6-2018

Association between PM 2.5 in Minnesota and Influencing Factors: Tree Space Area, Road Pollution, and Rainfall

Junyang Shen
St. Cloud State University, and.friends@hotmail.com

Association between PM 2.5 in Minnesota and Influencing Factors:

Tree Space Area, Road Pollution, and Rainfall

By

Junyang Shen

A Thesis

Submitted to the Graduate Faculty of

St. Cloud State University

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in Geography and Planning

June 2018

Committee Members:
Mikhail Blinnikov, Chairperson
Jeffrey Torguson
Jeffrey Cheng

Abstract

The purpose of this paper is trying to use air monitoring data of Particulate Matter (PM 2.5) from 19 monitoring sites in Minnesota, to determine the correlations between PM 2.5 and the influencing factors, such as road traffic, tree space area, and rainfall. The study will be based on pollutant data which were from Environment Protection Agency (EPA) and Minnesota Pollution Control Agency (MPCA), then through regression analysis and Pearson correlation analysis to determine the correlations of all variables. The correlation analysis results between PM 2.5 concentration and three variables (tree space area, traffic volume, and rainfall) showed that tree space area ratio had a negative, traffic volume had a positive and rainfall had a negative, correlation with PM 2.5 in Minnesota urban. The air traffic volume had a positive correlation with PM 2.5 in airport areas.

In this study, GIS system is a useful tool for geostatistical analysis. It can be used for Normalized Difference Vegetation Index (NDVI) analysis, raster data geoprocessing, and kriging spatial analysis.

Acknowledgments

I would like to thank my committee members who supported my efforts in writing this thesis.

To my chair, Dr. Mikhail Blinnikov,

To Dr. Jeffrey Torguson,

To Dr. Jeffrey Cheng,

These professors have been instrumental in assisting me throughout my academic study at St. Cloud State University (SCSU). I would also like to extend my gratitude to the entire SCSU Geography Department, especially to Dr. Benjamin Richason, who guided me to use the Normalized Difference Vegetation Index (NDVI).

In addition to my academic supporting, I would like to add special recognition to my family and friends, who have supported me at all stages of my studies at SCSU.

Table of Contents

List of Tables

Table Page

List of Figures

Chapter I: Introduction

Background

Air pollution is a major problem worldwide and is having a seriously deleterious effect on human health. The common manifestations of air pollution effects include hard breathing and aggravation of existing respiratory diseases (Burnett et al., 2000). Even low levels of air pollution have bad influences on human health when people stay in this environment over long periods of time (Olmo et al., 2011). A high number of people are being exposed to low levels of air pollution; it will still cause a relatively high risk to public health.

At the same time, McCarty, Hafner, and Montzka (2006) studied the background concentrations of 18 air toxins in North America, and found that the background concentration of benzene, chloroform, formaldehyde, and fine particulate was higher than cancer benchmark values in the region, and these four also were higher than the cancer benchmarks recommended by the Environmental Protection Agency (EPA). Therefore, even in low traffic areas, the pollution from man-made sources could also raise the level of total air pollution concentration limits and cause hazardous illnesses to human health.

Based on the kinds of vehicle pollution exhaust on the road, the EPA and MPCA (Minnesota Pollution Control Agency) have rigid environmental-based criteria for some pollutants, such as ozone (O_3), particulate matter (PM), carbon monoxide (CO), and nitrogen oxides (NO_X). EPA and MPCA both use National Ambient Air Quality Standards (NAAQS) (Environmental Protection Agency, n.d.) as the standards. The purpose of these criteria is to avoid overexposure to humans, keep the high air quality level and protect people's health. This is because some results from medical studies have found associations between traffic-related pollution and health. For example, long-term exposure of the human body to combustion-related fine particulate air pollution has become a significant risk factor for lung cancer or lung disease mortality (Cordioli et al., 2014).

It is because air quality is so closely related to people's lives that a lot of scientists and country governments work on solving the air pollution problem. After the Great Smog of London in 1952, the main solution to the air pollution is reducing the sources of pollutant. It means reducing pollutant emissions from the source. With science and technology developing, many heavy air polluting industries have disappeared. But another pollution source is still here, and it is the vehicles (Adedeji, Oluwafunmilayo, & Oluwaseun, 2016). Therefore, when source control is not feasible, how to reduce pollution become important. Scientists started to focus on the green vegetation as one way of reducing non-point pollution.

Some studies have suggested that green vegetation could absorb or settle air pollutants, such as gas pollutants and particulate matters. The particulate matter pollution of urban areas is currently the major problem in air quality management (Huang et al., 2014). A lot of Chinese and developing countries’ news show the fact. Especially in China, atmospheric particulate matter pollution has affected people's daily lives (Chen et al., 2016). Although the USA is not commonly experiencing a lot of particulate matter pollution, it has a good environment to study the relationship between particulate matter pollution and traffic and vegetation without industrial pollutant effect.

In the study of the urban haze (serious particulate air pollution), PM 2.5 is the major part. PM 2.5 means fine particles with a diameter of 2.5μm or less. They usually have the complex chemical composition, small particle size, and long-term stay in the atmosphere (Seo et al., 2016). Those properties make the PM 2.5 more stable in air, so they can impact the environment significantly. PM 2.5 has the abilities to reduce visibility and changing cloud formation processes. PM 2.5 contains harmful compounds such as acids and heavy metals. Those can directly impair human health - especially respiratory functions (Grigg, 2011). Because PM 2.5 is directly related to human health, research on PM 2.5 reduction is necessary. Ji et al. (2013) pointed that green plants were natural enemies of PM 2.5 fine particles.

The study area here is in the state of Minnesota. The total area of Minnesota is approximately 225,163 km^2. According to the United States Census Bureau, the estimated population of Minnesota is 5,519,952 in 2016. The main source (80 %) of PM 2.5 in Minnesota monitoring site areas are vehicle emissions and road dust, which is also caused by vehicles. According to the Bureau of Transportation Statistics (2016), there were a total of 4.82 million registered vehicles in Minnesota. Most PM 2.5 monitoring sites are away from industrial pollution sources. Therefore, the study can focus on the vehicle pollutant in Minnesota.

This study was undertaken to analyze the green space ratio, traffic flow, and their impact on PM 2.5 concentrations in Minnesota using ArcGIS as a tool. The following analysis discusses their relationship and how their impact on PM 2.5 can be quantified.

Purpose

As people have known about the health and environmental risk of air pollution for a long time, a large number of air pollution models have been developed. The use of air pollution models is common in studies of public health. However, all have one major limitation: the use of fixed measuring stations or at EPA and MPCA established locations that approximate only part of the total daily information. If the subject under study spends most of his/her time in areas that are far from the point where the pollution is detected, the air pollution data are neither available nor reliable.

The aim of this paper is to address this limitation and study the relationship between urban vegetation, road pollution, and particulate contamination in Minnesota and determine if a model of PM 2.5 concentration and tree space ratio and road pollution can be produced. This model should be able to provide a snapshot of the level of pollution on a daily basis and at any place in Minnesota without relying on existing pollution monitoring stations.

Objectives

To address research aim, the specific research objectives are presented as follows:

(1) To gather the data of the particulate concentration in selected counties or cities;

(2) To calculate or gather the data of the tree space ratio of the selected county or city;

(3) To calculate or gather the data about traffic flow in selected counties, cities or selected area;

(4) To analyze the data and figure out the association;

(5) To apply GIS for the visualization of the association's effect on the particulate concentration in Minnesota.

Chapter II: Literature Review

The major particulate matter pollution is resulting from the burning of fossil fuels by transportation and industrial sources. The effect tends to appear in densely populated metropolitan areas in developing countries and industrial areas in both developed and developing countries. For example, developing countries: some cities in China and Mongolia (World Bank, 2009); developed counties: coal mining areas in Australia (Milman, 2015). Lee, Olsen, Wuebbles, and Youn (2013) studied the PM 2.5 distribution in Europe and found the airport average was a little bit higher than the surrounding area without industrial pollution. And they thought aviation emissions would be sources of PM 2.5.

The Relationship between Vegetation and PM

In these studies, the major particulate matter is PM 10 and PM 2.5. PM 10 is fine particles with a diameter of 10μm or less; PM 2.5 is fine particles with a diameter of 2.5μm or less. There are several types of association between particulate matter and vegetation.

Vegetation structure effect on PM. Vos, Maiheu, Vankerkom, and Janssen (2013) studied the impact of the urban street vegetation structure on air pollution. In their study, they use ENVI-met model simulation urban structure and assessed the effectiveness of 19 different real-life urban street vegetation designs. This study did not show that urban street vegetation can reduce air pollution. Some models even show that urban vegetation structure actually impacts the flow of urban air, and makes air pollution worse. Therefore, they proposed vegetation structure may have a negative impact on the environment and build the model to point out vegetation species that can negatively affect air quality: “Low green barriers do have a negative effect on air quality at the leeward side of the canyon. Higher hedges do not increase concentrations on the windward side of the canyon.” The results also indicate that if decreasing tall vegetation such as having fewer trees, the pollutant concentrations would be lower. In conclusion, the distinct air quality improvement model is only one design. In almost all other designs, roadside urban trees have a detrimental effect on the local air quality. The value of the deposition speed is hardly affecting the air quality impact of trees in a street canyon. This shows that for urban vegetation along inner city roads, it is mainly the aerodynamic effect that determines the effect on the air quality and not the pollutant removal capacity (Gromke, 2011).

Vegetation effect on PM. According to Setälä et al. (2012), in northern conditions (Helsinki in Finland), there was no evidence that urban vegetation can remove significant amounts of common gaseous pollution (NO2 and VOCs). On the other hand, two of five common air contaminants or airborne particles show the passive effect (20% - 40% PM 2.5 and PM 10 removal) in tree-covered areas. But other parts of the sample don't show a significant difference between open area and tree-covered area. Alaviippola and Pietarila (2011) have the similar idea, and they also show air quality in tree-covered areas improved a little (15% - 20%) when compared to treeless areas. This result did not show a significant difference due to seasonal change. Owing to lack of enough measure time, Setälä et al. (2012) did not have enough results to show the effective influence of vegetation structure on air pollutant concentrations and influences of season and local wind conditions on pollution concentrations. They focused on the deciduous trees. The measurements were made during summer and winter; therefore, the big problem was the seasonality effect. However, in their study, they did not manage to analyze the difference between seasons.

Also, the grasses are affected by the season. According to Weber, Kowarik, and Säumel (2014), vegetation can well absorb particulate matter when it is close to the pollutant source and has large contact areas. Hence, roadside vegetation is expected to have a considerable effect on reducing environmental particulate pollution. They agree that “urban roadside vegetation consists of a variety of vegetative structures beyond trees, including lawns and other types of herbaceous vegetation which could contribute to the immobilization of PM (Litschke & Kuttler, 2008).” However, they focus on herbaceous roadside vegetation and assess the role of species traits. The authors analyzed randomly sampled wild herbaceous plant leaves on three sites in Berlin with low, medium, and high traffic densities. Site selection was based on results of vehicle counts by local authorities. In total, they sampled 16 species of plants and used a microscope with magnification x200 to determine number and size of particles and type of particles. The authors analyzed 16 kinds of plants. These plants have different sizes, different shapes and leaf areas, and different leaf roughness. They also analyzed the difference between a single plant and mixing plants for PM deposition. They found that with higher traffic density, there was more deposited particulate matter. The plants, which have rough leaves, can adsorb more PM. Although they point out that many plants do not show differences in PM deposition, mixed plants can increase the immobilization of a wide range of PM. Furthermore, tall-growing herbs with leaves can accumulate PM more significantly than low-growing species. They also found planting trees and grasses together can enhance air filtration capacity of roadside vegetation beyond trees. However, if there is only grass, the PM value is not significantly different.

In addition, there is another type of urban vegetation, like evergreen trees. They are not affected by the seasonality of climate (Freer-Smith, El-Khatib, & Taylor, 2004). The evergreen species’ effects on PM accumulation have been studied (Freer-Smith, Beckett, & Taylor, 2005; Cavanagh, Zawar-Reza, & Wilson, 2009; Sæbø et al., 2012). However, in most of these studies, the time was limited and usually near the end of the growing season. Przybysz, Sæbø, Hanslin, and Gawroński (2014) studied three evergreen plants, which were Taxus baccata, Pinus sylvestris, and Hedera helix. They pointed out that evergreen plants may have better ability to deal with the problem because their leaves will not fall down in winter. They also agreed that deciduous plants had better results when the leaves were present because they had bigger leaf surfaces. At the same time, Armbrust (1986) pointed out that the rainfall can wash off 30% - 40% PM pollutant. It can help plants to have more area to absorb PM and it does not form secondary pollution.

PM effect on vegetation. There are a lot of air-pollutants coming from road traffic in the urban environment, and many of these pollutants can be absorbed by urban vegetation. However, at the same time, these pollutants can also affect the plant's growth, phenology and leaf surface characteristics. Przybysz et al. (2014) and Honour et al. (2008) found that particulate matter can affect vegetation negatively. For instance, Honour et al. (2008) built fumigation system to simulate the real environment exhaust, such as vehicles and industries. In the experiment, they monitored the reproduction and growth of plants and found PM 2.5 would increase plants leaves aging and death when the leaves surface was covered too much PM 2.5 because PM 2.5 would hinder the respiration of leaves. However, they found that particulate matter is water-soluble, and the bad effect will disappear after rainfall wash off the pollutant.

Temperature Effect on PM 2.5

Lai (2018) studied the influence of the urban heat island (UHI) effect on particulate matter (PM; including PM 2.5 and PM 10). In his study, he found that PM accumulation was associated with anthropogenic emissions. For example, there was more evident during the peak period of anthropogenic emissions than during the trough period of anthropogenic emissions. And some UHI area, which is caused by increasing temperature, will have the little PM value decrease than surrounding area. However, he did not find an obvious direct relationship between the temperature or heat island effect and PM. Wang and Ogawa (2015) found a negative correlation between temperature and PM 2.5 value. However, Chen et al. (2017) found a Positive association between temperature and PM 2.5 value.

Pollution Model

Model for PM measurement. Jeanjean, Leigh, and Monks (2016) suggested that green infrastructure can reduce PM2.5 traffic emissions on a city scale, by a combination of dispersion by trees and deposition on buildings, trees, and grass. They used a CFD model to simulations of PM 2.5 concentrations and validated the effect. At the same time, Jeanjean, Leigh, and Monks (2016) found that wind speed can affect the reduction in PM 2.5 level's effect. When wind speed is over 4.6 m/s, PM 2.5 will show an increasing trend in road area, and a decreasing trend in the open area. When wind speed is around 1.0 m/s, PM 2.5 will show no significant change. However, Chen et al. (2017) found a negative correlation between wind speed and PM 2.5 value. Wang and Ogawa (2015) did find the significant relation between wind and PM 2.5 in Japan. Based on Armbrust (1986), Beckett, Freer-Smith, and Taylor (2000) and Van Heerden, Krüger, and Kilbourn Louw (2007) figured out both the effects of tree aerodynamics and the deposition capabilities of trees and grass. "The results display that the aerodynamic dispersive effect of trees on PM 2.5 concentrations results in a 9.0% reduction" (Brantley, Hagler, Baldauf, & Deshmukh, 2014). "In contrast, a decrease of PM 2.5, by 2.8% owing to deposition on trees (11.8 ton per year) and 0.6% owing to deposition on grass (2.5 ton per year), was also observed (Jeanjean, Leigh, & Monks, 2016)". In the CFD model, the air flow is a steady flow. Therefore, the measurement of PM 2.5 will think about time-dependent effects such as fluctuations in wind speed or direction. They also mentioned that PM 2.5 will not change when the atmosphere height is over 500 meters.

GIS model for pollutants (PM et al.)

First model. Seo et al. (2016) studied associations between the level of PM 10 and allergic diseases to evaluate whether PM 10 have the impact on allergic diseases or not. They used GIS to visualize the relationship and built a geographically weighted regression model (GWR). In their measurement, there are not enough monitoring stations for 424 sub-districts, hence they used spatial interpolation technique to deal with air data, and keep them correspond with the disease prevalence data. In data analysis, they used multiple regression models to assess the impact of PM 10 on the prevalence of each allergic disease. At the same, they also compared them with the average values of temperature, wind speed, and precipitation to make sure all effect factors existed. Upon dealing with allergic disease prevalence, they use an ordinary least square (OLS) regression analysis to compare with pollutants data. In spatial heterogeneity analysis, they performed geographically weighted regression (GWR) analysis. According to the statistics of data and spatial correlations, they found that “only the prevalence of atopic dermatitis was significantly associated with the level of PM 10 interpolated at the sub-district level”, but “the prevalence of asthma [and] allergic rhinitis had the relatively small associations of air pollution (Seo et al., 2016).”

Second model. Tashayo and Alimohammadi (2016) set 17 air pollution monitoring stations and use hierarchical fuzzy inference system (HFIS) to develop optimized their own system based on the pollutants test data (daily PM 2.5, annual mean PM 2.5, daily nitrogen dioxide (NO_2), and annual mean NO_2) and the modeling parameters (transportation, land use, and meteorological conditions). According to the authors, the article uses a three steps procedure to build HFIS. The first step is “Preprocessing techniques are used for sample extraction” (Tashayo & Alimohammadi, 2016). They divide the data into two groups: one is meteorological data used for probabilistic preprocessing, and another is used for land use and traffic data used to geographic analysis preprocessing in GIS. Around the air pollution monitoring stations, they use buffer analysis to prepare the traffic and land use data. Each buffer was the PM 2.5 and NO_2 emission sources. The second step is initializing HFIS. They used HFIS with hybrid distribution to overcome exponential rule explosion and a loss of accuracy. Firstly: to set domains of variables as fuzzy regions. Secondly: to use data pairs to generate fuzzy rules. Thirdly, to assign each rule a degree. And then, to select the highest degree rule based on the same premises and different conclusions. The third step is that "A multiple objective particle swarm optimization (MOPSO) is applied for simultaneous optimization of the accuracy and complexity of the system built in step 2" (Tashayo & Alimohammadi, 2016). In this step, they use the "Wang-Mendel" method to get the initial knowledge base. They needed not only the number of rules and root-mean-square error but also specific objective function to calculate the data. According to the measurement and analysis results, their model indicates the capability of modeling non-linear relationships between influence parameters and air pollution. This study built an optimized HFIS to give us "a flexible framework for combining qualitative, quantitative, and non-homogeneous data" (Tashayo & Alimohammadi, 2016).

Third model. Holford et al. (2010) built a framework by using generalized linear models (GLMs) and generalized linear mixed models (GLMMs) to define the association between traffic densities and health response variable. Through Geographic Information Systems (GIS), they show the model with a pattern of highways, and the model also includes pollutant levels, wind direction, meteorology, and topography. This model is similar to a distance-weighted approach. They separate their pollutants sources into three types. The first is point sources of exposure. They regard the pollutant source as a single point and plot its longitude/latitude coordinates onto a plane in GIS. The coordinate system they used is UTM or universal transverse Mercator projection. The second is line sources of exposure. They considered traffic-related air pollution is from a locus of points on lines representing roadways. It shows a line for highway segments. The third is area sources of exposure. It means that "Area sources of environmental exposure can arise from the spreading of a potential pollutant over a region. (Holford et al., 2010)" In order to estimate their dispersion function in general linear model (GLM), they introduce the regresses, specify function type, and introduce the intensity and other covariates and then transfer the data into GIS. In their study, they used two examples to show and explain their model. One is estimating the distribution of traffic-related pollution, and another is effects of traffic on respiratory symptoms. Based on these examples, they have enough supporting data, information, analysis, and discussion to effectively support the modeling. They successfully built a model to estimate effects of exposure to traffic-related pollution using publicly available data.

Fourth model. Cordioli et al. (2014) used self-reported and GIS-derived residential area, which is exposed to environmental pollution, to set case-control and study lung. In their research, there were 649 objects of inquiry from the Province of Modena. They give the information about their residential history and perceived exposure. And then, the authors used GIS to evaluate land use patterns for each residence. Most of the residences were proximal to major roads and exposed to industrial pollution. The cancer data were from the Mirandola Health District, 2009 - 2010. In GIS, the coordinate system is UTM 32, datum ED 50. The GIS data visualization is working on ArcGIS (v.9.3), and the post-elaboration is used by Stata (SE v.12 and R v.2.13.1). For the road traffic, it separated three parts: major, secondary, and minor. The statistical result shows that most residences reported being on a "quiet road". Land use is based on the neighborhood with three concentric buffers within a radius of 250, 500 and 1000 m. The Chisquare test is highly significant. It shows that the zone of residence is predominantly: rural and residential. The pollution data (total suspended particles and volatile organic compounds) were from the Regional (Italian) Environmental Protection Agency. The industrial has the higher pollution. According to the authors, “the results of their work showed moderate to good agreements between GIS-derived proxies of exposure to environmental risks and self-reported evaluation from a questionnaire. Nevertheless, it is important to provide evidence of the validity of these new tools with respect to ‘classical’ epidemiological methods (Cordioli et al., 2014).”

Fifth model. Chen et al. (2012) used a two-stage approach to build a spatiotemporal model. Before building the model, they assumed that pollutant levels of a metropolitan area, spatial and temporal variations in concentrations may be decomposed separately. Based on this assumption, they get the following two stages. Stage one is modeling for daily mean temporal trend. Use their formula to get the average of the K monitoring station. Stage two is using spatial interpolation for land-use regression. Based on their model, they get the average daily PM 2.5 concentrations are 28.35$\mu g/m^3$. Then they use simple LUR and kriging method for site-specific annual-average PM 2.5 predictions. They compared for their performance in daily PM 2.5 concentration predictions. After analysis and discussion of the result, they prove that “the proposed method successfully divides the dates of the study period into approximately temporal-invariant homogeneous subgroups with separate LUR models for daily PM 2.5 concentration predictions, which had in general better performance in comparison with a single LUR model (Chen et al., 2012).”

Sixth model. Pantaleoni (2013) has a useful model which could be used to study for the association between traffic flow and PM 2.5. In the study, the pollutant is carbon monoxide

(CO); however, the method to build a model with using GIS to compare pollutants to traffic flow can be used for any other pollutant. The following is the method:

The study location is selected at Nashville, Knoxville, Memphis, and Chattanooga in Tennessee. The reason they choose Carbon Monoxide as the pollutants test data is that 60% of emission sources are vehicles in the USA. Using Baker's method, she calculated the concentrations of CO. Also, she chose climatic data such as wind speed and temperature as dependent data. The data were from the National Climatic Data Center (2000 - 2008) for 10 stations. Traffic data are the annual average daily traffic (AADT) data, which were from the Tennessee Department of Transportation (TNDOT). These data were collected at 5,710 locations for each year, from 2000 to 2008 about 3,288 days. Also, the traffic count data rely on the daily traffic flow patterns for seven days per week. Software: ArcGIS. Use ArcGIS software to convert the coordinates to points, and plot on a Tennessee map. The validated data: daily CO records by Environment Protection Agency (EPA) sites from 2000 to 2008. There were 10 stations separated at study location. In this paper, the author "build the model on traffic flow data and combine it with climatic factors using a GIS interface" to "present an air pollution model capable of determining the level of carbon monoxide on a day to day basis" (Pantaleoni, 2013). Through linear regressions, Pearson correlation coefficients, and interval analysis, she got a relationship between elements. After analysis data, she uses ArcGIS to map the air pollution. In this research, she "builds the model on traffic flow data and combines it with climatic factors using a GIS interface" to "present an air pollution model capable of determining the level of carbon monoxide on a day to day basis." (Pantaleoni, 2013)

GIS Application in Air Pollution Study

GIS technology is often used in air pollution research. For instance, GIS technology is suitable for the dynamic monitoring and analysis of the atmospheric environment, the use of GIS technology and database management techniques to collect and sort out the main pollutants, the scope of the proliferation of pollutants, the surrounding terrain and the presence of air pollution risks of plant enterprises and their location information, the establishment of a geographic information database, and then obtain the concentration distribution of pollutants in the atmosphere by GIS spatial analysis and data display, and finally get the spatial distribution of pollutants and the situation of exceeding the standard. Seo et al. (2016) used GIS to analyze spatial correlations, visualize the relationship and built the normal model between PM10 and allergic diseases.

Sargazi et al. (2011) focused on the application of GIS for the modeling of the spatial distribution of air pollutants. In their research, the study area was around the city of Tehran in the northern part of Iran with the total area of 18,909 square kilometers. In 2008, they set 16 air pollution monitoring stations in Tehran to measure the hourly CO concentration. Then they average the hourly data of a day as the mean daily CO concentration. At the same time, they measure the 3-hourly wind speed and direction data of the five meteorological stations to decrease the error. They used four geostatistical methods to analyze the data in GIS. The four geostatistical methods were inverse distance weighting (IDW), thin plate splines (TPS), kriging and cokriging methods. Inverse distance weighting method estimates the value of an attribute at non sampled points. Thin plate splines calculate the observed value of the studied parameter. Using kriging o inset the value of a random field at an unobserved location from observations of its value at nearby locations. Cokriging is an extension of kriging. They also build their algorithm to solve their problem. They put the 3-hourly wind speed and direction data and daily into ArcGIS software, and use IDW, TPS, kriging, and cokriging to implement all data. Then they determine the optimum values of the parameters of interpolation methods through cross-validation technique. After the optimum values are determined, they determine the best method for hourly and daily interpolation. Then show the spatial distribution maps of carbon monoxide concentration in Tehran. Finally, the wind speed and direction data layer are overlaid on the spatial distribution map of carbon monoxide concentration in the GIS environment.

NDVI and Green Space Ratio

According to Oštir (2006), the remote sensing is the detection, recognition, or evaluation of objects by means of distant sensing or recording devices. From the history of remote sensing developed, digital remote sensing developed rapidly from aerial photography and photo interpretation in the past 30 years. Information, which extracted visually from remote sensing, is widely used in forestry, urban planning and Vegetation Disease Research (Franklin, 2001). Remote sensing can give the information which is to prove the importance and complexity of forest preservation and sustainable forest management (Pagiola, Landell-Mills, & Bishop, 2002). One of the methods is to evaluate a normalized difference vegetation index (NDVI) in local forest management (Weier & Herring, 2000).

"NDVI is actually a simple graphic indicator that can be used to analyze remote sensing measurements, whether the target observed contains live green vegetation or not" (Jovanović, Milanović, & Zorn, 2018). NDVI could simply and quickly identify vegetated areas and their condition, and it is the best-known and most-used index for detecting live green plant canopies by using multispectral remote sensing data (Jovanović, Milanović, & Zorn, 2018). NDVI allowed comparisons between images acquired at different times.

According to Kriegler, Malila, Nalepka, and Richardson (1969), they were the earliest persons, who formulated the normalized difference spectral index, and the NDVI is calculated from these individual measurements as follows:

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{Red}}{\mathrm{NIR} + \mathrm{Red}}$$

NIR is the color infrared band of imagery.

Red is the red band of color imagery.

In the formula, red stands for the spectral reflectance measurements acquired in the red (visible) regions, and NIR stands for the spectral reflectance measurements acquired in the near-infrared regions. By design, the NDVI itself thus varies between -1.0 and +1.0. It should be noted that NDVI is functionally, but not linearly, equivalent to the simple infrared/red ratio (NIR/VIS). The advantage of NDVI over a simple infrared/red ratio is therefore generally limited to any possible linearity of its functional relationship with vegetation properties. In general, the urban area NDVI is from 0 to 0.3 and with vegetation the range is from 0.3 to 1.

According to Jovanović, Milanović, and Zorn (2018), NDVI of dense vegetation canopy tends to have positive values which are from 0.3 to 0.8. In detail, water, clouds, and snowfields show as negative values of this index or very low positive. Soils generally tend to generate rather small positive NDVI values from 0.1 to 0.2. When soils and sands are dry or rocky, they tend to have very low values of NDVI, less than 0.1. Shrub and grasslands have values from 0.2 to 0.3. Temperate and tropical rainforests generally have higher values from 0.6 to 0.8. In their research, the difference between NDVI and the official forest area estimates is from 0.2 % to 4.7%. When compared with official forest area estimates, the NDVI results show a mere +0.12 km2 (+0.2%) difference. In their opinion, the ± 5% margin of error is allowed for this method, “but they also allow room for further analysis and investigation (Jovanović, Milanović, & Zorn, 2018).” The reason is that the aerial photos were generally taken during spring and summer time; however, official forest area estimates are made at the end of the year. By contrast, NDVI values may be expected to be higher than the real. At the same time, they compared between NDVI and CORINE land cover classes. They found the NDVI data were more similar to the real forest area estimates.

Chapter III: Methodology

Study Area

The study area for this project is the state of Minnesota, located in the north of the United States of America (Figure 1). The state is located between 43 - 50°N latitudes and 89 - 98°W longitudes. The total area of Minnesota is approximately 225,163 km^2. According to Minnesota Pollution Control Agency (MPCA), there are 24 air monitoring sites located in 19 cities of MN in 2016. Among them, Minneapolis and Duluth have two monitoring points each, while St Paul has three monitoring sites. All monitoring sites belong to Minnesota Pollution Control Agency. At each site, there were no surrounding buildings over 20 meters higher than the monitors, and there was less industrial facilities pollutant effect to ensure getting an effective relationship between PM concentrate and traffic condition. In Minnesota, all selected sites do not have much industrial pollution surrounding them. The tables 1 and 2 listed the details for each monitoring site. According to the coordinates from MPCA 2018 Annual Air Monitoring Network Plan, the monitoring site locations can be plotted in Figure 1.

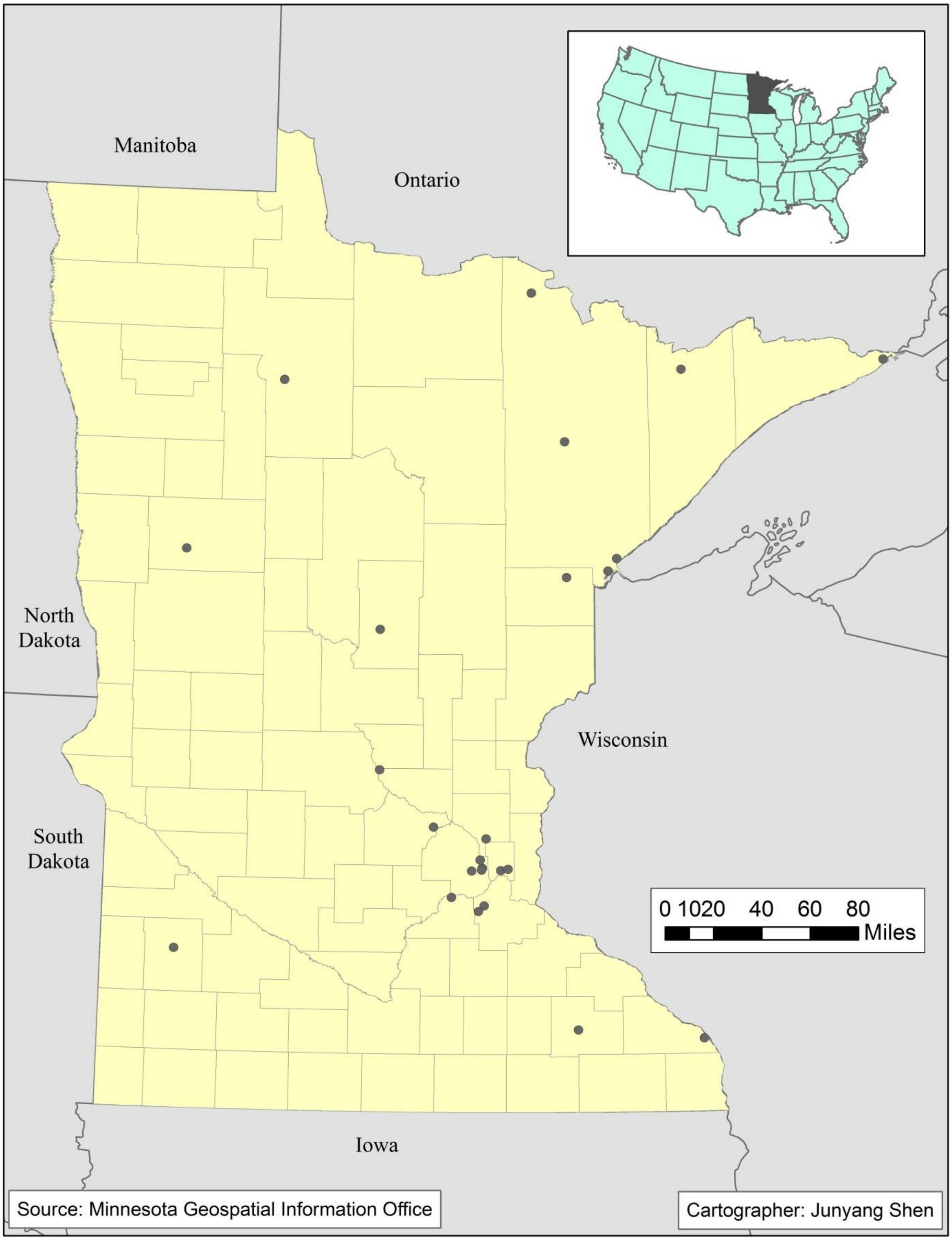

Figure 1: Monitoring sites and a location map for Minnesota within the United States

Table 1

Monitoring site locations and ID

County	City	MPCA Site ID	Latitude	Longitude	Setting
ANOKA	Blaine	6010	45.1407	-93.2220	Suburban
BECKER	Detroit Lakes	2013	46.8499	-95.8463	Rural
BELTRAMI	Red Lake	2304	47.8782	-95.0292	Rural
CARLTON	Cloquet	7417	46.1737	-92.5117	Rural
COOK	Grand Portage	7810	47.9701	-89.6910	Rural
CROW WING	Brainerd	3204	46.3921	-94.1444	Rural
DAKOTA	Apple Valley	470	44.7387	-93.2373	Suburban
	Lakeville	480	44.7061	-93.2858	Suburban
HENNEPIN	St. Louis Park	250	44.9481	-93.3429	Suburban
	Minneapolis	962	44.9652	-93.2548	Urban
	Minneapolis	963	44.9535	-93.2583	Urban
LAKE	Ely	7001	47.9466	-91.4956	Rural
LYON	Marshall	4210	44.4559	-95.8363	Rural
OLMSTED	Rochester	5008	43.9949	-92.4504	Suburban
RAMSEY	Saint Paul	868	44.9507	-93.0985	Urban
	Saint Paul	871	44.9593	-93.0359	Urban
ST LOUIS	International Falls	NA	48.4128	-92.8292	NP
	Virginia	1300	47.5212	-92.5363	Urban
	Duluth	7550	46.8182	-92.0894	Suburban
	Duluth	7554	46.7437	-92.1660	Suburban
SCOTT	Shakopee	505	44.7894	-93.5125	Suburban
SHERBURNE	Saint Cloud	3052	45.5497	-94.1335	Suburban
WINONA	Winona	NA	43.9373	-91.4052	Rural
WRIGHT	St. Michael	3201	45.2092	-93.6690	Suburban

NP is National Park.

Table 2

Monitoring Sites (for PM 2.5) in Minnesota, sampling frequency, and duration

MPCA Site ID	Collection Frequency	Measurement Scale	Monitor Objective Type	Sample Duration
250	Every 3rd Day	Neighborhood	Population Exposure	24 hours
470	Every Day	Neighborhood	Population Exposure	1 hour
	Every 3rd Day	Neighborhood	Population Exposure	24 hours
480	Every Day	Middle Scale	Source Oriented	1 hour
505	Every 3rd Day	Neighborhood	Population Exposure	24 hours
868	Every 3rd Day	Neighborhood	Population Exposure	24 hours
871	Every Day	Neighborhood	Population Exposure	1 hour
	Every Day	Neighborhood	Population Exposure	1 hour
	Every 3rd Day	Neighborhood	Population Exposure	24 hours
	Every 6th Day	Neighborhood	Population Exposure	24 hours
962	Every Day	Middle Scale	Source Oriented	1 hour
963	Every Day	Neighborhood	Population Exposure	1 hour
	Every 3rd Day	Neighborhood	Population Exposure	24 hours
1300	Every Day	Neighborhood	Population Exposure	1 hour
2013	Every Day	Urban Scale	Population Exposure	1 hour
2304	Every Day	Neighborhood	Population Exposure	1 hour
3052	Every Day	Neighborhood	Population Exposure	1 hour
3201	Every Day	Neighborhood	Population Exposure	1 hour
3204	Every Day	Urban Scale	Population Exposure	1 hour
4210	Every Day	Urban Scale	Population Exposure/ Regional Transport	1 hour
5008	Every Day	Neighborhood	Population Exposure	1 hour
	Every 3rd Day	Neighborhood	Population Exposure	24 hours
6010	Every Day	Urban Scale	Population Exposure	1 hour
	Every 3rd Day	Urban Scale	Population Exposure	24 hours
7001	Every Day	Regional	General / Background	1 hour
7417	Every Day	Neighborhood	Population Exposure	1 hour
7550	Every 3rd Day	Neighborhood	Population Exposure	24 hours
	Every 6th Day	Neighborhood	Population Exposure	24 hours
7554	Every Day	Neighborhood	Population Exposure	1 hour
	Every 3rd Day	Neighborhood	Population Exposure	24 hours
7810	Every Day	Neighborhood	Population Exposure	1 hour

Note: the measurement scale information is explained below. Middle Scale (100 - 1,000 m) defines the concentration typical of areas up to several city blocks in size, with dimensions ranging from about 100 to 1,000 meters. Neighborhood Scale (1 - 4 km) defines concentrations within some extended area of the city that has relatively uniform land use with dimensions in the one to four kilometers range. Generally, these stations represent areas with moderate to high population densities. Urban Scale (4 - 50 km) defines the overall, citywide conditions with dimensions on the order of four to 50 kilometers. This scale represents conditions over an entire metropolitan area and is useful in assessing city-wide trends in air quality. Regional Scale/Background (50 - 1,000 km) usually represents a rural area of reasonably homogeneous geography and extends from tens to hundreds of kilometers.

