolitan Areas—A Localized Spatio-Temporal Environmental Inequality Analysis." Environmental Research 134 (October): 315–24. https://doi.org/10.1016/j.envres.2014.07.017.

Parmenter, Barbara M., and Janet Lau. 2013. "Estimating and Mapping Reliability for American Community Survey Data." Tufts GIS Center.

Potash, Eric, Rayid Ghani, Joe Walsh, Emile Jorgensen, Cortland Lohff, Nik Prachand, and Raed Mansour. 2020. "Validation of a Machine Learning Model to Predict Childhood Lead Poisoning." JAMA Network Open 3 (9): e2012734. https://doi.org/10.1001/jamanetworkopen.2020.12734.

Pulido, Laura. 2000. "Rethinking Environmental Racism: White Privilege and Urban Development in Southern California." Annals of the Association of American Geographers 90 (1): 12–40.

Rabin, Richard. 2008. "The Lead Industry and Lead Water Pipes 'A MODEST CAMPAIGN.'" American Journal of Public Health 98 (9): 1584–92. https://doi.org/10.2105/AJPH.2007.113555.

Raddatz, Liv, and Jeremy Mennis. 2013. "Environmental Justice in Hamburg, Germany." The Professional Geographer 65 (3): 495–511. https://doi.org/10.1080/00330124.2012.700500.

Riva, Michele Augusto, Alessandra Lafranconi, Marco Italo D'Orso, and Giancarlo Cesana. 2012. "Lead Poisoning: Historical Aspects of a Paradigmatic 'Occupational and Environmental Disease.'" Safety and Health at Work 3 (1): 11–16. https://doi.org/10.5491/SHAW.2012.3.1.11.

Rufat, Samuel, Eric Tate, Christopher G. Burton, and Abu Sayeed Maroof. 2015. “Social Vulnerability to Floods: Review of Case Studies and Implications for Measurement.” International Journal of Disaster Risk Reduction 14 (December): 470–86. https://doi.org/10.1016/j.ijdrr.2015.09.013.

Salm Ward, Trina, Naoyo Mori, Timothy Patrick, Mary Madsen, and Ron Cisler. 2010. “Influence of Socioeconomic Factors and Race on Birth Outcomes in Urban Milwaukee.” WMJ : Official Publication of the State Medical Society of Wisconsin 109 (October): 254–60.

Sampson, Robert J., and Alix S. Winter. 2016. “The Racial Ecology of Lead Poisoning: Toxic Inequality in Chicago Neighborhoods, 1995-2013.” Du Bois Review: Social Science Research on Race 13 (2): 261–83. https://doi.org/10.1017/S1742058X16000151.

Schlenker, Thomas L, Renee Baxmann, Peter McAvoy, John Bartkowski, and Amy Murphy. 2001. “Primary Prevention of Childhood Lead Poisoning Through Community Outreach.” Wisconsin Medical Journal 100 (8): 8.

Schwarz, Kirsten, Adam Berland, and Dustin L. Herrmann. 2018. “Green, but Not Just? Rethinking Environmental Justice Indicators in Shrinking Cities.” Sustainable Cities and Society 41 (August): 816–21. https://doi.org/10.1016/j.scs.2018.06.026.

Shen, Shiran Victoria. 2018. “Using Machine Learning to Find Environmentally At-Risk Communities.” SSRN Scholarly Paper ID 3329557. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3329557.

Shrader-Frechette, Kristin. 2002. Environmental Justice: Creating Equality, Reclaiming Democracy. Oxford University Press.

Small, Virginia. 2019. “Environmental Justice Advocacy Growing in Milwaukee.” Shepherd Express. March 5, 2019. https://shepherdexpress.com/api/content/43e8a1ec-3f84-11e9-b6cb-120e7ad5cf50/.

