7-18-2013

A Spatial Analysis of West Nile Virus in Texas, 2012

Allison G. Bradshaw
University of Connecticut - Storrs, allison.bradshaw@uconn.edu

A Spatial Analysis of West Nile Virus in Texas, 2012

Allison G. Bradshaw

B.A., College of William and Mary, 2007

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Arts

At the

University of Connecticut

2013

APPROVAL PAGE

Master of Arts Thesis

A Spatial Analysis of West Nile Virus in Texas, 2012

Presented by

Allison G. Bradshaw, B.A.

Major Advisor ______________________________

Debarchana Ghosh, Ph.D.

Associate Advisor______________________________

Jeffrey Osleeb, Ph.D.

Associate Advisor______________________________

Chuanrong Zhang, Ph.D.

University of Connecticut

2013

ACKNOWLEDGMENTS

I would first like to thank my advisor, Debarchana Ghosh, who has spent countless hours providing feedback and guidance over the past year. I would also like to thank my associate advisors, Jeffrey Osleeb and Chuanrong Zhang, for their insightful comments and suggestions. Additionally, I owe my family and friends my gratitude for their support and encouragement.

Table of Contents

I. Introduction 1
II. Literature Review 5
Background of West Nile Virus 5
Disease transmission 6
Geographic Distribution 6
The Human Ecology model 7
Environmental factors 8
Population factors 13
Behavioral factors 14
Adaptation of the Human Ecology of Disease Triangle for WNV 15
The spatial analysis of West Nile Virus 16
III. Data and Methods 18
Study area 18
Dependent Variable 22
Independent Variables 27
Methods 36
Analyzing the incidence of West Nile Virus 37
Model Selection 43
Model Interpretation 44
IV. Results 46
Hot Spot Analysis 46
Correlations and Model Testing 49
Final Model 50
Dallas 51
Houston 52
V. Discussion and Conclusion 54
Spatial Distribution of West Nile Virus and Health Disparity Indicators 54
Significant Factors in Determining WNV Incidence 57
Dallas 58
Houston 58
Temporal Effects 59
Challenges to Prevention and Control 62
Conclusion 63
References 65

Table of Figures

Figure 1. West Nile Virus Cases by State, 2012; source: CDC.. 2
Figure 2. WNV incidence rate in Texas, by ZIP code, 2012; source: TX DSHS......................... 3
Figure 3. WNV transmission cycle, from CDC(Center for Disease Control and Prevention 2013) 6
Figure 4. Human Ecology of Disease triangle, modified from Meade (Meade and Emch 2010).. 8
Figure 5. Human Ecology of Disease model, adapted for WNV 16
Figure 6. Map of Texas, with study areas highlighted...19
Figure 7. Dallas study area, including 161 ZIP codes included in study 20
Figure 8. Houston study area, including 140 ZIP codes included in study 21
Figure 9. Annual cases of WNV in Texas, 2002-2012. Source: CDC 23
Figure 10. WNV cases in Texas, 2012, by age range...24
Figure 11. WNV incidence rate in Texas, by ZIP code 25
Figure 12. Incidence rates for Dallas study area, by ZIP code...26
Figure 13. Incidence rates in Houston study area, by ZIP code...27
Figure 14. Social Environment variables 28
Figure 15. Built Environment variables...29
Figure 16. Natural Environment variables 30
Figure 17. Policy variable...30
Figure 18. Mean value for each weather variable, by month 2012..33
Figure 19. Number of stations each weather variable was collected from 34
Figure 20. Average percent of land covered by different land cover types................................35
Figure 21. Mean elevation values..35
Figure 22. Mean values for demographic and socioeconomic variables...................................36
Figure 23. Hot Spot Analysis for Texas, 2012 38
Figure 24. Hot Spot Analysis for Dallas study area 39
Figure 25. Hot Spot Analysis for Houston study area 40
Figure 26. Dallas incidence frequency 41
Figure 27. Dallas log of incidence frequency...41
Figure 28. Houston incidence frequency 42
Figure 29. Houston log of incidence frequency..42
Figure 30. Final set of variables for each study area 44
Figure 31. Hot Spot Analysis for Dallas study area 47
Figure 32. Hot Spot Analysis for Houston study area 48
Figure 33. Hot Spot Analysis for Dallas study area - population over 70 years old....................56
Figure 34. Hot Spot Analysis for Dallas study area - renter-occupied homes 57
Figure 35. Dallas precipitation, 2012 and 30-year average 60
Figure 36. Dallas WNV Cases and Precipitation, 2012..60
Figure 37. Houston precipitation, 2012 and 30-year average 61
Figure 38. Houston WNV cases and Precipitation, 2012 61

