Distribution Agreement

Environmental Drivers of Rocky Mountain Spotted Fever in Tennessee

Bryant Jones

MPH
Department of Environmental Health

Dr. Uriel Kitron, PhD
Committee Member

Dr. Julie Clennon, PhD
Committee Member

Dr. Paige Tolbert, PhD
Committee Member

Environmental Drivers of Rocky Mountain Spotted Fever in Tennessee

By

Bryant Jones

Bachelor of Science in Environmental Studies
Emory University
2012

Thesis Committee Chair: Uriel Kitron, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Environmental Health
2016

Abstract

Environmental Drivers of Rocky Mountain Spotted Fever in Tennessee
By Bryant Jones

Introduction: *Rickettsia rickettsii* is one of the most pathogenic ricketsial strains infecting humans and can be found throughout North and South America as well as some areas in the Eastern hemisphere. The primary vector is the American dog tick (*Dermacentor Variabilis)* which is distributed throughout the Eastern U.S. and contributes to the particularly high incidence of RMSF in Tennessee. This study aims to utilize GIS technologies and data available from various sources to estimate the association between presence of RMSF with environmental variables.

Methods: Passive surveillance data of Rocky Mountain spotted fever cases throughout the state of Tennessee were used to document both RMSF cases and non-cases at the census tract level. Case locations were determined using the address of the infected individual as a proxy for actual tick bite location. Clusters of high prevalence census tracts were used to create a dichotomous outcome variable for use in logistic regression analysis. Analysis was performed using the environmental factors as independent exposure variables and clustered versus non-clustered census tracts as the outcome.

Results: Cluster analysis of positive cases showed areas of High-High clustering west of the capital. Significant environmental variables for the logistic regression model included land cover, elevation, and average annual precipitation. Geological classes and soil order were not deemed significant but were retained for the final model based on a priori knowledge. The predictive risk map shows the majority of the state with the 25 to 50% probability range while areas of higher risk are scattered throughout the state.

Discussion: Research has shown that the distribution of tick vectors is associated with environmental variables. This project serves as a starting point to better understand the role of environmental variables in the distribution of Rocky Mountain spotted fever. More in-depth studies on the subject could lead to new health policy initiatives regarding tick borne disease in Tennessee.

Environmental Drivers of Rocky Mountain Spotted Fever in Tennessee

By

Bryant Jones

Bachelor of Science in Environmental Studies
Emory University
2012

Thesis Committee Chair: Uriel Kitron, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Environmental Health
2016

Acknowledgements

This thesis would not have been possible without the help and support from Uriel Kitron, Dr. Julie Clennon and Dr. Abelardo Moncayo. A special thanks to Dr. Bill Size as well for his insight into geological categorizations and Dr. Abelardo Moncayo, without whom, the project would not have been possible. Thank you for all your time and consideration in providing guidance throughout this research experience. I feel fortunate to have had the opportunity to work with such talented and knowledgeable mentors.

<u>**Table of Contents**</u>

Objectives

- **Define Environmental Variables associated with tick abundance and disease risk**
- **Define ecological drivers of risk for Rocky Mountain Spotted Fever**
- **Quantify and compare selected ecological drivers of risk**
- **Model probability of disease using case incidence as a proxy for tick presence**

Chapter 1: Introduction

Rocky Mountain spotted fever in Tennessee

Rocky Mountain spotted fever (RMSF) is a tick borne disease caused by a gram - negative, intracellular bacterium, *Rickettsia rickettsii*, that infects endothelial cells lining blood vessels. It is a highly virulent human infection, and without proper treatment, can lead to fatal outcomes even in younger, healthy people [8].

First reported in 1896, RMSF has been notoriously difficult to diagnose due to its non-specific clinical presentation. Common symptoms include fever, headache, myalgia, gastrointestinal distress, and rash for some individuals. This tell-tale rash, used as an indicator for the disease, does not generally appear until multiple days into the illness, and at times, does not manifest at all [38]. The rash generally begins near the extremities and can spread to the center of the body, presenting with small, pink, non-itchy spots on the wrists, forearms, and ankles [28]. Many times, a patient is treated even if the disease is only probable, or even possible because early treatment is necessary to prevent severe complications ranging from vasculitis to neurological deficits.

Rickettsia rickettsii is one of the most pathogenic ricketsial strains infecting humans and can be found throughout North and South America as well as some areas in the Eastern hemisphere [25]. The disease has been reported throughout the contiguous U.S. but there are five states that bear the brunt of infection: North Carolina, Oklahoma, Arkansas, Missouri, and Tennessee. These states account for over 60% of all cases within the U.S. [33].

Moncayo et al. stated that "Tennessee historically reports one of the highest incidence rates of Rocky Mountain spotted fever in the nation, with the disease being the most common tick-borne illness reported throughout the state" [20]. The present study aims to examine the effect of several key environmental risk factors for RMSF within Tennessee to help understand patterns of RMSF infection and to better inform surveillance and health policy initiatives.

The Vectors

There are three main vectors of Rocky Mountain spotted fever in the contiguous United States. The primary vector is the American dog tick (*Dermacentor Variabilis)* which is distributed throughout the Eastern U.S. and contributes to the particularly high incidence of RMSF in Tennessee. The Brown dog tick (*Rhipicephalus sanguineus*) and the Rocky Mountain wood tick (*Dermacentor andersoni*) are secondary vectors. The Rocky Mountain wood tick is distributed throughout the Northwestern U.S. at elevations between 4,000 and 10,500 feet while the Brown dog tick can be found throughout the contiguous U.S. [15].

Ticks act as both the vector and reservoir for RMSF. Ticks go through three different life stages: larvae, nymph, and adult and each state must take a blood meal to progress. Ticks can become infected in multiple ways: by feeding on an infected animal (the main route of transmission), through fertilization, or through transovarial passage of the pathogen from an infected tick to her progeny [38]. Generally, RMSF is transmitted to humans via the bite of an infected tick. Laboratory studies have shown that it is possible to become infected by inhaling contaminated aerosol, but the mode of transmission has only been observed in laboratory settings and is unlikely to play a significant role in nature [8]. When the tick is attached to a host and feeding, the pathogen is released through the salivary glands in a process that reactivates *R rickettsii* from a dormant state into a pathogenic one [8]. This process is usually initiated after 6 -10 hours of attachment but can take up to twenty-four hours [33]. Once the tick has finished feeding, it will detach and fall to the ground to prepare for the next life stage.

Generally, RMSF incidence peaks during the summer months with the highest number of cases occurring in June and July, though seasonality can vary somewhat in different regions depending on the vectors involved and climate [33].

Using GIS and mapping to estimate disease risk

Spatial models and maps have been used extensively in analyzing associations between the environment and disease [13, 15, 17, 19, 24, 39]. Estimating the risk of tick borne disease in certain areas can be enhanced by using geographical information systems (GIS) and spatial analysis. These methods have been used most commonly for Lyme disease [15, 19, 22]. GIS is particularly useful in this regard as it allows for risk analysis to be expanded to larger regions that are less well-defined.

Disease hotspots and areas of high risk can be defined using passive surveillance data of patient residences. Also, creating models from this data can help to drive active surveillance as well as disease interventions. An advantage of being able to use GIS to analyze passive surveillance data is that it decreases the laboratory costs and time-intensiveness of field sampling for ticks. With current GIS and modelling capabilities, environmental variables can be assessed over large areas. Relevant environmental conditions can be determined at a smaller scale and then extrapolated out to encompass entire states or countries. This study aims to utilize GIS technologies and data available from various sources to estimate the association between the presence of RMSF in these areas with environmental variables. Environmental variables that are significantly associated with RMSF distribution across the state can be assessed to determine where greater surveillance and support are needed. With further study, this could be lead to the implementation of methods aimed at controlling the spread of Rocky Mountain spotted fever. In this way, risk maps for RMSF will help us understand the patterns of RMSF in Tennessee to better protect human health.

Chapter 2: Materials and Methods

Study Area

The Southeastern state of Tennessee is located between the Great Smokey Mountains to the east and the Mississippi river to the West, and covers an area of 41,234 square miles. According to the 2010 census, the state population was 6,346,105, with 153.9 persons per square mile. Passive surveillance throughout the state of Tennessee was used to determine the incidence of RMSF.

Geospatial Analysis

Epidemiological data

Passive surveillance data of Rocky Mountain spotted fever cases throughout the state of Tennessee was received from the Tennessee Department of Health and was used to document both RMSF cases and non-cases at the census tract level. Case locations were determined using the address of the infected individual as a proxy for actual tick bite location. This method was also used to determine locations of non-cases from the same dataset (individuals who had been tested for RMSF but received negative results). Address locations were geolocated within ArcGIS [11] software and given latitude (X) and longitude (Y) coordinates for use in analysis. In order to create a dichotomous dependent variable for use in logistic regression, prevalence of cases per 100,000 people were calculated by census tract. Prevalence per census tract was then used to determine clusters of High-high prevalence. Clustered census tracts and non-clustered census tracts were used as the dichotomous outcome for logistic regression analysis.

A RMSF case was defined as an individual with a case status of confirmed or probable disease as defined by National Notifiable Disease Surveillance System (NNDSS). Confirmed cases include any case meeting the clinical case definition with confirmatory laboratory evidence and probable cases include any case meeting the clinical case definition with supportive laboratory evidence. The NNDSS defines a case of RMSF as follows: "Any reported fever and one or more of the following: rash, eschar, headache, myalgia, anemia, thrombocytopenia, or any hepatic transaminase elevation" [30]. Suspected cases were not included in the analysis as they include cases with laboratory evidence but no clinical information.

Exploratory Data Analysis

Case prevalence was determined at the census tract level by spatially joining case data and census tract data to determine the number of cases per 100,000 individuals. Point density and Kernel density were also performed using the spatial analyst toolbox in ArcGIS 10.3. Point density calculates a magnitude per unit area from point features that fall within a neighborhood around each cell. Kernel density performs similarly but uses a kernel function to smooth out the surface of each point. Cluster analysis was performed to determine areas of high-high and low-

low clustering. For this, Anselin local Moran's I was performed to determine where high and low values clustered spatially.

Meteorological data

Annual average mean temperatures (degrees F) and total precipitation (inches) for fifteen consecutive years (2000-2015) were derived from over 50 weather stations within Tennessee and the surrounding states. Data were obtained from NOAA Climate Data Online (CDO) [5]. Inverse Distance Weighting (IDW) was applied to annual mean temperature and total precipitation to create an interpolated continuous surface which was then clipped to the state of Tennessee using geoprocessing tools in ArcGIS.

Soil data

Soils data was derived from State Soil Geographic Database (STATSGO) generalized soils coverage dataset provided by the USDA National Resources conservation Service (NRCS) [21]. Soil order data was pulled from the STATSGO dataset and vector database was converted raster format. USDA classification of soils is split into 12 orders, 6 of which are present in Tennessee.   The 6 major soil order categories within the state include: Alfisols, Entisols, Inceptisols, Mollisols, Ultisols, and Vertisols.