Data

The PM 2.5 daily data were from the United States Environmental Protection Agency (EPA) and the PM 2.5 monitoring sites 2016 information was from Minnesota Pollution Control Agency (MPCA). The Minnesota daily and annual summaries in MN come directly from the AQS (Air Quality System) Data Mart which contains ambient air pollution data collected by MPCA and reported by EPA and MPCA. AQS is updated practically every day as reporting agencies have data ready to submit. The data from AQS will be output in to ".csv" format, and they can be downloaded in https://www.epa.gov/outdoor-air-quality-data/download-daily-data. The study is using data on PM 2.5. There are three reasons why I chose this pollutant. First, the main source of PM 2.5 in Minnesota is vehicle emissions. Second, PM 2.5 is not only a health hazard, but it also would increase plants' aging and death (Honour et al., 2008). Third, the association is related to traffic count data. According to the Bureau of Transportation Statistics, there were a total of 4.82 million registered vehicles in Minnesota in 2016. For PM 2.5; the 24 sites complete daily data 2016 will be used, but there are some daily data missing. Generally, they have 330 data points per year.

The traffic data come from Minnesota Department of Transportation (MNDOT). The data are from 2016. The following methods were used to gather the traffic data by MNDOT: "More than 2600 total Short Duration Counts sites that collect traffic volume; make up the majority of all count locations, ATR (Continuous) devices with loops in the pavement that collect traffic volume and sometimes vehicle classification and/or speed data. More than 15 weigh in Motion System, WIM (Continuous) sites that collect vehicle weight, type, speed, and volume." The

Figure 2 and Figure 3 display all traffic segments and traffic counting sites. However, there were only about 9000 traffic segment and 8000 traffic counting sites used in 2016 and around each 1000 traffic segments and counting sites were used in the selected area.

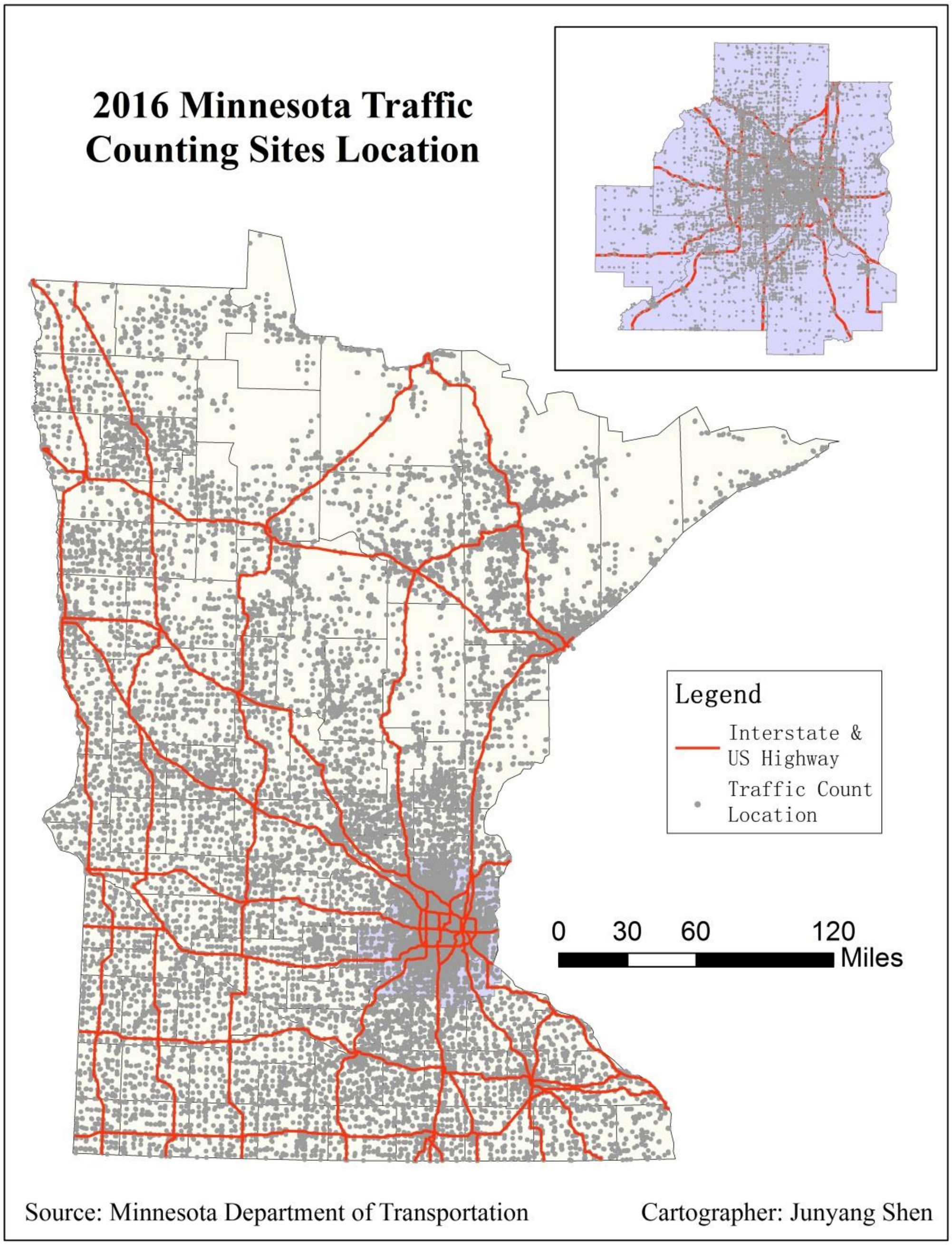

Figure 2: 2016 Minnesota traffic counting sites

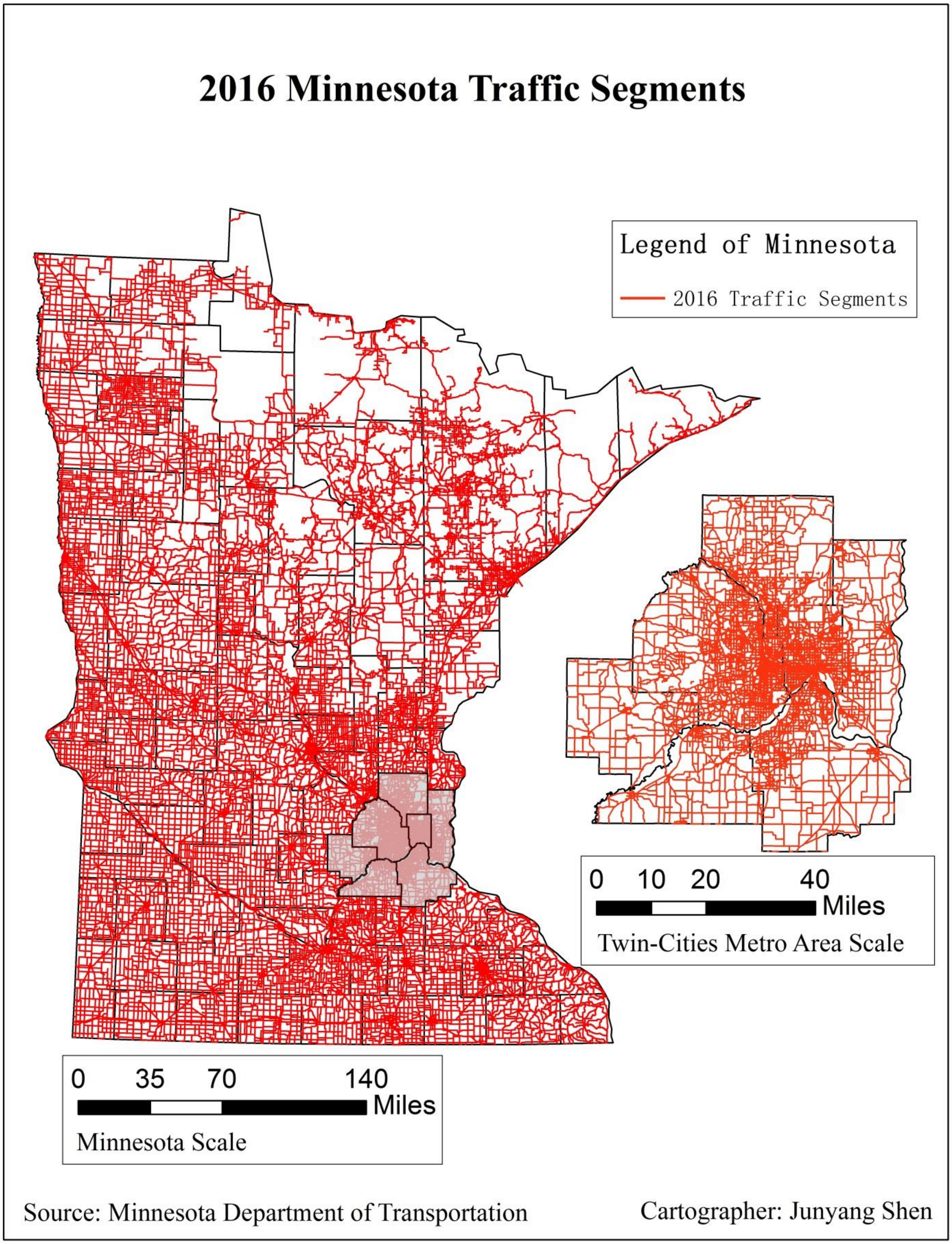

Figure 3: 2016 Minnesota traffic segments

Rainfall data is from Minnesota Department of Natural Resources (MNDNR), and the time frame of data is 2016. All their daily climate data are collected and reported by National Weather Service Reporting Stations.

One of tree area ratio data is from Multi-Resolution Land Characteristics (MRLC) consortium which belongs to Earth Resources Observation and Science (EROS) Center. They have NLCD 2011 Land Cover data and NLCD 2011 USFS Tree Canopy analytical data. I use them to calculate the tree area ratio. From 2011 to 2016, land use may have a lot of changing, therefore this data it is not a perfect match. The best way is to use 4 bands orthorectified aerial image to calculate NDVI, and then to analyze the tree area. The useful aerial images are from 2017, and they are more closed to 2016 than 2011. There are two selected areas NAIP (National Agriculture Imagery Program) four bands aerial images from MNGEO (Minnesota Geospatial Information Office). (MNGEO sent the data to Dr. Richason, and I got the data from him.)

There are reasons for choosing 2016 data in this study. Firstly, there are six PM 2.5 monitoring sites added after 2012. Two of them were built in the summer of 2013 and started recording in the August of 2013. Four of them started recording in the summer of 2015. Secondly, MNDOT had changed their data management system in 2015, so there were no data available for 2015. The ATR data and the WIM data were available at 2013, 2014 and 2016, but there were eight selected areas which did not have data in 2013 and 2014.

Selection of the Surrounding Area in Monitoring Sites

For the surrounding areas near each site, according to MPCA Air Monitoring Unit Supervisor, MPCA’s target would be “neighborhood scale” for all PM 2.5 monitors in their network (R. Strassman, personal communication, August 28, 2017). The monitoring objective was to assess population exposure at the neighborhood scale or urban scale. Neighborhood means the data from a single monitor should be representative of an area from 1-4 kilometers, and urban scale is from 4-50 kilometers. To meet these scales in 1998 they selected sites in residential areas with uniform land use to meet EPA's siting criteria. For example, the monitor in St. Cloud is located on the roof of the Talahi Community School (neighborhood scale). PM 2.5 concentrations do not vary greatly over larger spatial areas and reflect the air mass that is over the region. But he's confident the monitor at Talahi provides representative data for the city of St. Cloud. Also, other neighborhoods can comprise the entire urban or suburban area for smaller cities. For the Twin Cities Metro area, multiple PM 2.5 monitors rise and fall in unison to reflect diurnal traffic patterns and the air mass over the region. Therefore, Rick Strassman said it is possible to compare these data to show the entire Metro area. Also, other neighborhoods can display the entire urban or suburban area. Table 1 shows the study counties and cities. Table 2 shows the selected measurement scale for each site.

Outline of the Methodology

The following steps are taken for the analysis of data. These steps are further analyzed as follows (Figure 4):

1. Traffic count data are collected at 1,000 locations for the selected areas in 2016, and then using Kriging in ArcGIS to estimate the other sites' AADT data.
2. The traffic count data are weighted based on the daily Automatic Traffic Recorder (ATR) for each day.
3. Using ArcGIS to select the study area in NLCD 2011, then to analyze and calculate the tree space area. And to use NDVI 2017 to compare to the result.
4. Using regression and Pearson correlation to determine the relationship between PM 2.5, tree space area, traffic value, and rainfall.

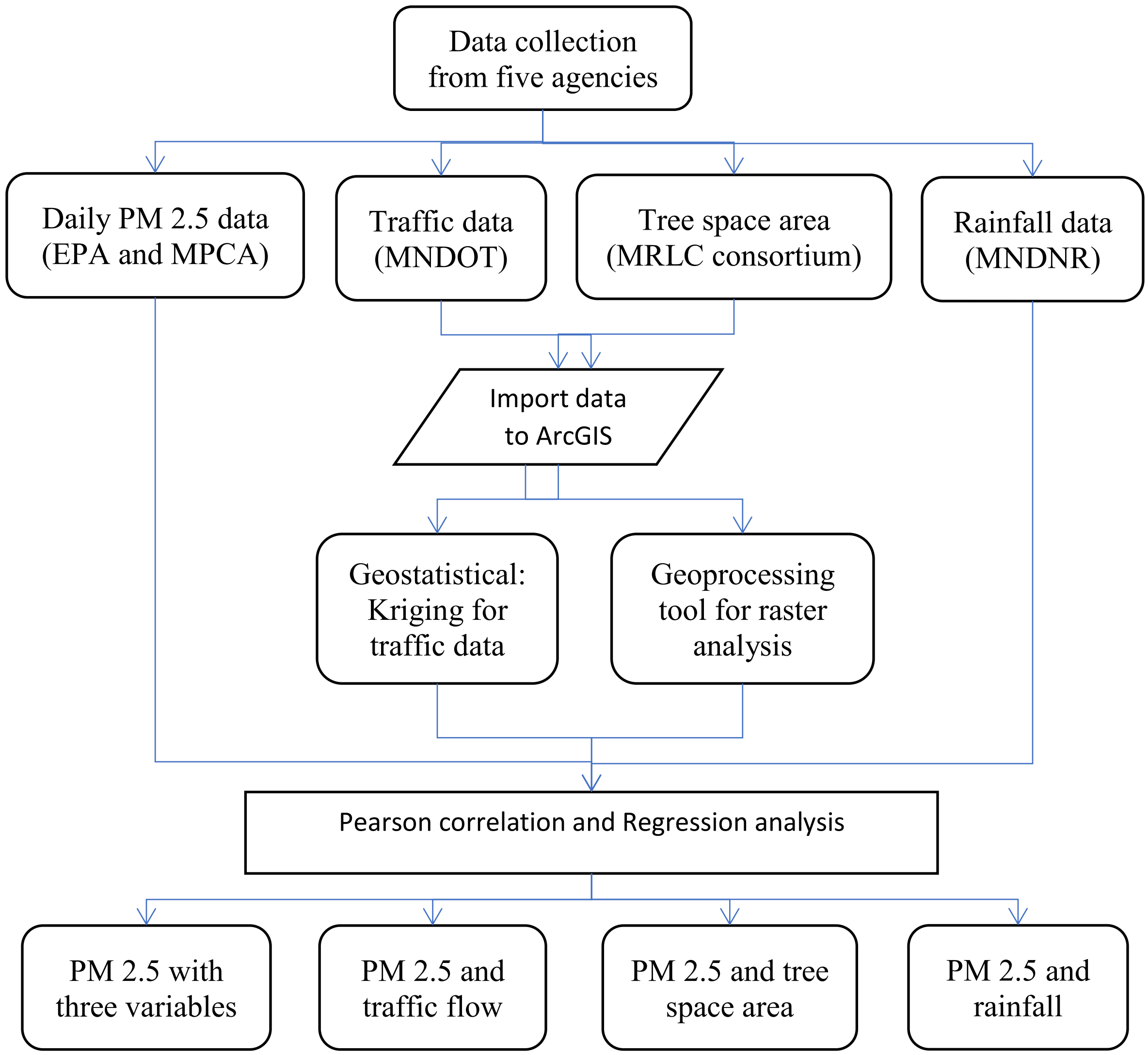

Figure 4: The procedure for data analysis

PM Data Preprocessing

According to MPCA, PM 2.5 value data can be downloaded as Excel spreadsheets. Some monitoring sites have two or three kinds of methods to gather data, so data reports were separated as three duration (as blow list) so there were two or three ".csv" for one monitoring sites. Therefore, PM 2.5 data need to be combined and averaged when they used two or three different reports on the same site and same date.

1. The data were recorded every day for most monitoring sites;
2. The data were recorded once in every 3rd-day for some monitoring sites: Winona site, Minneapolis-963 site, Duluth-7554 site, Duluth-7550 site, St. Paul-871 site, St. Paul-868 site, St. Louis Park site, Apple Valley site, Blaine site;
3. The data were recorded once in every 6th-data for few monitoring sites: Ely site, International Fall site, Rochester site;

There was a measurement change in Shakopee site. This site had changed their measurement from federal reference model to federal equivalent model.

Traffic Data Processing

The Minnesota Department of Transportation (MNDOT) has data on the annual average daily traffic (AADT) flow along Minnesota's road network. Over 2,600 traffic count stations are available for the year 2016. The data are saved in Excel table format. Each station presents the number of vehicles and their positions' coordinates such as latitude and longitude. Then through the ArcGIS® software, the coordinates are converted to points and shown on a Minnesota map (Figure 2). However, the available stations were not covered all counting sites and segments; therefore, it needs to be estimated the other sites' data. Through the ArcGIS® software, using geostatistical analysis method kriging to predict other sites values (Figure 5).

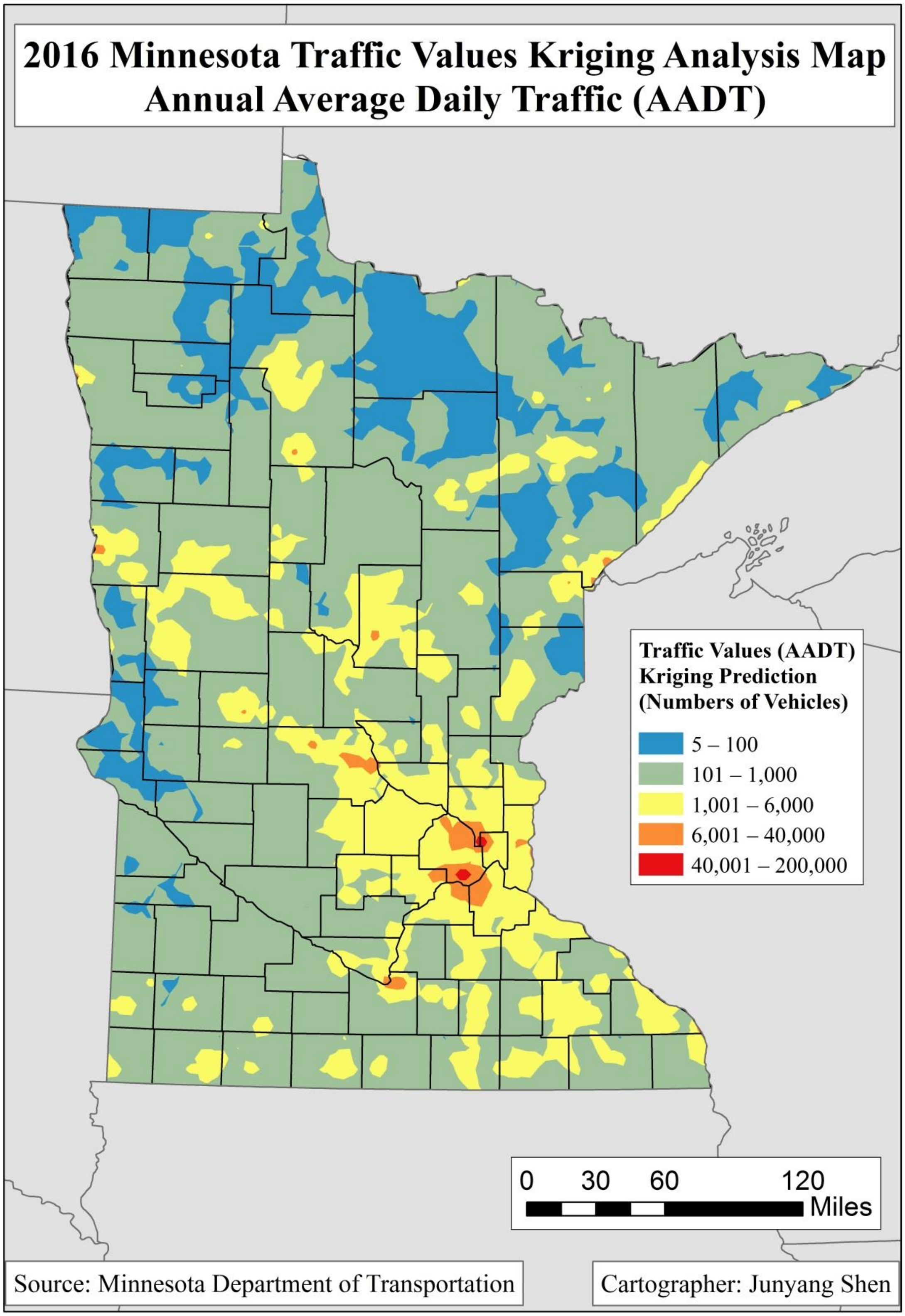

Figure 5: 2016 Minnesota Traffic Values (AADT) Kriging Analysis Map

The traffic data are from MNDOT, and they assume that there are not any differences from Monday to Sunday; however, the data shows that the workdays from Monday to Friday have busier streets traffic than the weekends (Pantaleoni, 2013). In this research, the selected areas need to use the daily data to compare with the PM 2.5 data. Thus, I need to sum the daily data of the selected buffer area. Unfortunately, MNDOT does not report the daily count of traffic flow for any specific day of the week for any station; they only have several tens of stations owning daily data. Therefore, to estimate the selected daily data, the AADT data can be modified to account for variation of traffic volume through the week and to calculate daily traffic value. To create the selected area daily traffic data, I use the Minnesota daily Automatic Traffic Recorder (ATR) data to estimate them. The ATR data are ".txt" files showing the average number of vehicles per day of the week for 86 recording station in Minnesota. These records are used to generate the weights of date in selected sites and applied to the AADT count locations. The number of vehicles per day (N) of is calculated as follows equation: N = AADT x R

Where R = the ratio of the number of the daily vehicles divided by the annual average. According to the Minnesota Seven Day Summary, the number of vehicles on Saturday and Sunday is lower than average, while it is higher than average from Monday to Friday.

About the velocity of the vehicles, it is hard to assess it. Although there are speed limit regulations on each road, drivers may move at lower or higher speeds depending on their individual preferences or the real traffic conditions at the time (such as traffic jams). There are a lot of factors influencing the velocity of vehicles, which are hard to control. Therefore, the velocity of vehicles was assumed to equal the posted speed limit at each AADT location. In Minnesota, the speed limits are shown below:

16 km/h in alleys

48 km/h on streets in urban districts

89 km/h on other roads

105 km/h on expressways

105 km/h on urban interstate highways

113 km/h on rural interstate highways

On the other side, the vehicle miles traveled (VMT) will be used in this research. VMT can be calculated as the AADT multiplied by the length of the road segment. Average Vehicle miles traveled per unit (AVMTU) is calculated as the average daily miles of vehicle travel divided by the total area in the selected monitoring site measurement zone. In order to determine the amount of traffic a state may have, the VMT can be summed for all road segments (Figure 3).

Tree Space Area and GIS

As monitoring sites were selected, their locations were recorded using a global positioning system (GPS) device by MPCA. Those sites' locations were then plotted in a GIS. Using GIS, a selected buffer was created around each site location. Areas in the vicinities of monitoring sites were digitized using the aerial imagery (which is from MNGEO). This allowed for calculations of how much area within the selected area of each site was covered by measurement scale. General selected buffer of all sites are as follows (Table 3 & Appendix).

The tree space area can be calculated by NLCD (2011). The NLCD is a raster dataset which is able to be imported in a geodatabase. According to the Multi-Resolution Land Characteristics (MRLC) Consortium, each individual value represents the area or proportion of that 30m cell covered by tree canopy. Therefore, when there are trees in a cell, it will show green. Also, each cell has the pixel value to show the percentage of the tree canopy. The file pixel values range is from 0 to 100 percent. 0 is that there is not tree covered in a cell area, and it is a non-tree area. 100 is that the cell area is filled full of trees. The Appendices Figures shows the tree canopy area which is analyzed, calculated and plotted by ArcGIS. Also, the 4 bands air photos were used to calculate NDVI in ArcGIS, to compare to the NLCD 2011. There are two air photos of two monitoring sites area covered in this study: Apple Valley - Westview School and Voyageurs National Park.

Table 3

Monitoring sites buffer radius

County	City	AQS Site ID	Buffer Radius(km)	setting
ANOKA	Blaine	270031002	10	Suburban
BECKER	Detroit Lakes	270052013	10	Rural
BELTRAMI	Red Lake	270072304	4	Rural
CARLTON	Cloquet	270177417	4	Rural
COOK	Grand Portage	270317810	2	Rural
CROW WING	Brainerd	270353204	10	Rural
DAKOTA	Apple Valley	270370470	4	Suburban
	Lakeville	270370480	1	Suburban
HENNEPIN	St. Louis Park	270532006	2	Suburban
	Minneapolis	270530962	1	Urban
	Minneapolis	270530963	2	Urban
LAKE	Ely	270750005	10	Rural
LYON	Marshall	270834210	10	Rural
OLMSTED	Rochester	271095008	4	Suburban
RAMSEY	Saint Paul	271230868	2	Urban
	Saint Paul	271230871	2	Urban
ST. LOUIS	International Falls	271370034	4	NP
	Virginia	271377001	4	Urban
	Duluth	271377550	4	Suburban
	Duluth	271377554	4	Suburban
SCOTT	Shakopee	271390505	4	Suburban
SHERBURNE	Saint Cloud	271453052	4	Suburban
WINONA	Winona	271699000	10	Rural
WRIGHT	St. Michael	271713201	4	Suburban

Chapter IV: Results

Mean PM 2.5 Value

The average PM 2.5 from all of the monitoring sites taken over the entire study area was 4.29 micrograms per cubic meter of air (μg/m^3). The highest average PM 2.5 for any individual site was 6.65μg/m^3 recorded by the Brainerd Lakes Regional Airport (270353204) monitoring site. That was 2.36μg/m^3 above the mean. The Detroit Lakes (270052013) monitoring site recorded the lowest average PM 2.5 of 2.44μg/m^3. That site was 1.85μg/m^3 below the overall mean. Therefore, the difference between the average PM 2.5 recorded by the highest and the lowest sites was 4.21μg/m^3.

Figure 6 shows the value map created using the mean PM 2.5 from each site. There is a noticeable pattern of higher PM 2.5 along the urban, suburban and roadside areas. These areas of high value coincide with high traffic value, specifically urban centers, and roadside locations, which are found mainly in those areas. The highest mean PM value is found at Brainerd Lakes Regional Airport’s open area. It is a rural site, but it is also a roadside site. It is less than one mile away from the Highway 210 and is about three miles’ distance from the city of Brainerd business district. The areas with lower pollution were found in more rural areas where large expanses are tree covered and with less traffic flow. The range between the lowest and highest mean PM 2.5 values of 4.21μg/m^3 is shown on the map (Figure 6).

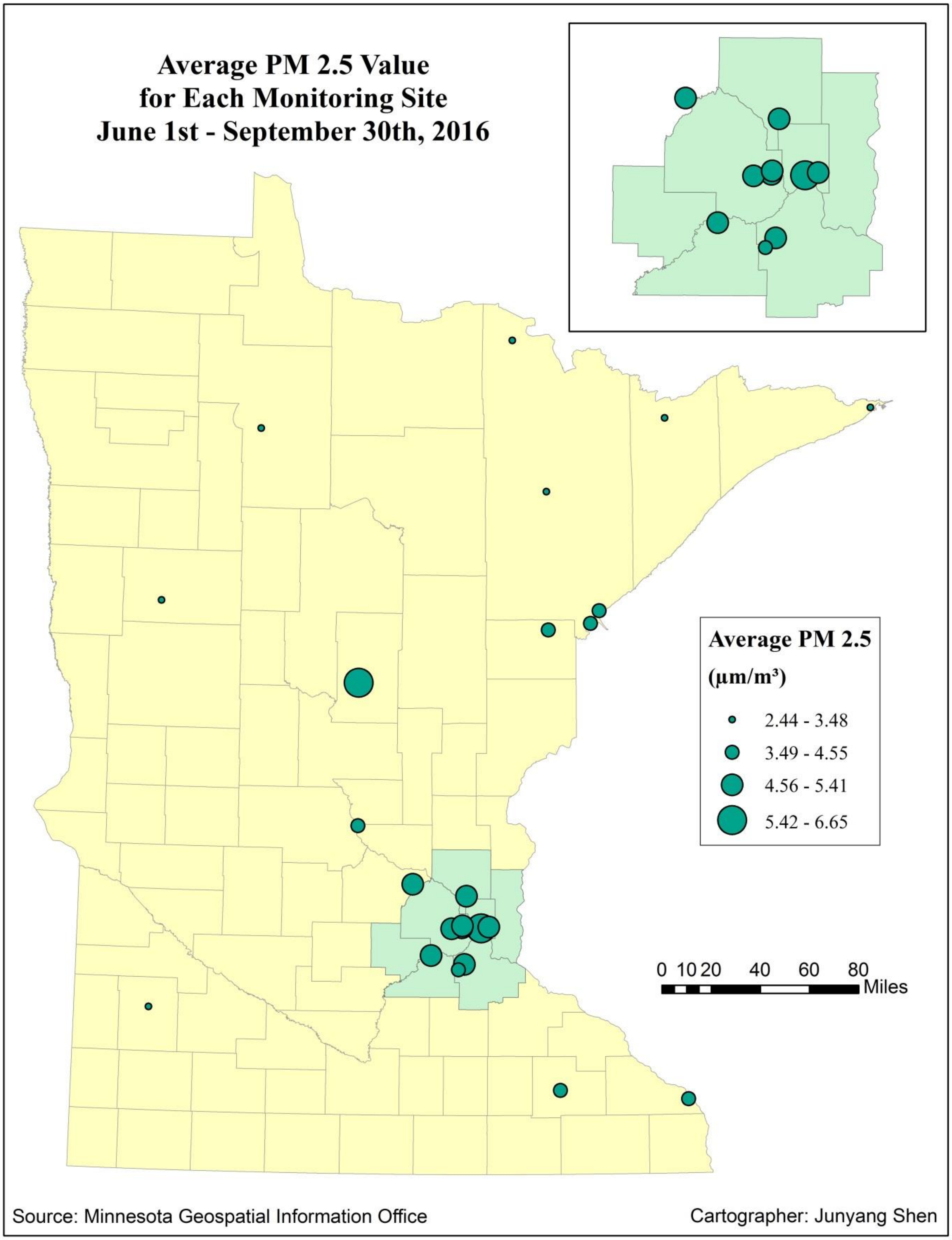

Figure 6: PM 2.5 pattern created by the exact mean PM 2.5 values provided by the monitoring sites

The average PM 2.5 from all urban and suburban monitoring sites which were taken over the entire duration of the study area was 5.68µg/m^3. That was 1.39µg/m^3 above the mean. The average PM 2.5 from all rural and national park monitoring sites which were taken over the entire duration of the study area was 3.74µg/m^3, which is 0.55µg/m^3 below the overall mean.

Temperature and Wind Speed

According to the literature review, the different places have different effects on temperature and wind speed.

Temperature. There is no significant relationship between temperature and PM 2.5 in Minnesota (Table 4 and Figure 7). The results of Pearson correlation coefficient test (r^2=0.004, r = - 0.064, p= 0.706 > 0.05) indicated that no correlation exists between temperature and PM 2.5.

Table 4

Pearson correlation between temperature and PM 2.5

Correlations			
		PM 2.5	Temperature
PM 2.5	Pearson Correlation	1	-.064
	Sig. (2-tailed)		.706
Temperature	Pearson Correlation	-.064	1
	Sig. (2-tailed)	.706	

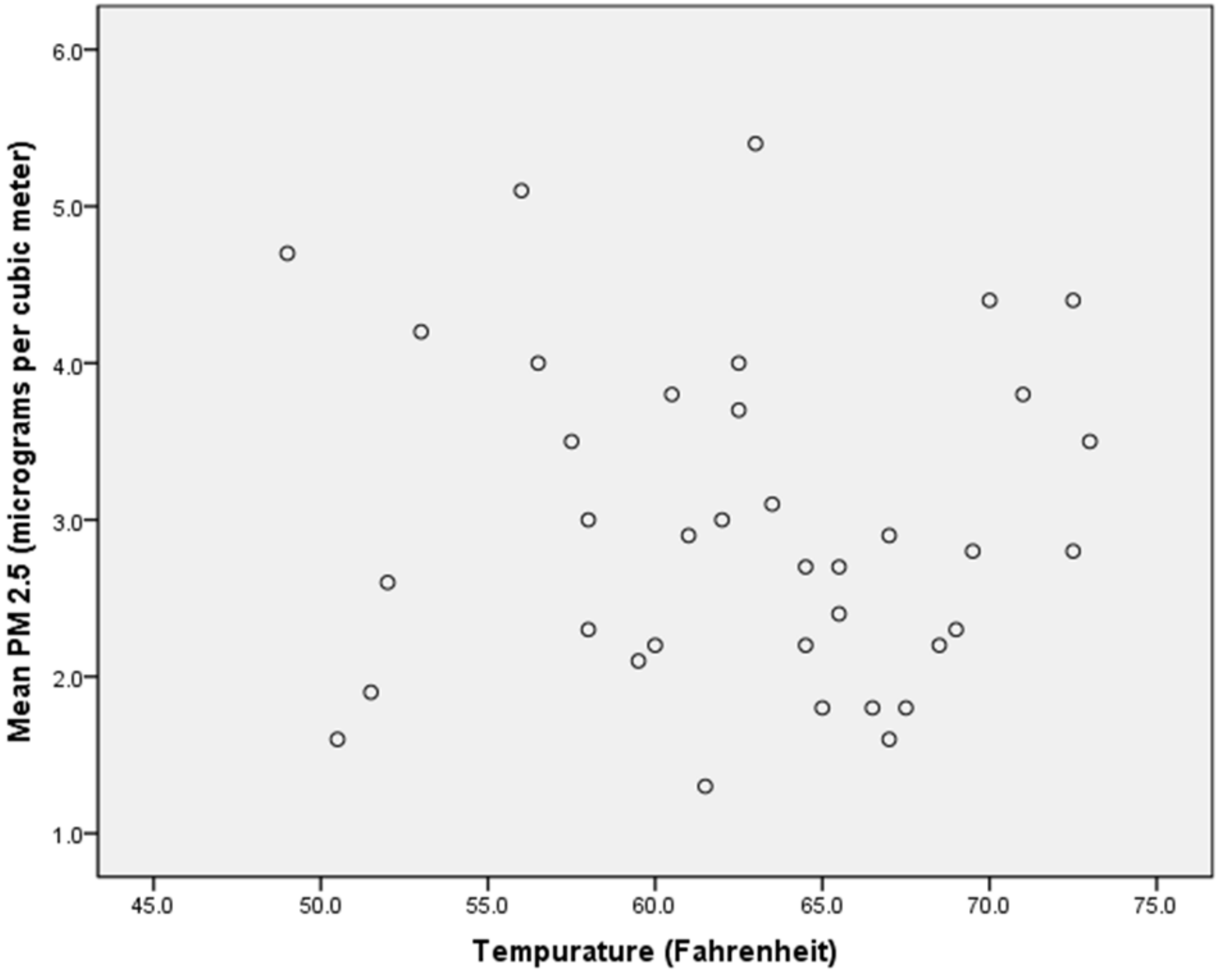

Figure 7: A scatter plot is showing the mean PM 2.5 against the temperature surrounding the monitoring site locations. The Pearson's r correlation coefficient was - 0.064, and the significant value p was over than 0.05. Those were considered no correlation, and the data were randomly separated on the plot.