I. Introduction

In 2012, the Center for Disease Control and Prevention (CDC) reported the second-worst outbreak of West Nile Virus (WNV) in the United States since it was first detected in the country in 1999. All 48 contiguous states, plus the District of Columbia and Puerto Rico, had confirmed cases of the disease, including 286 deaths out of 5,674 cases in humans (Center for Disease Control and Prevention 2013). This potentially serious vector-borne disease has become endemic in the United States, and does not have a specific treatment or cure. It is typically spread through infected mosquitoes, which feed on infected birds and pass the virus on to other mammals which they subsequently bite. The CDC estimates that approximately 80 percent of people who are infected with WNV show no symptoms at all. Milder symptoms include fever, headache, and body aches, while more serious symptoms can include high fever, headache, neck stiffness, stupor, and muscle weakness. These symptoms can last for a few days, or in severe cases, may be permanent. The more severe cases can develop neuroinvasive disease, which manifests as meningitis, encephalitis, or acute flaccid paralysis, and therefore present a risk to the long term health of the patient (Center for Disease Control and Prevention 2013). As there is no treatment or cure for this disease, health officials have focused on prevention methods as a mechanism of control, which can include personal behavior modifications or vector control. Because of the potential seriousness of the disease, understanding the factors that contribute to more severe outbreaks in certain areas is crucial to disease control.