According to the USDA soil taxonomy guide, the Alfisol order includes soils that are naturally fertile with high base saturation and a clay-enriched subsoil horizon. Entisols and inceptisols are both young soils that show little or no profile development or a weak, but noticeable, profile development, respectively. Mollisols are very dark-colored, very fertile, soils of grasslands. Ultisols have low base status with a clay-enriched subsoils while Vertisols are defined as very clayey and can shrink and crack while dry or expand when wet [29]. Soil order was classified into ordinal categories based on fertility: entisols and vertisols (0), inceptisols (1), ultisols (2), mollisols (3), and alfisols (4). For analysis, the majority of each reclassified soil order was determined for each census tract.

Elevation data

A digital elevation model for Tennessee was obtained from the USGS National Elevation Dataset (NED) at a spatial resolution of 1 arc-second [36]. Elevation data ranged from 2016.4 m in the eastern part of the state to 28.9 m in the western part.

Geological data

USGS geology data for the state was obtained from tngis.org [35]. The data contain seventeen major types of geological classes within Tennessee: siltstone, silt, shale, sandstone, sand, quartzite, limestone, greywacke, gravel, dolostone, conglomerate, claystone, clay or mud, chert, calcarenite, and black shale. Geological classes were classified in order of increasing drainage: water (0); igneous rock (1) which includes granite, diorite, and gabrro; metamorphic rock (2) which includes quartzite, migmatite, tectonic breccia, gneiss, and metasedimentary rock; sedimentary rock (3) which includes chert, limestone, shale, sandstone, dolostone, conglomerate, claystone, black shale, calcarenite, siltstone, and greywacke; and Sediment (4) which includes clay or mud, gravel, sand and silt. For analysis, the majority of each reclassified geological class was determined for each census tract.

Land cover data

Land cover data for the state of Tennessee was obtained from the USGS National Land cover Database (NLCD) [38]. The state of Tennessee comprises 15 of the 20 categories within the NLCD including: open water, developed open space, developed low intensity, developed medium intensity, developed high intensity, barren land, deciduous forest, evergreen forest, mixed forest, shrub/scrub, herbaceous, cultivated crops, woody wetlands, and emergent herbaceous wetlands. Land cover was grouped into five ordinal categories: water/wetlands (0) which includes water, woody wetlands, and emerging herbaceous wetlands; developed (1) which includes low intensity, medium intensity, and high intensity developed land; developed open space (2) which includes developed open space and barren land; cultivated/pasture land (3) which includes hay/pasture and cultivated cropland; and deciduous/evergreen forest (4) which includes deciduous, evergreen, and mixed forest. NLCD vector data was converted to raster format for analysis. For analysis the majority of each reclassified land cover type was determined for each census tract.

Statistical method

Confirmed and probable cases from January, 2000 to June, 2015 were used to calculate prevalence by census tracts. All analyses were performed using STATA statistical software and ArcGIS 10.3.

Values from each environmental variable raster were extracted to case points using the Spatial Analyst tool in ArcMap 10.3. Categorical variables were grouped into ordinal, ranked categories. Pearson correlation analysis was performed on all variables included in the model and showed that average annual temperature was highly correlated with average annual precipitation with a coefficient of 0.815. For this reason, it was excluded from analysis.

In order to determine significant variables for inclusion in the model, stepwise backwards elimination was used. Geological class and soil order were not determined as significant, however, due to *a priori* knowledge from the literature, both these variables were included in the final model [1, 9, 15, 18, 26, 39]. Logistic regression was performed using the environmental factors as independent exposure variables and clustered census tracts (1) versus non-clustered census tracts (0) as the outcome. Confounding and interaction were not assessed in this analysis; we were only interested in examining the effects of environmental risk factors as exposure variables on the incidence of RMSF. The logistic regression equation was used to generate a risk map showing the probability of the presence of Rocky Mountain spotted fever within each census tract across Tennessee. The final model included an intercept and all levels of each environmental exposure.

Chapter 3: Results

Although cases of Rocky Mountain spotted fever were most common around major cities in Tennessee, the distribution of cases can be biased by population. Therefore, prevalence of cases were determined per 100,000 people in order to get a more accurate view of the proportion of the population exposed to RMSF. As shown in figure 1, an increased prevalence of cases per 100,000 can be seen below the Land Between the Lakes area (the outcrop along the north border of TN) to the West and Northwest of Nashville, TN. Increased prevalence can also be seen in the Big South Fork area in the northern portion of the state, east of Nashville.

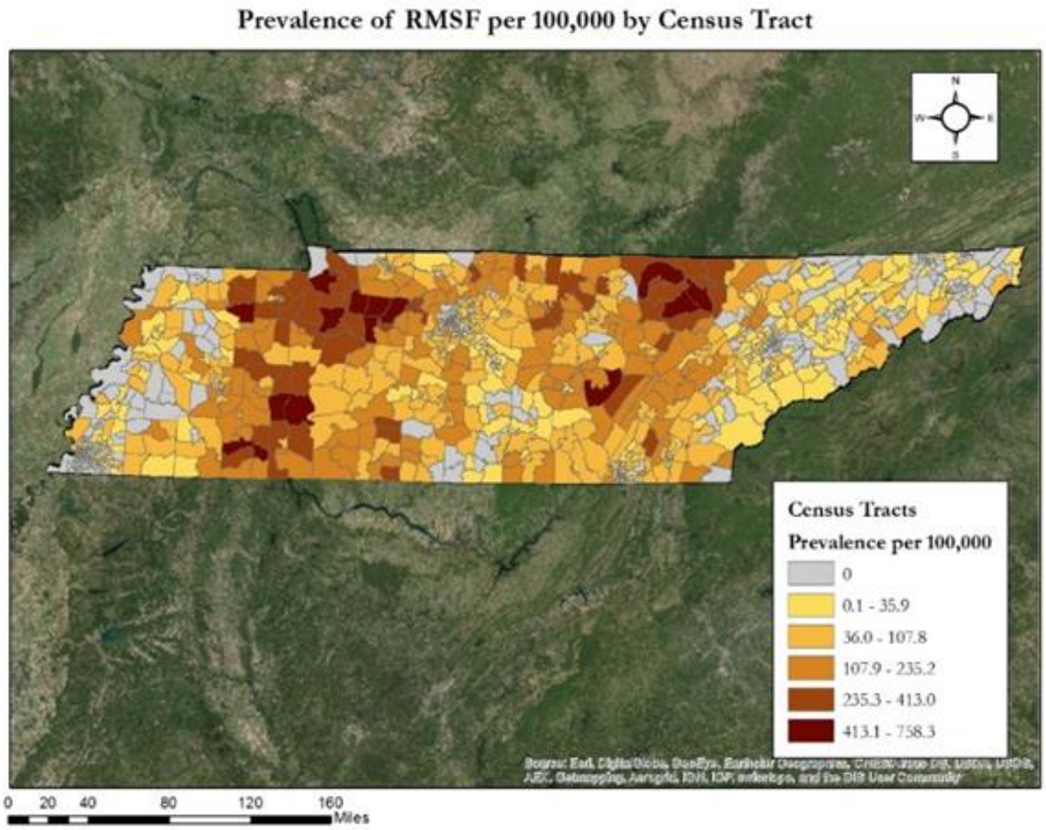

Figure 1. Prevalence of RMSF per 100,000 by census tract

Anselin local Moran's I was used to determine statistically significant clusters for each census tract. Cluster analysis of positive cases showed areas of High-High clustering north of the capital. Areas of Low-Low clustering were seen in the southwestern section of the state around the city of Memphis (Figure 2.).

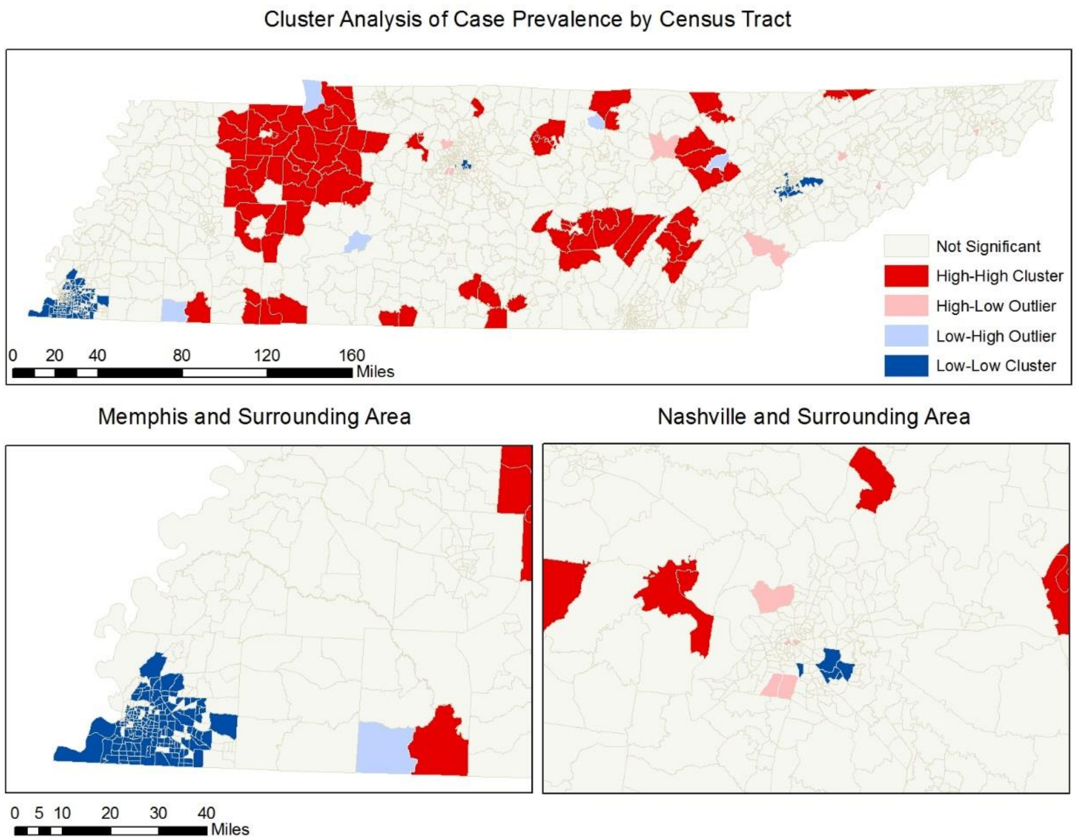

Figure 2. Cluster analysis of census tracts using Anselin Local Moran's I.

Descriptive Statistics

Precipitation

Continuous precipitation data combined with RMSF case data showed that 23% of positive cases occurred in areas that received over 450 inches of rainfall a year. Mean minimum and maximum were 223.6 and 473.3 inches respectively. An odds ratio of 1.06 was found for average annual precipitation with a p-value of <0.001 when adjusting for all other variables. 95% Confidence intervals for precipitation were 1.041 and 1.086.