Wind speed. There is no significant relation between wind speed and PM 2.5 in Minnesota (Table 5 and Figure 8). The results of Pearson correlation coefficient test (r^2=0.0014, r = - 0.038, p= 0.681 > 0.05) indicated that no correlation exists between wind speed and PM 2.5.

Table 5

Pearson correlation between wind speed and PM 2.5

Correlations			
		PM 2.5	Wind Speed
PM 2.5	Pearson Correlation	1	-.038
	Sig. (2-tailed)		.681
Wind Speed	Pearson Correlation	-.038	1
	Sig. (2-tailed)	.681	

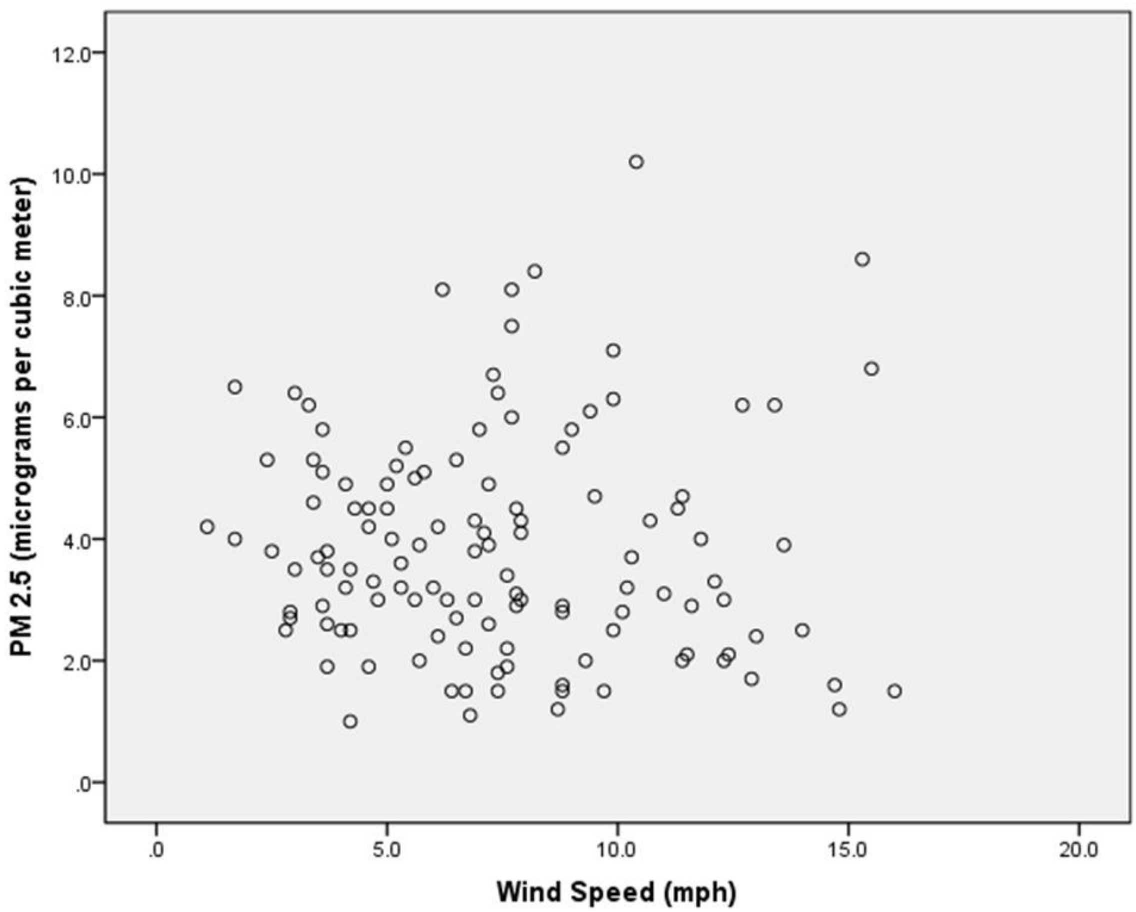

Figure 8: A scatter plot is showing the mean PM 2.5 against the wind speed surrounding the monitoring site locations. The Pearson's r correlation coefficient was - 0.038, and the significant value p was over than 0.05. Those were considered no correlation, and the data were randomly separated on the plot.

Tree Space Area Ratio

The mean Tree Covered Pixel Ratio (TCPR) from all monitoring sites buffer of the study was 31.74%. The highest TCPR for any individual site was 76.07%, which is belonging to Grand Portage Band (270317810) monitoring site. That was 44.33% above the mean. This monitoring site is located at the Grand Portage Band offices in Grand Portage, and the land use of the surrounding community is an undeveloped forest. The Minneapolis - Near Road I-35/ I-94 (270530962) monitoring site buffer had the lowest TCPR, which was 2.71%. That site was 29.03% below the overall mean. This monitoring site is a near-road monitoring site, and it is closed to the I-35/I-94 commons near downtown Minneapolis. Therefore, the difference of the TCPR from the highest site to the lowest site was 73.36%.

According to 2011 Minnesota Tree Canopy raster data, the ArcGIS can be used to calculate the Tree Covered Pixel Ratio (TCPR). Table 6 shows the Tree Covered Pixel Ratio for each monitoring site.)

Table 6

Tree Covered Pixel Ratio (TCPR), & Setting. (NP is a national park.)

County	City	AQS Site ID	TCPR	Setting
ANOKA	Blaine	270031002	25.66%	Suburban
BECKER	Detroit Lakes	270052013	30.20%	Rural
BELTRAMI	Red Lake	270072304	37.75%	Rural
CARLTON	Cloquet	270177417	69.59%	Rural
COOK	Grand Portage	270317810	76.07%	Rural
CROW	Brainerd	270353204	46.09%	Rural
DAKOTA	Apple Valley	270370470	23.04%	Suburban
	Lakeville	270370480	25.30%	Suburban
HENNEPIN	St. Louis Park	270532006	20.71%	Suburban
	Minneapolis	270530962	2.71%	Urban
	Minneapolis	270530963	4.375%	Urban
LAKE	Ely	270750005	53.15%	Rural
LYON	Marshall	270834210	4.56%	Rural
OLMSTED	Rochester	271095008	48.96%	Suburban
RAMSEY	Saint Paul	271230868	3.92%	Urban
	Saint Paul	271230871	18.96%	Urban
ST LOUIS	International Falls	271370034	69.07%	NP
	Virginia	271377001	30.01%	Urban
	Duluth	271377550	40.29%	Suburban
	Duluth	271377554	33.86%	Suburban
SCOTT	Shakopee	271390505	16.49%	Suburban
SHERBURNE	Saint Cloud	271453052	13.15%	Suburban
WINONA	Winona	271699000	56.62%	Rural
WRIGHT	St. Michael	271713201	11.32%	Suburban

Figure 9 shows the TCPR value map, which is created by using the TCPR from each site. There is a noticeable pattern of lower ratio along the urban, suburban and roadside area. The lowest ratio is found at Minneapolis - Near Road I-35/ I-94. It is a near-road site. The higher ratio areas were found in more rural areas, park areas, and national park areas where there are more trees covered pixels. The range between the lowest and highest ratio value of 73.36% is on the map.

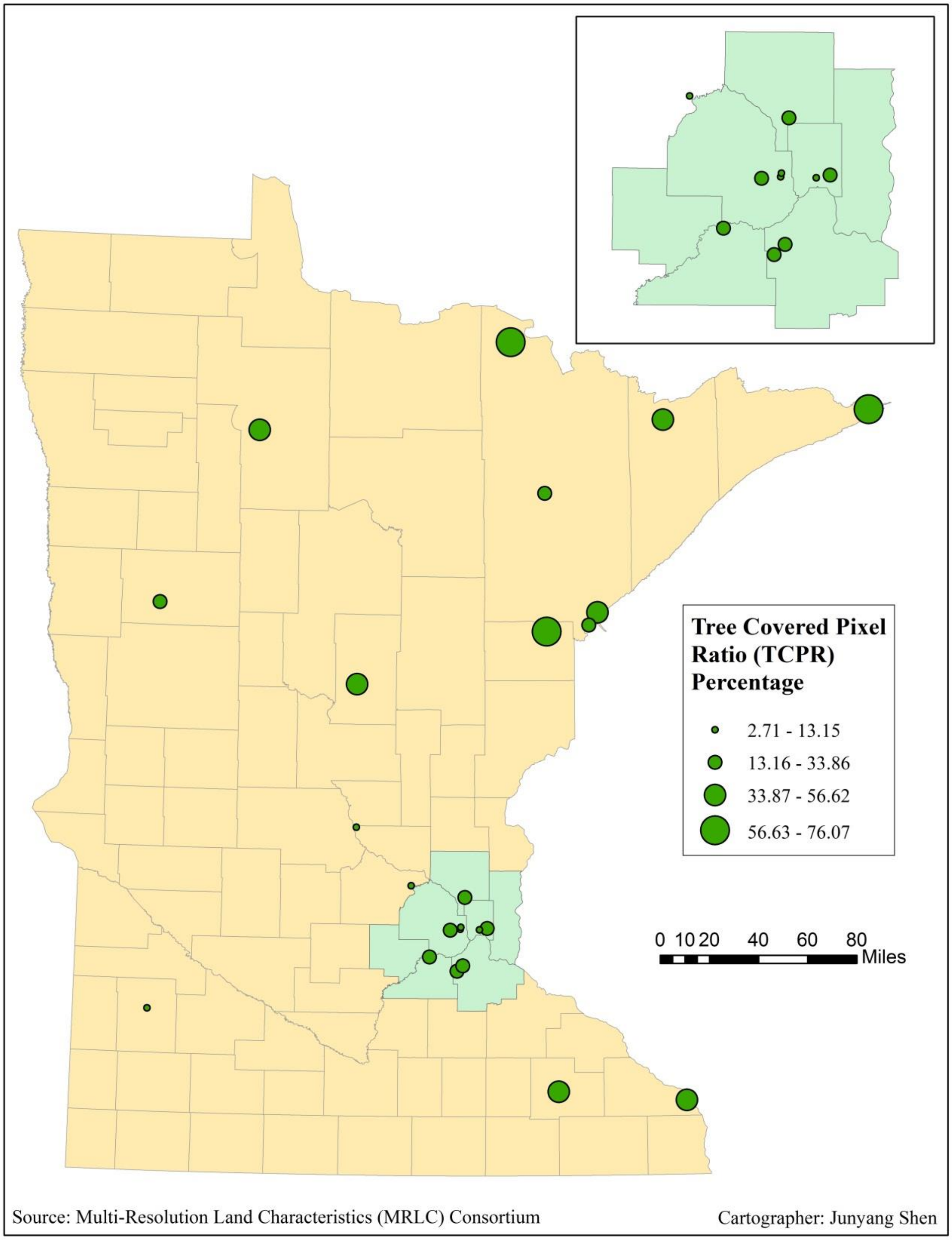

Figure 9: Tree Covered Pixel Ratio of each monitoring site

Tree Space Area Ratio (TSAR) is different from the Tree Covered Pixel Ratio. Each pixel is a 30 meters x 30 meters area, and it has a tree covered percentage value. After calculating, the result of the Tree Space Area Ratio is shown in Table 7.

Table 7

Tree Space Area Ratio (TSAR), & Setting. (NP is a national park.)

County	City	AQS Site ID	TSAR	Setting
ANOKA	Blaine	270031002	12.45%	Suburban
BECKER	Detroit Lakes	270052013	14.41%	Rural
BELTRAMI	Red Lake	270072304	16.47%	Rural
CARLTON	Cloquet	270177417	34.60%	Rural
COOK	Grand Portage	270317810	40.31%	Rural
CROW WING	Brainerd	270353204	21.56%	Rural
DAKOTA	Apple Valley	270370470	10.66%	Suburban
	Lakeville	270370480	10.54%	Suburban
HENNEPIN	St. Louis Park	270532006	8.14%	Suburban
	Minneapolis	270530962	0.80%	Urban
	Minneapolis	270530963	1.35%	Urban
LAKE	Ely	270750005	25.92%	Rural
LYON	Marshall	270834210	0.99%	Rural
OLMSTED	Rochester	271095008	14.97%	Suburban
RAMSEY	Saint Paul	271230868	1.30%	Urban
	Saint Paul	271230871	7.60%	Urban
ST LOUIS	International Falls	271370034	34.71%	NP
	Virginia	271377001	11.61%	Urban
	Duluth	271377550	18.46%	Suburban
	Duluth	271377554	16.04%	Suburban
SCOTT	Shakopee	271390505	8.26%	Suburban
SHERBURNE	Saint Cloud	271453052	6.58%	Suburban
WINONA	Winona	271699000	35.35%	Rural
WRIGHT	St. Michael	271713201	5.93%	Suburban

Figure 10 shows the TSAR value map, which is created by using the TSAR from each site. There is a noticeable pattern which is very similar to the TCPR, and the lower ratio is along the urban, suburban and roadside area. The lowest ratio is found at Minneapolis - Near Road I-35/ I-94. The highest ratio is found at Grand Portage Band. The higher ratio areas were found in more rural areas, park areas, and national park areas where there are more trees covered. The range between the lowest and highest ratio value of 39.51% is on the map.

The average Tree Space Area Ratio (TSAR) from all monitoring sites buffer of the study was 14.96%. The highest TSAR for any individual site was 40.31%, which is belonging to the Grand Portage Band of Lake (270317810) monitoring site. That was 25.35% above the mean. This monitoring site is located at the Grand Portage Band offices in Grand Portage, and the land use of the surrounding community is an undeveloped forest. The Minneapolis - Near Road I-35/ I-94 (270530962) monitoring site buffer had the lowest TSAR, which was 0.8%. That site was 14.16% below the overall mean. This monitoring site is a near-road monitoring site, and it is closed to the I-35/I-94 commons near downtown Minneapolis. Therefore, the difference of the TCPR from the highest site to the lowest site was 39.51%.

According to the 2017 NDVI calculation, the tree space area ratio of the Apple Valley monitoring site is 10.86%, and the NLCD 2011 result is 10.66%. The tree space area ratio of the Voyageurs National Park (from NDVI) is 34.11%, and the NLCD 2011 result is 34.71%. The figure 11 and figure 12 show the results.

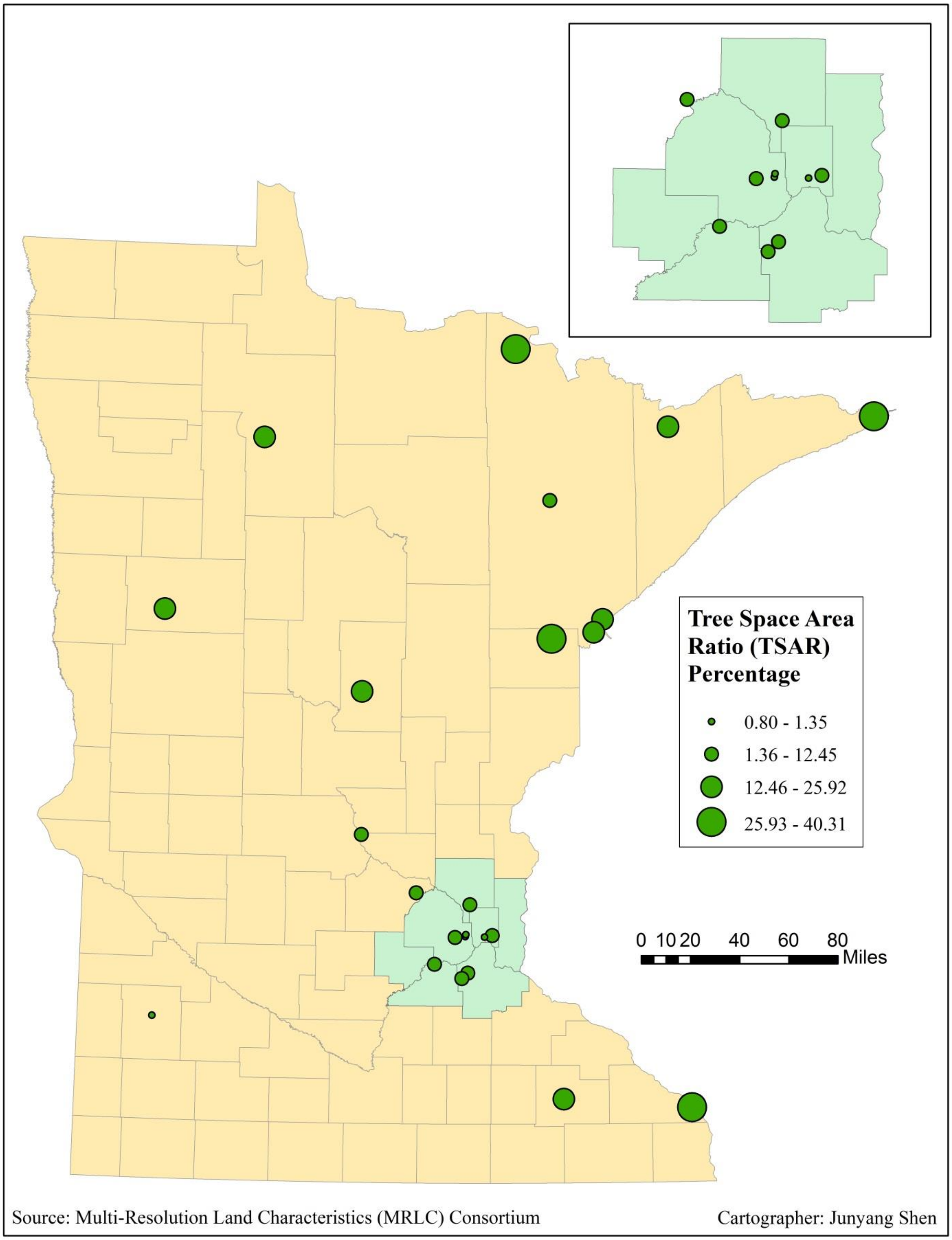

Figure 10: Tree Space Area Ratio of each monitoring site

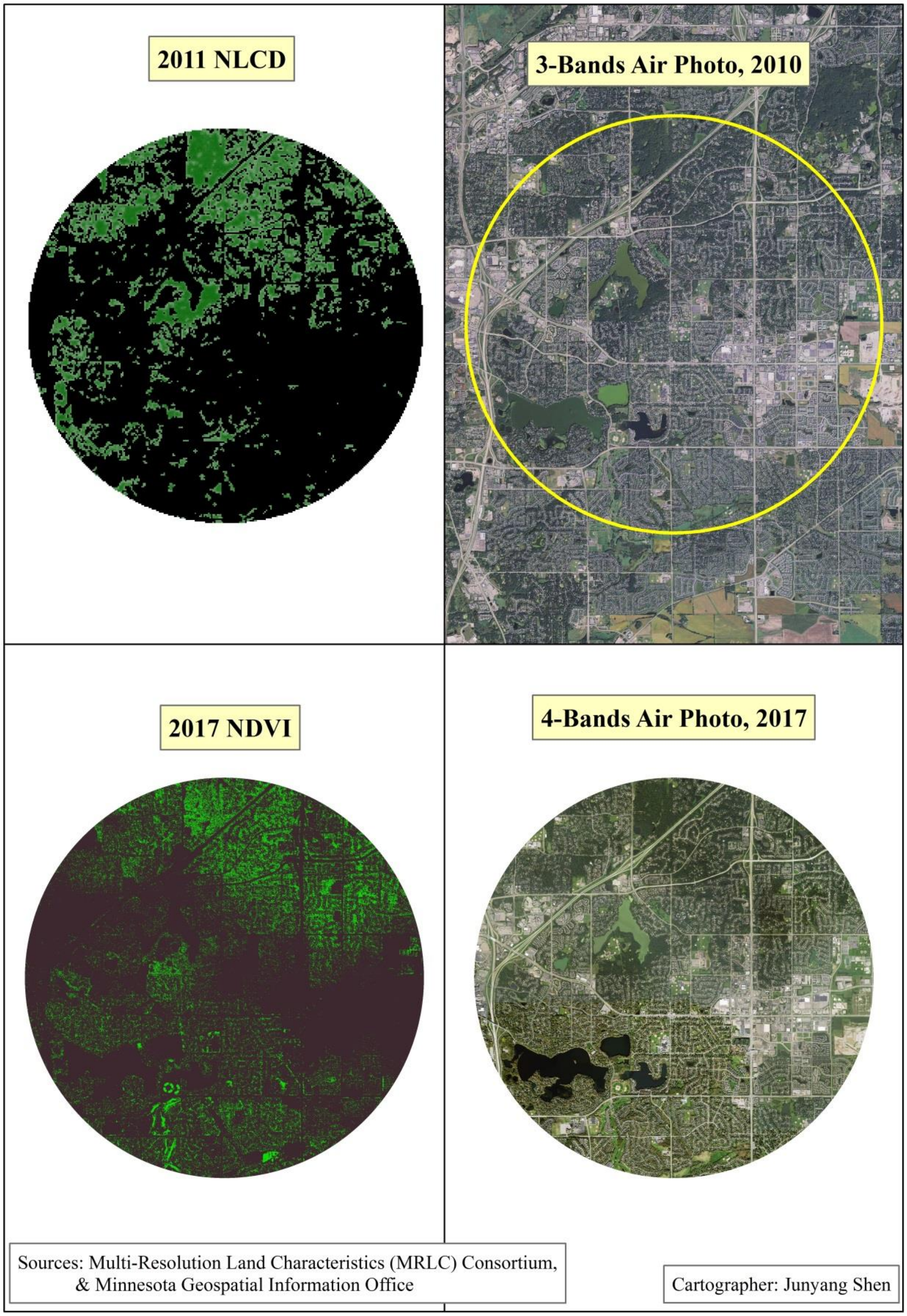

Figure 11: Apple Valley: 2011 NLCD, 2017 NDVI, Air Photo 2010 & Air Photo 2017

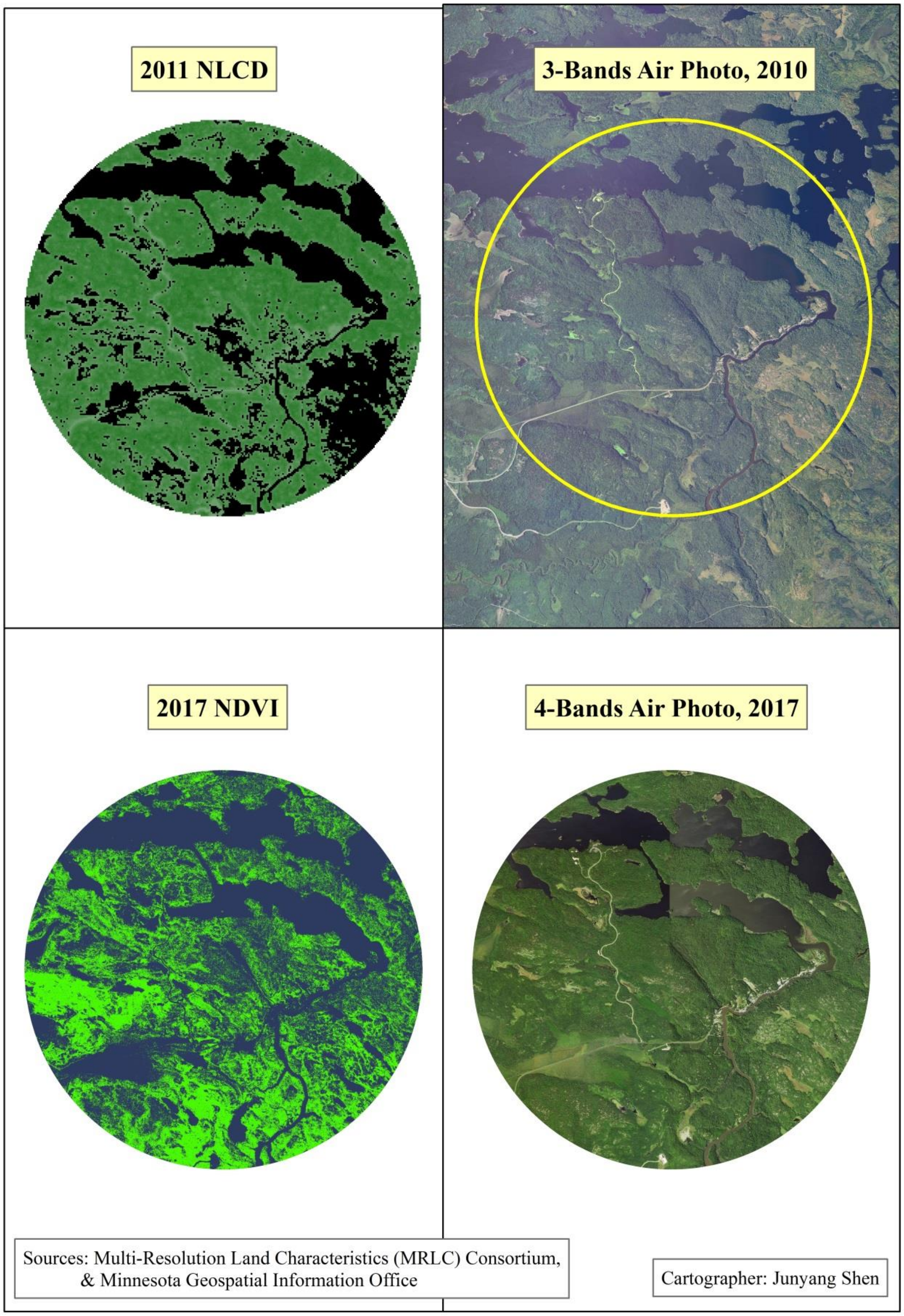

Figure 12: Voyageurs: 2011 NLCD, 2017 NDVI, Air Photo 2010 & Air Photo 2017

PM 2.5 Values and Tree Space Area Ratio

The mean PM 2.5 data were plotted and analyzed to determine if there was a relationship between the percentage of the tree space area covered the monitoring site and the mean PM 2.5 over the study period. A Pearson correlation coefficient was computed to assess the strength and direction of that relationship. The results of that test (r^2=0.164, r = - 0.406, n= 24, p= 0.049) indicated that a medium negative correlation exists between those variables (r= - 0.406). The r^2 value of 0.164 means that there is only about 16% of the variation in the mean temperatures can be explained by differences in the percentage of tree space cover near the monitoring sites. There were 24 samples used in the calculation (n). The p-value of less than 0.05 indicates that correlation is significant in this finding. Table 8, Table 8, Table 10, and Figure 13 summarize those results.

Table 8

Pearson Correlations (PM 2.5 & TSAR)

		Mean PM 2.5	TSAR
Mean PM 2.5	Pearson Correlation	1	-.406*
	Sig. (2-tailed)		.049
TSAR [a]	Pearson Correlation	-.406*	1
	Sig. (2-tailed)	.049	

*. Correlation is significant at the 0.05 level (2-tailed).
a. TSAR is tree space area ratio.

Table 9

Regression Linear Model Summary (PM 2.5 & TSAR)

Model	R	R Square	Adjusted R Square	Std. The error of the Estimate
1	.406[a]	.164	.153	.92249

a. Predictors: (Constant), TSAR [b]
b. TSAR is tree space area ratio.

This table provides the r and r^2 values. The r value represents the simple correlation and is 0.406, which indicates a medium degree of correlation. The r^2 value indicates how much of the total variation in the dependent variable, mean PM 2.5, can be explained by the independent variable, tree space area ratio. In this case, 16.4% can be explained, which is small. Although the p-value of less than 0.05 indicates that correlation is significant in this finding, statistical significance does not imply practical significance.

Table 10

Regression Linear Model Coefficients (PM 2.5 & TSAR)

Model [a]		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5.087	.310		16.399	.000
	TSAR [b]	-.045	.016	-.406	-2.747	.012

a. Dependent Variable: Mean PM 2.5
b. TSAR is tree space area ratio.

Also, based on the coefficients table, it provides us with the necessary information to predict PM 2.5 value from the relationship between PM 2.5 values and tree space area ratio, as well as determine whether TSAR (tree space area ratio (%)) contributes significantly to the model. Furthermore, we can use the values in the “B” column under the “Unstandardized Coefficients” column to present the regression equation as below:

Mean PM 2.5 = 5.087 - 0.045 x TSAR + error

TSAR is tree space area ratio.

When holding constant mean PM 2.5 value of 5.087µg/m^3, one percentage of tree space area ratio increase produces a 0.045µg/m^3 decrease in mean PM 2.5.

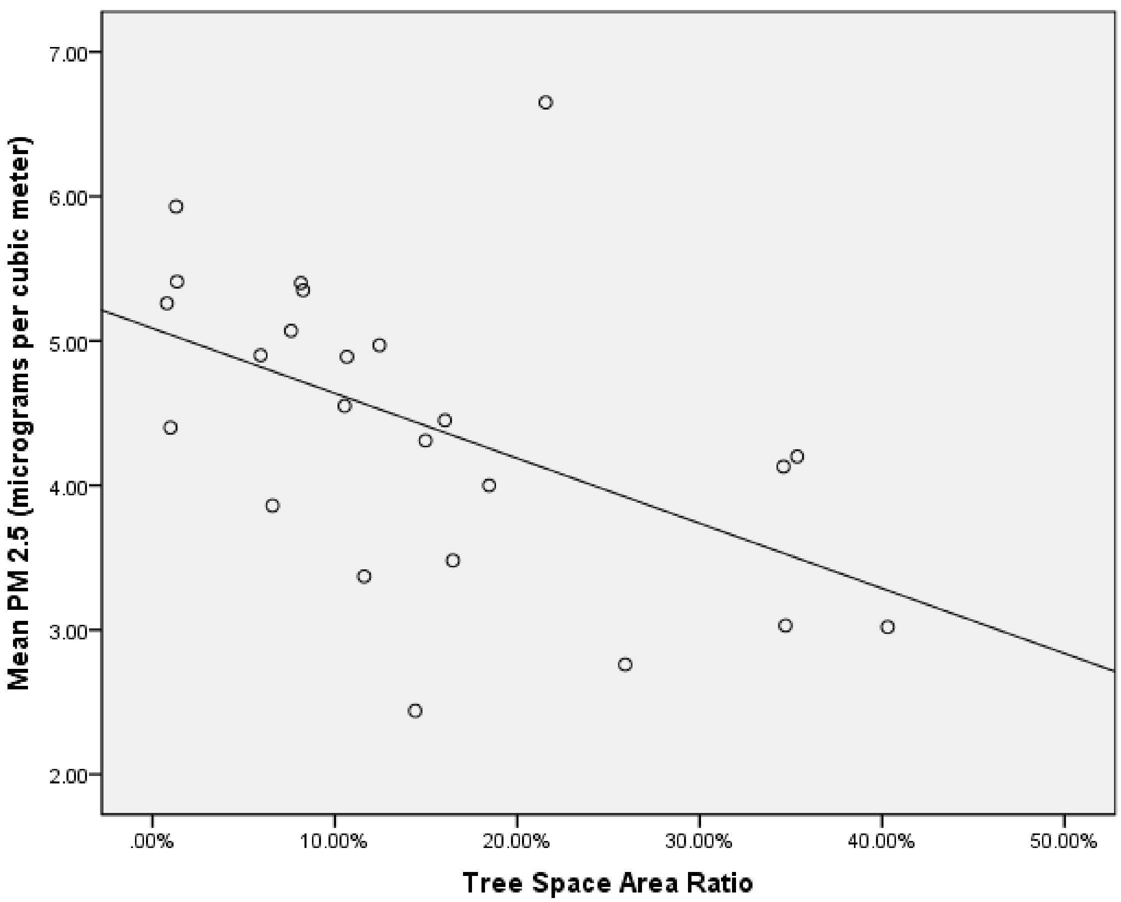

Figure 13: A scatter plot is showing the mean PM 2.5 against the tree space ratio surrounding the monitoring site locations. The Pearson’s r correlation coefficient was - 0.406, which is considered a medium negative correlation.

According to Figure 13, there are two mean PM 2.5 values that are possible outliers in the data. They are the Brainerd Lakes Regional Airport (270353204) monitoring site, and the Detroit Lakes - FWS Wetland Management District (270052013) monitoring site.

Brainerd Lakes Regional Airport (270353204) monitoring site: mean PM 2.5 was 6.65µg/m^3, and tree space area ratio is 21.56%.

Detroit Lakes - FWS Wetland Management District (270052013) monitoring site: mean PM 2.5 was 2.44µg/m^3, and tree space area ratio is 14.41%.

After removing the outlier data, a Pearson correlation coefficient was computed to assess the strength and direction of that relationship. The results of that test (r^2=0.497, r = - 0.705, n= 22, p= 0.000) indicated that a strong negative correlation exists between those variables (r= - 0.705). 49.7% of the variation in the mean PM 2.5 value can be explained by differences in the percentage of tree space cover near the monitoring sites. There were 22 samples used in the calculation (n). The p-value of less than 0.001 indicates a high confidence level in this finding. Table 11, Table 12, Table 13, and Figure 14 summarize those results.

Table 11

Pearson correlation without outlier (PM 2.5 & TSAR)

		Mean PM 2.5	TSAR [a]
Mean PM 2.5	Pearson Correlation	1	-.705**
	Sig. (2-tailed)		.000
TSAR	Pearson Correlation	-.705**	1
	Sig. (2-tailed)	.000	

**. Correlation is significant at the 0.01 level (2-tailed).
a. TSAR is tree space area ratio.

Table 12

Regression Linear Model Summary without outlier (PM 2.5 & TSAR)

Model	R	R Square	Adjusted R Square	Std. The error of the Estimate
1	.705[a]	.497	.471	.63945

a. Predictors: (Constant), TSAR [b]
b. TSAR is tree space area ratio.

Table 12 provides the r and r^2 values. The r value represents the simple correlation and is 0.705, which indicates a high degree of correlation. The r^2 value indicates how much of the total variation in the dependent variable, mean PM 2.5, can be explained by the independent variable, tree space area ratio. In this case, 49.7% can be explained, which is a stronger relationship than that with outlier data. The p-value of less than 0.001 indicates that correlation is significant in this finding.

Table 13

Regression Linear Model Coefficients without outlier (PM 2.5 & TSAR)

Model [a]		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5.148	.217		23.712	.000
	TSAR [b]	-.051	.012	-.705	-4.441	.000

a. Dependent Variable: Mean PM 2.5
b. TSAR is tree space area ratio

The coefficients in the table 13 provide us with the necessary information of the relations between mean PM 2.5 and tree space area ratio. Furthermore, we can use the values in the “B” column under the “Unstandardized Coefficients” column to present the regression equation as:

Mean PM 2.5 = 5.148 - 0.051 x TSAR + error

TSAR is tree space area ratio.

When holding constant mean PM 2.5 value of 5.148µg/m^3, one percentage of tree space area ratio increase produces a 0.051µg/m^3 decrease in mean PM 2.5.

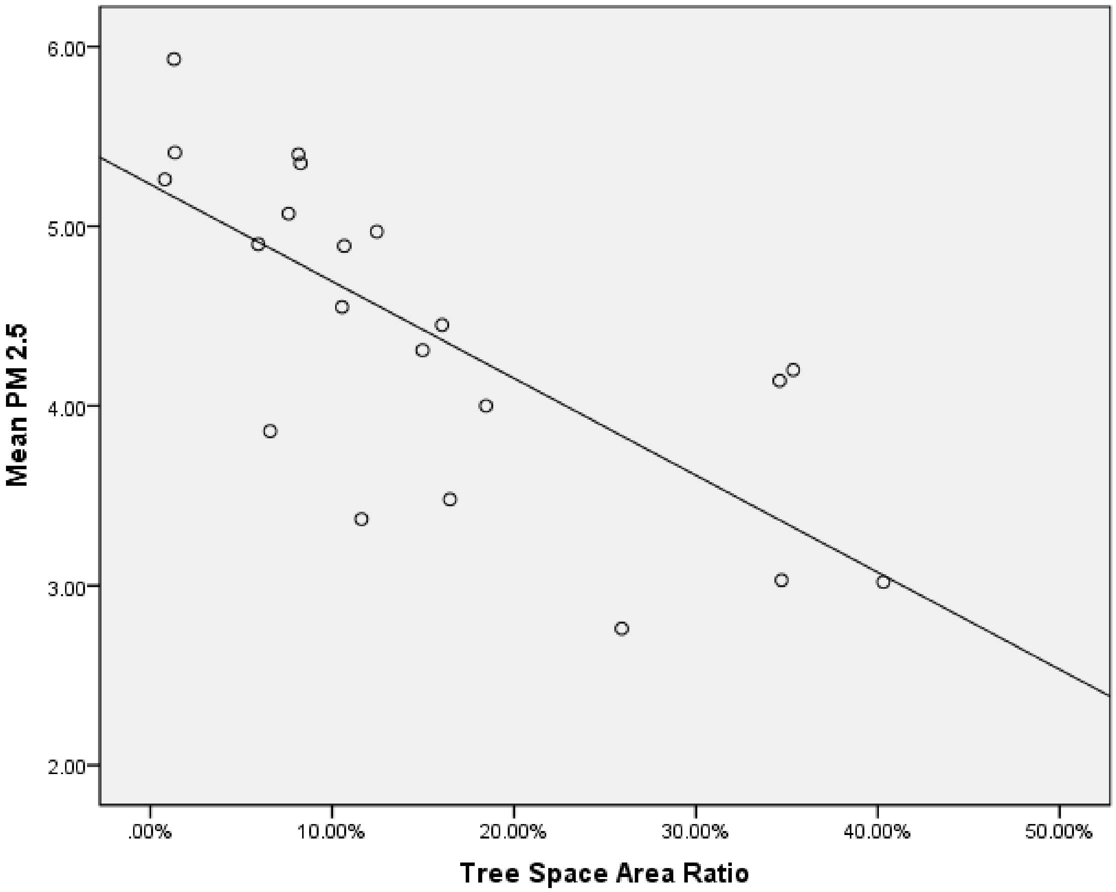

Figure 14: A scatter plot is showing the mean PM 2.5 without outlier site against the tree space ratio in the monitoring sites' locations. The Pearson's r correlation coefficient was -0.705, which is considered a strong negative correlation.