Sorensen, Lucy C., Ashley M. Fox, Heyjie Jung, and Erika G. Martin. 2019. “Lead Exposure and Academic Achievement: Evidence from Childhood Lead Poisoning Prevention Efforts.” Journal of Population Economics 32 (1): 179–218. https://doi.org/10.1007/s00148-018-0707-y.

Swift, Andrew, Lin Liu, and James Uber. 2008. “Reducing MAUP Bias of Correlation Statistics between Water Quality and GI Illness.” Computers, Environment and Urban Systems 32 (2): 134–48. https://doi.org/10.1016/j.compenvurbsys.2008.01.002.

Trotter, Joe William. 1985. Black Milwaukee: The Making of an Industrial Proletariat, 1915-45. University of Illinois Press.

UN-HABITAT, and World Health Organization. 2020. Integrating Health in Urban and Territorial Planning: A Sourcebook. Geneva: UN-HABITAT and World Health Organization.

US EPA, OP. 2014. “Environmental Justice.” Collections and Lists. US EPA. November 3, 2014. https://www.epa.gov/environmentaljustice.

US EPA. 2015. “Environmental Justice Timeline.” Collections and Lists. US EPA. April 15, 2015. https://www.epa.gov/environmentaljustice/environmental-justice-timeline.

US EPA, OW. 2015. “Lead and Copper Rule.” Other Policies and Guidance. October 13, 2015. https://www.epa.gov/dwreginfo/lead-and-copper-rule.

US EPA. 2016. “Basic Information about Lead in Drinking Water.” Overviews and Factsheets. US EPA. February 2, 2016. https://www.epa.gov/ground-water-and-drinking-water/basic-information-about-lead-drinking-water.

US EPA. 2020. “Revised Lead and Copper Rule.” Overviews and Factsheets. December 21, 2020. https://www.epa.gov/ground-water-and-drinking-water/revised-lead-and-copper-rule.

Vila, Peter M, Geoffrey R Swain, Dennis J Baumgardner, Sarah E Halsmer, Patrick L Remington, and Ron A Cisler. 2007. “Health Disparities in Milwaukee by Socioeconomic Status.” WISCONSIN MEDICAL JOURNAL, 9.

Waller, Lance A, Bradley P Carlin, Hong Xia, and Alan E Gelfand. 1997. “Hierarchical Spatio-Temporal Mapping of Disease Rates.” Journal of the American Statistical Association, 12.

Washington, Rodney, and Thomas Welcenbach. 2019. “Meeting Minutes: Transportation, Public Works and Transit Committee.” Milwaukee County. http://milwaukeecounty.granicus.com/player/clip/1855?view_id=2&meta_id=406313&redirect=true.

Weigand, Matthias, Michael Wurm, Stefan Dech, and Hannes Taubenböck. 2019. “Remote Sensing in Environmental Justice Research—A Review.” ISPRS International Journal of Geo-Information 8 (1): 20. https://doi.org/10.3390/ijgi8010020.

WHO. 2019. “Lead Poisoning and Health.” August 23, 2019. https://www.who.int/news-room/fact-sheets/detail/lead-poisoning-and-health.

Wisconsin DNR. n.d. “Lead Service Line (LSL) Replacements | Wisconsin DNR.” Wisconsin Department of Natural Resources. Accessed August 7, 2021. https://dnr.wisconsin.gov/aid/documents/EIF/leadServiceLineFunding.html.

Wolff, Lisa S., Dolores Acevedo-Garcia, S.V. Subramanian, Deanne Weber, and Ichiro Kawachi. 2010. “Subjective Social Status, a New Measure in Health Disparities Research: Do Race/Ethnicity and Choice of Referent Group Matter?” Journal of Health Psychology 15 (4): 560–74. https://doi.org/10.1177/1359105309354345.

Zandbergen, Paul A., and Drew A. Ignizio. 2010. “Comparison of Dasymetric Mapping Techniques for Small-Area Population Estimates.” Cartography and Geographic Information Science 37 (3): 199–214. https://doi.org/10.1559/152304010792194985.