Temperature

Continuous temperature data combined with RMSF case data showed that 53% of positive cases occurred within 56.71 to 58.69 annual average temperature. Minimum and maximum were 50.5 and 63.2 °F, respectively. During correlation analysis, temperature was found to be highly correlated with precipitation with a coefficient of 0.815. For this reason, temperature was excluded from the logistic regression analysis.

Soil

Soil order data combined with RMSF case data indicated that 53% of cases occurred within the ultisol soil order (clay-enriched subsoils). 26.5% of positive cases occurred in the alfisol soil category (naturally fertile with high base saturation), 15% occurred in the inceptisol category (young, with weak profile development) and the remaining 5% was split between the vertisol, mollisol, and entisol categories. Soil order had an odds ratio of 0.85 with a p-value of 0.325. 95% confidence intervals were 0.614 and 1.175. Although insignificant, soil order was retained in the model due to a priori knowledge

Elevation

Elevation data contained mean minimum and maximum values of 62.1 and 1,069.7 meters, respectively. An odds ratio of 0.99 was found for elevation when accounting for other variables in the model with a p-value of 0.004. 95% confidence intervals for elevation were 0.9903 and 0.9908.

Geology

Geology data combined with RMSF case data showed 21.2% of positive cases occurring within the Clay or Mud category. 12.3% of cases occurred within the limestone category, 11.6% within the shale category, 10% within the chert category 9.7% within dolostone, 9.5% within sand, 7.9% within silt, 6.5% within calcarenite, and the remaining 6% within all other categories. An odds ratio of 1.50 was found for geologic class with a p-value of 0.22. 95% confidence intervals for geologic class were 0.778 and 2.910. Like soil order, geologic class was retained in the final model, despite its statistical insignificance, due to a priori knowledge.

Land cover

Land cover data combined with RMSF case data showed 36.9% of positive cases occurred in the developed, open space category. 22.7% occurred in low intensity developed land, 13.4% occurred in deciduous forest, 13.6% in hay/pasture, 3.9% in herbaceous, and 3.1% in medium intensity developed land. Land cover had an odds ratio of 7.61 with a p-value of <0.001. The confidence intervals for this variable were 3.612 and 16.020.

Table 1. Significant and a priori environmental variables in the logistic regression model.

Variable	Beta	Std. Error	P-Value	Odds Ratio	95% C.I Lower	Upper
Land Cover	2.029	0.379	<0.001	7.61	3.612	16.020
Soil Order	-0.163	0.165	0.325	0.85	0.614	1.175
Geologic Class	0.409	0.336	0.224	1.50	0.778	2.910
Elevation	-0.006	0.002	0.004	0.99	0.990	0.998
Precipitation	0.061	0.011	<0.001	1.06	1.041	1.086
Constant	-36.617	5.068				

Modelling Results

Risk map

The predictive risk map generated from the logistic regression model is shown in figure 3. Higher probabilities show areas of increased RMSF presence. No areas showed greater than a 75% probability for RMSF presence within the state. Based on clustering of high case prevalence by census tract, the area between Memphis and Nashville (western portion of central TN) shows the highest probability for RMSF. Scattered areas of medium-high (50 – 75%) risk can be found throughout the rest of the state with areas of low (10%-25%) risk covering the majority of the central part of the state. The area of lowest probability is towards the eastern section of the state where the Appalachian Mountains run along the border.

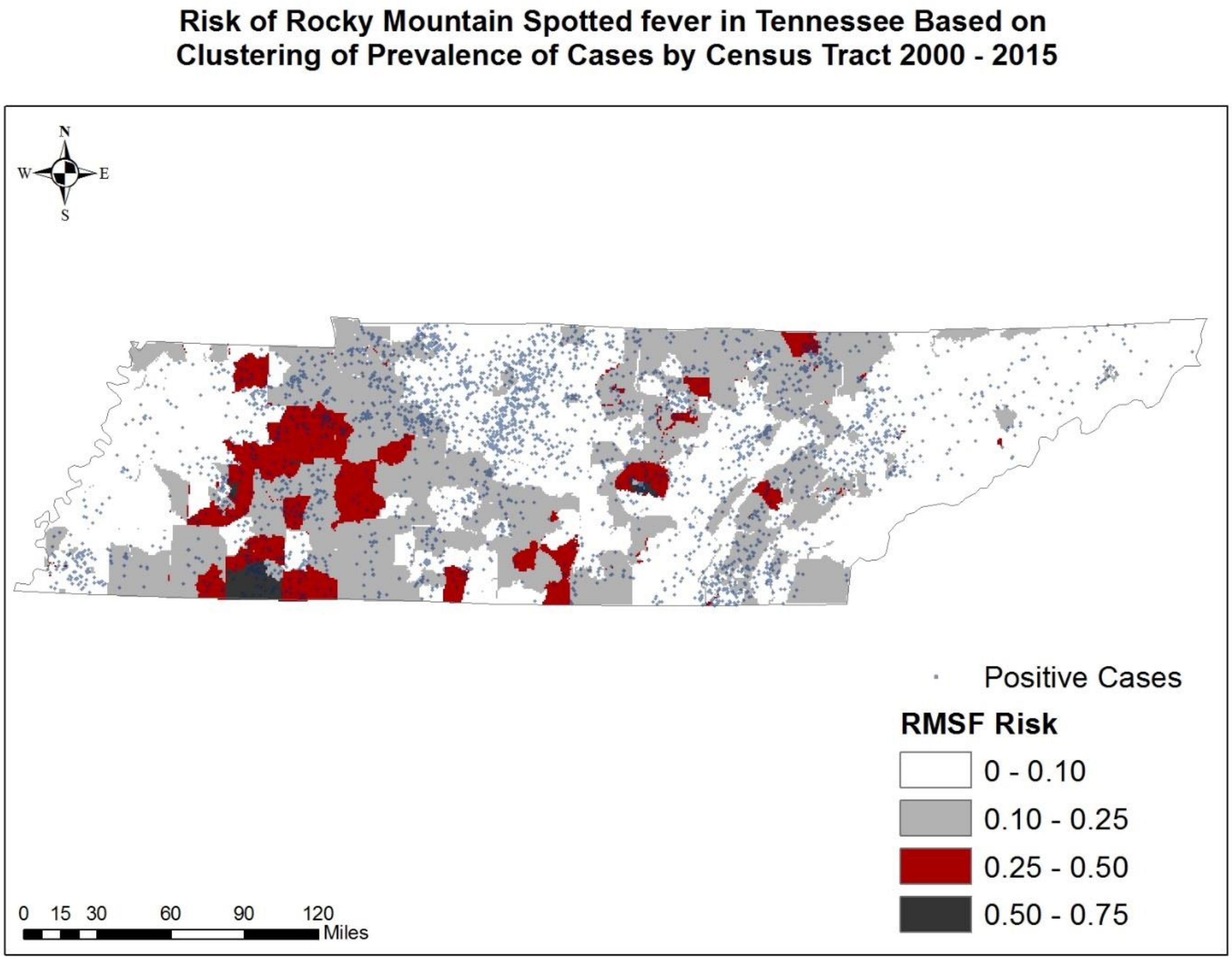

Figure 3. Risk map showing probability of RMSF presence in Tennessee

Chapter 4: Discussion

The primary goal of this project was to assess risk of Rocky Mountain spotted fever in relation to six environmental variables across the state of Tennessee. Environmental variables included Soil order, land cover type, geology type, elevation, average annual precipitation, and average annual temperature. This study also aimed to address a gap in the literature for RMSF as the majority of research and analysis has been concerned with other tick borne diseases such as Lyme disease. Even though RMSF has a large presence in the southeast and is the number one tick borne disease in Tennessee, there is very little literature associated with its environmental drivers of risk.

Research has shown that the distribution of tick vectors is associated with environmental variables [9, 15, 34]. Much work has been done on the tick vector, *Ixodes scapularis*, the primary vector of Lyme disease, and has shown that abundance and infection prevalence is associated with variables such as precipitation, temperature, humidity, vegetation, and soil type [12, 15, 18, 19, 22]. For this reason, several of these environmental drivers of risk were quantified for the state of Tennessee.

Multivariate analysis showed land cover, elevation, and average annual precipitation categories to be significantly associated with RMSF presence. Stepwise backwards elimination was performed in order to determine individual variables of significance. Although multivariate analysis did not prove significant for geologic class and soil order, they were retained in the model based on a priori knowledge of their role in vector distribution.

It is possible that soil orders and geologic classes were found to be insignificant due to effect modification. In the instance of this model, the magnitude of effect of soil order or geological class could be influenced by type of land cover.   Since confounding and interaction terms were not assessed in this study, further studies are needed to assess the effect of these variables as confounders and determine the how they might modify results.

The predictive risk model developed for this study showed that the majority of significant clusters of high prevalence were found in areas with high average annual precipitation and in lower elevation. The majority of these significant clusters also were found in the ultisols soil order in deciduous and mixed forest types. Geologic classes with higher drainage capacity, such as sediment and sedimentary rock can be seen around the areas of high clusters. Shrub layer and vegetation found in forested areas provides cover and areas where questing for hosts can occur. Since the tick vectors that can spread RMSF are sensitive to hot and dry conditions, they require habitat that can provide moist, cool coverage.

Tick sampling could be used to validate the model created in this study. Sampling sites can be selected based on areas of high (and low) RMSF probability and vulnerable populations. This study was limited in that it was not able to take into account variables such as host abundance and distribution, as well as other abiotic and biotic factors that influence tick distribution. The inclusion of other variables such as host density and forest fragmentation will increase the robustness of the study.

There are inherent disadvantages to working with epidemiological data because true tick bite locations cannot be assessed based on passive surveillance using addresses as proxies for tick bite locations. Despite this, this model can be used as a starting point for assessing areas that can be prioritized for active surveillance and also as an educational guide to help inform the public of areas at risk for Rocky Mountain spotted fever. More in-depth studies could lead to new health policy initiatives regarding tick borne disease in Tennessee and could help to increase the amount of knowledge of this pathogen within the state.

Chapter 5: References

[1] Altizer, S., Ostfeld, R. S., Johnson, P. T., Kutz, S., & Harvell, C. D. (2013). Climate change and infectious diseases: from evidence to a predictive framework. *science*, *341*(6145), 514-519.

[2] Burgdorfer, W. (1969). Ecology of tick vectors of American spotted fever. *Bulletin of the World Health Organization, 40*(3), 375-381. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2554644/ http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2554644/pdf/bullwho00224-0045.pdf

[3] Campbell, A., & MacKay, P. R. (1979). Distribution of the American dog tick, Dermacentor variabilis (Say), and its small-mammal hosts in relation to vegetation types in a study area in Nova Scotia. *Canadian Journal of Zoology*, *57*(10), 1950-1959. doi:10.1139/z79-258

[4] Clarke, K. C, McLafferty S. L, & Tempalski B. J. On Epidemiology and Geographic Information Systems: A Review and Discussion of Future Directions. Emerg Infect Dis. 1996, Jun. Retrieved from http://wwwnc.cdc.gov/eid/article/2/2/96-0202. DOI: 10.3201/eid0202.960202

[5] Climate Data Online. Retrieved April 12, 2016, from https://www.ncdc.noaa.gov/cdo-web/

[6] Cooney, J. C., & Burgdorfer, W. (1974). Zoonotic potential (Rocky Mountain spotted fever and tularemia) in the Tennessee Valley region. I. Ecologic studies of ticks infesting mammals in Land Between the Lakes. *American Journal of Tropical Medicine and Hygiene*, *23*(1), 99-108.