Traffic Value and Mean PM 2.5

The vehicle miles traveled (VMT) is the number of miles traveled by vehicles in the study area. The average daily vehicle miles traveled per unit (miles per square mile) indicates that the average daily number of miles traveled by vehicles per square mile in the study area. The average daily vehicle miles traveled per unit (miles per square mile) (AVMTU) for the study area of each monitoring site is shown in Table 14.

Based on the average vehicles miles per square mile, the data can be classified into two groups:

Group A: over 10,000 miles per square mile is belonging to the urban and suburban area.

Group B: less than 10,000 miles per square mile is belonging to the rural and national park area.

Table 14

The Average Vehicle Miles Traveled per Unit

Setting	Site Name	Mean PM 2.5 ($\mu g/m^3$)	AVMTU ($miles/mile^2$)
Rural & National Park	Ely - Fernberg Road	2.76	30
	Voyageurs NP	3.03	66
	Grand Portage Band	3.02	1168
	Red Lake Nation	3.48	1404
	Fond du Lac Band	4.13	1696
	Great River Bluffs State Park	4.2	2214
	Marshall - Southwest Minnesota Regional Airport	4.4	2498
	Detroit Lakes - FWS Wetland Management District	2.44	3336
	Brainerd Lakes Regional Airport	6.65	5282
Urban & Suburban	Virginia City Hall	3.37	10265
	St. Michael Elementary School	4.9	21116
	Duluth - Laura MacArthur School	4.45	21537
	Saint Cloud - Talahi School	3.86	26472
	Duluth - U of M	4	26568
	Shakopee - B.F. Pearson School	5.35	28052
	Rochester - Ben Franklin School	4.31	37778
	Blaine - Anoka County Airport (National Core)	4.97	52019
	Apple Valley - Westview School	4.89	64684
	Saint Paul - Harding High School	5.07	92518
	St. Louis Park - City Hall	5.4	107242
	Lakeville - Near Road I-35	4.55	121128
	St. Paul - Ramsey Health Center	5.93	220242
	Minneapolis - Andersen School	5.41	248038
	Minneapolis - Near Road I-35/I-94	5.26	382763

The highest AMVTU for any individual site was 382,763 miles/mile^2, which belongs to Minneapolis - near the Road I-35/I-94 monitoring site. This monitoring site is located along the I-35/I-94 W commons near downtown Minneapolis. The Ely - Fernberg Road monitoring site and Voyageurs NP monitoring site had the lowest AMVTU, which was both less than 100 miles/mile^2. The Ely - Fernberg Road monitoring site is located on a remote hilltop about 19 miles east of Ely, and it is close to the Boundary Waters Canoe Area Wilderness. The Voyageurs NP monitoring site is located on a rocky outcrop near the Ash River Interpretive Center on the southeast side of the Voyageurs National Park.

The average vehicles mile traffic per unit data were plotted and analyzed to determine if there was a relationship between AMVTU around monitoring sites and the mean PM 2.5 over the study period (Figure 15). According to the Figure 15, there is a monitoring site data considering as an obvious outlier (The outlier data is the upper left point). This site is the Brainerd Lakes Regional Airport (270353204) monitoring site. After removing this site (which is shown in Figure 16), a Pearson correlation coefficient was computed to assess the strength and direction of that relationship. The results of that test ($r = 0.644$, $n= 23$, $p= 0.001$) indicated that a strong positive correlation exists between those variables ($r= 0.644$). Table 15 summarizes those results.

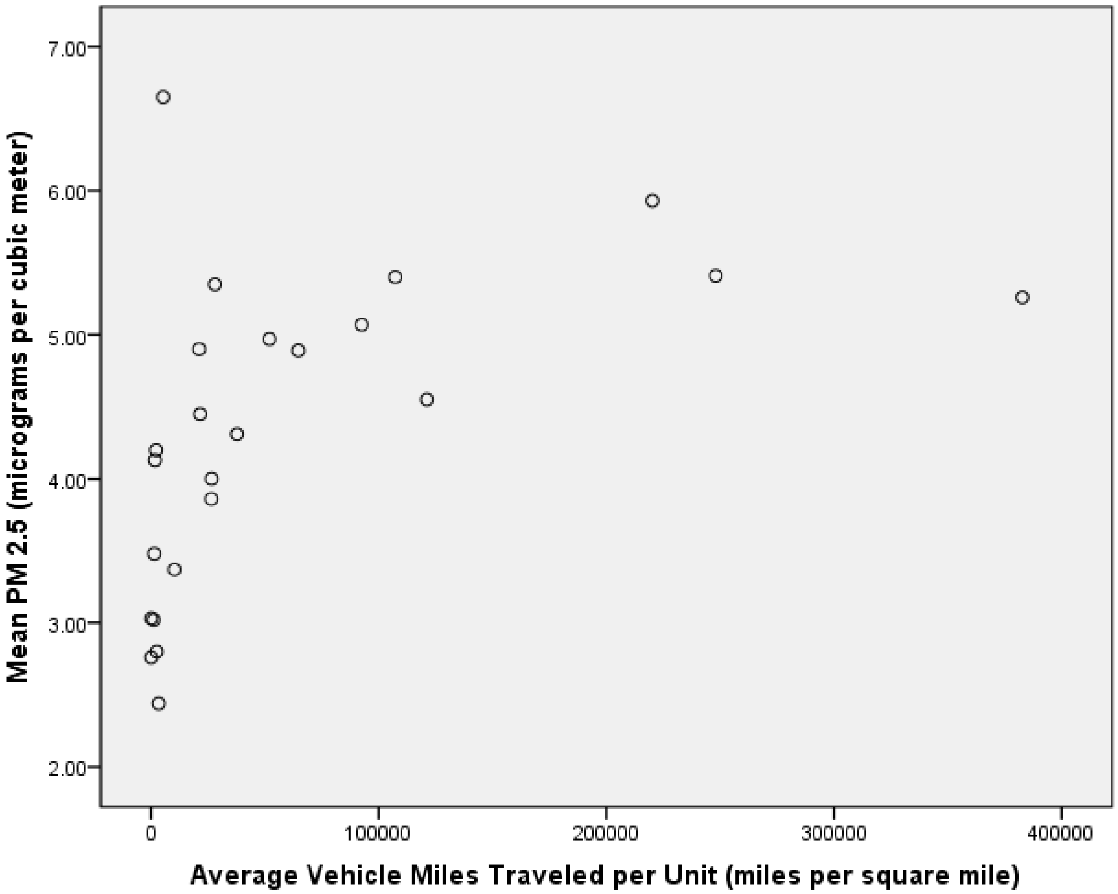

Figure 15: A scatter plot is showing the Average Vehicle Miles Traveled per Unit against the mean PM 2.5 surrounding the monitoring site locations.

Table 15

Pearson correlation without outlier (PM 2.5 & AMVTU)

		AVMTU (Miles/km^2)	Mean PM 2.5
AVMTU (Miles/ km^2)	Pearson Correlation	1	.644**
	Sig. (2-tailed)		.001
Mean PM 2.5	Pearson Correlation	.644**	1
	Sig. (2-tailed)	.001	

**. Correlation is significant at the 0.01 level (2-tailed).

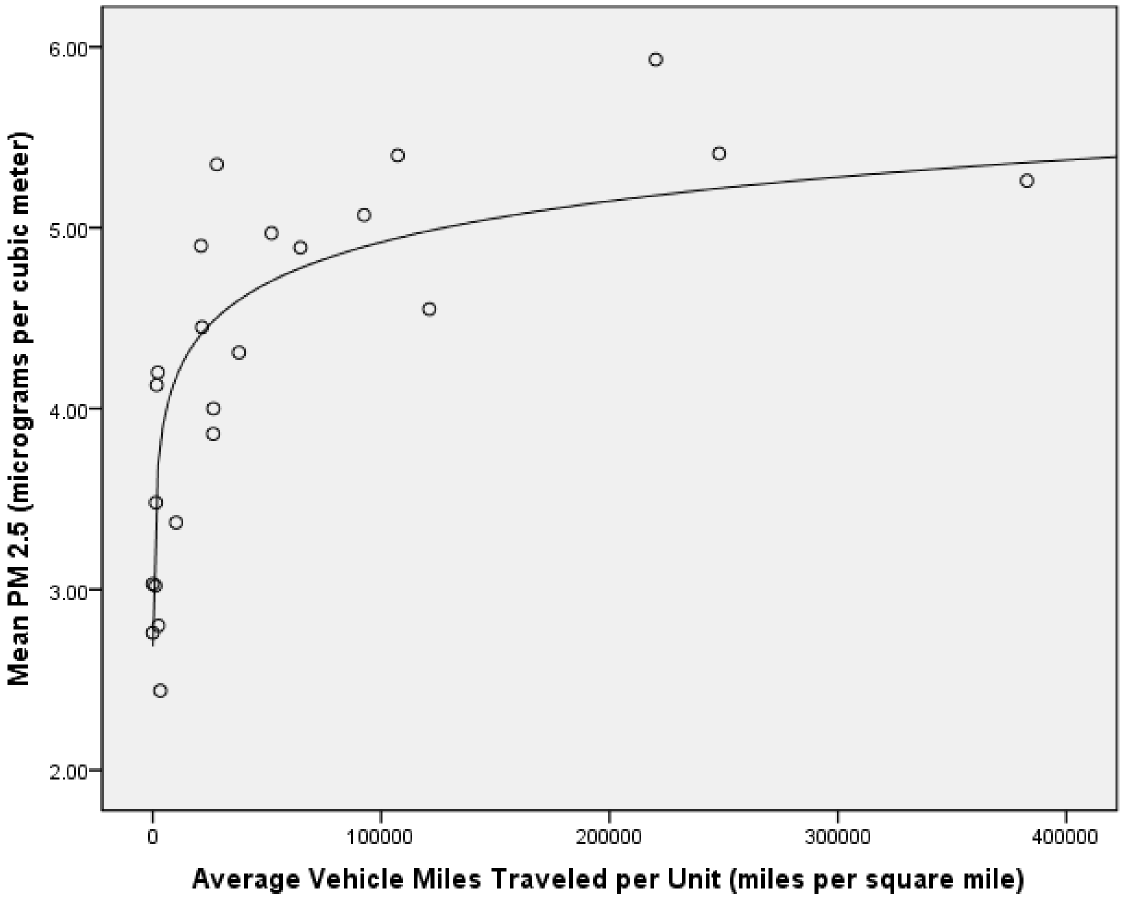

Figure 16: A scatter plot is showing the Average Vehicle Miles Traveled per Unit against the mean PM 2.5 surrounding the monitoring site locations. The Pearson's r correlation coefficient was 0.644, which is considered a strong positive correlation.

Table 16

Regression non-linear Model Coefficients without outlier (PM 2.5 & AVMTU)

Model Summary and Parameter Estimates (Dependent Variable: Mean PM 2.5)							
	Model Summary					Parameter Estimates	
Equation	R Square	F	df1	df2	Sig.	Constant	b1
Logarithmic	.664	41.539	1	21	.000	1.144	.328

The independent variable is AVMTU (miles/mile^2).

Table 16 provides the r^2 values. The r^2 value indicates how much of the total variation in the dependent variable, mean PM 2.5, can be explained by the independent variable, tree space area ratio. In this case, 66.4% can be explained. The p-value of less than 0.001 indicates that correlation is significant in this finding.

Also, the coefficients (table 16) provide us with the necessary information of the relations between mean PM 2.5 and AVMTU. Furthermore, we can use the values in the "Parameter Estimates" data to present the regression equation as:

Mean PM 2.5 = 1.144 + 0.328 x ln (AMVTU) + error

AMVTU is average vehicle miles traveled per unit.

When AMVTU is increasing, it produces a logarithmic growth in mean PM 2.5.

Rainfall and Mean PM 2.5

The average non-rainy day PM 2.5 from all of the monitoring sites taken over the entire duration of the study area was 4.59 micrograms per cubic meter of air (µg/m^3). That was 0.3µg/m^3 above the mean value (4.29µg/m^3). The average rainy day PM 2.5 from all rural and national park monitoring sites which were taken over the entire duration of the study area was 4.04µg/m^3. That was 0.25µg/m^3 below the overall mean.

Out of 24 monitoring sites, 23 sites showed the decrease. general PM 2.5 value after rainfall. Only one site showed the increase general PM 2.5 value after rainfall. However, this site (St. Louis Park - City Hall) is an outlier datum because the PM 2.5 data are recorded by every third-day duration and not every day. There is lack of data on some days. Figure 17 shows those results about 24 monitoring sites, and it also shows that rainfall can reduce PM 2.5 value. There is a correlation between rainfall and PM 2.5 value change shown in Table 17.

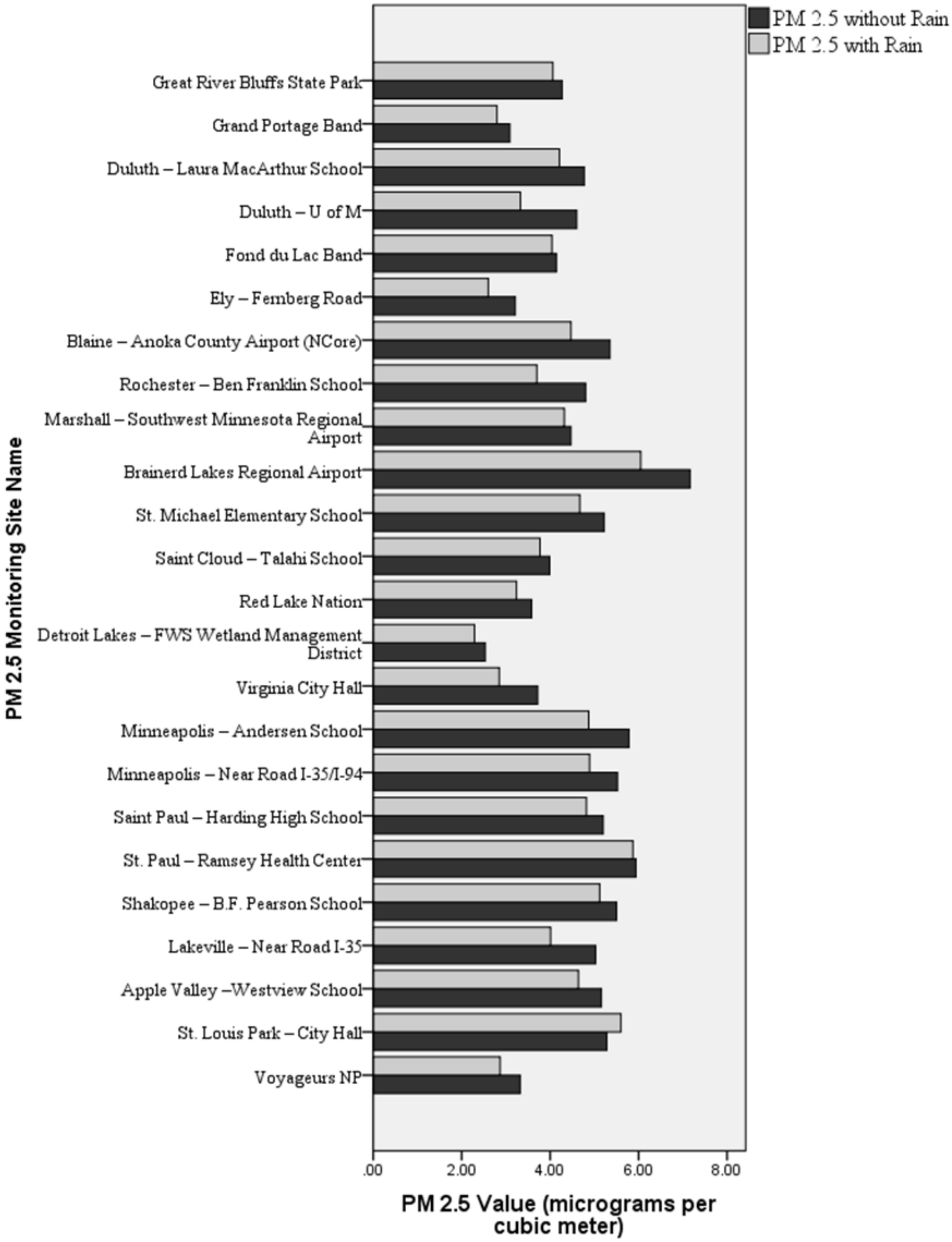

Figure 17: A bar chart is showing the rainy day PM 2.5 value and non-rainy day PM 2.5 value surrounding each monitoring site location.

Table 17

Pearson correlation between Rainfall and PM 2.5 Value Reduce

		Mean PM 2.5 Reduce	Mean Rainfall
Mean PM 2.5 Reduce	Pearson Correlation	1	.622**
	Sig. (2-tailed)		.006
Mean Rainfall	Pearson Correlation	.622**	1
	Sig. (2-tailed)	.006	

**. Correlation is significant at the 0.01 level (2-tailed).

Note: mean PM 2.5 change with rainfall is also called mean PM 2.5 reduction. Mean PM 2.5 change has two calculation forms. One indicates PM 2.5 on the day before the rainy day minus the PM 2.5 during the rainy day when the previous day is not a rainy day, and another indicates PM 2.5 on a rainy day minus the PM 2.5 during the second rainy day when the two days are both rainy days.

The results of that test (r^2=0.387, r = 0.622, n= 24, p= 0.006) indicated that a medium positive correlation exists between those variables (r= 0.622).

The Pearson Correlation, r = 0.622, indicated that a strong positive correlation exists between those variables. The r^2 value of 0.387 means that there is 38.7% of the variation in the mean PM 2.5 reduction which can be explained by differences in the total rainfall near the monitoring sites. There were 24 samples used in the calculation (n). The p-value of less than 0.01 indicates that correlation is significant in this finding. Table 17, Table 18, and Figure 18 summarizes the results.

Table 18

Regression model summary and Coefficients (PM 2.5 & Rainfall)

Model Summary				
Model	R	R Square	Adjusted R Square	Std. The error of the Estimate
1	.622[a]	.387	.358	.2284494837

a. Predictors: (Constant), Mean Rainfall

Coefficients [a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.005	.173		.286	.778
	Mean Rainfall	1.031	.335	.577	3.080	.006

a. Dependent Variable: Mean PM 2.5 reduce

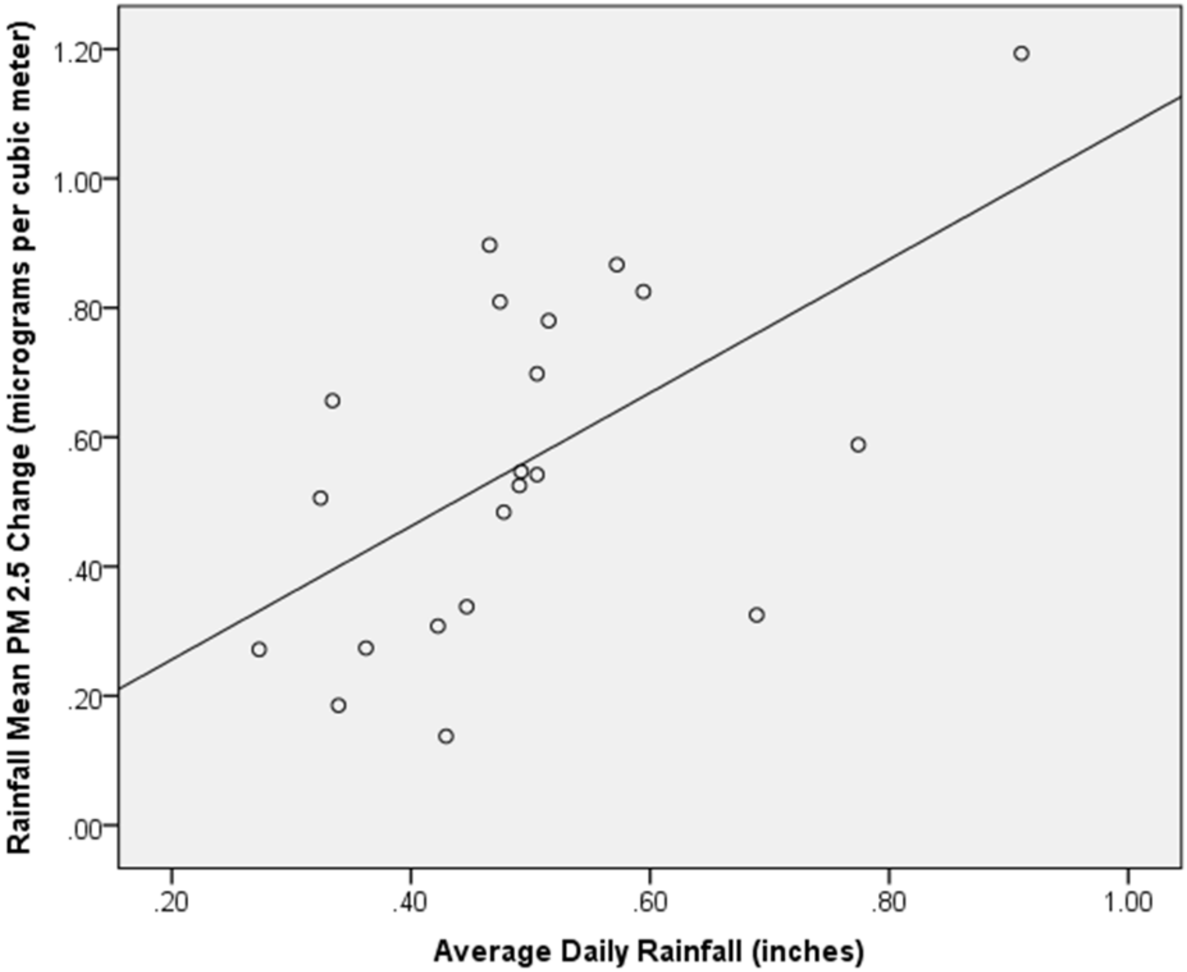

Figure 18: Rainfall and mean PM 2.5 reduction in rainfall

The Coefficients table 18 provides us with the necessary information to predict mean PM 2.5 value from total rainfall. The output p-value (0.000) for total rainfall is less than 0.050 so it can be assumed that the regression coefficient is significant. The output p-value (0.778) for the constant is greater than 0.050 so it can be assumed that the regression coefficient is not significant. The reason is that the real constant number should be greater than zero.

Mean PM 2.5, TSAR, AVMTU, and Rainfall

According to the previous separate correlation analysis between elements, the mean PM 2.5 has an association with TSAR and AVMTU. There is a negative correlation linear relationship between PM 2.5 and TSAR, and a positive correlation natural logarithmic relationship between PM 2.5 and AVMTU. In each correlation test, the Brainerd Lakes Regional Airport (270353204) monitoring site is shown as an outlier datum, so it is still out of the data when using multiple regression analysis. Table 19 and 20 show the results of linear regression analysis of three sets of correlations.

Table 19

Regression model summary and ANOVA analysis (three variables)

Model Summary				
Model	R	R Square	Adjusted R Square	Std. The error of the Estimate
1	.843[a]	.710	.665	.55079

a. Predictors: (Constant), TSAR, ln(AVMTU) , Rainfall

ANOVA [a]						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	14.143	3	4.714	15.540	.000[b]
	Residual	5.764	19	.303		
	Total	19.907	22			

a. Dependent Variable: Mean PM 2.5
b. Predictors: (Constant), TSAR, ln(AVMTU), Rainfall
c. TSAR is tree space area ratio.
d. AVMTU is Average Vehicle Miles Traveled per Unit.

The model (table 19) provides the r and adjusted r^2 values. The r value represents the simple correlation and is 0.843, which indicates a high degree of correlation. The adjusted r^2 value indicates how much of the total variation in the dependent variable, mean PM 2.5, which can be explained by the independent variable, TSAR, AVMTU, and rainfall. In this case, about 66.5% can be explained, which is large.

The ANOVA analysis (table 19) indicates that the regression model predicts the dependent variable significantly well. Here, F (3, 22) = 15.540, p = 0.000, which is less than 0.05, and indicates that, overall, the regression model significantly predicts the outcome variable.

Table 20

Regression Linear Model Coefficients (three variables)

Model [a]		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.179	.862		2.529	.020
	Rainfall	-.035	.019	-.248	-1.873	.077
	TSAR	-.010	.015	-.129	-.710	.486
	Ln(AVMTU)	.305	.070	.799	4.347	.000

a. Dependent Variable: Mean PM 2.5
b. TSAR is tree space area ratio.
c. VMTU is Average Vehicle Miles Traveled per Unit.

The Coefficients (table 20) provides us with the necessary information to predict mean PM 2.5 value from two variables. The output p-value for total rainfall is greater than 0.050 so it can be assumed that the regression coefficient is not significant. The output p-value for TSAR is greater than 0.050 so it can be assumed that the regression coefficient is not significant. The output p-value (p= 0.000) for ln (AVMTU) is less than 0.050 so it can be assumed that the regression coefficient is significant. Although it could not build the bicorrelation with those three variables based on the p-value to predict PM value, it still can give the general idea of the relationship. Also, the p-values show the three variables had different level correlations: AMVTU > total rainfall > tree space area ratio. Figure 19 shows the difference between regression analysis result and EPA data.

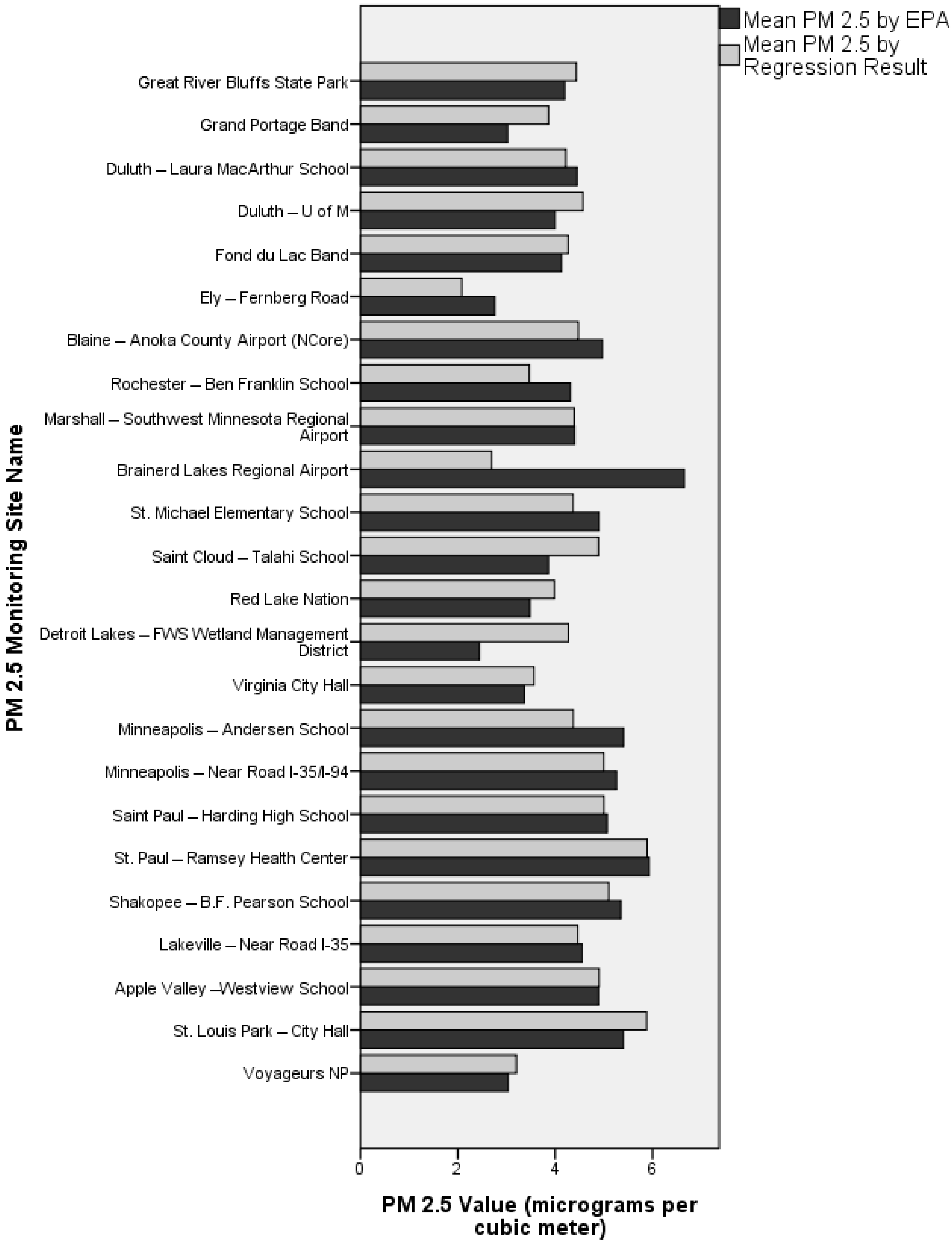

Figure 19: Regression analysis result and EPA mean PM 2.5

Airport Effect

The average PM 2.5 from the monitoring sites, which are close to an airport, taken over the entire duration of the study area was 5.34 micrograms per cubic meter of air (µg/m^3). It is higher than average PM 2.5 (4.29µg/m^3) from all sites. The result shows in Figure 20.

Brainerd Lakes Regional Airport (270353204) monitoring site: 6.65µg/m^3.

Blaine - Anoka County Airport (National Core) (270031002) monitoring site: 4.97µg/m^3.

Marshall - Southwest Minnesota Regional Airport (270834210) monitoring site: 4.40µg/m^3.

Figure 21 - 23, which were from 2015 Minnesota Geospatial Image Service, showed the surrounding environment of the three monitoring sites locations.

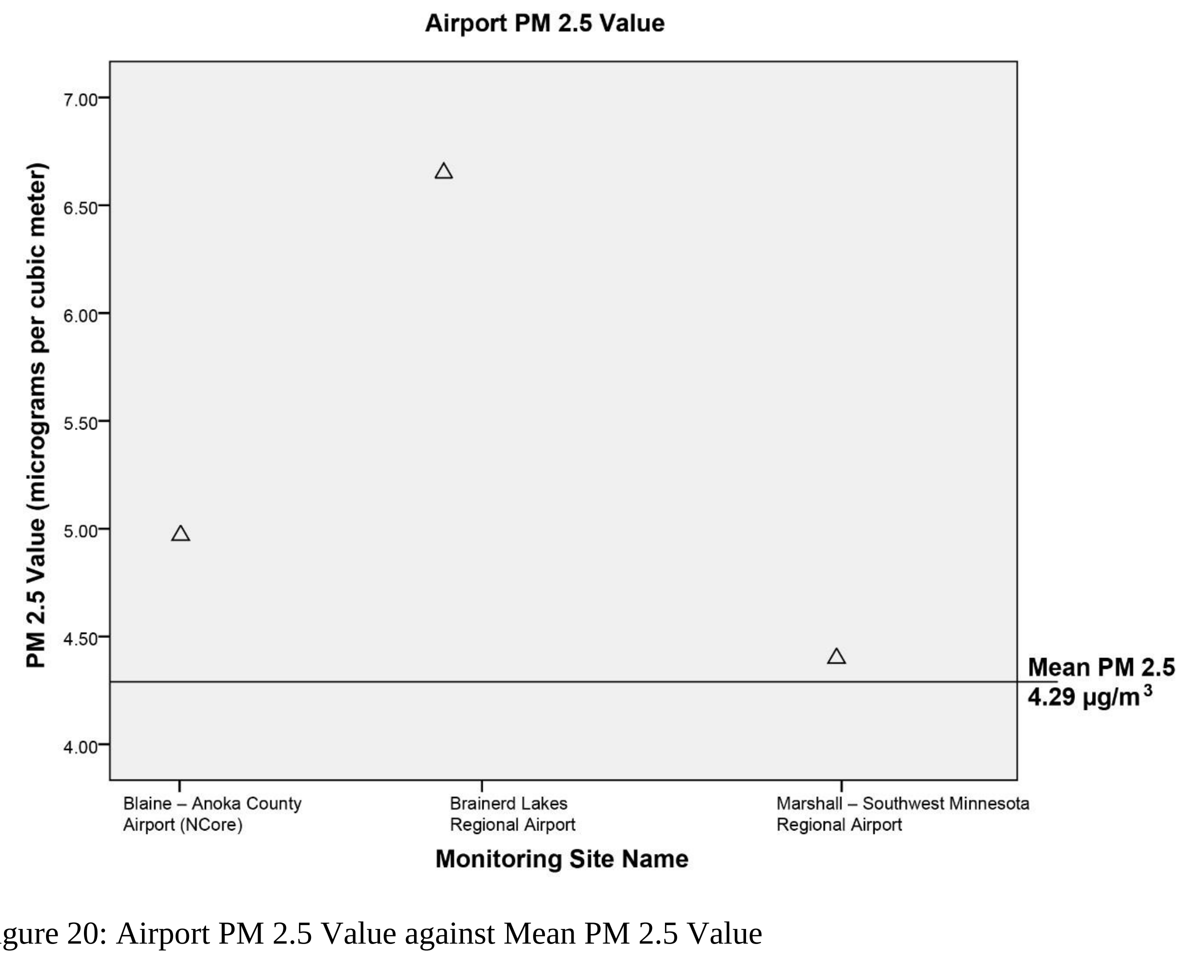

Figure 20: Airport PM 2.5 Value against Mean PM 2.5 Value

Figure 21: Brainerd Lakes Regional Airport

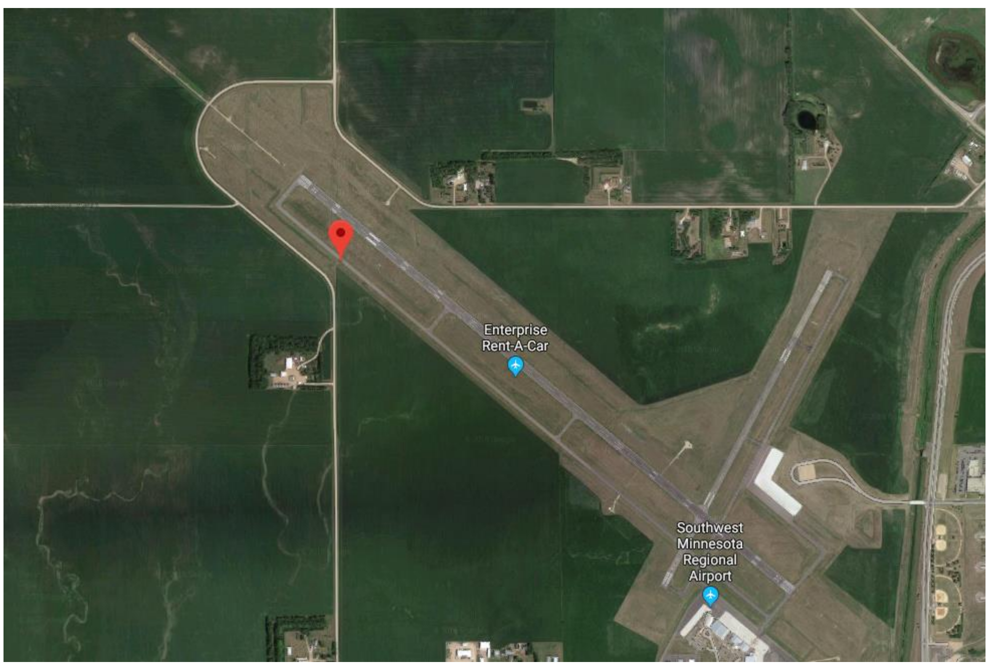

Figure 22: Marshall - Southwest Minnesota Regional Airport

Figure 23: Blaine – Anoka County Airport (NCore)

In this study, one Local aircraft operation, which aircraft operates in the local traffic pattern or within sight of the airport, was used in air traffic statistical analysis. According to the Federal Aviation Administration (FAA), the average daily air traffic data (2016) of those three airports were shown below.