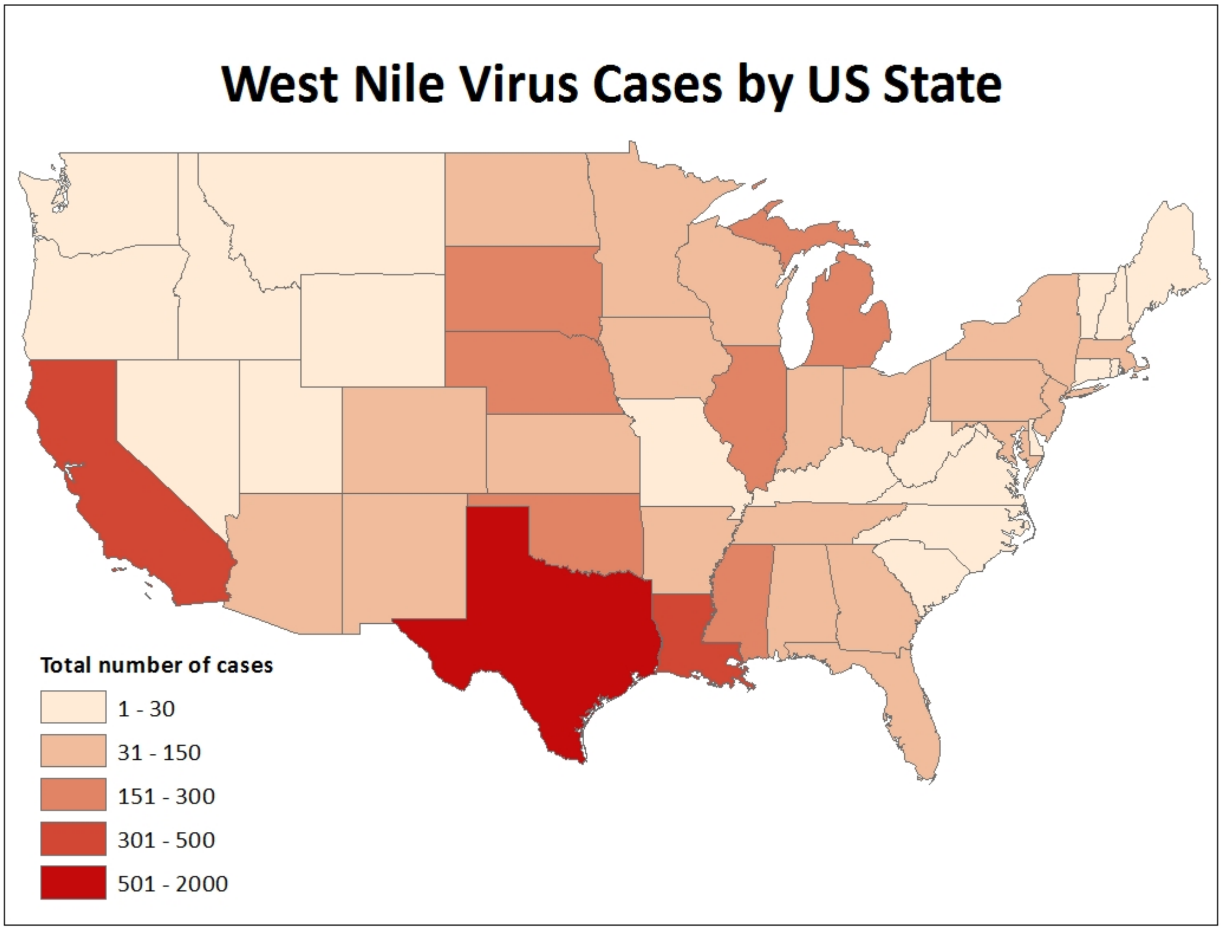

Figure 1. West Nile Virus Cases by State, 2012; source: CDC

The 2012 outbreak of West Nile Virus was particularly severe in the state of Texas. Of the 5,674 cases reported to the CDC, 1,868 – over one-fifth – were recorded in Texas. This was a significant increase over past annual activity in the state. Nationally, of the 286 deaths associated with WNV, 89 – or nearly one-third – occurred in Texas, showing a distinct severity to the cases there. Just as the disease was not uniformly dispersed across the country, it was also unevenly distributed across the state of Texas. This presented an opportunity to further investigate the factors contributing to areas of higher incidence rates of West Nile Virus.