[7] Daniel, M., Kolár, J., Zeman, P., Pavelka, K., & Sádlo, J. (1998). Predictive map of Ixodes ricinus high-incidence habitats and a tick-borne encephalitis risk assessment using satellite data. *Experimental & applied acarology*, *22*(7), 417-433

[8] Dantas-Torres, F. (2007). Rocky Mountain spotted fever. *The Lancet Infectious Diseases, 7*(11), 724-732. Retrieved from http://www.thelancet.com/journals/laninf/article/PIIS1473-3099(07)70261-X/abstract

[9] Del Fabbro, S., Gollino, S., Zuliani, M., & Nazzi, F. (2015). Investigating the relationship between environmental factors and tick abundance in a small, highly heterogeneous region. *Journal of Vector Ecology*, *40*(1), 107-116.

[10] Dodds, D. G., Martell, A. M., & Yescott, R. E. (1969). Ecology of the American dog tick, Dermacentor variabilis (Say), in Nova Scotia. *Canadian Journal of Zoology, 47*(2), 171-181. doi:10.1139/z69-039

[11] ESRI 2001. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems

[12] Feria-Arroyo, T. P., Castro-Arellano, I., Gordillo-Perez, G., Cavazos, A. L., Vargas-Sandoval, M., Grover, A., ... & Esteve-Gassent, M. D. (2014). Implications of climate change on the distribution of the tick vector Ixodes scapularis and risk for Lyme disease in the Texas-Mexico transboundary region. *Parasit Vectors*, *7*(1), 199.

[13] Furlanello, C., Neteler, M., Merler, S., Menegon, S., Fontanari, S., Donini, A., ... & Chemini, C. (2003). GIS and the random forest predictor: Integration in R for tick-borne disease risk assessment. In *Proceedings of DSC* (p. 2).

[14] Geographic distribution of ticks that bite humans. (2015, June 1). Retrieved January 18, 2016, from http://www.cdc.gov/ticks/geographic_distribution.html

[15] Glass, G. E., Schwartz, B. S., Morgan, J. M., Johnson, D. T., Noy, P. M., & Israel, E. (1995). Environmental risk factors for Lyme disease identified with geographic information systems. *American Journal of Public Health*, *85*(7), 944–948.

[16] Guerra, M., Walker, E., Jones, C., Paskewitz, S., Cortinas, M. R., Stancil, A., … Kitron, U. (2002). Predicting the Risk of Lyme Disease: Habitat Suitability for*Ixodes scapularis* in the **North Central United States**. *Emerging Infectious Diseases*, *8*(3), 289–297. http://doi.org/10.3201/eid0803.010166

[17] Hay, S. I., Packer, M. J., & Rogers, D. J. (1997). Review article The impact of remote sensing on the study and control of invertebrate intermediate hosts and vectors for disease. *International Journal of Remote Sensing*, *18*(14), 2899-2930.

[18] Khatchikian, C. E., Prusinski, M., Stone, M., Backenson, P. B., Wang, I. N., Levy, M. Z., & Brisson, D. (2012). Geographical and environmental factors driving the increase in the Lyme disease vector Ixodes scapularis. *Ecosphere*, *3*(10), 1-18.

[19] Kitron, U., & Kazmierczak, J. J. (1997). Spatial analysis of the distribution of Lyme disease in Wisconsin. *American journal of Epidemiology*, *145*(6), 558-566.

[20] Moncayo, A. C., Cohen, S. B., Fritzen, C. M., Huang, E., Yabsley, M. J., Freye, J. D., . . . Dunn, J. R. (2010). Absence of Rickettsia rickettsii and occurrence of other spotted fever group rickettsiae in ticks from Tennessee. *American Journal of Tropical Medicine and Hygiene*, *83*(3), 653-657. doi:10.4269/ajtmh.2010.09-0197

[21] Natural Resources Conservation Service Soils. Retrieved April 12, 2016, from http://www.nrcs.usda.gov/wps/portal/nrcs/site/soils/home/

[22] Ogden, N. H., St-Onge, L., Barker, I. K., Brazeau, S., Bigras-Poulin, M., Charron, D. F., ... & Michel, P. (2008). Risk maps for range expansion of the Lyme disease vector, Ixodes scapularis, in Canada now and with climate change. *International journal of health geographics*, *7*(1), 1.

[23] Openshaw, J. J., Swerdlow, D. L., Krebs, J. W., Holman, R. C., Mandel, E., Harvey, A., . . .

McQuiston, J. H. (2010). Rocky mountain spotted fever in the United States, 2000-2007: interpreting contemporary increases in incidence. *American Journal of Tropical Medicine and Hygiene*, *83*(1), 174-182. doi:10.4269/ajtmh.2010.09-0752

[24] Ostfeld, R. S., Glass, G. E., & Keesing, F. (2005). Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends in ecology & evolution*, *20*(6), 328-336.

[25] Perlman, S. J., Hunter, M. S., & Zchori-Fein, E. (2006). The emerging diversity of Rickettsia. *Proceedings of the Royal Society of London B: Biological Sciences, 273*(1598), 2097-2106. doi:10.1098/rspb.2006.3541

[26] Randolph, S. (2002). Predicting the risk of tick-borne diseases. *International journal of medical microbiology*, *291*, 6-10.

[27] Randolph. S. E., (2008). The impact of tick ecology on pathogen transmission dynamics. In: Alan S. Bowman and Patricia A. Nuttall (eds.) *Ticks*. pp. 40-72. [Online]. Cambridge: Cambridge University Press. Available from: Cambridge Books Online http://dx.doi.org/10.1017/CBO9780511551802.003

[28] Rocky Mountain Spotted Fever. (2014, October 23). Retrieved January 18, 2016, from http://www.cdc.gov/ticks/tickbornediseases/rmsf.html

[29] Soil Survey Staff. 2015. Illustrated guide to soil taxonomy. U.S. Department of Agriculture, Natural Resources Conservation Service, National Soil Survey Center, Lincoln, Nebraska

[30] Sonenshine, D. E., Elisberg, B. L., Atwood, E. L., & Lamb Jr, J. T. (1999). THE ECOLOGY OF TICK VECTORS OF ROCKY MOUNTAIN SPOTTED FEVER IN VIRGINIA, USA. In A. Corradetti (Ed.), *Proceedings of the First International Congress of Parasitology* (pp. 1033-1034): Pergamon.

[31] "Spotted Fever Rickettsiosis (Rickettsia Spp.) 2010 Case Definition." *National Notifiable Diseases Surveillance System (NNDSS)*. Web.

[32] State & County QuickFacts. (2015, December 1). Retrieved January 18, 2016, from http://quickfacts.census.gov/qfd/states/47000.html

[33] Statistics and Epidemiology. (2013, September 5). Retrieved January 18, 1016, from http://www.cdc.gov/rmsf/stats/

[34] Tabachnick, W. J. (2010). Challenges in predicting climate and environmental effects on vector-borne disease episystems in a changing world. *The Journal of experimental biology*, *213*(6), 946-954.

[35] Tennessee GIS Clearinghouse. Retrieved April 12, 2016, from http://www.tngis.org/

[36] The National Map: Elevation. Retrieved April 12, 2016, from http://nationalmap.gov/elevation.html

[37] The USGS Land Cover Institute. Retrieved April 12, 2016, from http://landcover.usgs.gov/uslandcover.php

[38] Thorner, A. R., Walker, D. H., & Petri, W. A.. (1998). Rocky Mountain Spotted Fever. *Clinical Infectious Diseases*, *27*(6), 1353–1359. Retrieved from http://www.jstor.org/stable/4481728

[39] Wimberly, M. C., Baer, A. D., & Yabsley, M. J. (2008). Enhanced spatial models for predicting the geographic distributions of tick-borne pathogens. *International Journal of Health Geographics*, *7*(1), 1.