Brainerd Lakes Regional Airport: average 104 aircraft /day.

Blaine - Anoka County Airport (National Core): average 79 aircraft /day.

Marshall - Southwest Minnesota Regional Airport: average 63 aircraft /day.

Table 21

Pearson correlation coefficient (PM 2.5 & air traffic)

		Mean PM 2.5	Air traffic
Mean PM 2.5	Pearson Correlation	1	.989
	Sig. (2-tailed)		.096
Air traffic	Pearson Correlation	.989	1
	Sig. (2-tailed)	.096	

The Pearson's r correlation coefficient was 0.989, which is considered as a very strong positive correlation. However, there were only three samples. Therefore, the Pearson correlation was not significant (Table 21).

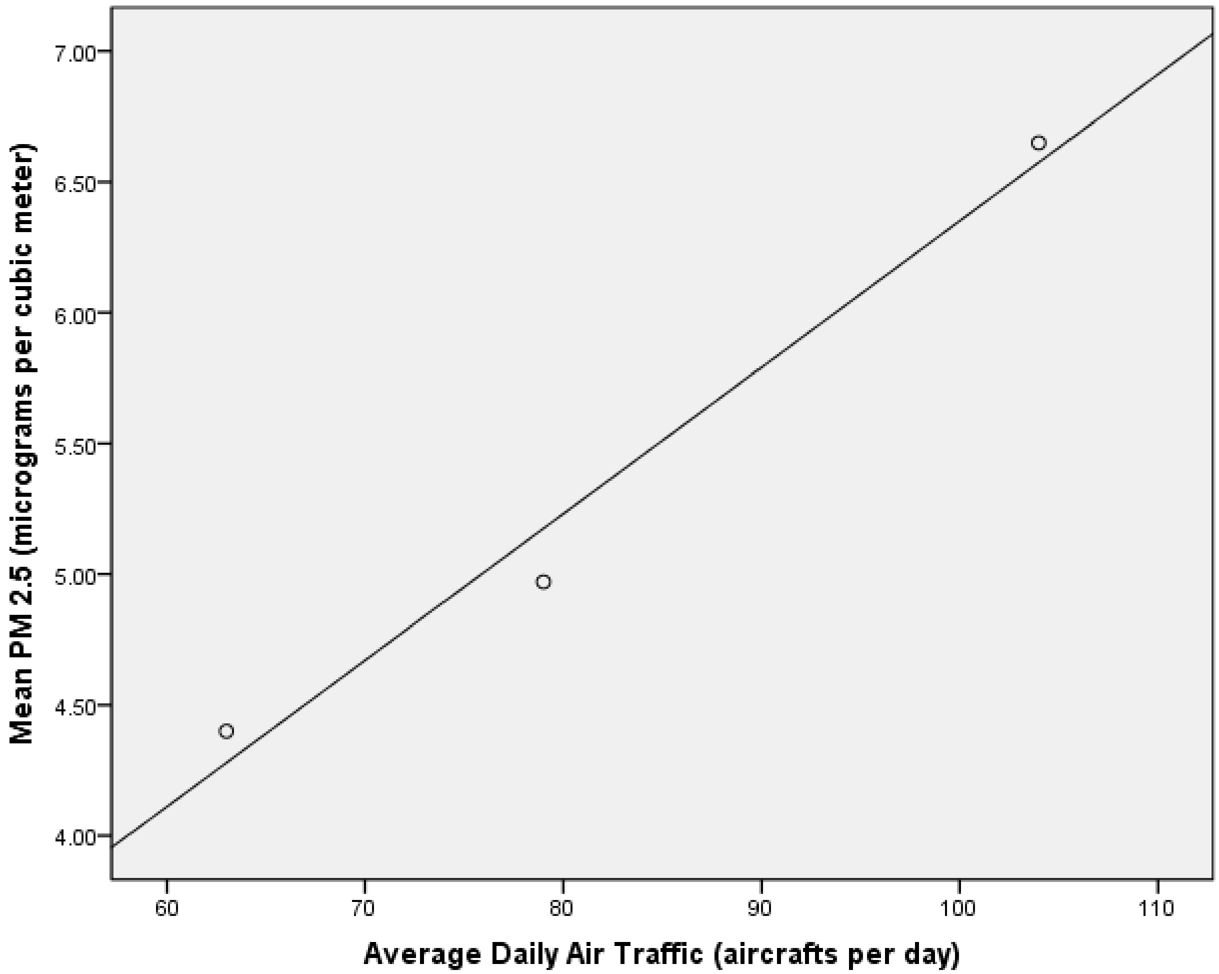

Figure 24: Mean PM 2.5 value and average daily air traffic (airport area)

The figure 24 shows that the mean PM 2.5 value increases when the average daily air traffic increases. However, the correlation is not a simple linear. Through the non-linear regression (curve estimation), it indicates a significantly quadratic relation (p (Sig.) = 0.000 < 0.01). The result is shown in Table 22 and Figure 25.

Table 22

Regression non-linear Model Coefficients (PM 2.5 & air traffic)

Model Summary and Parameter Estimates								
Dependent Variable: Mean PM 2.5								
	Model Summary					Parameter Estimates		
Equation	R Square	F	df1	df2	Sig.	Constant	b1	b2
Quadratic	1.000	.000	2	0	.000	5.989	-.074	.001

The independent variable is Air traffic (aircrafts per day).

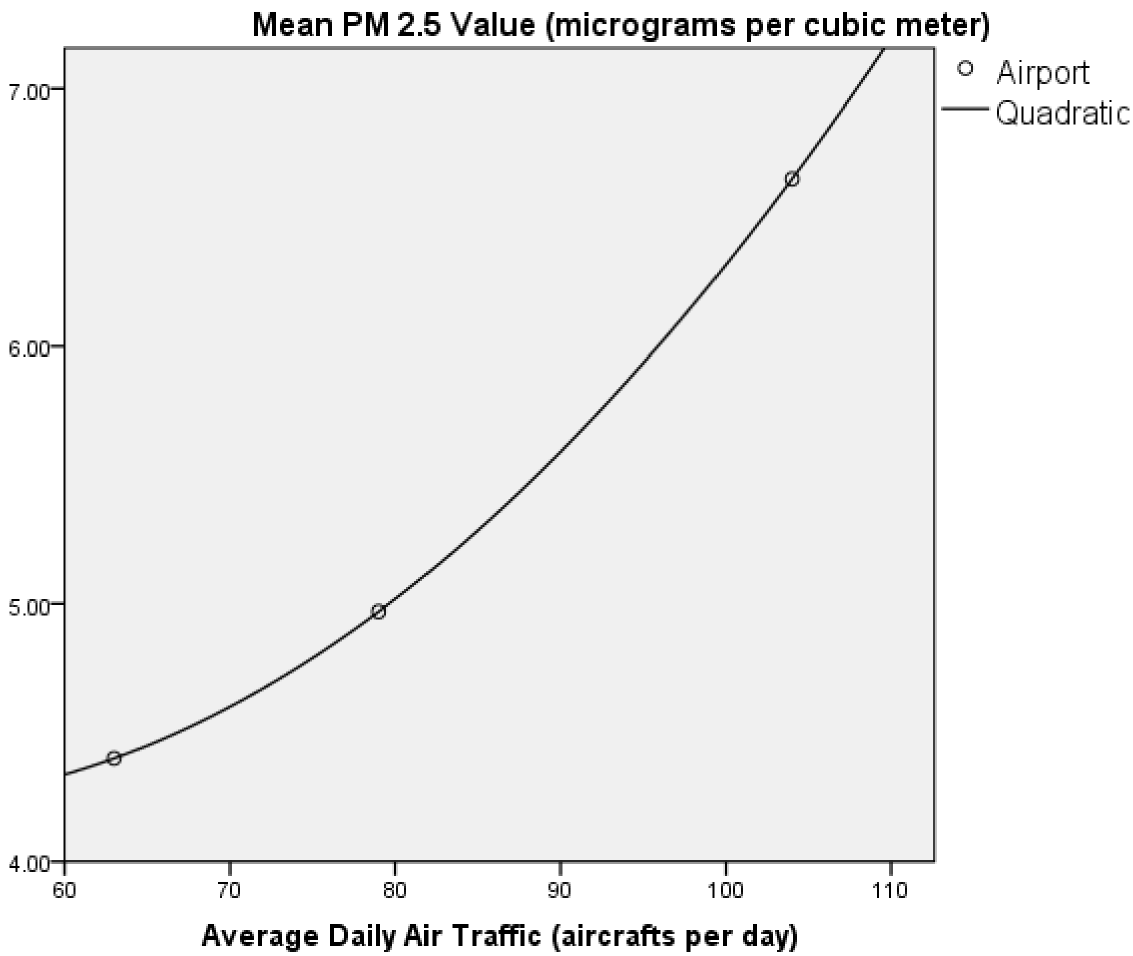

Figure 25: The non-linear relation between mean PM 2.5 value and average daily air traffic

Urban and Rural

According to AMVTU, the data can be classified into two groups, and the classification in AMVTU is similar to the classification in population by the U.S. Census Bureau. The U.S. Census Bureau identifies as an urban place, which has populations of 2,500 or more. A rural place has the population of less 2,500.

Group A: over 10,000 miles per square mile is belonging to the urban and suburban area.

Group B: less than 10,000 miles per square mile is belonging to the rural and national park area.

Urban PM 2.5 with three variables. The average PM 2.5 from all of the urban monitoring sites was 5.68µg/m^3, which was greater than the total average PM 2.5 (4.29µg/m^3). The results of Pearson correlation between PM 2.5 and three variables were shown in Table 23.

Table 23

Pearson correlations between Mean PM 2.5 and three variables for urban

		Mean PM 2.5	TSAR	Ln(AVMTU)
Mean PM 2.5	Pearson Correlation	1	-.863**	.617*
	Sig. (2-tailed)		.000	.025
TSAR	Pearson Correlation	-.863**	1	
	Sig. (2-tailed)	.000		
Ln (AVMTU)	Pearson Correlation	.617*		1
	Sig. (2-tailed)	.025		

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

		Mean PM 2.5 Change	Mean Rainfall
Mean PM 2.5 Change	Pearson Correlation	1	.759**
	Sig. (2-tailed)		.001
Mean Rainfall	Pearson Correlation	.759**	1
	Sig. (2-tailed)	.001	

**. Correlation is significant at the 0.01 level (2-tailed).

Mean PM 2.5 and TSAR. The Pearson correlations between Mean PM 2.5 and TSAR (r^2 = 0.754, r = - 0.863, p= 0.000) indicated that a strong negative correlation. The absolute r-value of 0.863 was greater than 0.705, which was all sites correlation. The r^2 value of 0.754 shows 75.4% of the variation in the mean PM 2.5 can be explained by differences in the percentage of tree space cover near the monitoring sites. Table 24 and Figure 26 summarize the results (PM 2.5 & TSAR).

Table 24

Regression coefficients between Mean PM 2.5 and TSAR for urban

Coefficients [a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5.680	.149		38.165	.000
	TSAR	-.080	.014	-.863	-5.666	.000

a. Dependent Variable: Mean PM 2.5

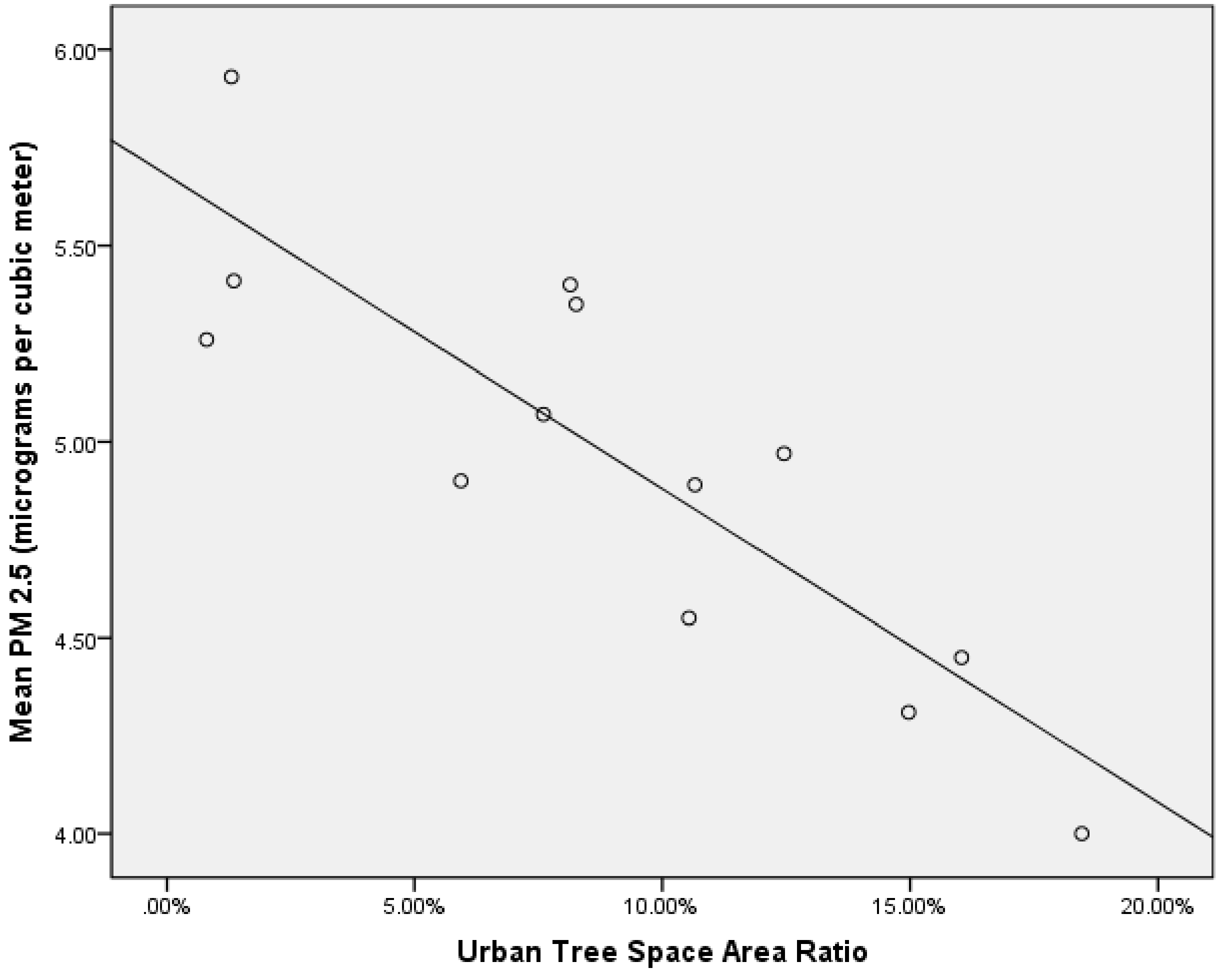

Figure 26: Mean PM 2.5 value and urban TSAR for urban

Mean PM 2.5 and traffic data. The Pearson correlations between Mean PM 2.5 and ln (AMVTU) (r^2=0.381, r = 0.617, p= 0.025) indicated that a strong positive correlation. The absolute r-value of 0.617 was less than 0.0.664, which was all sites correlation. The r^2 value of 0.0.381 shows 38.1% of the variation in the mean PM 2.5. Table 25 and Figure 27 summarize the results (PM 2.5 & ln (AMVTU)).

Table 25

Regression coefficients between Mean PM 2.5 and ln (AMVTU) for urban

Coefficients [a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.217	1.443		.844	.041
	Ln (AMVTU)	.335	.129	.617	2.603	.025

a. Dependent Variable: Mean PM 2.5

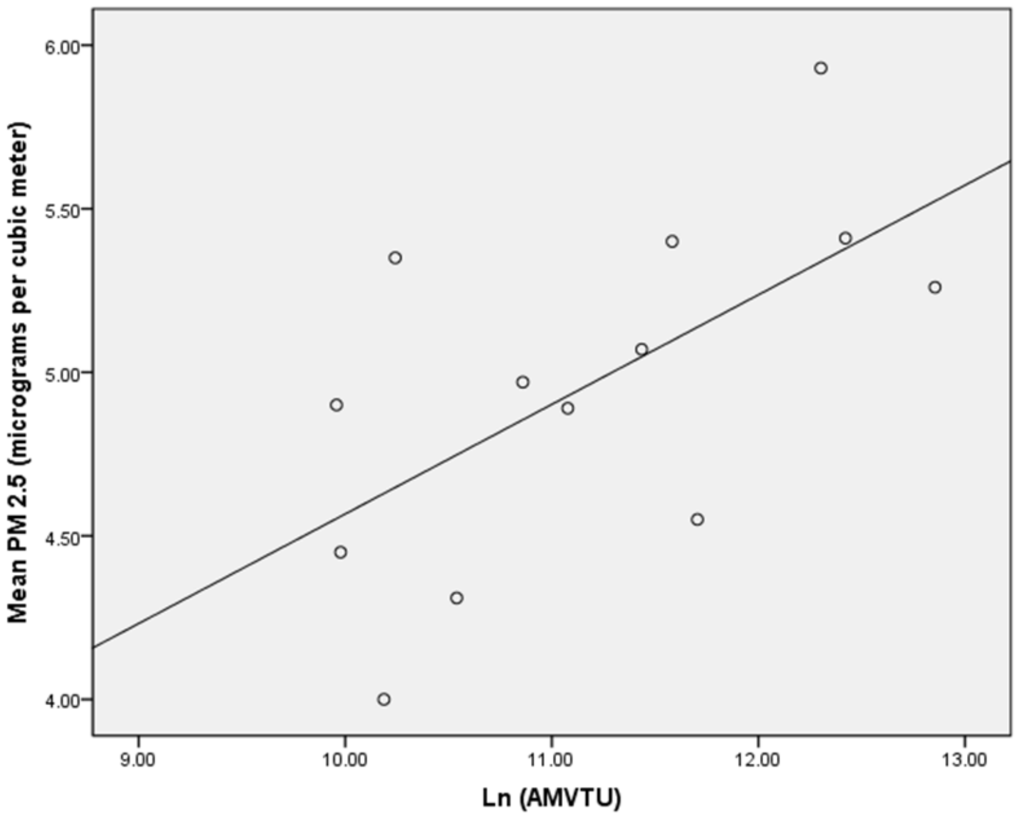

Figure 27: Mean PM 2.5 value and ln (AMVTU) for urban

Mean PM 2.5 change and rainfall. The Pearson correlations between Mean PM 2.5 and ln (AMVTU) (r^2= 0.577, r = 0.759, p= 0.001) indicated that a strong positive correlation. The r-value of 0.759 was greater than 0.622, which was all sites correlation. The r^2 value of 0.577 shows 57.7% of the variation in the mean PM 2.5 change with rainfall. Table 26 and Figure 28 summarize the results (PM 2.5 & rainfall).

Table 26

Regression coefficients between PM 2.5 change and rainfall for urban

Coefficients [a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.070	.199		.353	.730
	Mean Rainfall	1.426	.362	.658	3.025	.001

a. Dependent Variable: PM 2.5 Change with Rainfall

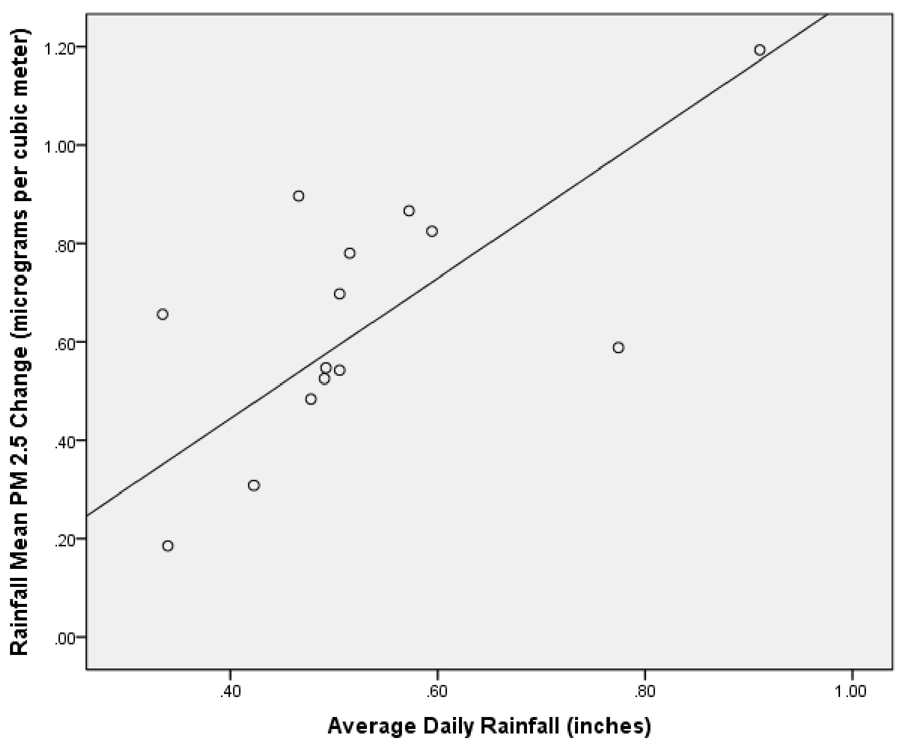

Figure 28: Mean PM 2.5 change with rainfall and rainfall for urban

Mean PM 2.5 and three variables. The r value represents the simple correlation and is 0.788, which indicates a high degree of correlation. However, the r value is less than 0.843, which is all sites correlation. The adjusted r^2 value indicates how much of the total variation in the dependent variable, mean PM 2.5, which can be explained by the independent variable, TSAR, AVMTU, and rainfall. In this case, about 51.8% can be explained, which is large. The ANOVA table indicates the model is significant because the significant value, 0.011, is less than 0.05. The coefficients table shows that there is only traffic value is significant, which is similar to the results of all sites coefficients. (Table 27)

Table 27

Regression model summary, ANOVA, and coefficients (PM 2.5 & three variables)

Model Summary				
Model	R	R Square	Adjusted R Square	Std. The error of the Estimate
1	.788[a]	.621	.518	.47796

ANOVA [a]						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.118	3	1.373	6.009	.011[b]
	Residual	2.513	11	.228		
	Total	6.631	14			

Coefficients [a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.186	2.146		.087	.932
	TSAR	-.020	.034	-.153	-.574	.577
	Traffic	.393	.165	.602	2.381	.036
	Rainfall	.908	.801	.227	1.134	.281

a. Dependent Variable: Mean PM 2.5

b. Predictors: (Constant), rainfall, traffic [ln(AMVTU)], TSAR

Rural PM 2.5 with three variables. The average PM 2.5 from all of the rural monitoring sites was 3.74µg/m^3, which was less than the total average PM 2.5 (4.29µg/m^3). In this study of rural sites, the two airport site had been removed as outliers: Marshall - Southwest Minnesota Regional Airport and Brainerd Lakes Regional Airport. The results of Pearson correlation between PM 2.5 and three variables were shown in Table 28.

Table 28

Pearson correlations between Mean PM 2.5 and three variables for rural

		Mean PM 2.5	TSAR	Ln(AVMTU)
Mean PM 2.5	Pearson Correlation	1	.432	.366
	Sig. (2-tailed)		.333	.419
TSAR	Pearson Correlation	.432	1	
	Sig. (2-tailed)	.333		
Ln (AVMTU)	Pearson Correlation	.366		1
	Sig. (2-tailed)	.419		

		Mean PM 2.5 Change	Mean Rainfall
Mean PM 2.5 Change	Pearson Correlation	1	-.119
	Sig. (2-tailed)		.800
Mean Rainfall	Pearson Correlation	-.119	1
	Sig. (2-tailed)	.800	

Mean PM 2.5 and TSAR. The Pearson correlations between Mean PM 2.5 and TSAR ($r = 0.432$, $p= 0.333 > 0.05$) indicated that the correlation is not significant. Table 28 and Figure 29 summarize the results (PM 2.5 & TSAR).

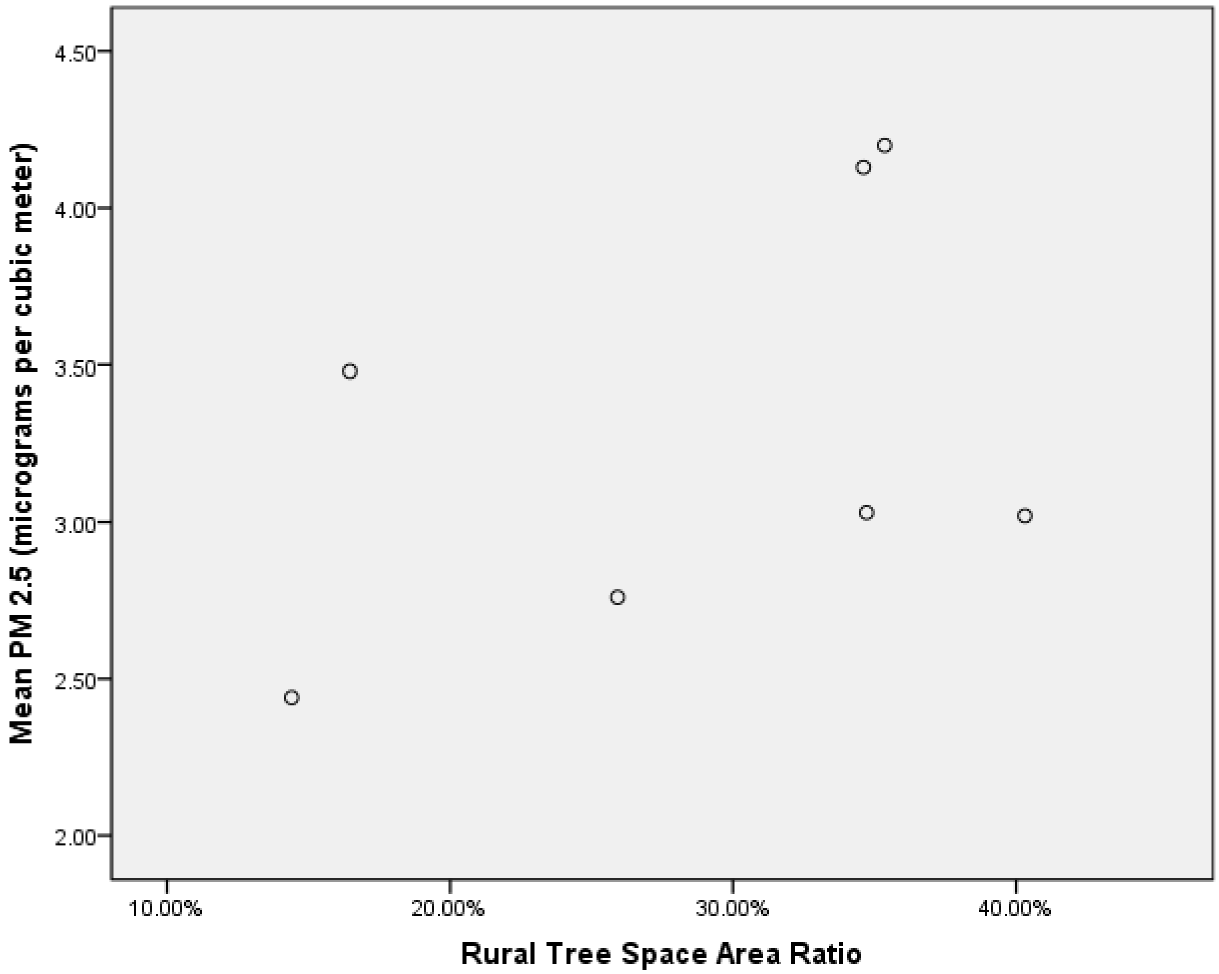

Figure 29: Mean PM 2.5 value and rural TSAR

Mean PM 2.5 and traffic data. The Pearson correlations between Mean PM 2.5 and ln (AMVTU) ($r = 0.366$, $p= 0.415 > 0.05$) indicated that the correlation is not significant. Table 28 and Figure 30 summarize the results (PM 2.5 & ln (AMVTU)).

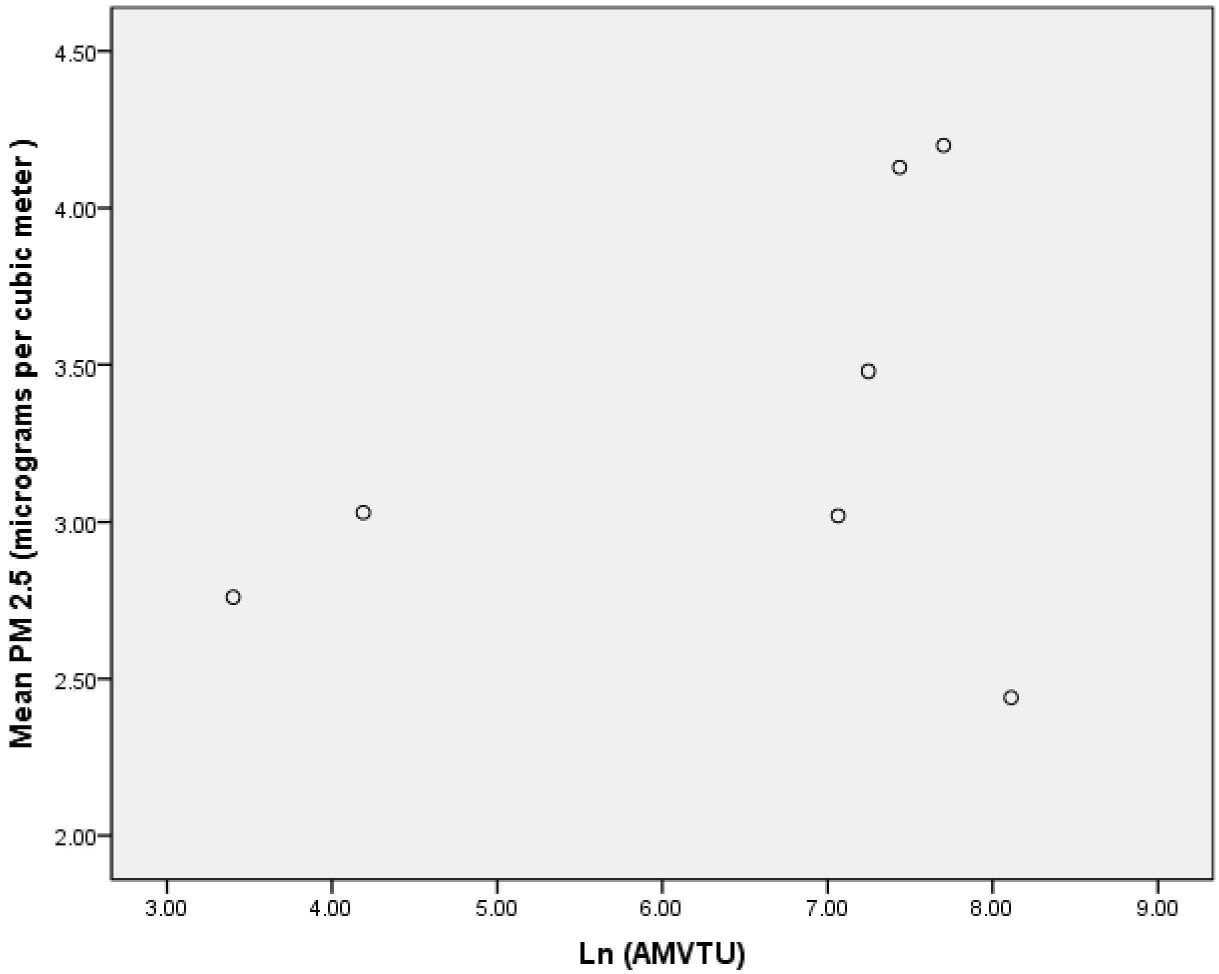

Figure 30: Mean PM 2.5 value and ln (AMVTU) for rural

Mean PM 2.5 change and rainfall. The Pearson correlations between Mean PM 2.5 and ln (AMVTU) (r = - 0.119, p= 0.800 > 0.05) indicated the correlation is not significant. Table 28 and Figure 31 summarize the results (PM 2.5 & rainfall).

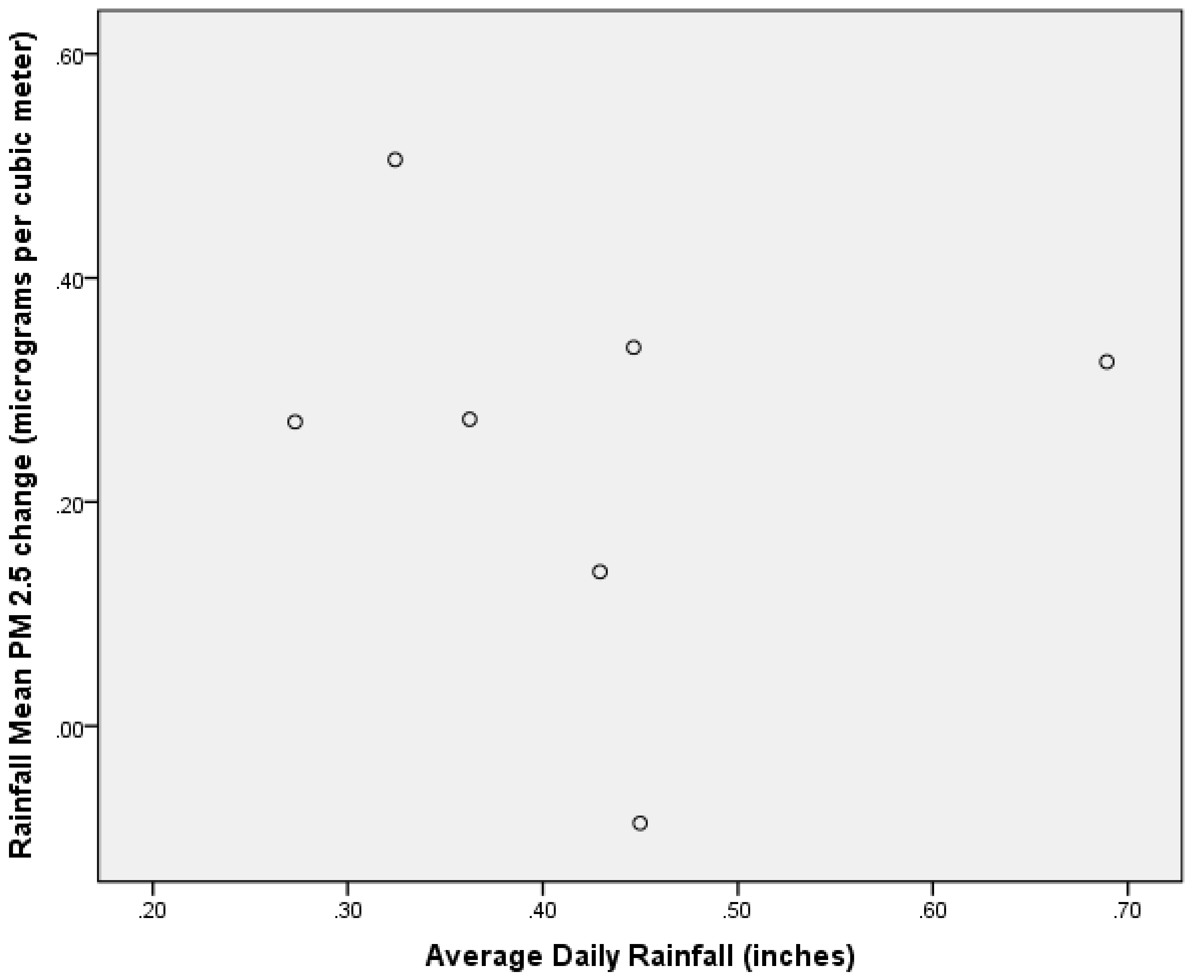

Figure 31: Mean PM 2.5 change with rainfall and rainfall for rural

Chapter V: Discussion

The results of this study show that there is an association between PM 2.5 and tree space area ratio, traffic volume, and rainfall in Minnesota urban areas. PM 2.5 values from the selected areas are negatively associated with increased percentages of tree space area; conversely, they are positively associated with increased traffic volume. PM 2.5 values from the selected areas decrease during periods of rainfall.

Tree Space Area Ratio Effects on PM 2.5

Setälä et al. (2012), Alaviippola and Pietarila (2011) found the tree-covered areas had about 20% less PM 2.5 than non-tree-covered areas. However, they did not delve into the correlations between them. According to Weber, Kowarik, and Säumela (2014), PM values are not significantly different if there are only grass-cover and no tree. In accordance with their findings, the results of this study show that PM 2.5 values from the selected areas are moderately negatively correlated with increased percentages of tree space area. The average PM 2.5 values for urban/ suburban areas were higher than the average PM 2.5 values for rural areas. This negative correlation indicates that tree-covered areas can reduce PM 2.5. While there are some rural monitoring sites with PM 2.5 values greater than or equal to some suburban area monitoring sites, it is reasonable to attribute these phenomena to near airport and industrial zones.