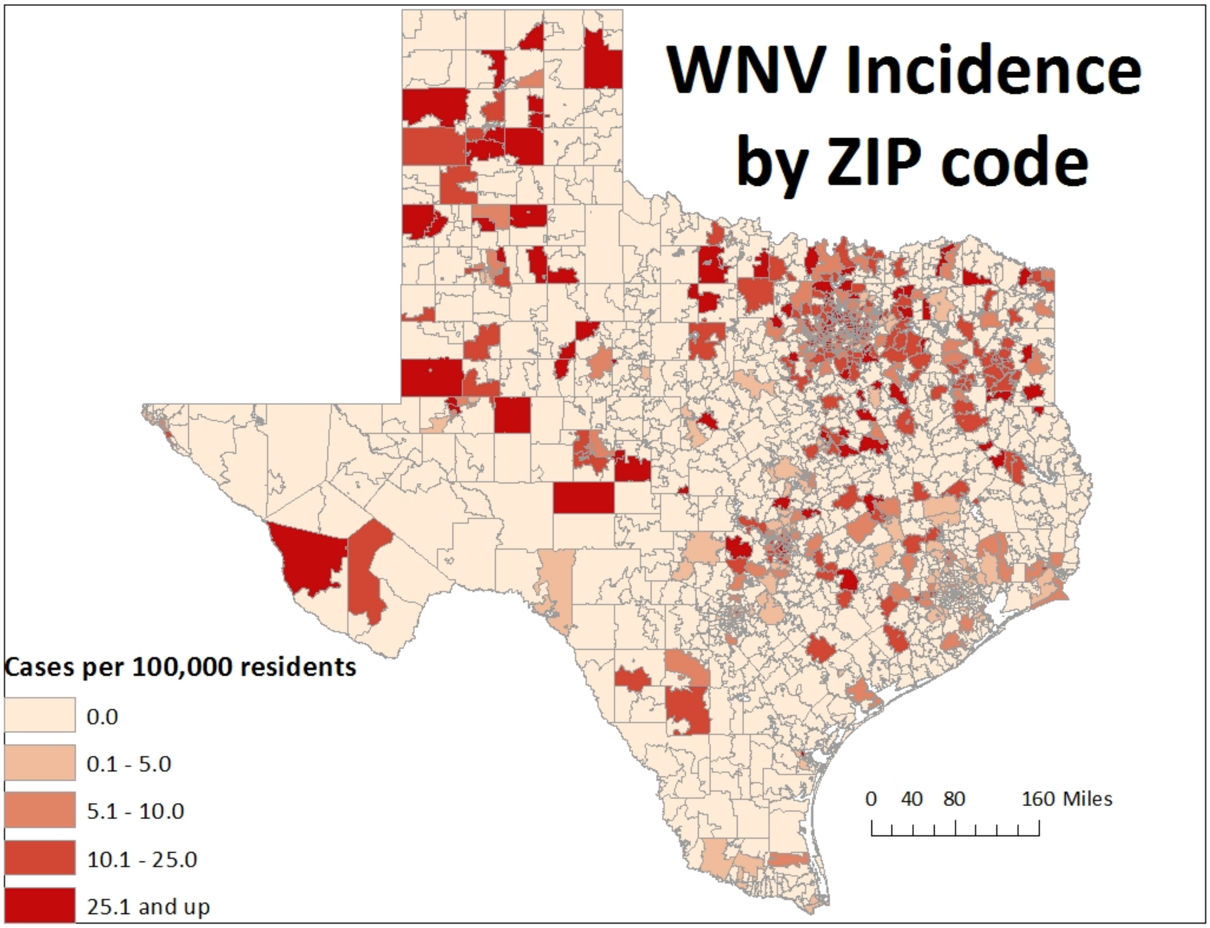

Figure 2. WNV incidence rate in Texas, by ZIP code, 2012; source: TX DSHS

Since its introduction to the United States, the pool of literature investigating the epidemiology of West Nile Virus has grown significantly. This has included studies of disease transmission, as well as the geographic spread of West Nile Virus. Other studies have also sought to pinpoint the factors that influence the incidence rate of WNV during previous outbreaks. The literature ranges from very specific to very broad, addressing West Nile Virus through the ecology of disease framework. As the pattern of disease shifts each year, these studies can be used to structure the analysis of the new information. In turn, this can allow health officials to remain aware of the risks of West Nile Virus and provide better information, prevention methods, and perhaps treatment of the disease in the future.

This thesis aims to provide an overall analysis of the West Nile Virus outbreak in Texas in 2012 through statistical and spatial analysis. The research will address the following specific questions:

1) In 2012, what was the spatial distribution of West Nile Virus is Texas?
2) Were areas of high or low incidence correlated with known health disparity indicators?
3) What natural, built, and social environmental or policy factors affect the incidence rate of WNV in Dallas and Houston?

In order to address and discuss these research questions, the thesis is structured through the following chapters. After the introduction in Chapter 1, Chapter 2 provides a review of the literature surrounding West Nile Virus, specifically in the United States. This review will place the study of the disease in the broader framework of the theory of human ecology of disease. Chapter 3 will then investigate the data and methods used in this study. A description of the variables considered and their sources will supply a frame of reference for the statistical model that was developed. Chapter 4 will review the results of the methods, and the final chapter will discuss the findings for the study areas included and consider broader concerns with the study of the disease along with limitations and future work.