Appendix I: Maps

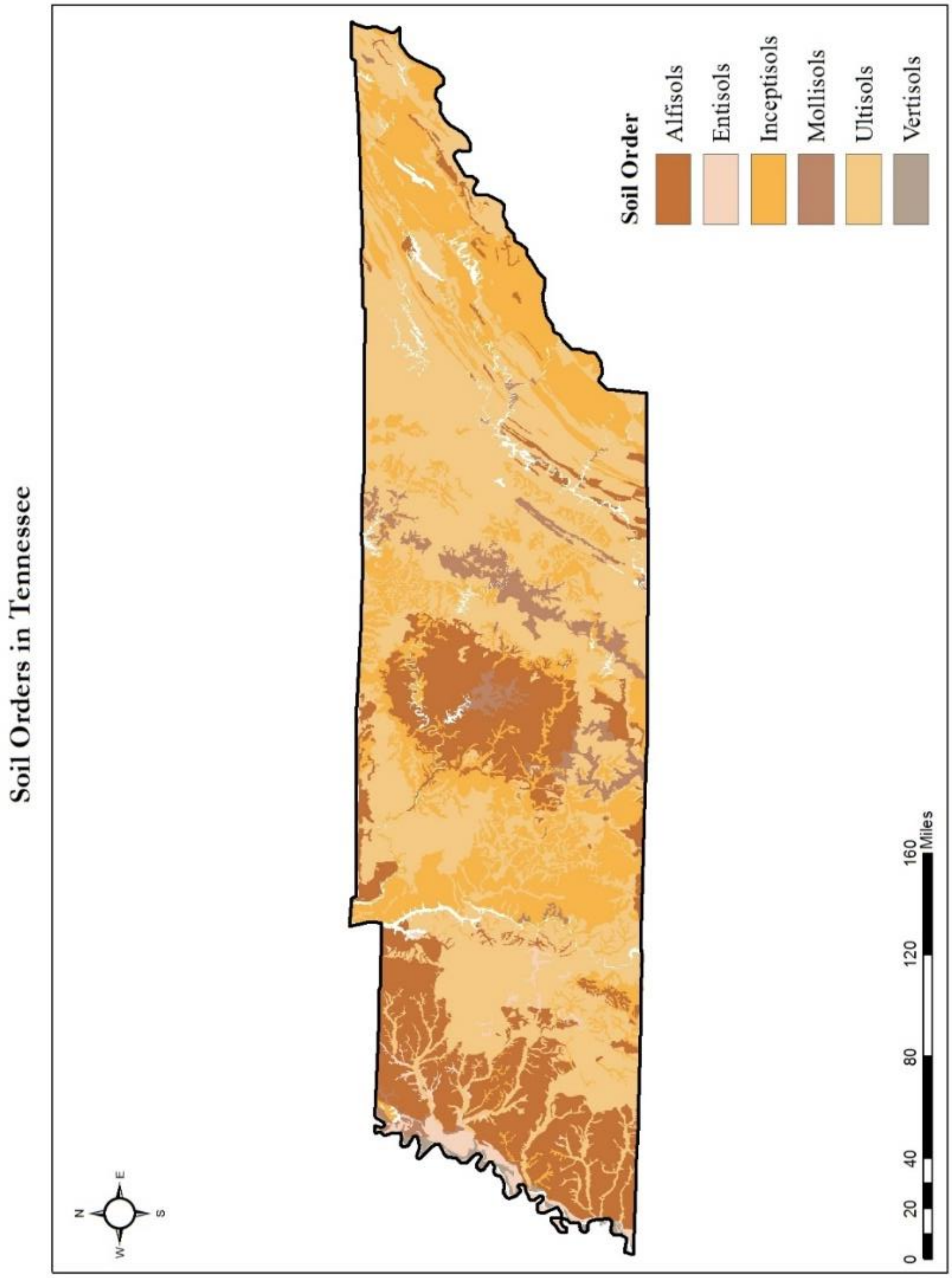

Figure 4. Soil orders of Tennessee

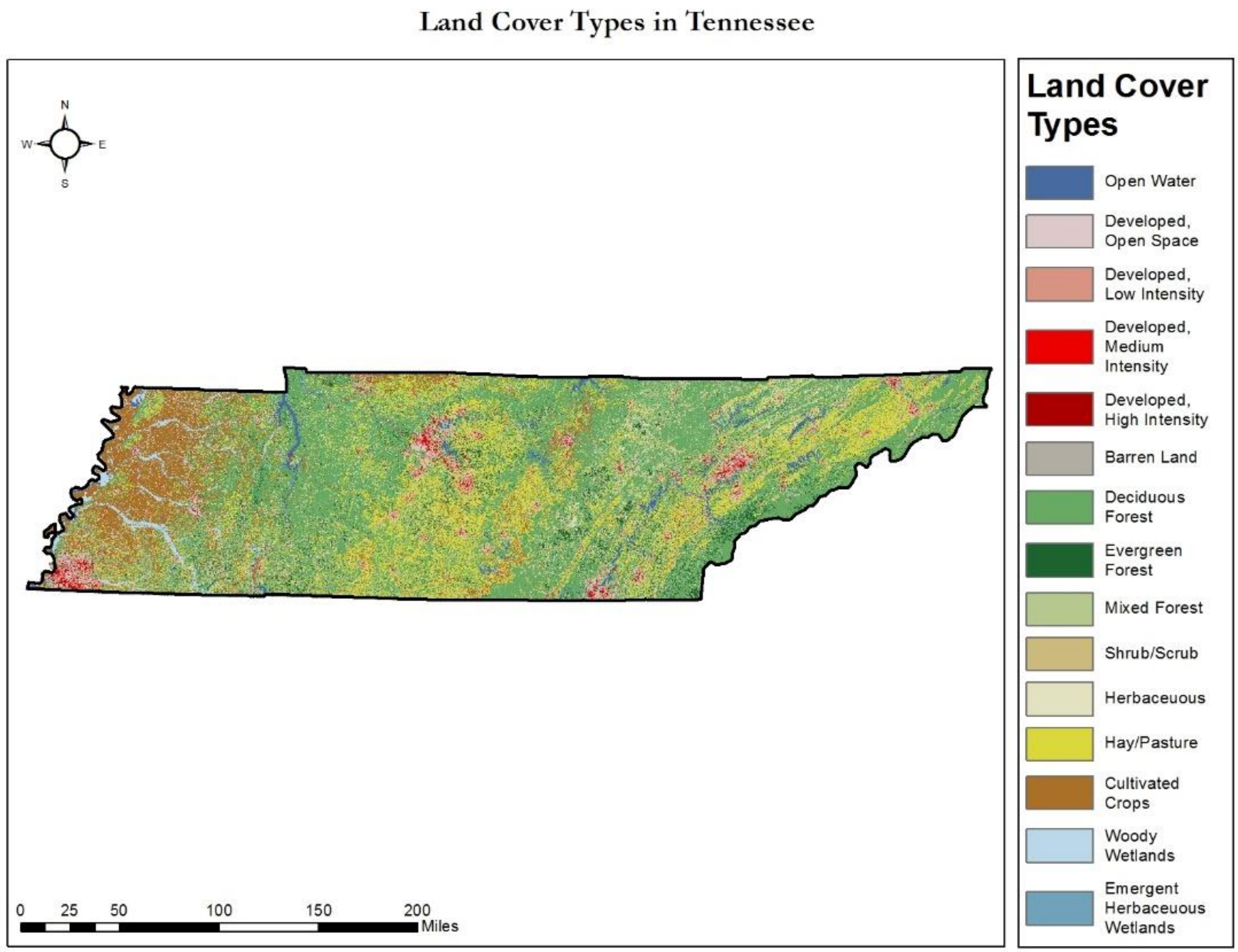

Figure 5. Land cover in Tennessee

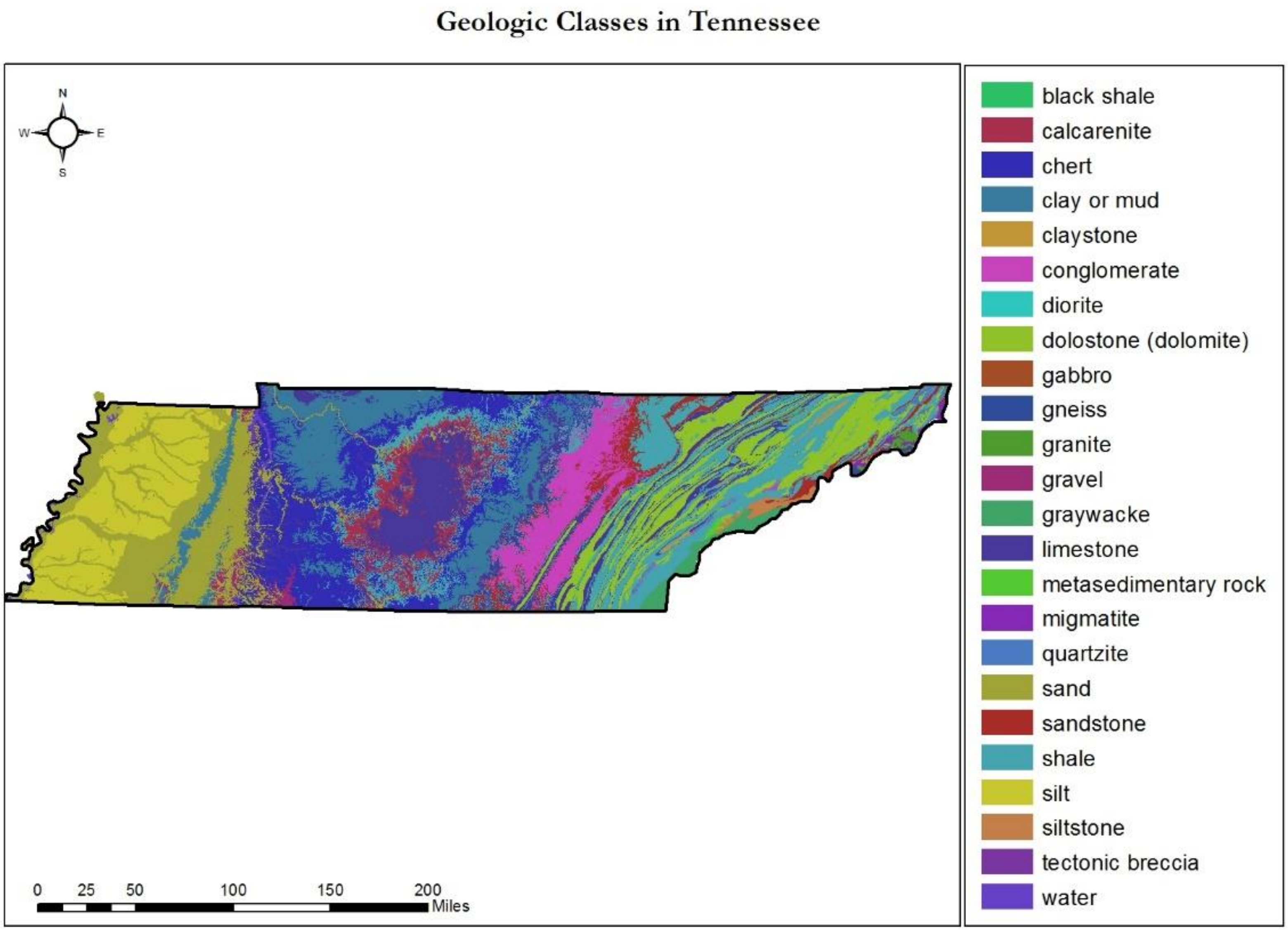

Figure 6. Geologic Classes in Tennessee

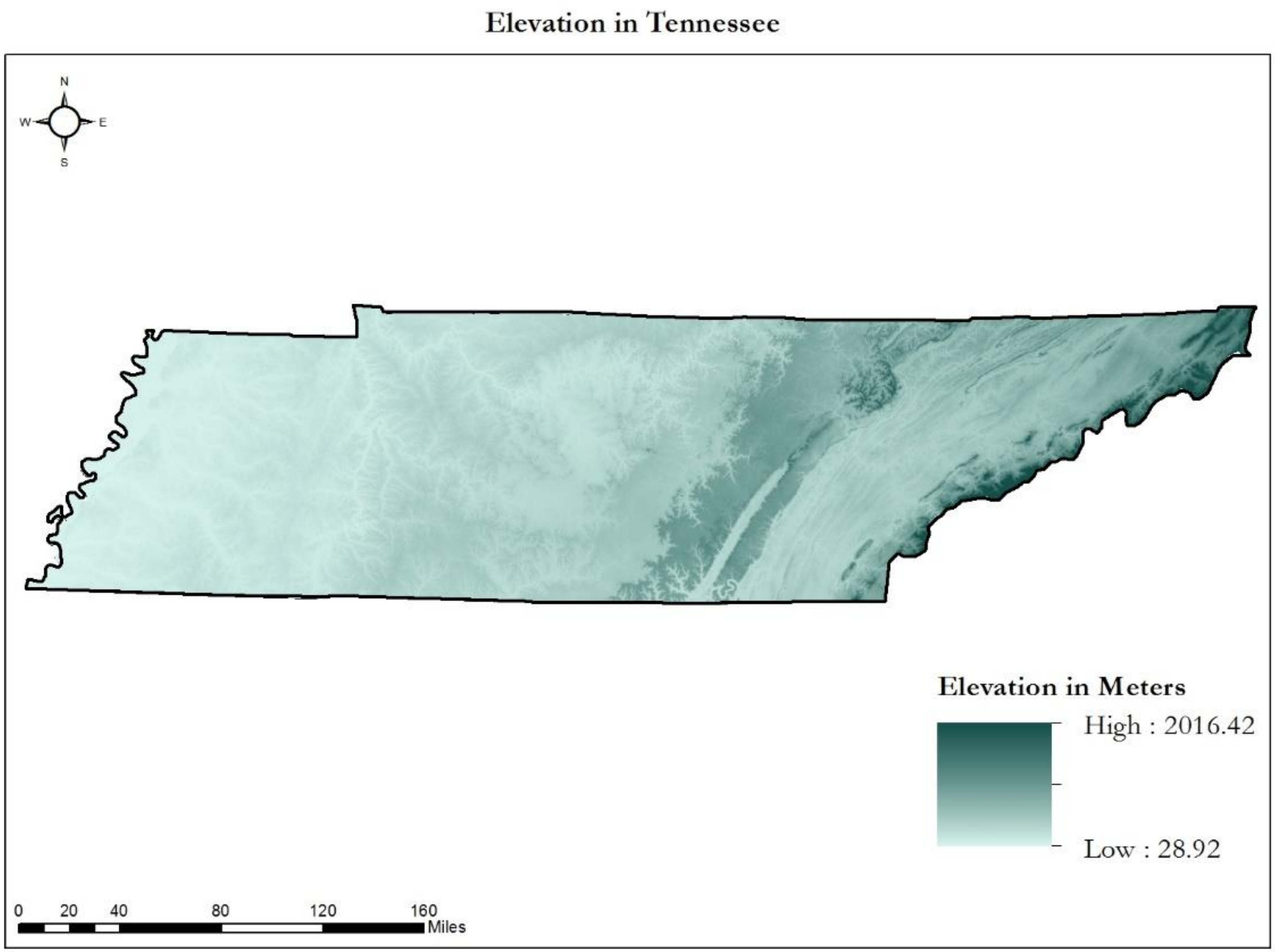

Figure 7. Elevation in Tennessee

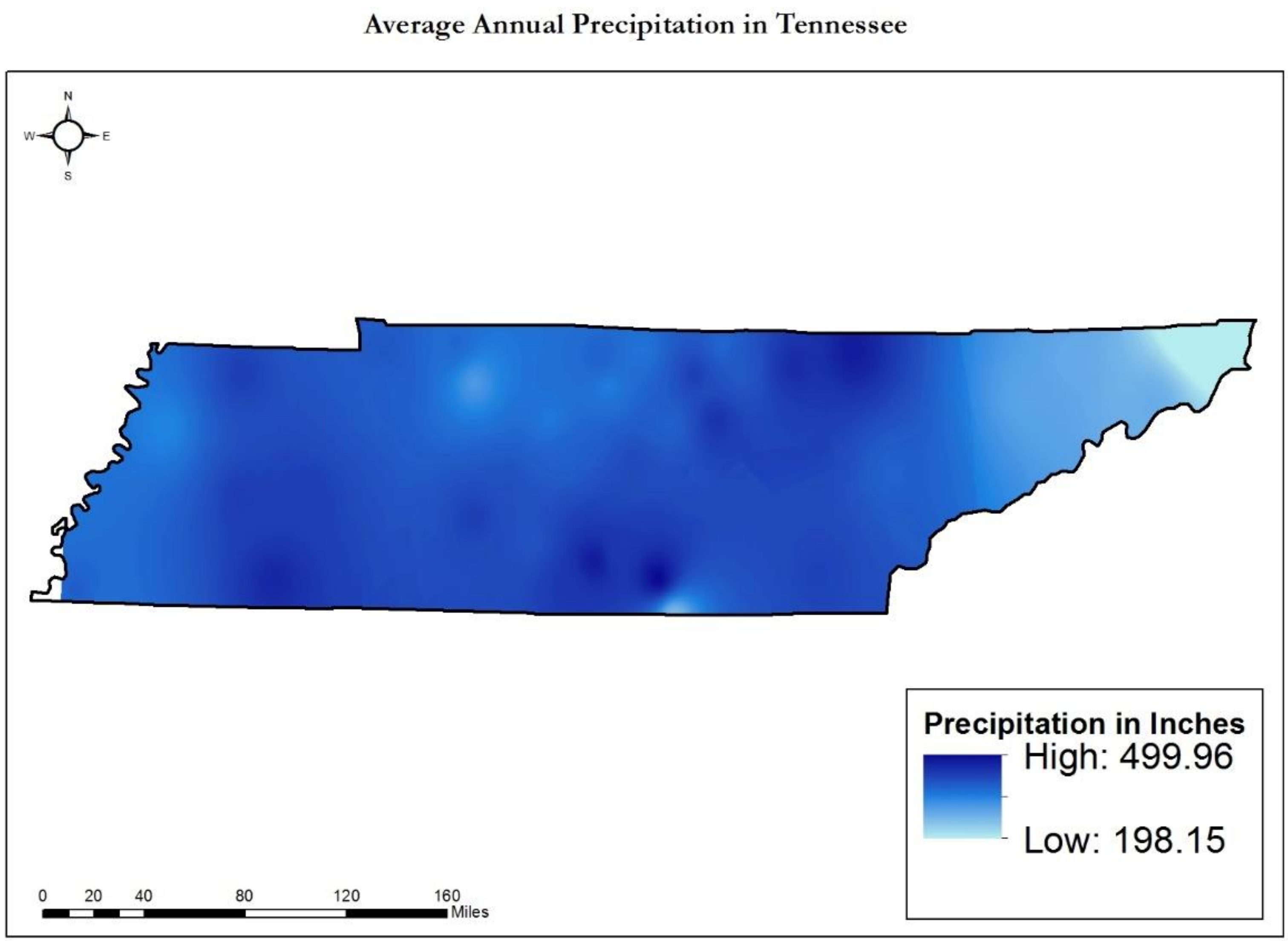