Traffic Volume Effects on PM 2.5

As the traffic volume (average vehicle miles traveled per unit) increases, the mean PM 2.5 increases logarithmically. Pantaleoni (2013) developed a model showing the relationship between traffic flow and carbon monoxide emissions in Tennessee. She found that there was a positive correlation between traffic volume and this vehicle pollutant. Although, Minnesota has a different traffic pattern from Tennessee, and the current study used a different vehicle pollutant (PM 2.5), the relationship between PM 2.5 and traffic volume is similar to the CO air pollution model described in Pantaleoni's research. It should be noted, however, that road dust could also be a contributing factor for the PM 2.5 values recorded in this study. Edvardsson and Magnusson (2009) found that PM 2.5 was positively correlated with road dust. Also, there was one monitoring site with especially higher PM 2.5 values than other sites due to the aforementioned airport effect.

Rainfall Effects on PM 2.5

Przybysz et al. (2014) and Honour et al. (2008) found that particulate matter is water-soluble, and rainfall can reduce the effect of particulate matter on vegetation. Therefore, it is assumed that rainfall can directly reduce PM 2.5 from the atmosphere before it is absorbed or settled by vegetation. The results of the current study support these findings as PM 2.5 values from the selected areas decreased during periods of rainfall. In fact, only one monitoring site produced the opposite consequence. The reason is that PM 2.5 data from this site (St. Louis Park - City Hall) were recorded every three days rather than every day. Since a lot of rainfall data were missing, it is hypothesized that this was the reason for the conflicting results.

In this study, 14 out of the 24 monitoring sites had a mean PM 2.5 value that matched EPA data. 14 out of the 24 monitoring sites were urban sites. However, regression analysis did not show bicorrelation between PM 2.5, tree space area, traffic value, and rainfall data (two of the p-values were greater than 0.05). When analyzing urban and rural separately, the urban area had a similar result buy the rural area did not show the relation. The data of rural area were random. The data of rural area did not show a direct relationship between PM 2.5 and three variables.

Airport Effect

There are two rural monitoring sites located in the airport area: the Brainerd Lakes Regional Airport monitoring site (6.65µg/m^3) and the Marshall - Southwest Minnesota Reginal Airport monitoring site (4.4µg/m^3). The Brainerd Lakes Regional Airport monitoring site is located between the airport terminal and the airport takeoff runway. The Marshall - Southwest Minnesota Reginal Airport monitoring site is located at the open field on the edge of the airport runway. There is another suburban monitoring site, Blaine - Anoka County Airport (NCore), that is also located in the airport area, but it is situated at the edge and closed to the commercial area. The aviation fuel is a possible reason for the higher measured PM 2.5 values. According to the analysis of the relationship between PM 2.5 and air traffic, the strong positive correlation was found in the airport. PM 2.5 values increase with increasing aircrafts number.

Industrial Effect

There is one monitoring site in the rural area, but the surrounding area has several large wood products industries. It is the Fond du Lac Band monitoring site (4.13µg/m^3). However, the industrial pollution and higher truck ratio of the traffic flow can explain the higher PM 2.5 value.

Outliers

When the PM 2.5 values were analyzed by the daily average, some contained value beyond two standard deviations from the mean. One outlier data occurs at the Brainerd Lakes Regional Airport monitoring site. When analyzing PM 2.5 and tree space area, or PM 2.5 and traffic value, the PM 2.5 standard deviation was 1.04. The Brainerd Lakes Regional Airport monitoring site average PM 2.5 was 6.65µg/m^3 higher than all other monitoring sites (mean PM 2.5 = 4.29µg/m^3). This monitoring site was located in a rural area, which had a higher tree space area ratio (21.56%) and had a lower traffic value (5282 miles/mile^2). The mean PM 2.5 value for this site should be less than 4.29μg/m^3. However, the PM 2.5 value of this site was 6.65μg/m^3 higher than 4.29μg/m^3. Since it was located in an airport area, the airport effect present at that location may explain why data were not included in this study. Another outlier occurs at the Detroit Lakes - FWS Wetland Management District monitoring site. According to the PM 2.5 data from 2010 to 2015, and the 2017 summer section, the average PM 2.5 for this site was 4.40μg/m^3. 2016 summer data (2.44μg/m^3) were much lower than 4.40μg/m^3. According to EPA, the Lower Detection Limit (LDL) in the PM 2.5 Federal Reference Method is 0.80μg/m^3. The 2016 summer section had 18% of PM 2.5 data, which were lower than the LDL. Therefore, this site was not used in current study.

Limitations

In addition to the limitations, which arose upon the outlier data, there are numerous other factors which must be addressed. The first limitation is in regard to the location of the monitoring sites and traffic recording equipment. Although the traffic volume of all selected monitoring site areas can be calculated by using MNDOT traffic data, the MNDOT traffic data are not well matched with the MPCA monitoring sites. Therefore, some data are estimated by using the nearby automatic traffic recorder and average annual daily traffic data. Data collection duration was another limitation. The collection frequency of most data in this study is daily. It is acceptable for studying the relation between PM 2.5 and traffic volume data. However, it is not good enough for the relation between PM 2.5 and rainfall data. The best duration should be hourly. The hourly data can identify the PM 2.5 change before, during, and after rainfall. The hourly data will be useful to determine the relationship between PM 2.5 and rainfall. For instance, it was estimated that there was no rain on the 9th of July 2016 but rain on the next day. The rain duration was from 5 a.m. to 8 a.m. Then, the PM 2.5 value showed a downward trend in these three hours. However, PM 2.5 would begin to slowly rise to normal levels after the rain. Since the frequency of the data is daily, it is only possible to compare the difference between July 9th and 10th. Therefore, the recorded data of PM 2.5 reduction with rain obtained here should be less than the actual situation. If the data record was hourly, relatively accurate changes in PM 2.5 data could be obtained. The time frame of tree canopy data is the third limitation. In this study, the time frame of PM 2.5 data, traffic data, and rainfall data is from May 2016 to October 2016, but the time frame of the National Land Cover Database (NLCD) from Multi-Resolution Land Characteristics (MRLC) Consortium is from May 2011 to October 2011. 2011 NLCD may not completely match the real land use of 2016. 2016 NLCD is not available. According to the MRLC Consortium, NLCD is made by Landsat imagery (which is used to calculate NDVI) and it has a higher accuracy in the context of 2011. In this study, 2017 NDVI imageries of two monitoring site areas were compared with 2011 NLCD, and the differences of two areas between 2017 and 2011 are respectively 1.7% and 1.8%, which are less than 5% difference requirement. According to Jovanović, Milanović, and Zorn (2018), 2011 NLCD is acceptable. Therefore, the 2011 NLCD can be used in 2016, but the best choice is to use the entire 2017 NDVI or 2016 NLCD to replace the 2011 NLCD. The last limitation is traffic type. MNDOT did not have separate data for different vehicle types, such as passenger vehicles, trucks, and commercial vehicles. Different vehicles had different emissions, and they would also cause different road dust (Edvardsson & Magnusson, 2009). If the vehicle types can be identified, it will help to determine the correlation between PM 2.5 and traffic volume accurately.

Future Studies

This study attempted to examine the relationships between PM 2.5 and tree space area, traffic volume, and rainfall in Minnesota. There are still many uncertainties in our understandings of some of these very complicated relationships. Future studies could expand on this in several ways. One way would be to utilize the same data and incorporate more variables into the calculations to see whether other correlations appeared. For example, if I can identify the tree area change between summer and winter, it is possible to determine the PM 2.5 changes in the similar traffic pattern. Another way would be to change the size of the buffered area surrounding the monitoring site to determine whether the correlations identified remain. In this study, the buffer selected is based on the monitoring site measurement scale from MPCA, and it is acceptable for daily data. However, it would have a better choice of buffer size for this study when hourly data is used in this research. To build our associated monitoring equipment based on MPCA monitoring sites in the same selected areas is another way to compare with the existing data to search for patterns which this study overlooked. For instance, MPCA has a roadside site; I can build the tree-covered site in the same selected area. This new research may provide useful information.

Chapter VI: Conclusion

The study had not been able to successfully build a multivariate model to match Minnesota PM 2.5 pollution. However, results from this study show that PM 2.5 has the correlations with tree space area, traffic volume, and rainfall. The correlation analysis results between PM2.5 concentration and three variables (tree space area, traffic volume, and rainfall) showed that tree space area ratio had a negative, traffic volume had a positive and rainfall had a negative, correlation with PM2.5.

According to the results, several conceptual assumptions can be made:

In general, urban areas have higher mean PM 2.5 values than rural areas because of the lower tree covered area and higher traffic volume.

Areas with more tree coverage have lower PM 2.5 values in urban.

A higher traffic volume causes a higher PM 2.5 value.

Rainfall can reduce PM 2.5 in the air. The PM 2.5 value will decrease during rainfall.

The correlations between PM 2.5 and tree space area, traffic volume, and rainfall are significant in urban, but not in rural.

A higher air traffic volume causes a higher PM 2.5 value in airport areas.

References

Adedeji, O., Oluwafunmilayo, O., & Oluwaseun, T. (2016). Mapping of traffic-related air pollution using GIS Techniques in Ijebu-Ode, Nigeria. *Indonesian Journal of Geography*, 48(1), 76-86.

Alaviippola, B., & Pietarila, H. (2011). Ilmanlaadun alustava arviointi suomessa pienhiukkaset (PM 2.5). (Air quality assessment of fine particles (PM 2.5) in Finland.) *Helsinki: Finnish Meteorological Institute*.

Armbrust, D.V. (1986). Effect of particulates (dust) on cotton growth, photosynthesis, and respiration. *Agronomy Journal (USA)*, (6), 1078-1081.

Beckett K., Freer-Smith, P., &Taylor, G. (2000). Effective tree species for local air-quality management. *Journal of Arboriculture*, 26(1), 12-19.

Brantley, H., Hagler, G., Baldauf, R., & Deshmukh, P. (2014). Field assessment of the effects of roadside vegetation on near-road black carbon and particulate matter. *Science of the Total Environment*, 468-469, 120-129.

Burnett, R., Brook, J., Dann, T., Delocla, C., Philips, O., Cakmak, S., Vincent, R., Goldberg, M., & Krewski, D. (2000). Association between particulate and gas-phase components of urban air pollution and daily mortality in eight Canadian cities. *Inhalation Toxicology*, 12(4), 15-39.

Cavanagh, J., Zawar-Reza, P., & Wilson, J. (2009). Spatial attenuation of ambient particulate matter air pollution within an urbanized native forest patch. *Urban Forestry and Urban Greening*, 8(1), 21–30.

Chen, B., Lu, S., Zhao, Y., Li, S., Yang, X., Wang, B., & Zhang, H. (2016). Pollution Remediation by Urban Forests: PM 2.5 Reduction in Beijing, China. *Polish Journal of Environmental Studies*, 25(5), 1873-1881.

Chen, C., Wu, C., Yu, H., Chan, C., & Cheng, T. (2012). Spatiotemporal modeling with temporal-invariant variogram subgroups to estimate fine particulate matter PM 2.5 concentrations. *Atmospheric Environment*, 541-8.

Chen, C. H. (ed.) (2008). *Image processing for remote sensing*. New York.

Chen, Z., Xie, X., Cai, J., Chen, D., Gao, B., He, B., Cheng, N., & Xu, B. (2017). Understanding meteorological influences on PM2.5 concentrations across China: a temporal and spatial perspective. *Atmospheric Chemistry & Physics Discussions*, 1-30.

Cordioli, M., Ranzi, A., Freni Sterrantino, A., Erspamer, L., Razzini, G., Ferrari, U., Gatti, M.G., Bonora, K., Artioli, F., Goldoni, C.A., & Lauriola, P. (2014). Case Study: A comparison between self-reported and GIS-based proxies of residential exposure to environmental pollution in a case-control study on lung cancer. *Spatial and Spatio-Temporal Epidemiology*, 9, 37-45.

Edvardsson, K., & Magnusson, R. (2009). Monitoring of dust emission on gravel roads: Development of a mobile methodology and examination of horizontal diffusion. *Atmospheric Environment*, *43*(4), 889-896.

Environmental Protection Agency. (n.d.). National Ambient Air Quality Standards Table. *Environmental Protection Agency*, Retrieved from https://www.epa.gov/criteria-air-pollutants/naaqs-table

Freer-Smith, P.H., Beckett, K., & Taylor, G. (2005). Deposition velocities to Sorbus aria, Acer campestre, Populus deltoides × trichocarpa 'Beaupre', Pinus nigra and × Cupressocyparis leylandii for coarse, fine and ultra-fine particles in the urban environment. *Environmental Pollution*, 133(1), 157-167.

Freer-Smith P.H., El-Khatib, A., & Taylor, G. (2004). Capture of particulate pollution by trees: a comparison of species typical of semi-arid areas (Ficus nitida and Eucalyptus globulus) with European and North American species. *Water, Air, and Soil Pollution*, 155(1-4), 173–187.

Franklin, S. (2001). *Remote sensing for sustainable forest management*. New York.

Grigg, J. (2011). Air pollution and children's respiratory health - gaps in the global evidence. *Clinical and Experimental Allergy: Journal of the British Society for Allergy and Clinical Immunology*, 41(8), 1072-1075

Gromke, C. (2011). A vegetation modeling concept for building and environmental aerodynamics wind tunnel tests and its application in pollutant dispersion studies. *Environmental Pollution*, 159 (8-9), 2094-2099.

Holford, T., Ebisu, K., McKay, L., Gent, J., Triche, E., Bracken, M., & Leaderer, B. P. (2010). Integrated exposure modeling: a model using GIS and GLM. *Statistics in Medicine*, 29(1), 116-129.

Honour, S., Bell, J., Ashenden, T., Cape, J., & Power, S. (2008). Responses of herbaceous plants to urban air pollution: Effects on growth, phenology and leaf surface characteristics. *Environmental Pollution*, 157(The Behaviour and Effects of Nanoparticles in the Environment), 1279-1286.

Huang, R., Zhang, Y., Bozzetti, C., Ho, K., Cao, J., Han, Y., Daellenbach, K.R., Slowik, J.G., Platt, S.M., & Canonaco, F. (2014). High secondary aerosol contribution to particulate pollution during haze events in China. *Nature*, 514 (7521), 218–222.

Jeanjean, A., Leigh, R., & Monks, P. (2016). Modeling the effectiveness of urban trees and grass on PM2.5 reduction via dispersion and deposition at a city scale. *Atmospheric Environment*, 147, 1-10.

Ji, J., Wang, G., Du X., Jin, C., Yang H., Liu, J., Yang, Q., Josine, T., Li, J., & Chang C. (2013). Evaluation of absorbing haze PM 2.5 fine particulate matters with plants in Beijing-Tianjin-Hebei Region in China. *Scientia Sinica Vitae*, 43 (8), 694-699.

Jovanović, M., Milanović, M., & Zorn, M. (2018). The use of NDVI and CORINE Land Cover databases for forest management in Serbia. *Geografski Zbornik /Acta Geographica Slovenica*, 58(1), 109-123.

Kriegler, F., Malila, W., Nalepka, R., & Richardson, W. (1969). Preprocessing transformations and their effects on multispectral recognition. *Proceedings of the Sixth International Symposium on Remote Sensing of Environment*, II, 97-131.

Lai, L. (2018). The influence of urban heat island phenomenon on PM concentration: an observation study during the summer half-year in metropolitan Taipei, Taiwan. *Theoretical & Applied Climatology*, 131(1-2), 227-243.

Lee, H., Olsen, S., Wuebbles, D., & Youn, D. (2013). Impacts of aircraft emissions on the air quality near the ground. *Atmospheric Chemistry and Physics*, 13(11), 5505-5522.

Litschke, T., & Kuttler, W., (2008). On the reduction of urban particle concentration by vegetation - A review. *Meteorologische Zeitschrift*, 17, 229-240.

McCarty, M., Hafner, H., & Montzka, S. (2006). Background concentration of 18 air toxics for North America. *Journal of the Air & Waste Management Association*, 56(1), 3-11.

Milman, O. (2015). Call for action on pollution as emissions linked to respiratory illnesses double. *The Guardian*. Retrieved from https://www.theguardian.com/environment/2015/apr/02/call-for-action-on-pollution-as-emissions-linked-to-respiratory-illnesses-double

Olmo, N., Saldiva, P., Braga, A., Lin, C., Santos, U., & Pereirai, L. (2011). A review of low-level air pollution and adverse effects on human health: implications for epidemiological studies and public policy. *Clinics*, 66(4), 681-690.

Oštir, K. (2006). *Daljinsko zaznavanje (Remote sensing)*. Ljubljana: Založba.

Pagiola, S., Landell-Mills, N., & Bishop, J. (2002). *Selling forest environmental services*. London.

Pantaleoni, E. (2013). Combining a road pollution dispersion model with GIS to determine carbon monoxide concentration in Tennessee. *Environmental Monitoring and Assessment*, 185(3), 2705-2722.

Przybysz, A., Sæbø, A., Hanslin, H., & Gawroński, S. (2014). Accumulation of particulate matter and trace elements on vegetation as affected by pollution level, rainfall and the passage of time. *Science of the Total Environment*, 481, 360-369.

Sargazi, S., Taheri Shahraiyni, H., Habibi-Nokhandan, M., & Sanaeifar, M. (2011). Application of GIS for the modeling of spatial distribution of air pollutants in Tehran. *Proceedings of SPIE - The International Society for Optical Engineering*, 8181(Earth Resources and Environmental Remote Sensing/GIS Applications II)

Seo, S., Yoo, Y., Choung, J., Kim, D., Min, S., & Paul, C. (2016). GIS-based association between PM 10 and allergic diseases in Seoul: Implications for health and environmental policy. *Allergy, Asthma and Immunology Research*, 8(1), 32-40.

Setälä, H., Viippola, V., Rantalainen, A., Pennanen, A., & Yli-Pelkonen, V. (2013). Does urban vegetation mitigate air pollution in northern conditions? *Environmental Pollution*, 183, 104-112.

Sæbø, A., Popek, R., Nawrot, B., Hanslin, H., Gawronska, H., & Gawronski, S. (2012). Plant species differences in particulate matter accumulation on leaf surfaces. *Science of the Total Environment*, 427–428, 347–354.

Tashayo, B., & Alimohammadi, A. (2016). Modeling urban air pollution with optimized hierarchical fuzzy inference system. *Environmental Science and Pollution Research*, (19), 19417-19431.

Van Heerden P., Krüger G., & Kilbourn Louw, M. (2007). Dynamic responses of photosystem II in the Namib Desert shrub, Zygophyllum prismatocarpum, during and after foliar deposition of limestone dust. *Environmental Pollution*, 146(1), 34-45.

Vos, P., Maiheu, B., Vankerkom, J., & Janssen, S. (2013). Improving local air quality in cities: to tree or not to tree? *Environmental Pollution*, 183, 113-122.

Wang, J., & Ogawa, S. (2015). Effects of Meteorological Conditions on PM2.5 Concentrations in Nagasaki, Japan. *International Journal of Environmental Research and Public Health*, 12 (8), 9089-9101.

Weber, F., Kowarik, I., & Säumel, I. (2014). Herbaceous plants as filters: immobilization of particulates along urban street corridors. *Environmental Pollution*, 186, 234-240.

Weier, J. and Herring, D. (2000) Measuring Vegetation (NDVI & EVI). *NASA Earth Observatory*, Washington DC. Retrieved from https://earthobservatory.nasa.gov/Features/MeasuringVegetation/

World Bank. (2009). Mongolia: Air pollution in Ulaanbaatar - Initial assessment of current situations and effects of abatement measures (English). Retrieved from http://documents.worldbank.org/curated/en/866561468274261208/Mongolia-Air-pollution-in-Ulaanbaatar-initial-assessment-of-current-situation-and-effects-of-abatement-measures

Appendix

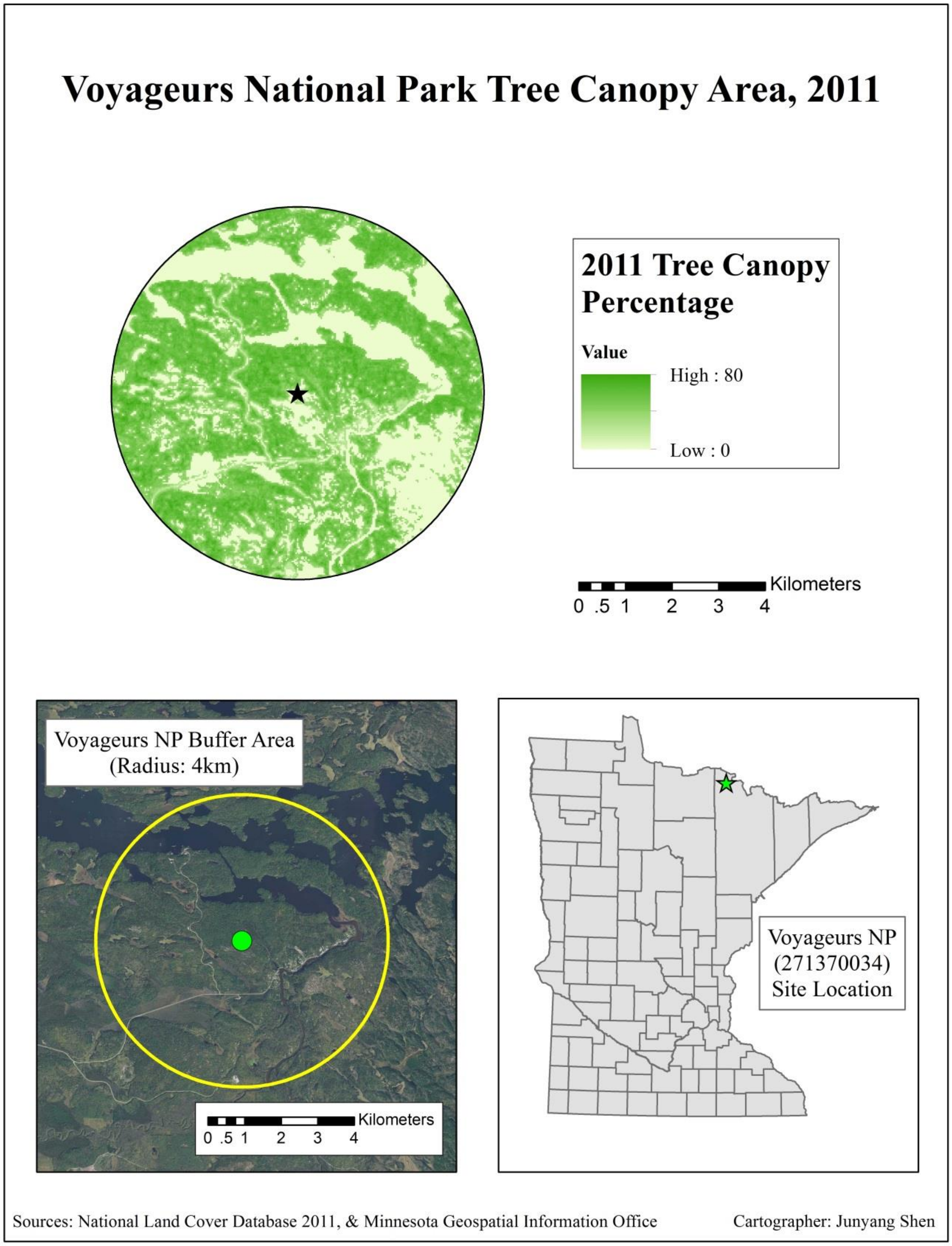

Figure I: 2011 Voyageurs National Park site: tree canopy, buffer size and monitoring site location

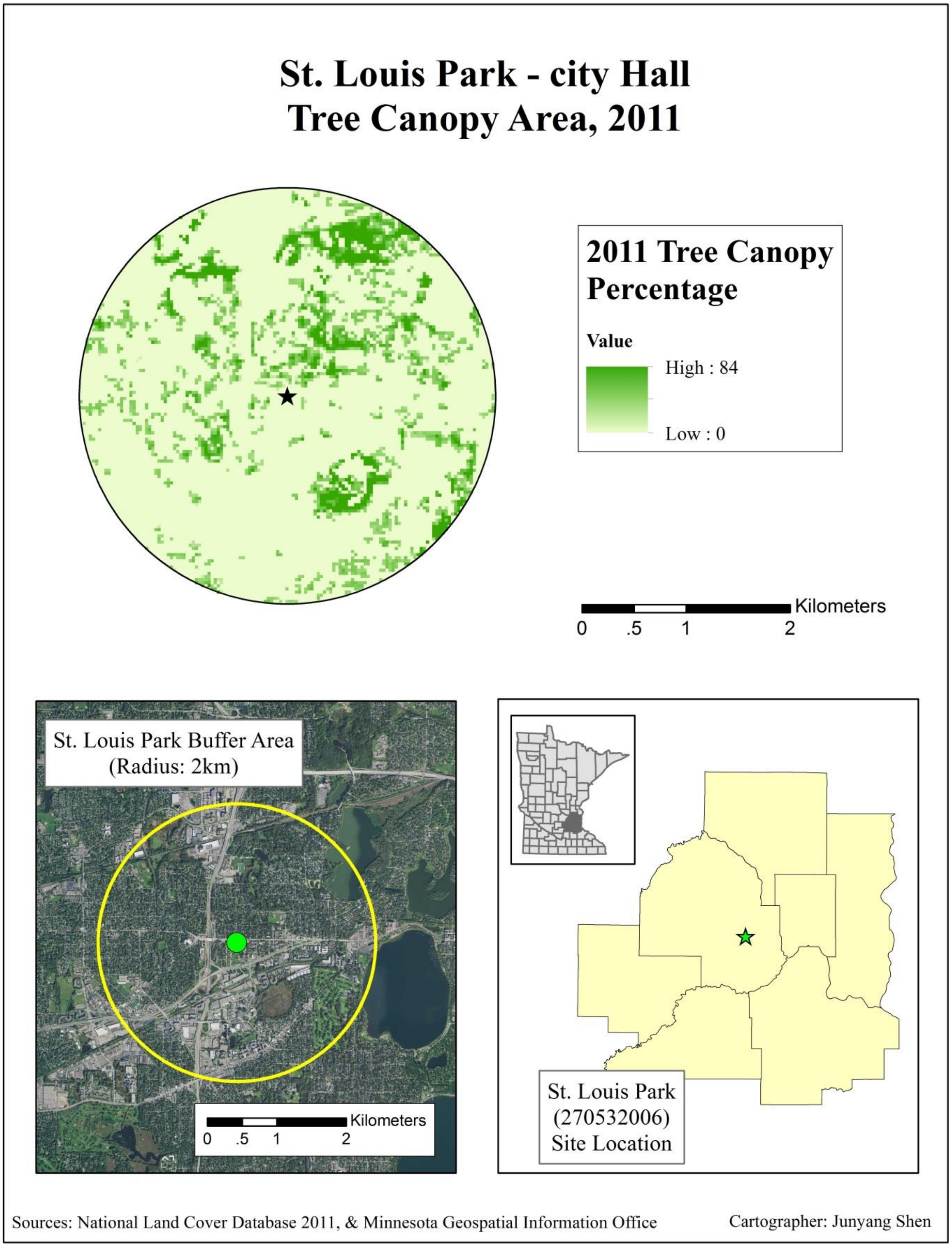

Figure II: 2011 St. Louis Park - City Hall site: tree canopy, buffer size and monitoring site location

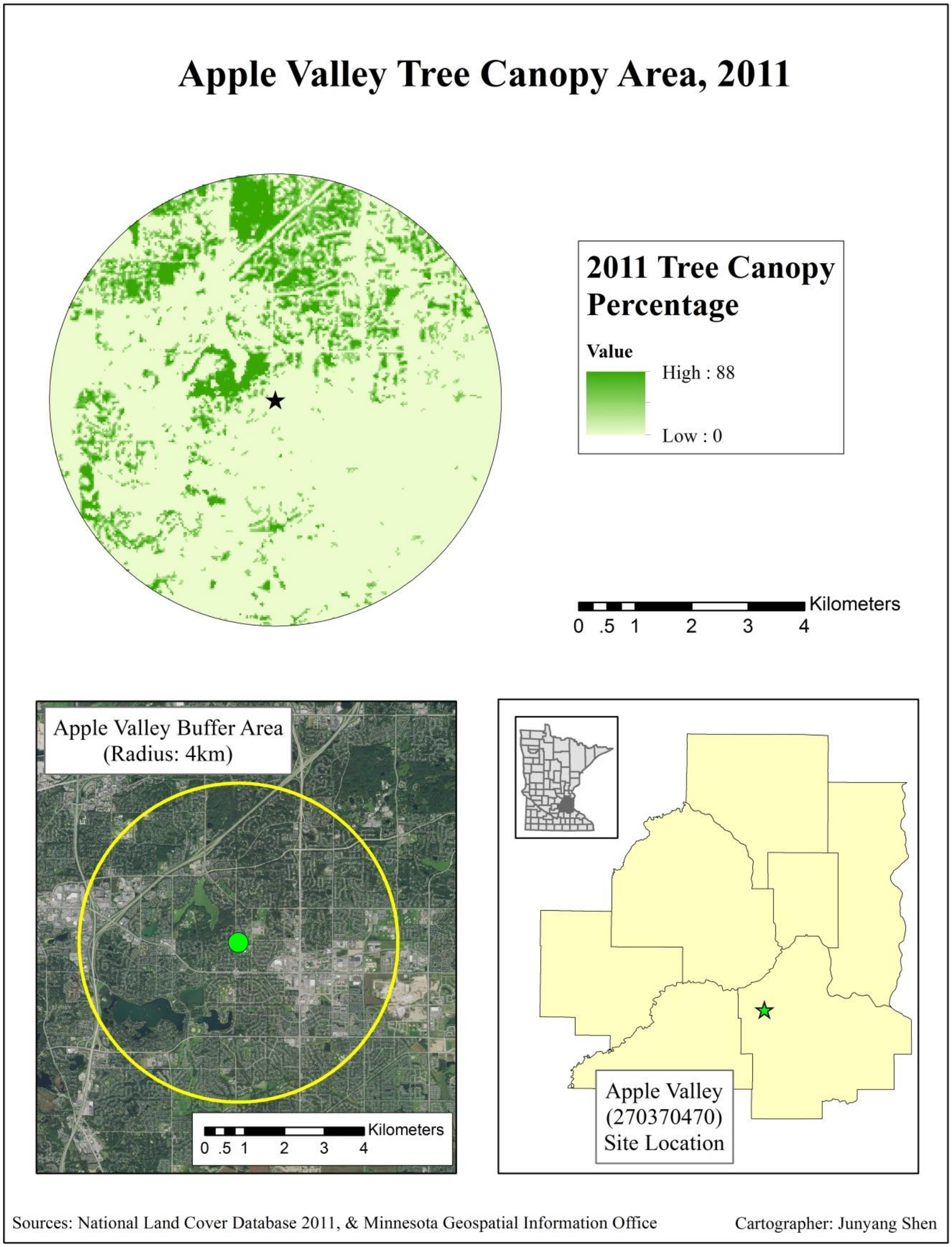

Figure III: 2011 Apple Valley - Westview School site: tree canopy, buffer size and monitoring site location

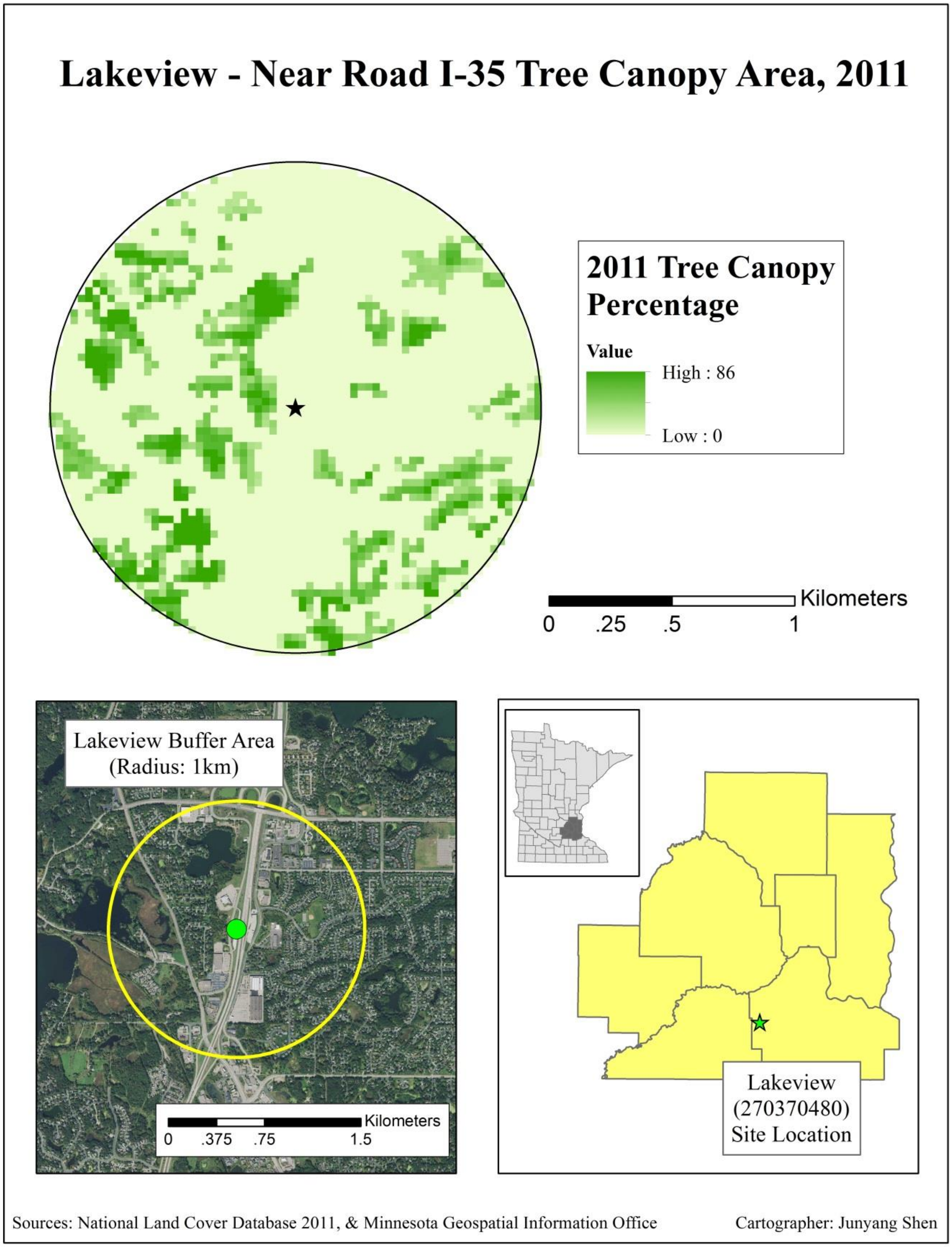

Figure IV: 2011 Lakeview - Near Road I-35 site: tree canopy, buffer size and monitoring site location

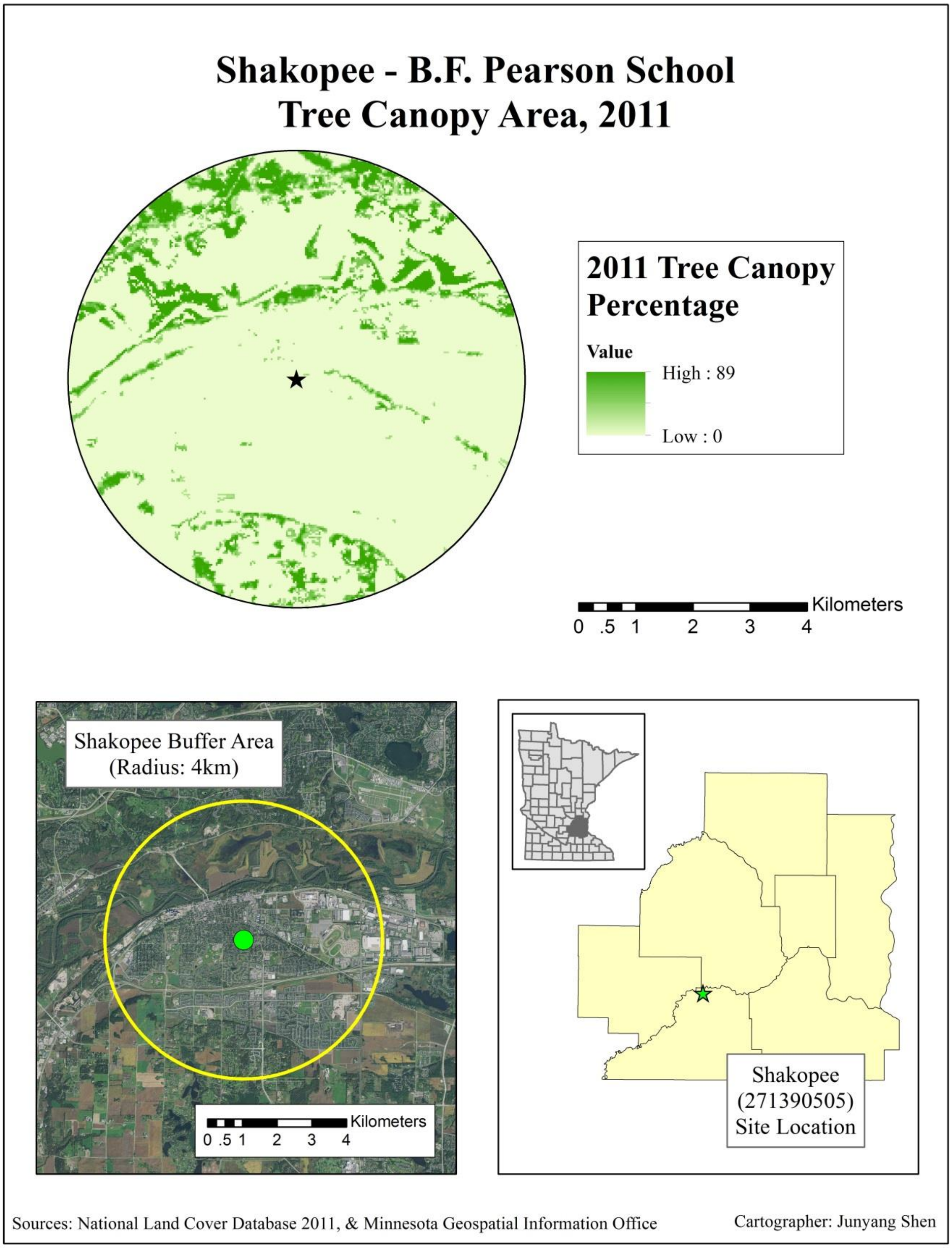

Figure V: 2011 Shakopee - B.F. Pearson School site: tree canopy, buffer size and monitoring site location