II. Literature Review

The confirmation of the first cases of West Nile Virus (WNV) in the United States in 1999 has led to a significant expansion of the literature investigating the disease, its causes, and its impacts. In order to place this study within the appropriate contextual framework, the following review of literature will first explore previous studies that have addressed the background of WNV, second, the human ecology of disease model will be used to examine contributing factors, and third, a summary of spatial methods used to study WNV. Based upon this review, the chapter will conclude by indicating the disease ecology model, the risk facts, and the methodology used in this thesis.

Background of West Nile Virus

The first cases of West Nile Virus were identified in Uganda and Egypt in 1937 and 1950, respectively (Smithburn et al. 1940, 471-492)(Melnick et al. 1951, 661-665). Early epidemiological tests indicated that while many patients were asymptomatic or presented with only mild febrile symptoms, the disease did attack the central nervous system (MURGUE et al. 2001, 117-126). Further outbreaks in the Mediterranean, specifically in Israel, led to the description of main symptoms as fever, rash, and headache, with less frequent presentations of back, limb, and abdominal pain, anorexia, and vomiting (Bernkopf, Levine, and Nerson 1953, 207-218). The gestation period is shown to be around 3-14 days, and diagnosis involves testing of serum specimens (Petersen and Marfin 2002, 173-179). Most cases of seropositive individuals are asymptomatic, making the extent of the disease incidence difficult to ascertain (Mostashari et al. 2001, 261-264). An early study of WNV in the U.S. show only around 20% of individuals who contract WNV develop West Nile fever, and even fewer developing meningitis, encephalitis, and neuroinvasive disease (Mostashari et al. 2001, 261-264). There were very few cases of meningitis or encephalitis and no fatalities in the first epidemics of WNV, and the majority of cases were in young children (Bernkopf, Levine, and Nerson 1953, 207-218). It was not until the late 1950s that more severe neurological symptoms were recorded, with a higher risk associated with advanced age (HAYES 2001, 25-37) (MURGUE et al. 2001, 117-126). Subsequent epidemics have shown higher rates of infection in the central nervous system and appear to occur more frequently (Sejvar 2003, 6).

Disease transmission

The disease was confirmed as an arbovirus in the *Flaviviridae* family, being transmitted through mosquitoes to avian, equine and human hosts (Philip and Smadel 1943, 49-50) (TAYLOR 1956, 579; 579-620.; 620.), similar to St. Louis encephalitis and Japanese encephalitis (Melnick et al. 1951, 661-665). In the United States, mosquitoes in the *Culex* genus, specifically *Cx. pipiens*, are the main vectors of transmission, though there are some regional differences and shifts as the disease spread westward (Campbell et al. 2005, 1167+). Birds have been determined to be the most critical amplifying host of WNV, specifically American crows (Komar et al. 0513), while humans and horses are thought to be incidental hosts (Gubler 2007, 1039-1046).

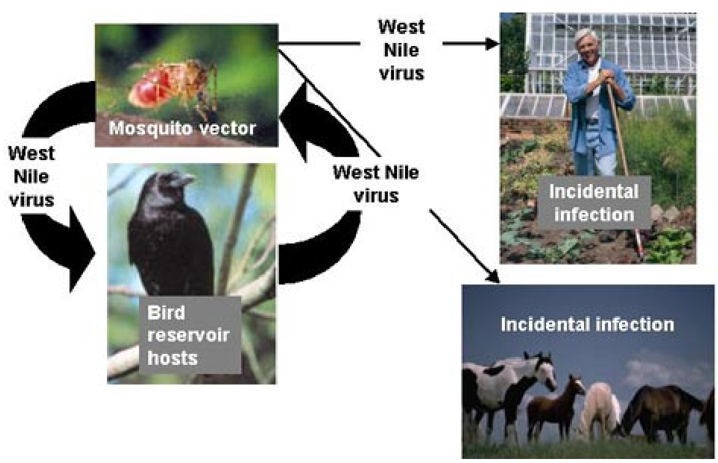

Figure 3. WNV transmission cycle, from CDC (Center for Disease Control and Prevention 2013)