Figure 8. Average Annual Precipitation in Tennessee

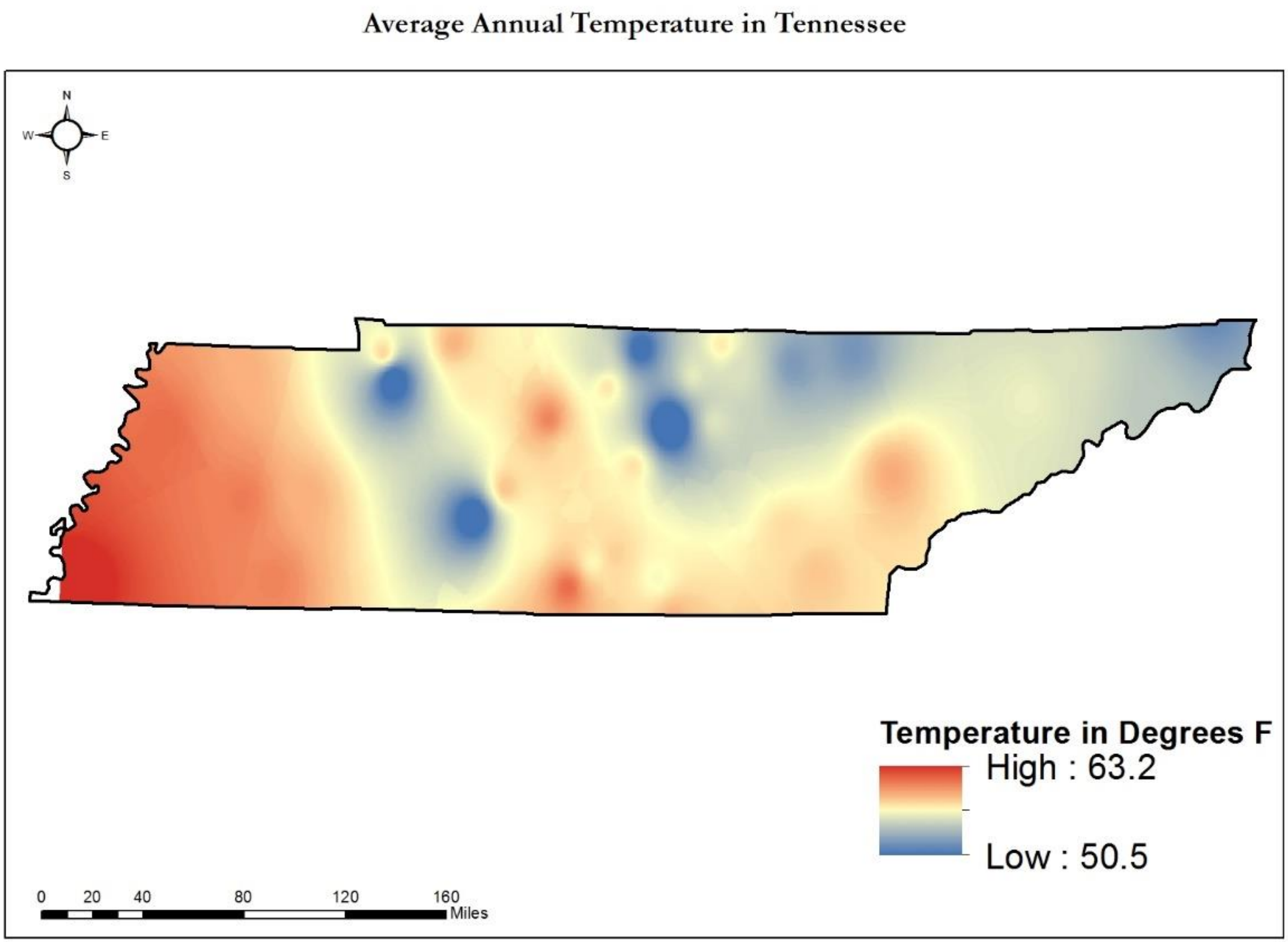

Figure 9. Average Annual Temperature in Tennessee

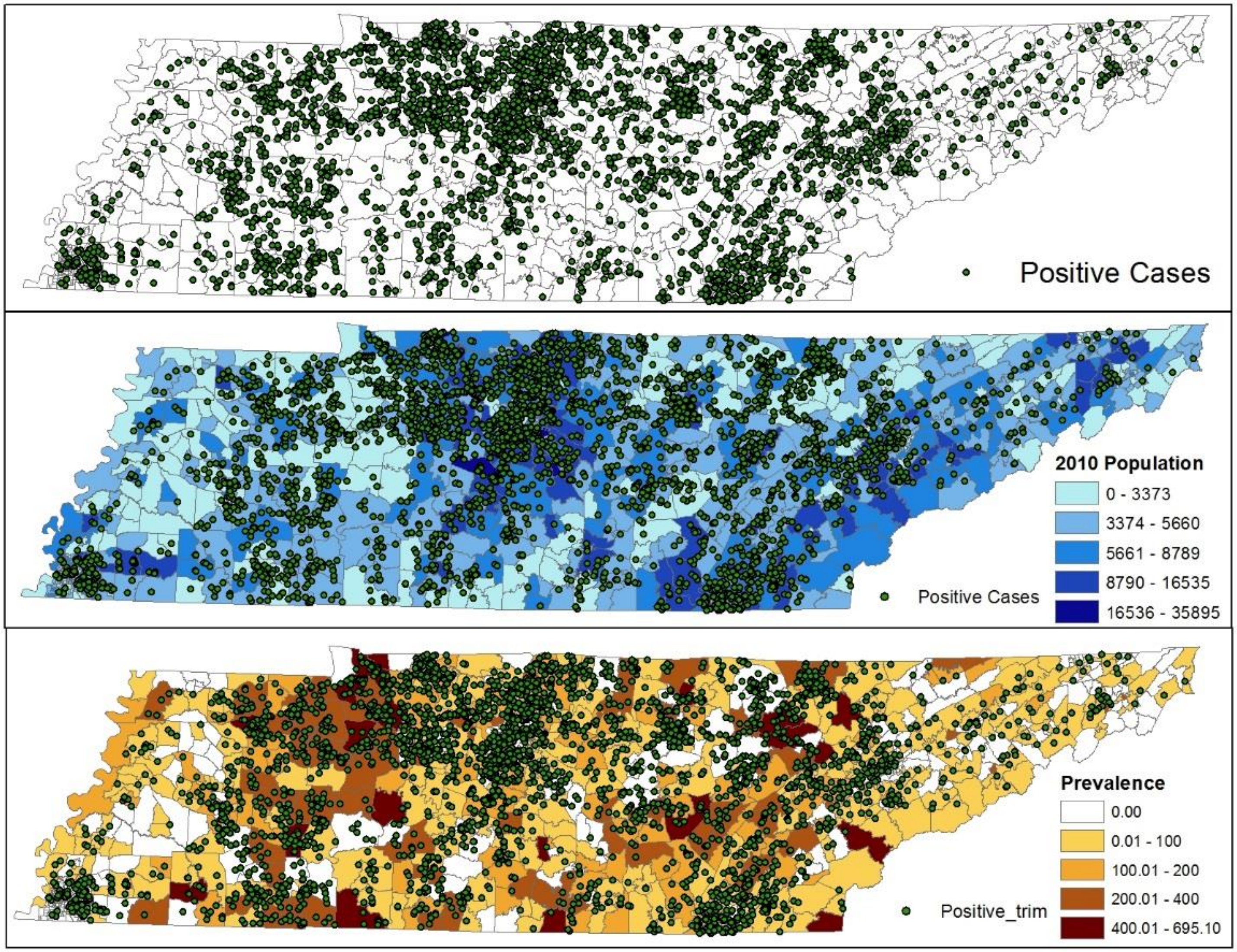

Figure 10. Positive RMSF cases in relation to population and prevalence by census tract

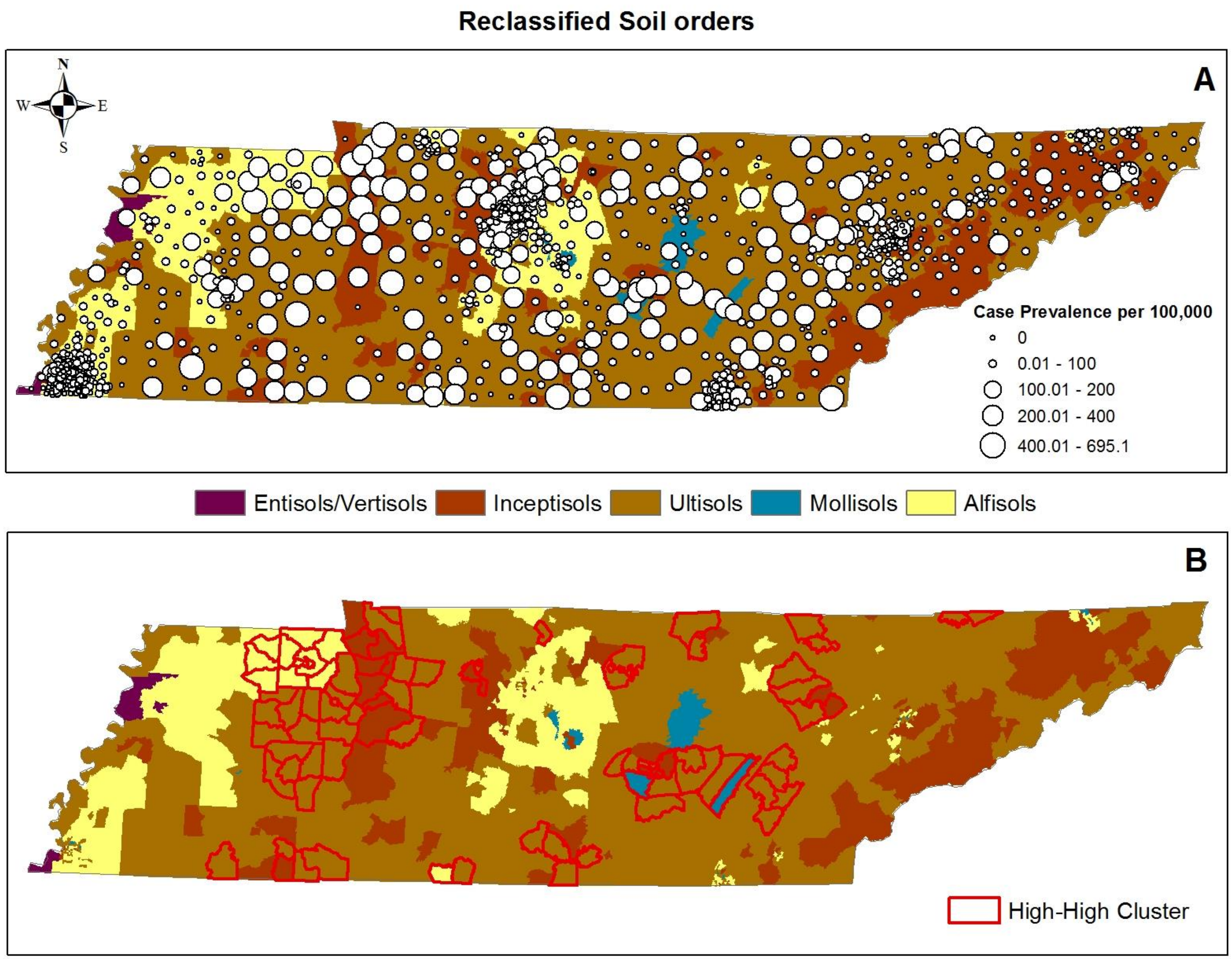