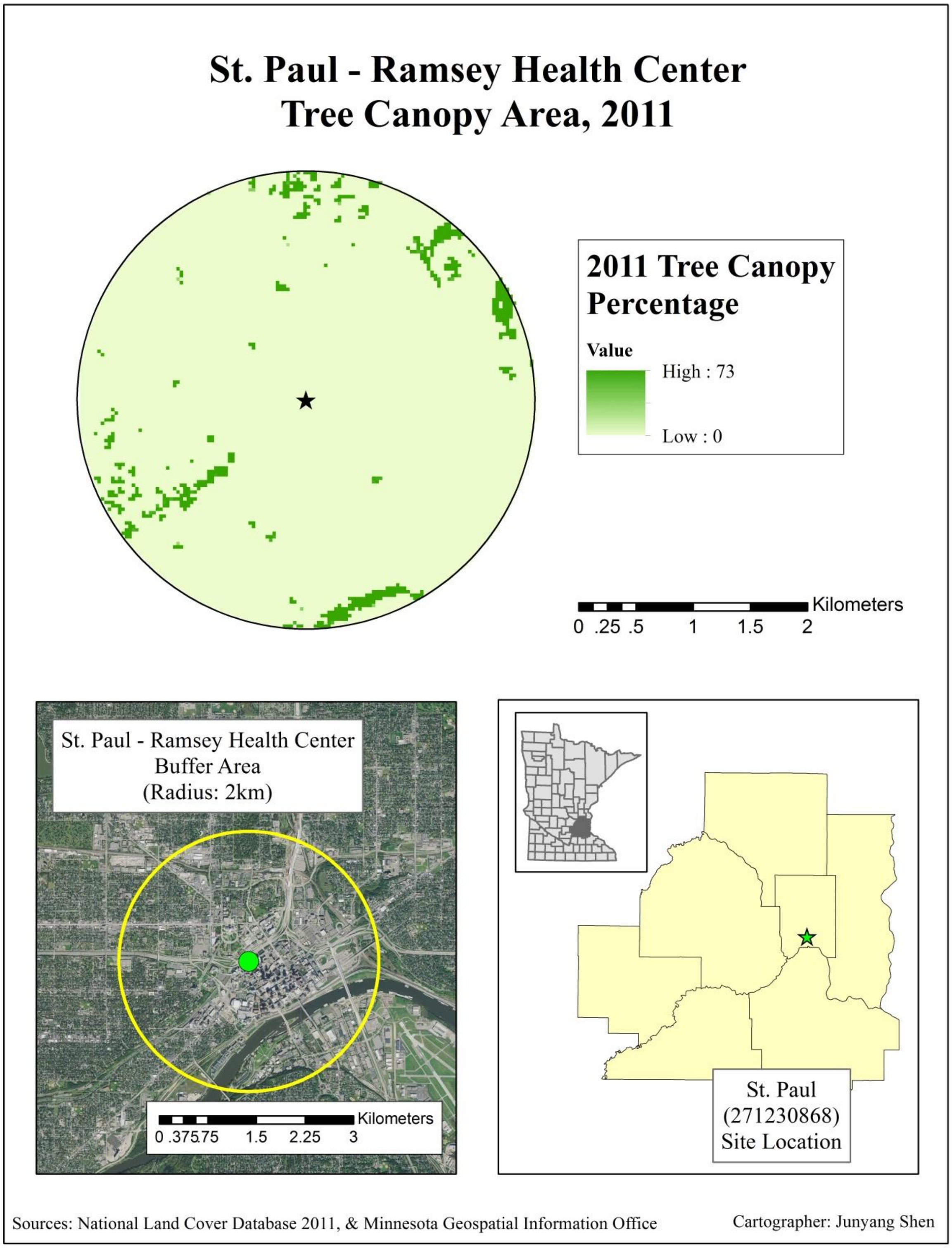

Figure VI: 2011 St. Paul - Ramsey Health Center site: tree canopy, buffer size and monitoring site location

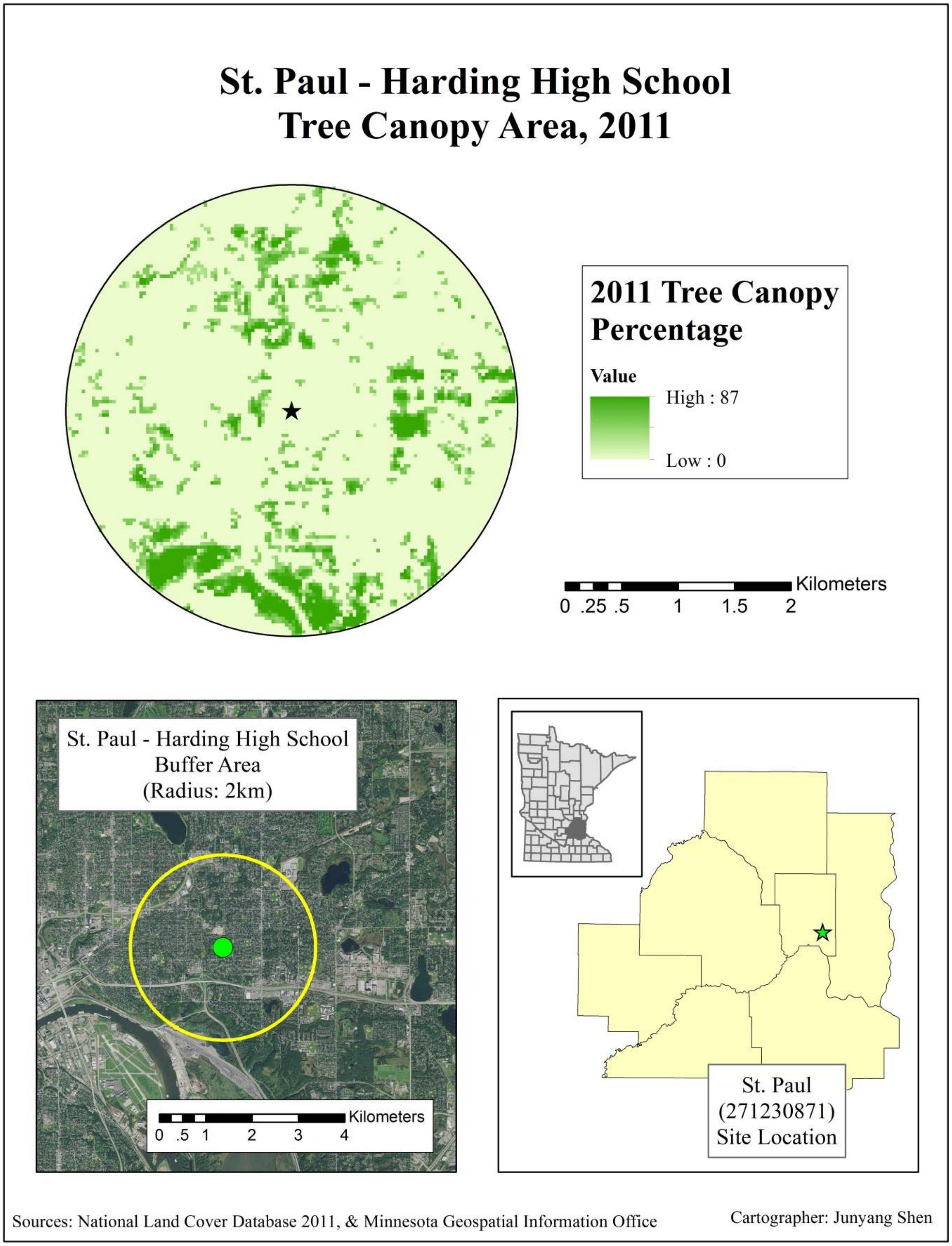

Figure VII: 2011 St. Paul - Harding High School site: tree canopy, buffer size and monitoring site location

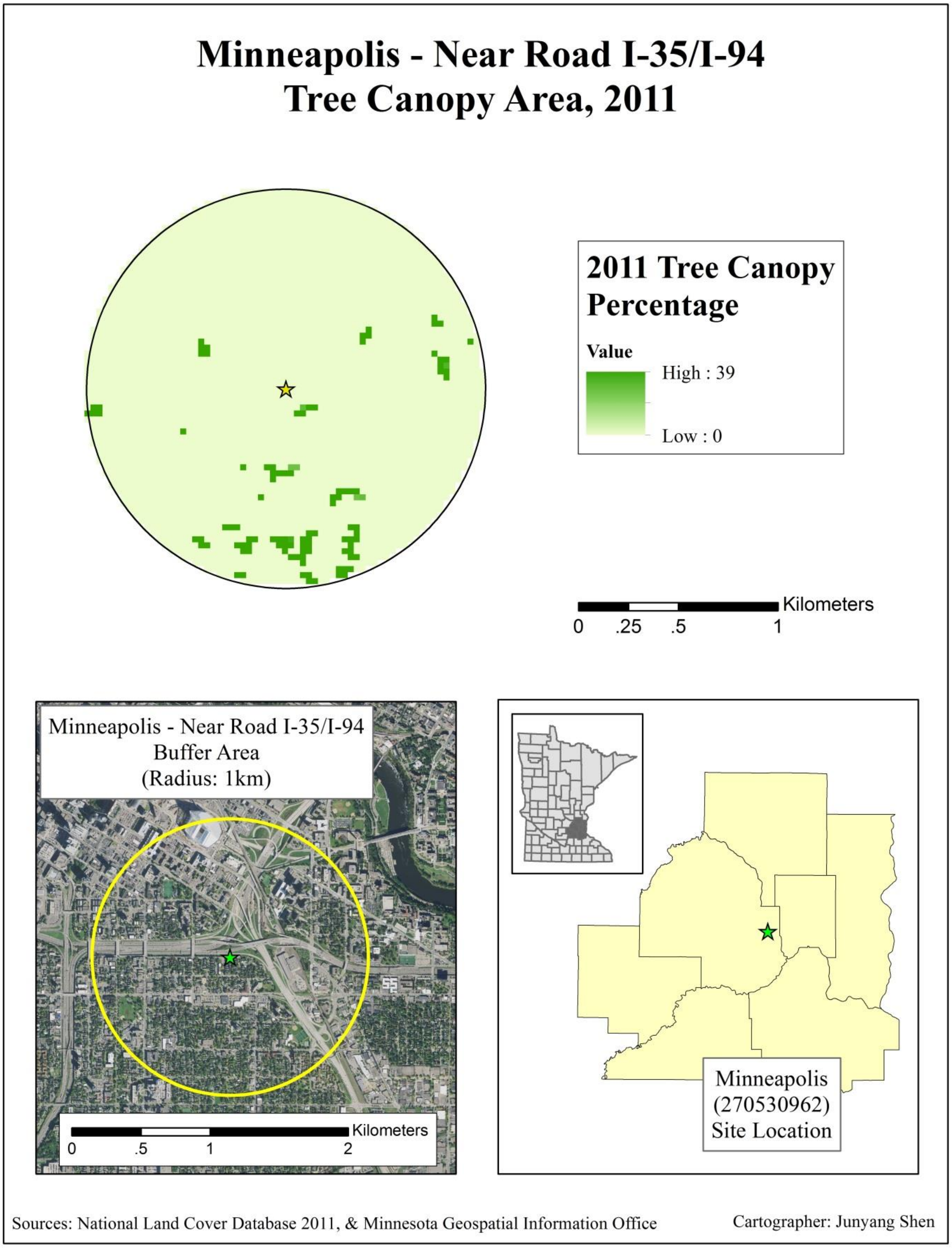

Figure VIII: 2011 Minneapolis - Near Road I-35/I-94 site: tree canopy, buffer size and monitoring site location

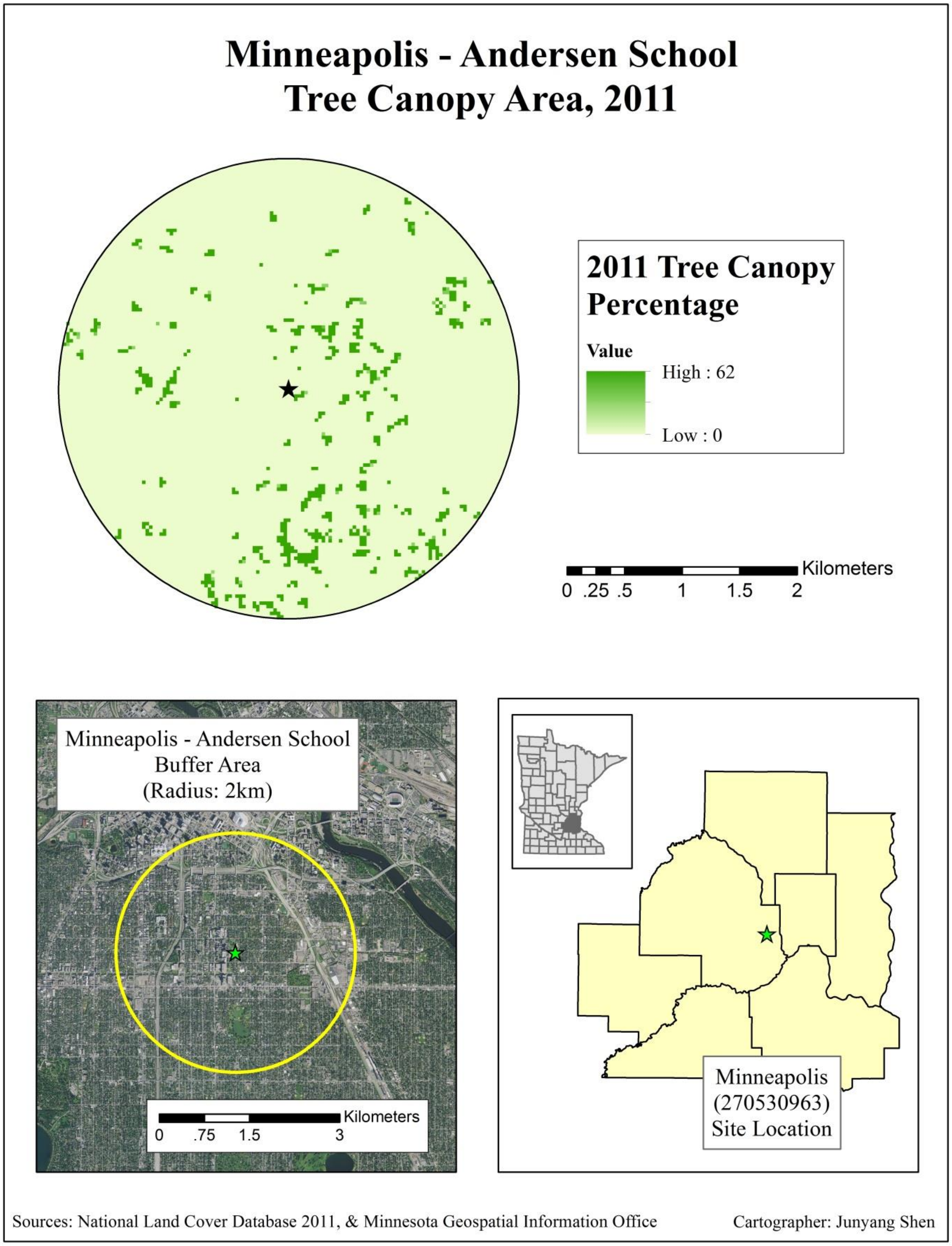

Figure IX: 2011 Minneapolis - Andersen School site: tree canopy, buffer size and monitoring site location

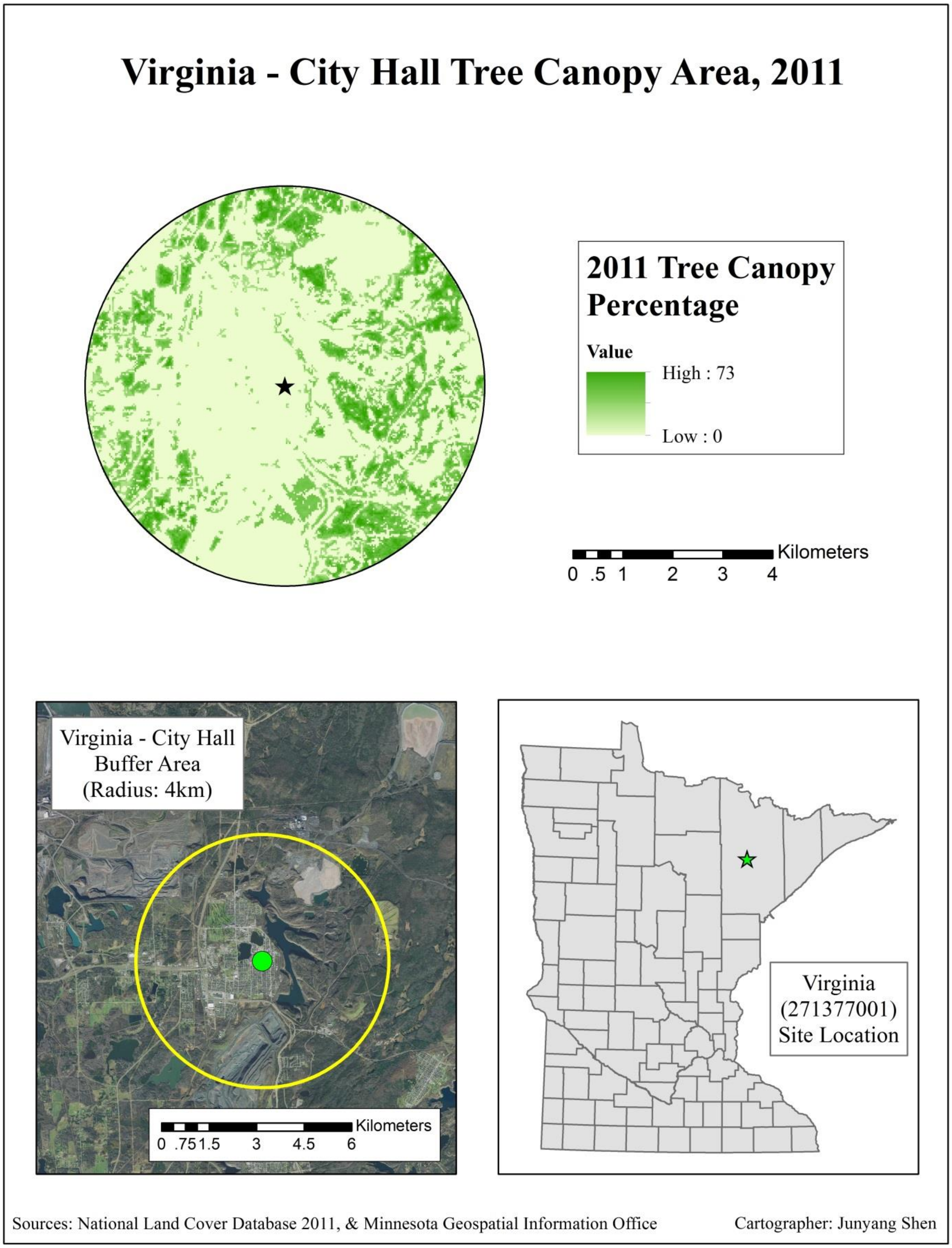

Figure X: 2011 Virginia - City Hall site: tree canopy, buffer size and monitoring site location

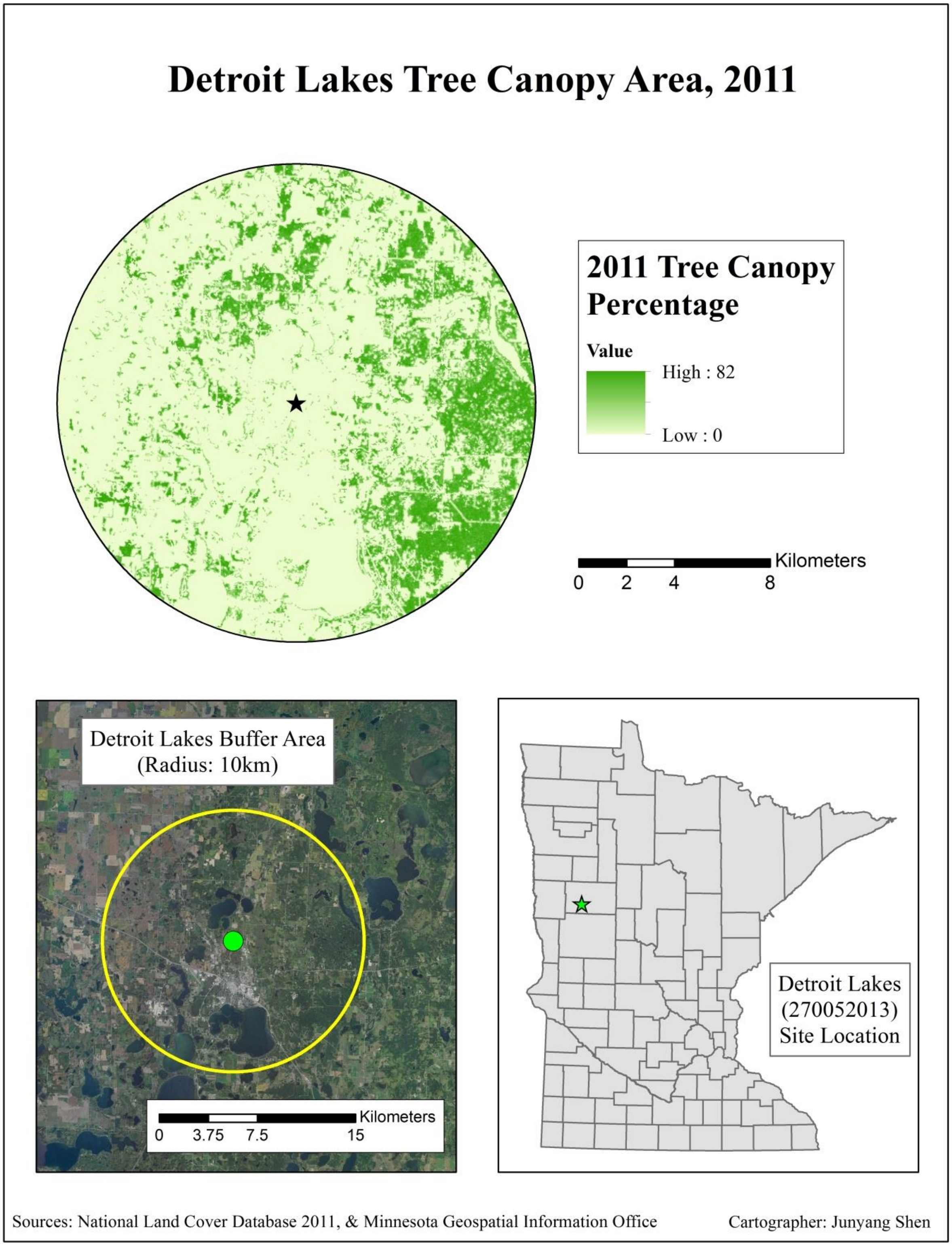

Figure XI: 2011 Detroit Lakes - FWS Wetland Management District site: tree canopy, buffer size and monitoring site location

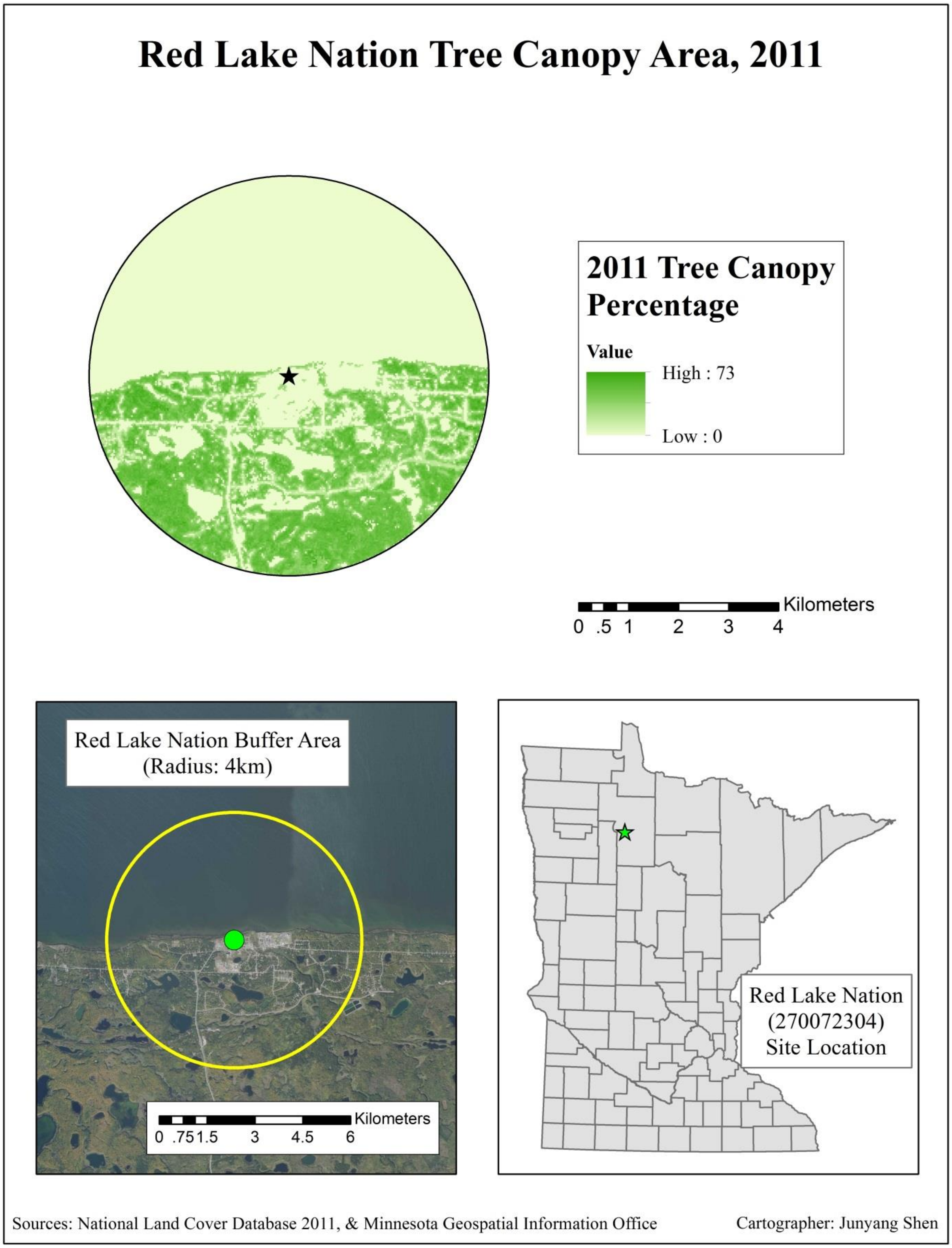

Figure XII: 2011 Red Lake Nation site: tree canopy, buffer size and monitoring site location

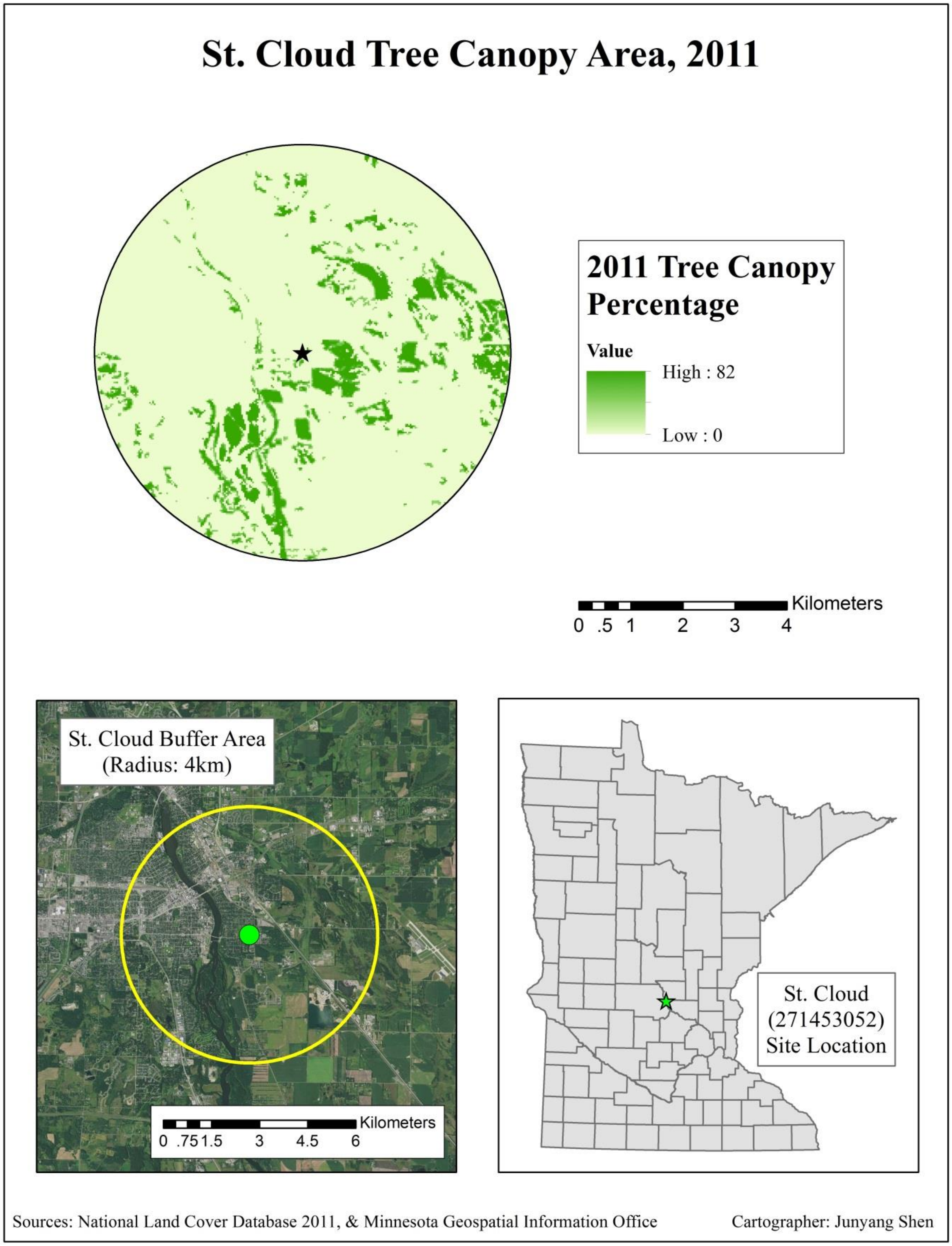

Figure XIII: 2011 St. Cloud - Talahi School site: tree canopy, buffer size and monitoring site location

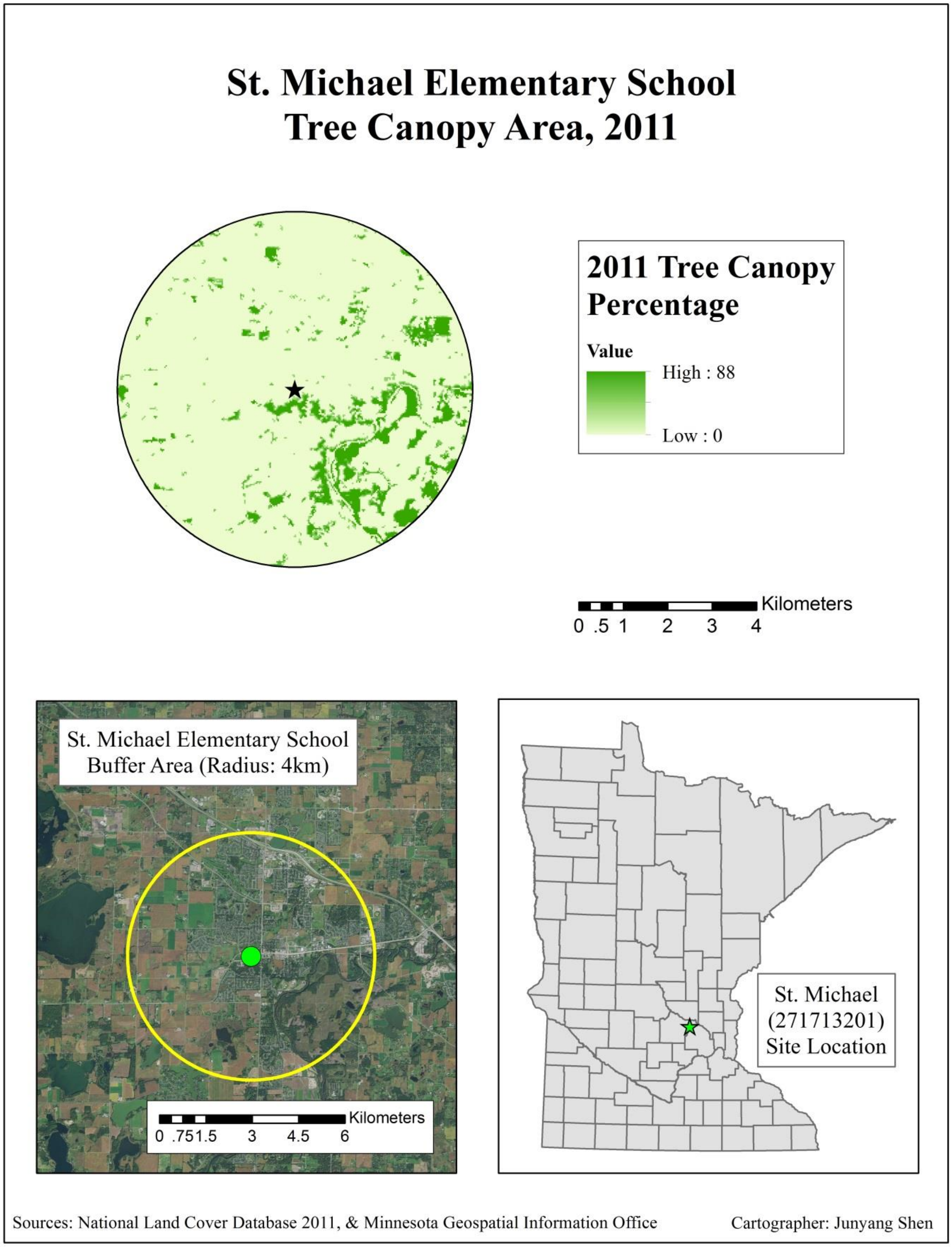

Figure XIV: 2011 St. Michael Elementary School site: tree canopy, buffer size and monitoring site location