Geographic Distribution

While there were several epidemics in the three decades following the initial isolation of WNV, most were in rural areas and had only mild health effects (HAYES 2001, 25-37). After the initial cases were identified in Uganda, Egypt, and Israel, there have been outbreaks in France in the 1960s, South Africa in the 1970s and 1980s, and sporadic cases elsewhere in Europe and India (MURGUE et al. 2001, 117-126). In the late 1990s epidemics in Romania, Russia, and the northeastern United States were the first to severely affect larger numbers of humans in urban environments (HAYES 2001, 25-37). In 1999, health officials in New York City identified West Nile Virus for the first time in the Western Hemisphere. A group of eight patients living in northern Queens were diagnosed with a suspected arbovirus, which was later confirmed as WNV (Nash et al. 2001, 1807-1814). By the end of the year, there had been 59 hospitalizations and seven deaths attributed to the disease (Campbell et al. 2002, 519-529). Within two years of WNV first appearing in North America, it had spread through much of the eastern half of the United States and into Canada (Campbell et al. 2002, 519-529), then into the Caribbean, Central America, and, as of 2006, deep into South America (Gubler 2007, 1039-1046; Artsob et al. 2009, 357-369). Over the past several years, there appears to be a significant cluster of cases in the US centered in the Northern Great Plains (Lindsey et al. 2008, 35).

The Human Ecology model

The human ecology of disease model is an effective tool in conceptualizing the risk factors of a particular health outcome and their spatial variations. By visualizing the relationships between environmental, population, and behavioral risk factors as three vertices with the health outcome at the center, it is easier to break down the contributing risk factors of a health outcome into pieces to be addressed separately, but within a more appropriate interdependent framework.

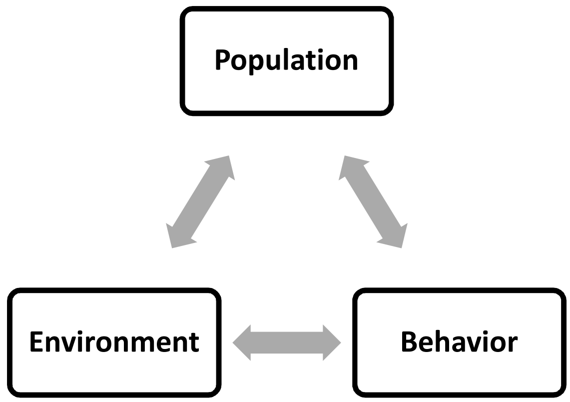

Figure 4. Human Ecology of Disease triangle, modified from Meade (Meade and Emch 2010)

This is particularly helpful when considering the contributing factors of West Nile Virus. Environmental factors can include both natural and built habitats, as well as social environments, and are specifically important when considering a vector-borne disease. Population characteristics make up the second vertex of the triangle, and include age, gender, and genetics of a person, all of which can contribute to a specific health outcome. Behavioral factors are the last vertex of the human ecology of disease triangle and, at the broadest level, include beliefs, technology, and other aspects of culture. These are generally modifiable and are therefore critical in the determination of a health outcome. Through this lens, previous studies of WNV can be placed in the broader context of health geography.

Environmental factors

The environmental factors (also called habitat factors in the human ecology of disease paradigm) that affect the presence of West Nile Virus are particularly nuanced based on the impact of the natural and built environment on the host or reservoir, the vector, and incidental infections. The studies addressing these environmental factors are perhaps the most deeply investigated of the three vertices, and can be broken down further into those that address the transmission of the disease and the effectiveness of the vector and host in the natural environment, and those that address built-environmental factors from a public health perspective.

Previous studies have identified a wide range of physical and biotic environmental factors that can influence the rate of incidence of West Nile Virus such as climate, specifically temperature and precipitation, elevation and slope, and land use and land cover.