Figure 11. Soil order reclassified into majority order type by census tract, overlaid with case prevalence per 100,000 (A) and high-high clustered census tracts (B).

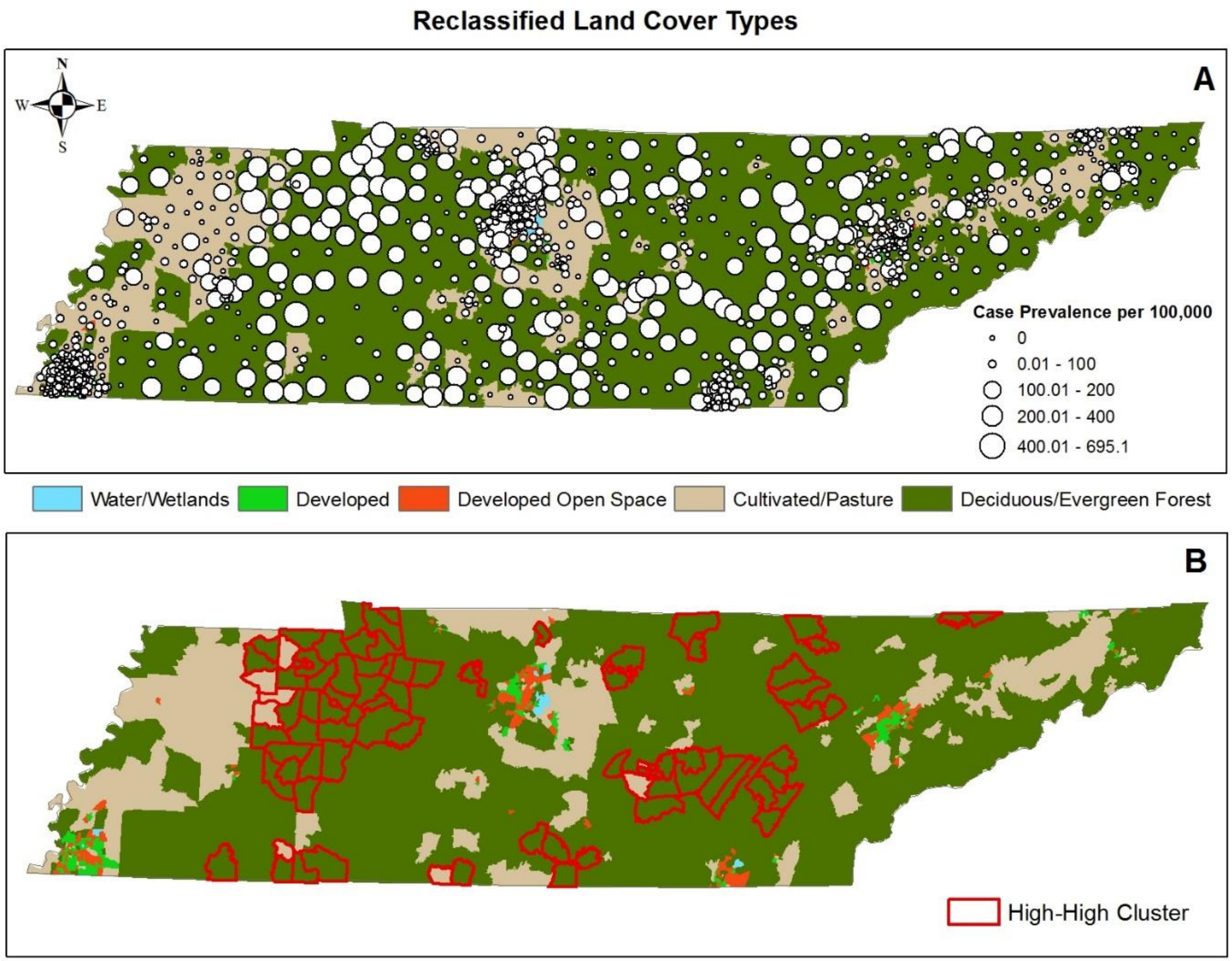

Figure 12. Land Cover reclassified into majority type by census tract, overlaid with case prevalence per 100,000 (A) and high-high clustered census tracts (B).