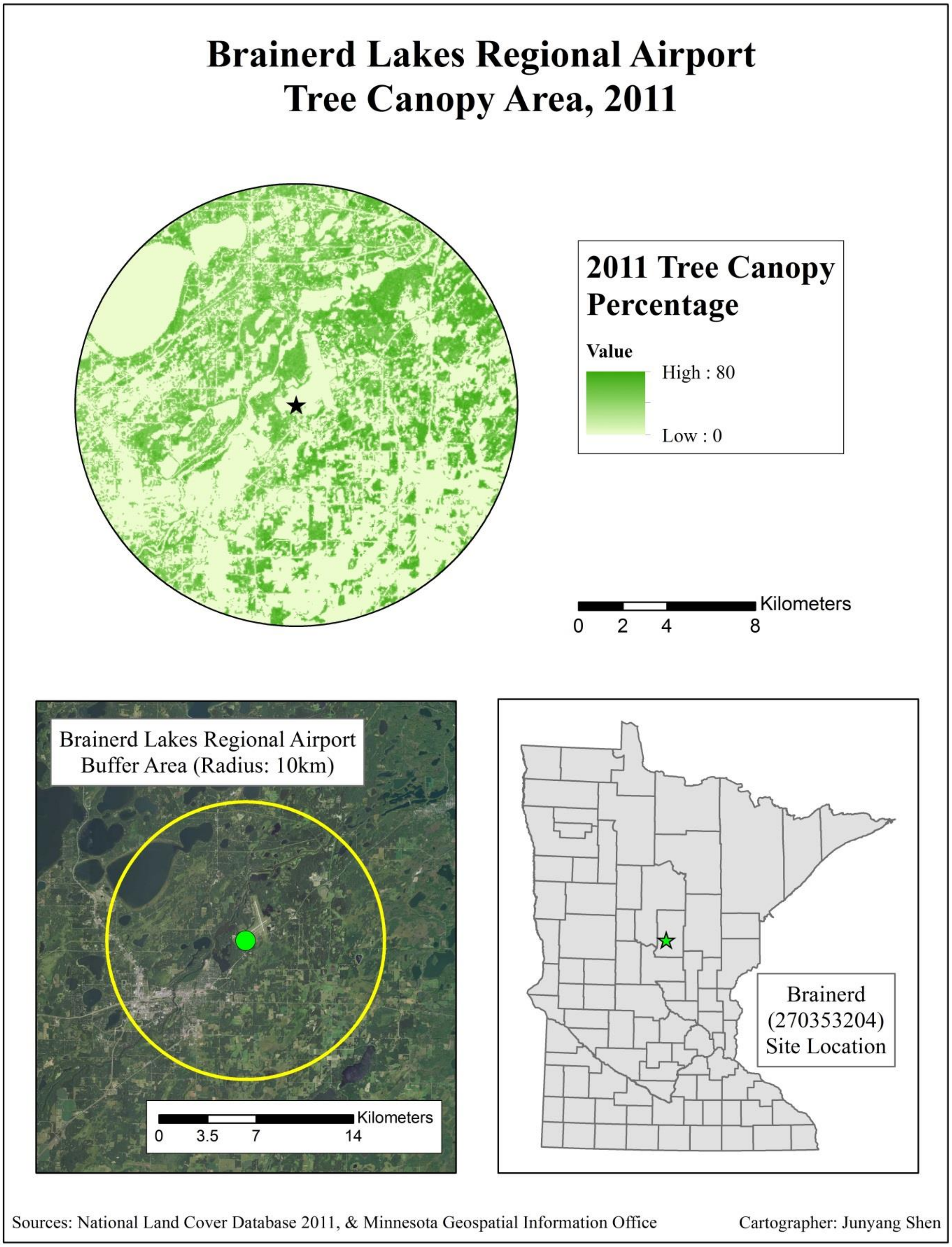

Figure XV: 2011 Brainerd Lakes Regional Airport site: tree canopy, buffer size and monitoring site location

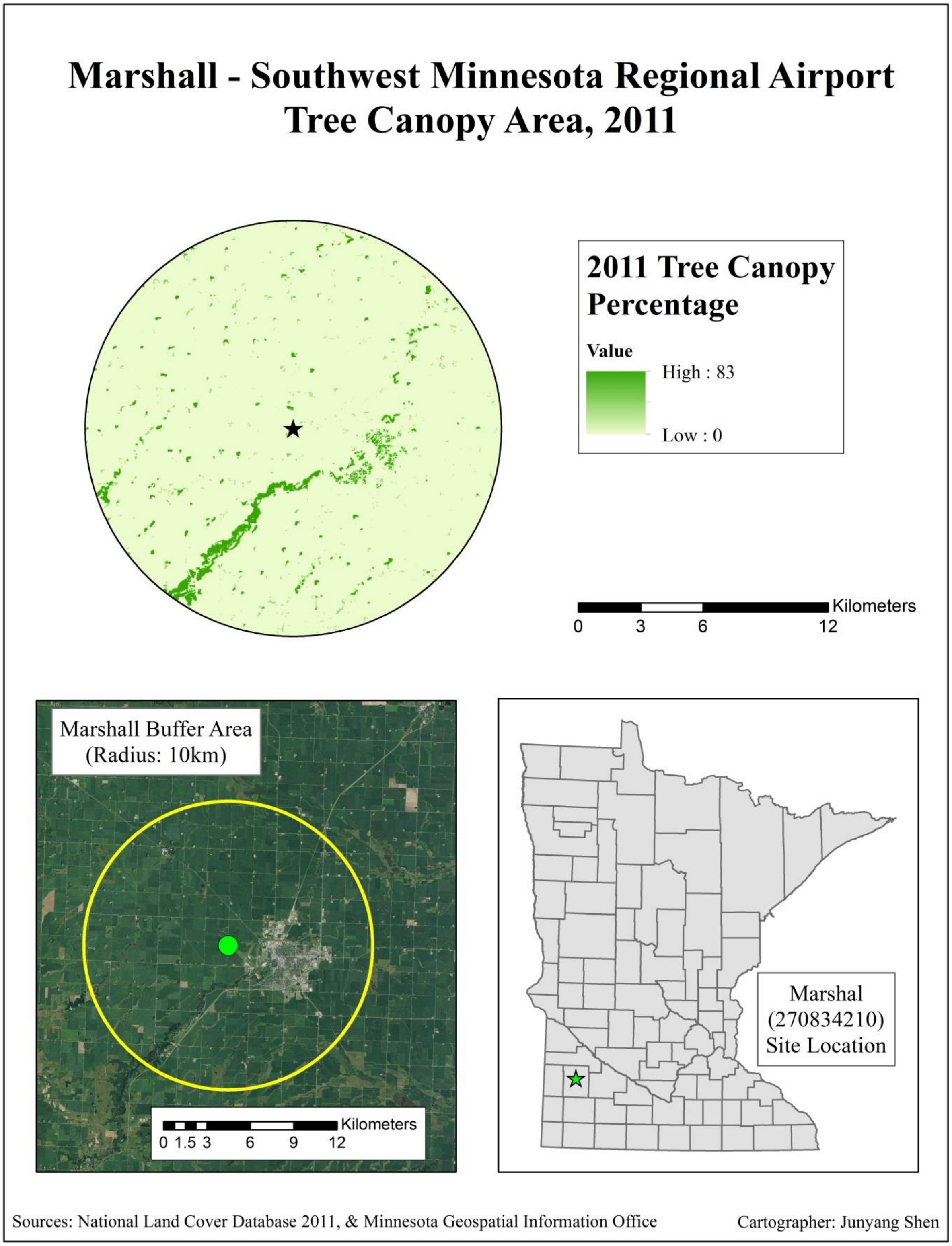

Figure XVI: 2011 Marshall - Southwest Minnesota Regional Airport site: tree canopy, buffer size and monitoring site location

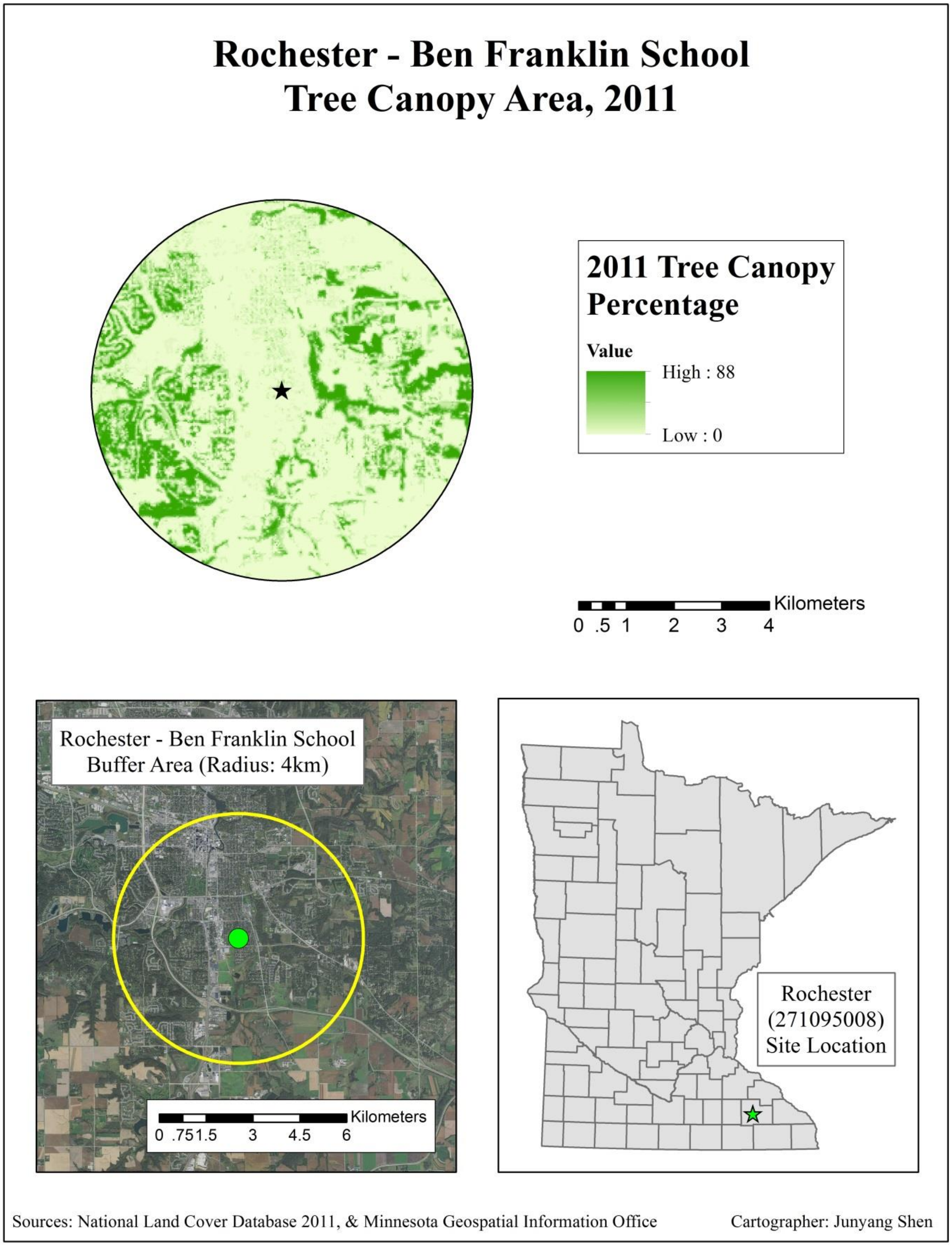

Figure XVII: 2011 Rochester - Ben Franklin School site: tree canopy, buffer size and monitoring site location

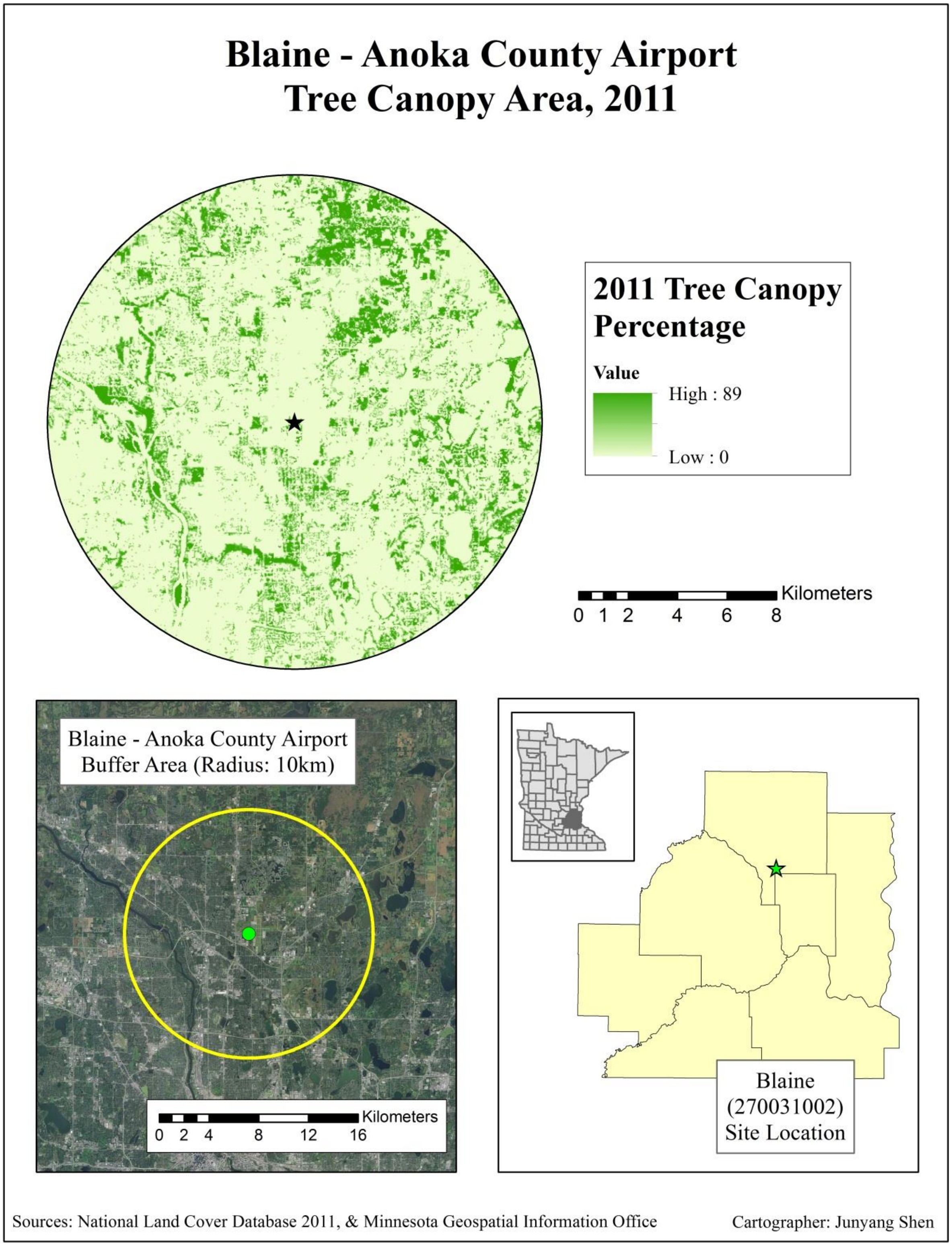

Figure XVIII: 2011 Blaine - Anoka County Airport site: tree canopy, buffer size and monitoring site location

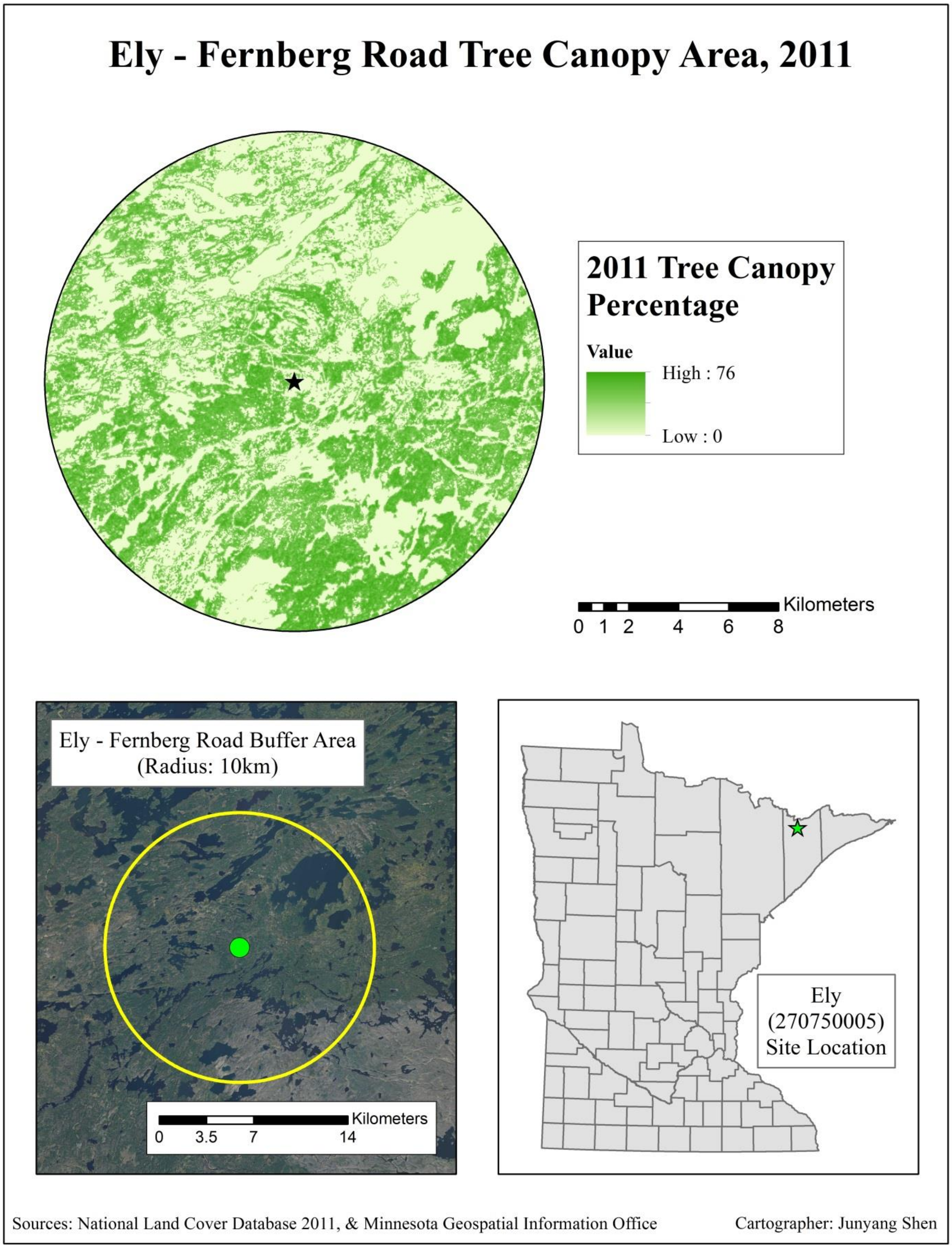

Figure XIX: 2011 Ely - Fernberg Road site: tree canopy, buffer size and monitoring site location

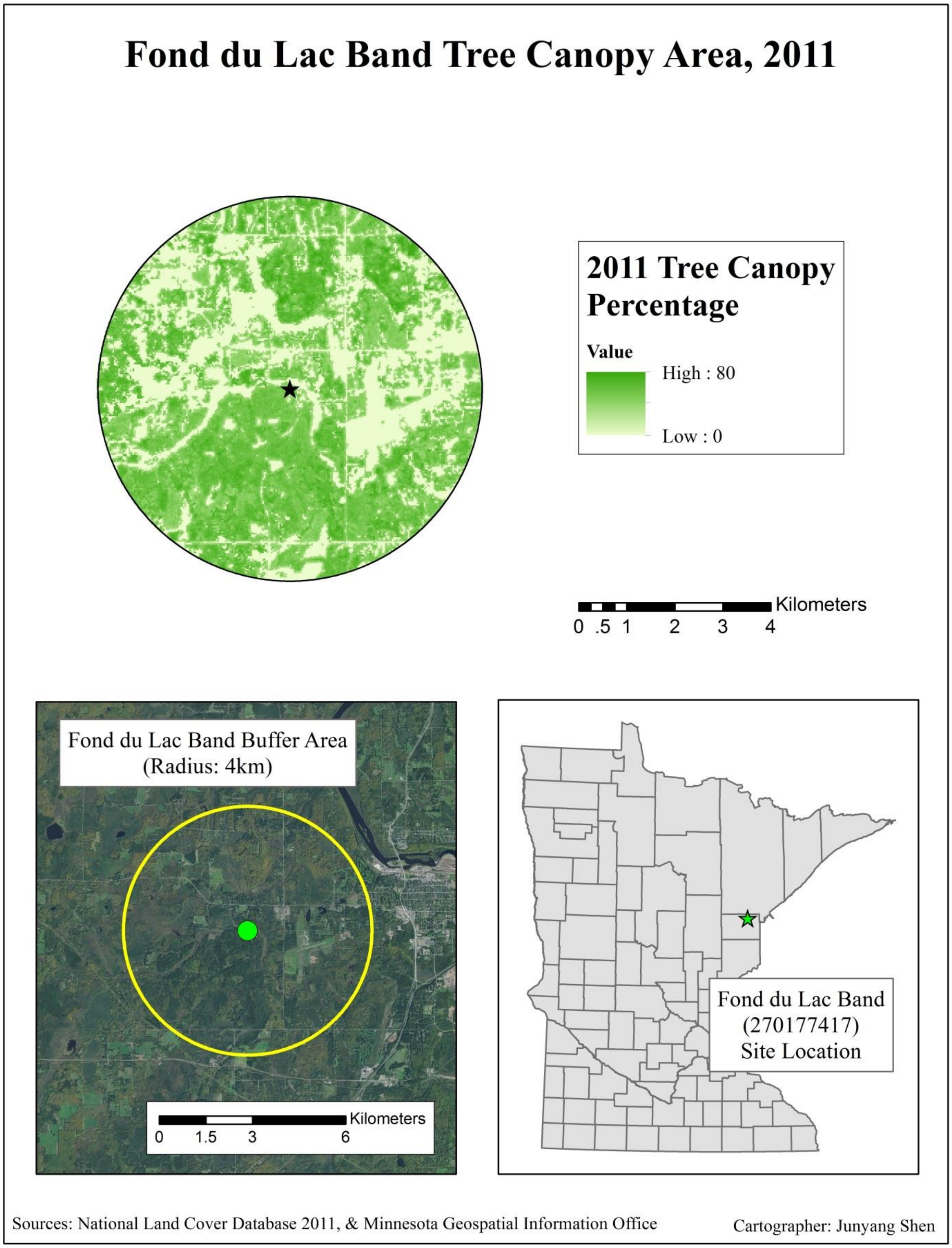

Figure XX: 2011 Fond du Lac Band site: tree canopy, buffer size and monitoring site location

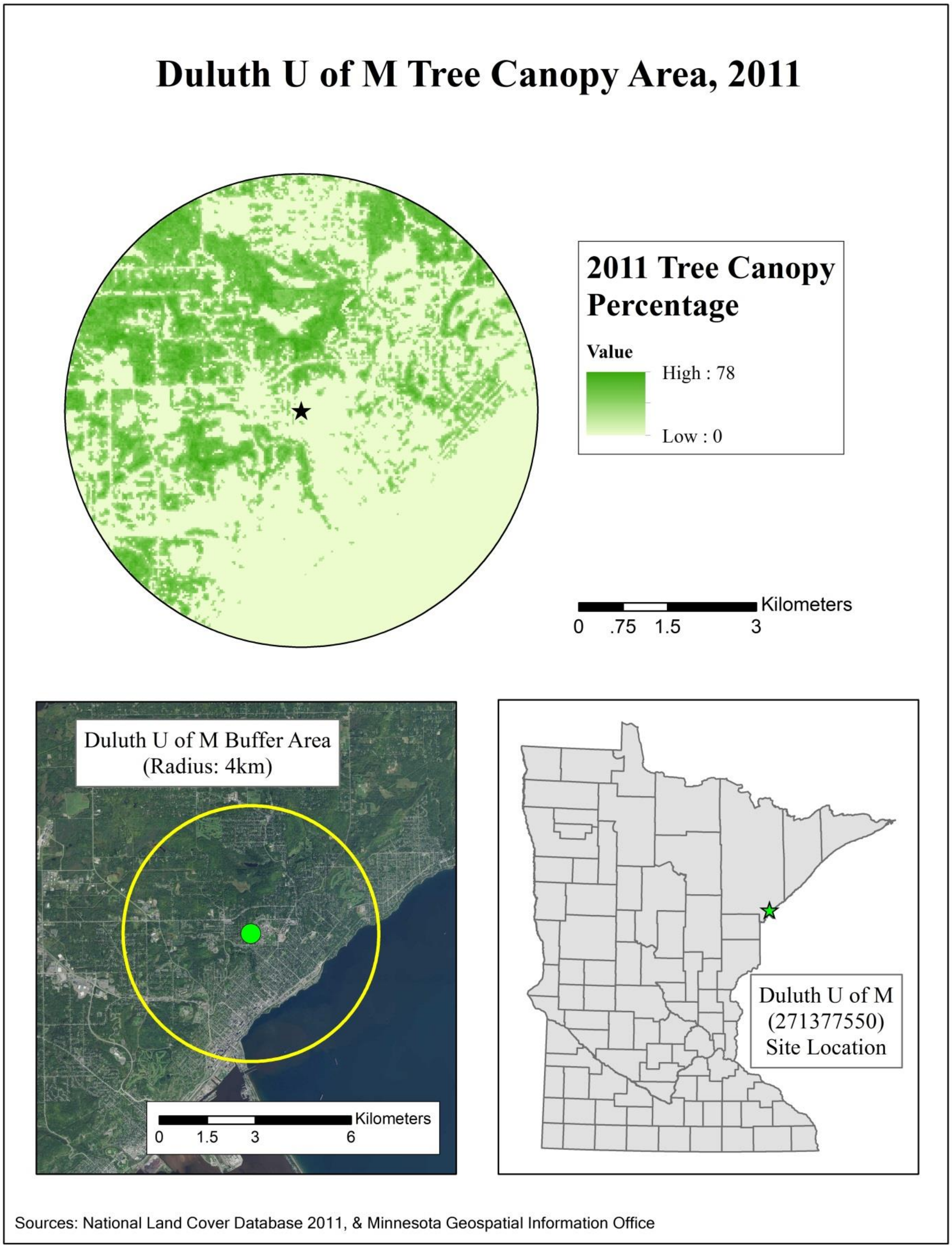

Figure XXI: 2011 Duluth U of M site: tree canopy, buffer size and monitoring site location

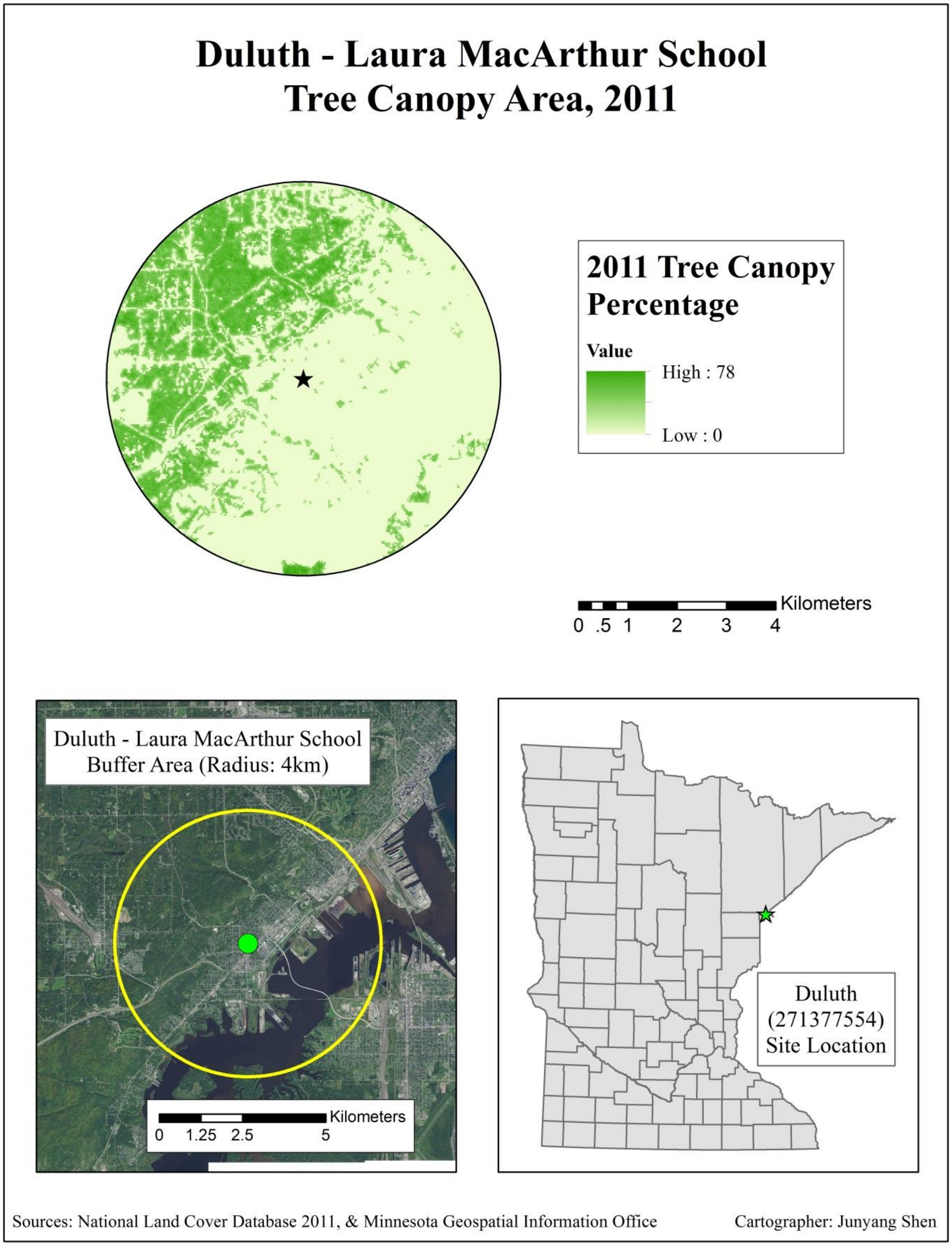

Figure XXII: 2011 Duluth - Laura MacArthur School site: tree canopy, buffer size and monitoring site location

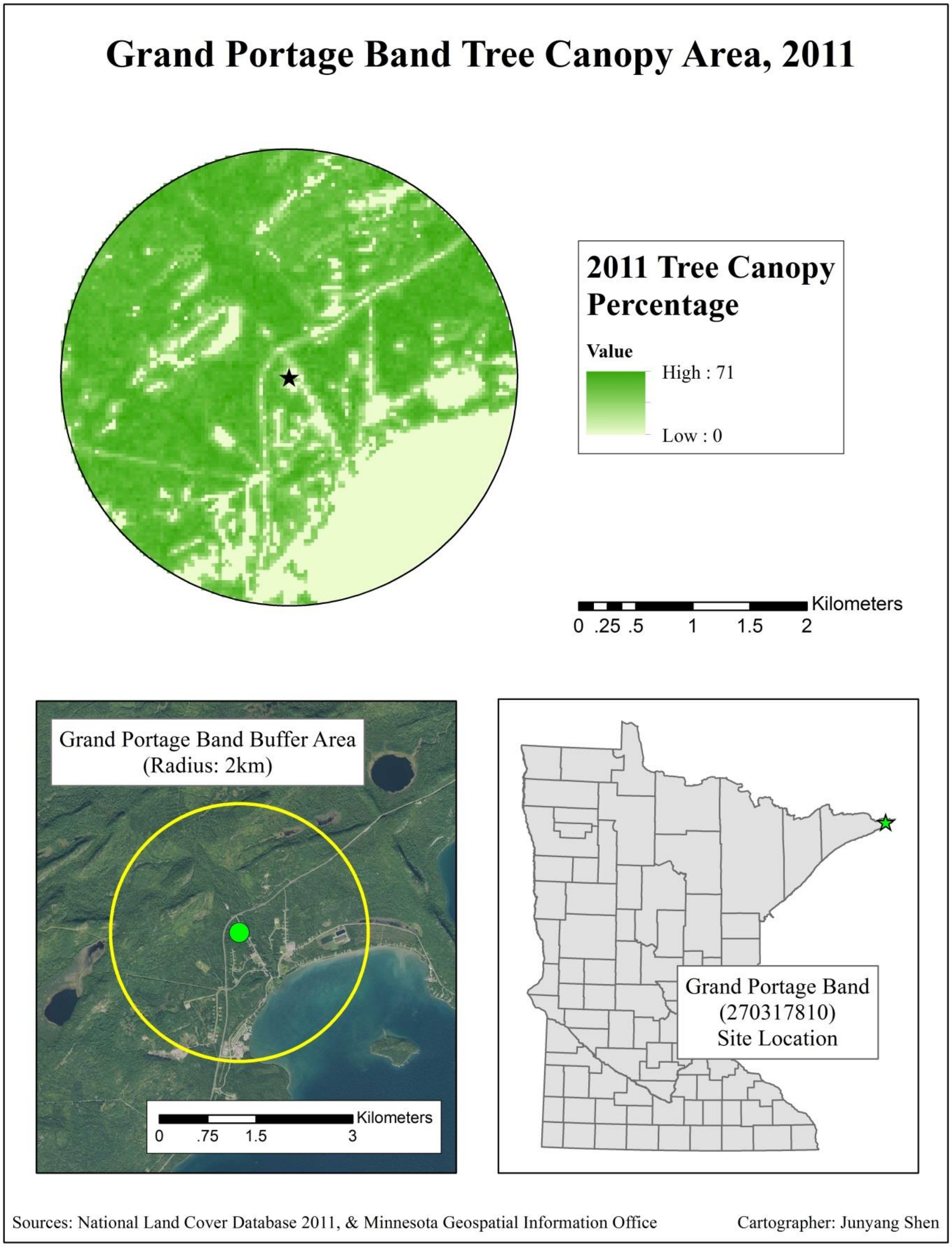

Figure XXIII: 2011 Grand Portage Band site: tree canopy, buffer size and monitoring site location

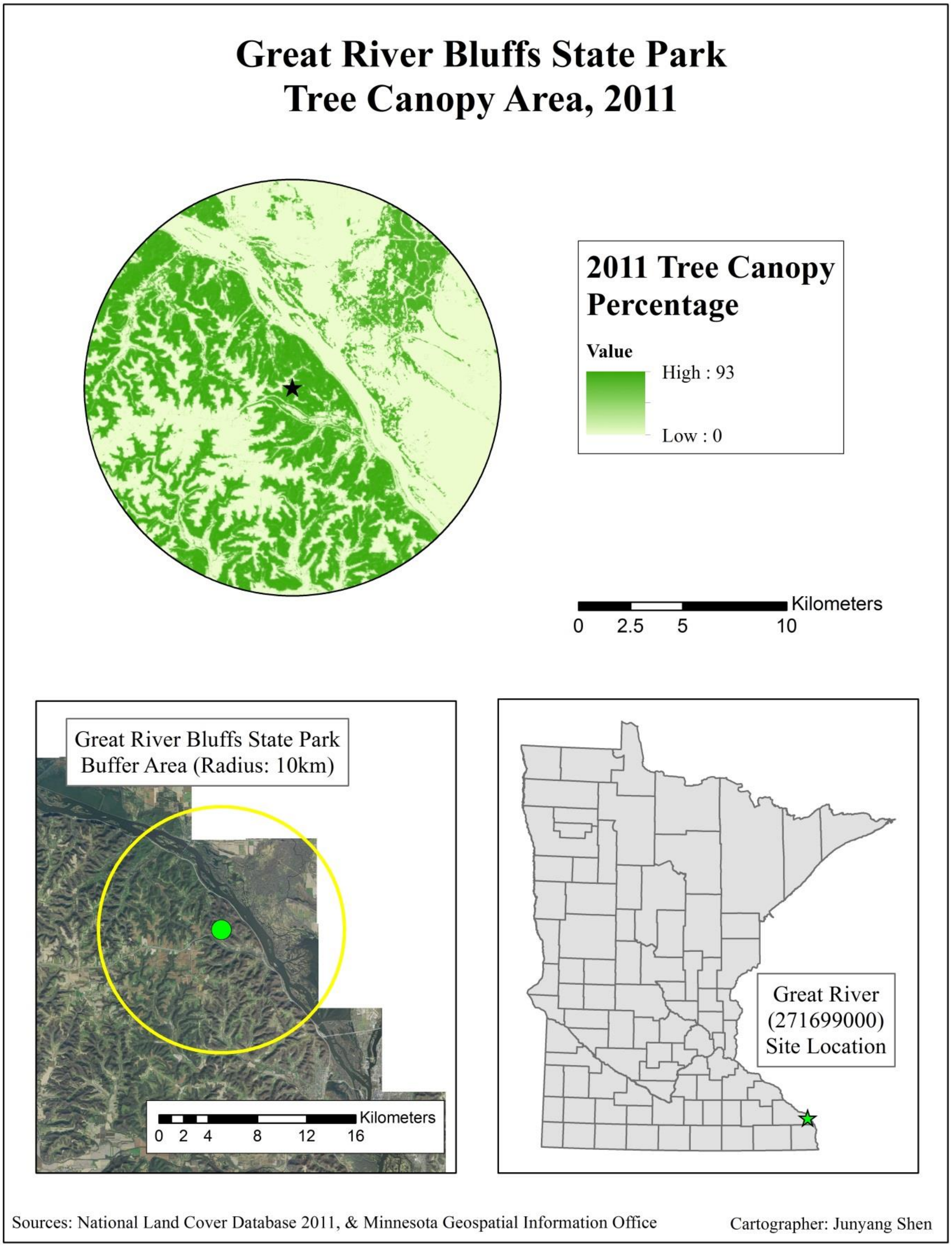

Figure XXIV: 2011 Great River Bluffs State Park site: tree canopy, buffer size and monitoring site location