Climate factors are typically related to the transmission of the disease through mosquitoes and birds, and their effectiveness as vector and host. Both temperature and precipitation have been shown to be important driving factors of WNV incidence (Brown et al. 2008, 1539+; Gibbs et al. 2006, 73-82; Epstein 2001, 367-371; KUNKEL et al. 2006, 168-173; Winters et al. 2008, 581-590; DeGroote et al. 2008, 1-16; Landesman et al. 2007, 337-343; Reisen 0218; Reisen et al. 2006, 344-355; Reisen, Fang, and Martinez 2006, 309-317; Tachiiri et al. 2006, 21-21; Wang G FAU - Minnis, Richard, B. et al. 0510; Ozdenerol, Bialkowska-Jelinska, and Taff 2008, 1-12; Nasci et al. 2001, 742; Shaman, Day, and Stieglitz 2005, 134-141; Crowder et al. 2013, e55006; Liu et al. 2009, 67-76). Temperature patterns that have been identified as contributing to higher rates of West Nile virus include warm winters (Epstein 2001, 367-371; Reisen et al. 2006, 344-355; Reisen, Fang, and Martinez 2006, 309-317; Nasci et al. 2001, 742) which can lead to continued transmission of the disease through birds and mosquitoes and chronic infection instead of mosquitoes dying off during the colder months; warmer temperatures in general (KUNKEL et al. 2006, 168-173; Reisen, Fang, and Martinez 2006, 309-317), which can affect when certain species of mosquitoes are present, as well as the inverse minimum temperature for the year and previous year (DeGroote et al. 2008, 1-16); and temperature during incubation (Dohm, Guinn, and Turell 2002, 221-225), which affects how long gestation lasts and what percentage of the exposed mosquitoes develop infection.

Precipitation also has several identifiable patterns correlating with higher WNV incidence rates. These include spring droughts (Epstein 2001, 367-371), which, combined with a warm winter can amplify the disease; the combination of drought 2-6 months prior to infection and land surface wetting one-half to one-and-one-half months prior to infection (Shaman, Day, and Stieglitz 2005, 134-141), which was associated with transmission to chickens and humans; higher than average precipitation levels (Landesman et al. 2007, 337-343), which can cause mosquito outbreaks if mosquitoes are limited by larval habitat; and annual precipitation from the year and the preceding year, especially drought events (DeGroote et al. 2008, 1-16; Landesman et al. 2007, 337-343; Wang G FAU - Minnis, Richard,B. et al. 0510), which can lead to changes in the food web. It is important to note that these patterns may be true for one geographic region, but not necessarily so for another. For example, human outbreaks of West Nile Virus are associated with above-average rainfall in the eastern United States, but were associated with below-average rainfall in the western United States in the previous year (Landesman et al. 2007, 337-343). Also, the timing of these climate events can affect the incidence rate of WNV. One study has shown that temperature in March, precipitation in July, heating degree days in August, and snowfall in September are all contributing factors to the incidence rate (Winters et al. 2008, 581-590).

The climate factors, represented by temperature and precipitation, affect the biotic environment in an area, thereby affecting the rate of WNV through the identified hosts and vectors. As noted above, many studies have addressed the impact of climate on mosquito and bird populations. In turn, there is also a pool of literature which investigates the impact of mosquito and bird populations on human WNV incidence (Mostashari et al. 2001, 261-264; Nash et al. 2001, 1807-1814; Gu et al. 2006, 91-98; Smith, Dushoff, and Mckenzie 2004, 1957-1964; Ruiz et al. 2004, 8-11; Ruiz et al. 2007, 10-11; Theophilides et al. 2006, 103-115; Theophilides et al. 2003, 843-854; Godsey MS Jr FAU - Blackmore, Mark,S. et al. 0607; Eidson et al. 2001, 615; Eidson et al. 2001, 662; Kulasekera et al. 2001, 722). These studies generally show the presence of WNV-positive mosquitoes, identified through mosquito traps, and dead birds are associated with higher rates of human WNV infection. The behavior of mosquitoes and birds also affects the biotic environment, in that the feeding patterns of mosquitoes and the migration of certain birds have been shown to affect the incidence rate of WNV in humans (Smith, Dushoff, and Mckenzie 2004, 1957-1964; Kilpatrick et al. 2006, e82). In that sense, when the biotic environment is altered by humans, that can be a significant factor. The most obvious example of this is mosquito abatement efforts, which can significantly impact the presence of mosquitoes and therefore the infection rate (Gu et al. 2006, 91-98; Ruiz et al. 2004, 8-11; Ruiz et al. 2007, 10-11). A well-known method of mosquito control is insecticide spraying. The debate surrounding these methods questions the overall environmental impact of the chemicals used (Elnaiem et al. 2008, 751-757). Another study investigates the impact of increased coalbed methane gas extractions in an area, which can create water pools that might lead to increased mosquito larval habitats (Zou, Miller, and Schmidtmann 2006, 1034-1041). The interaction between humans and the biotic environment are somewhat difficult to separate completely from the other vertices of the human ecology of disease triangle, but have been shown to be significant in identifying risk factors of higher WNV infection rates.