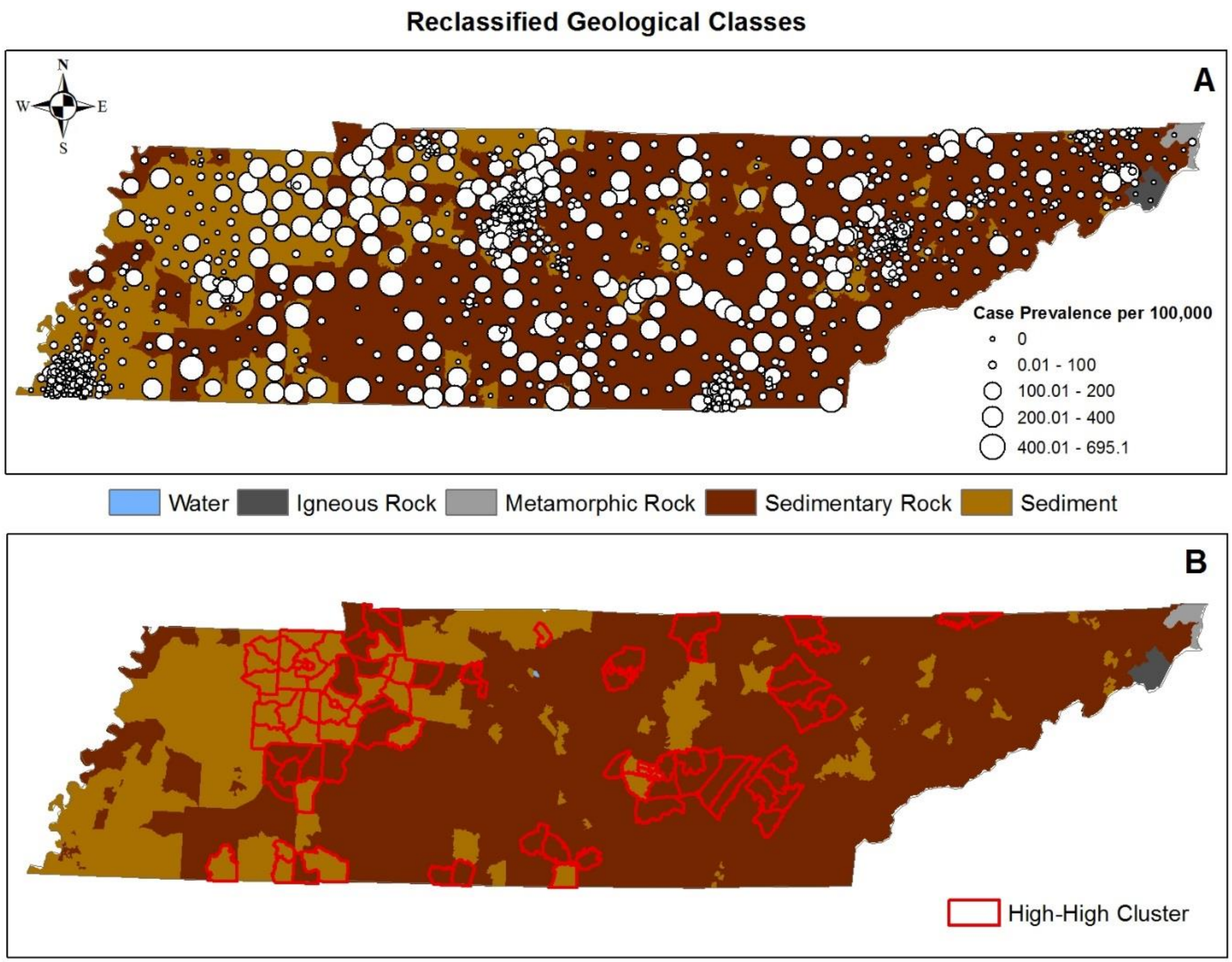

Figure 13. Geological Classes reclassified into majority type by census tract, overlaid with case prevalence per 100,000 (A) and high-high clustered census tracts (B).

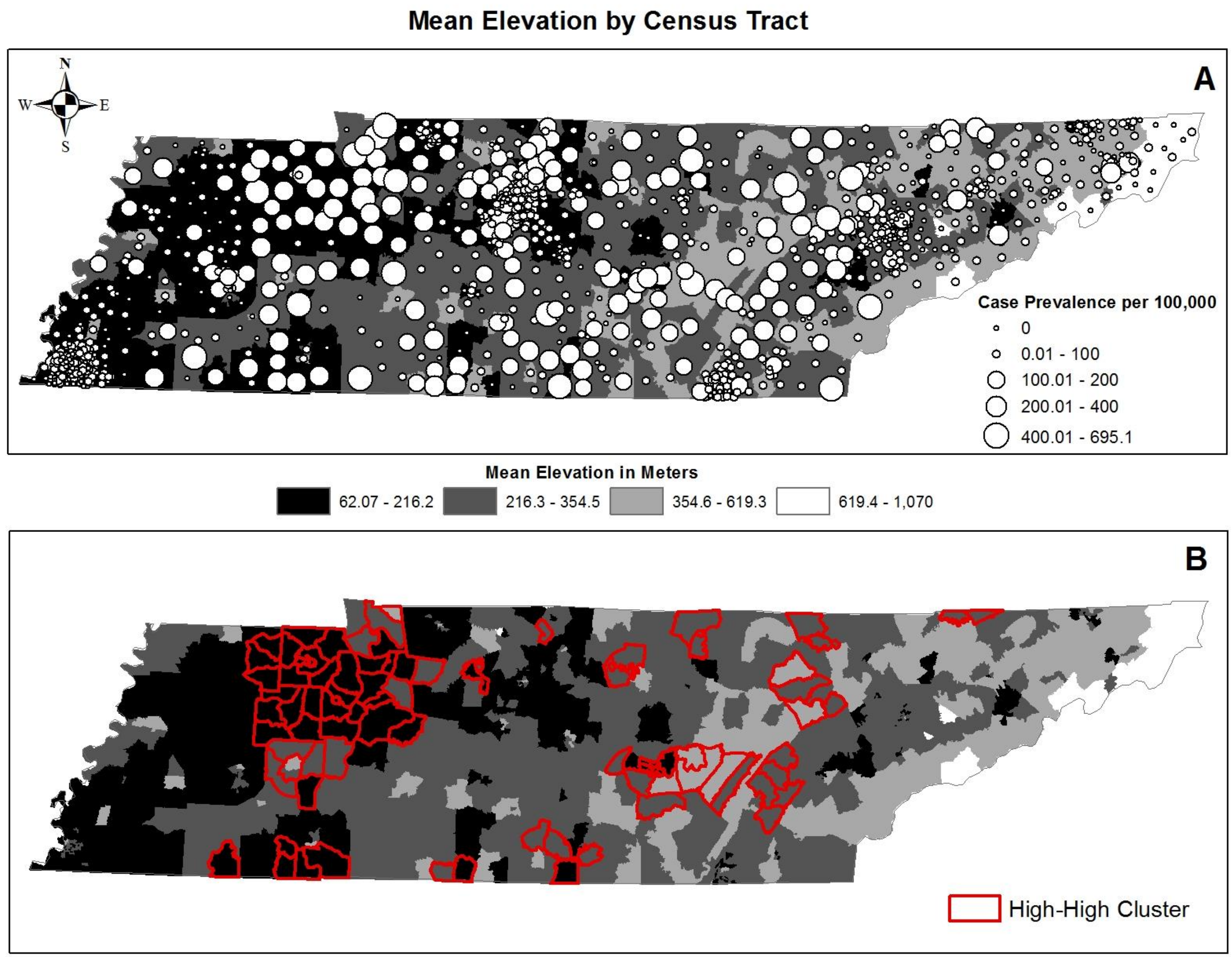

Figure 14. Elevation across the state reclassified into mean elevation by census tract, overlaid with case prevalence per 100,000 (A) and high-high clustered census tracts (B).

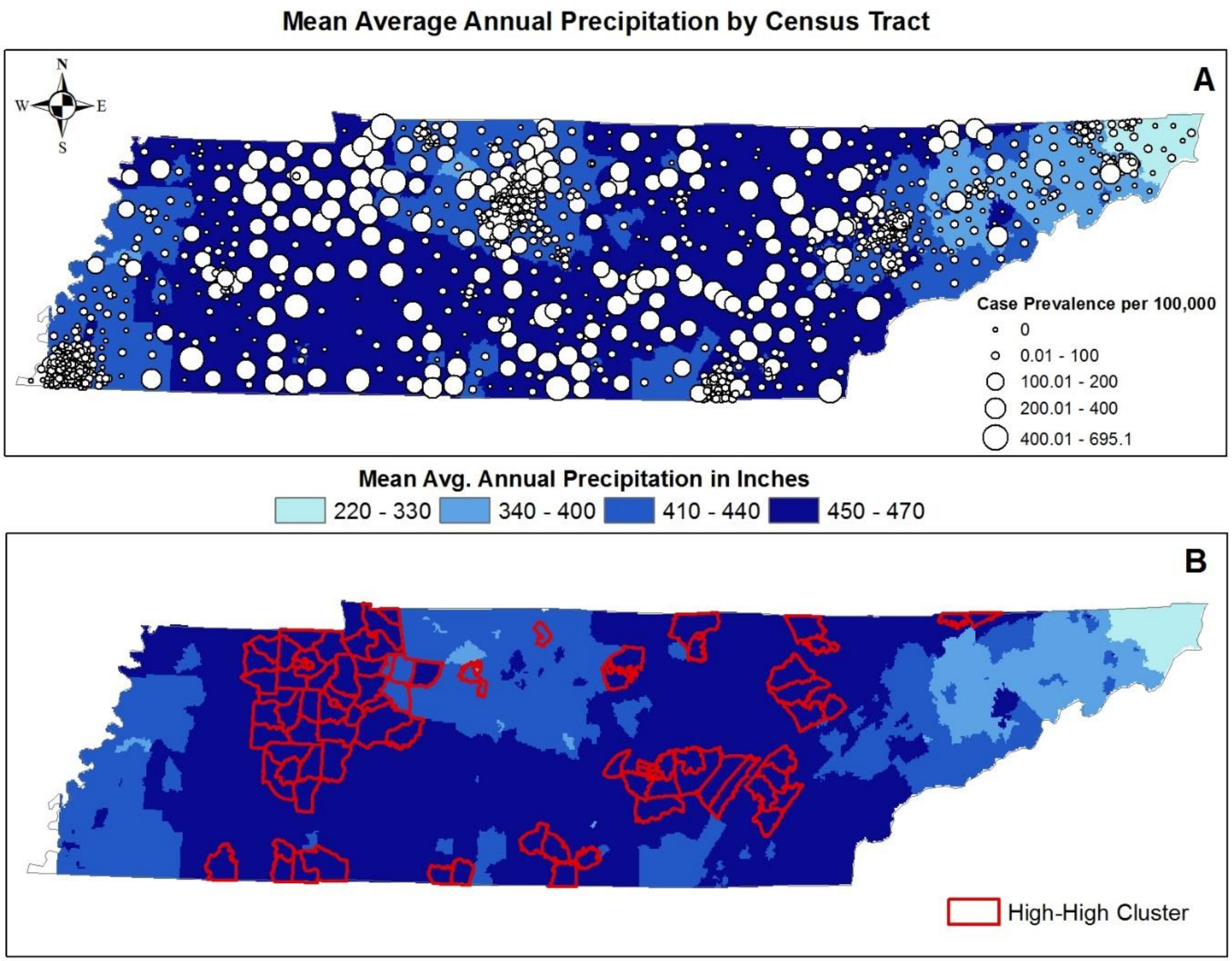

Figure 15. Average Annual Precipitation across the state reclassified into mean precipitation by census tract, overlaid with case prevalence per 100,000 (A) and high-high clustered census tracts (B).