Other physical or natural environmental factors that have been shown to affect the incidence of WNV include proximity to water, amount of vegetation cover, slope of land, elevation (Brown et al. 2008, 1539+; Gibbs et al. 2006, 73-82; Winters et al. 2008, 581-590; DeGroote et al. 2008, 1-16; Reisen 0218; Reisen et al. 2006, 344-355; Reisen, Fang, and Martinez 2006, 309-317; Ozdenerol, Bialkowska-Jelinska, and Taff 2008, 1-12; Crowder et al. 2013, e55006; Crowder et al. 2013, e55006; Liu et al. 2009, 67-76; Ruiz et al. 2004, 8-11; Ruiz et al. 2007, 10-11; Nolan et al. 2012, 1-5; Brownstein et al. 2002, 157-164; Chuang et al. 2011, 669-679; Cooke, Grala, and Wallis 2006, 36-54; Cooke, Grala, and Wallis 2006, 36-54). These studies have shown that certain landscape features are significant in determining WNV infection rates. Vegetation cover has consistently been found to be an important factor. Especially in urban areas, vegetation has been shown to have a positive association with human cases (Brownstein et al. 2002, 157-164), while in inner suburbs a more moderate vegetation cover leads to higher rates of WNV (Ruiz et al. 2004, 8-11; Ruiz et al. 2007, 10-11). The presence or proximity to water and wetlands, as well as stream density, has also been identified as a contributing factor, usually associated with mosquito breeding habitats. This was further specified by noting that proximity to slow-moving water was associated with higher rates of infection, especially in areas where vegetation lines the water, whereas living near fast-moving water decreases the odds of contracting WNV (Nolan et al. 2012, 1-5). Other specific landscape features associated with higher rates of WNV are rural land cover (Chuang et al. 2011, 669-679), orchard habitats (Crowder et al. 2013, e55006), less forest cover (Brown et al. 2008, 1539+), and rural agricultural settings (DeGroote et al. 2008, 1-16).

The built environment has also been shown to have an impact on the incidence rates of West Nile Virus through previous studies (Gibbs et al. 2006, 73-82; DeGroote et al. 2008, 1-16; Ozdenerol, Bialkowska-Jelinska, and Taff 2008, 1-12; Ruiz et al. 2004, 8-11; Ruiz et al. 2007, 10-11; Cooke, Grala, and Wallis 2006, 36-54; Su T FAU - Webb, James, P. et al. 0805; Han et al. 1999, 230; Harrigan et al. 2010, e15437; Gibney et al. 2012, 895-901). These factors range from large-scale public works to neighborhood upkeep, and can impact multiple stages of disease transmission. Significant features include underground storm drainage systems from landscaping curbs (Su T FAU - Webb, James,P. et al. 0805), which can lead to larger mosquito populations; road density (DeGroote et al. 2008, 1-16; Cooke, Grala, and Wallis 2006, 36-54), which can lead to standing pools of water or be correlated with vegetation cover; housing density (Gibbs et al. 2006, 73-82); presence of irrigation systems and animal feeding operations (DeGroote et al. 2008, 1-16); a high percentage of homes being rented or vacant (Ozdenerol, Bialkowska-Jelinska, and Taff 2008, 1-12), which could imply poor upkeep and lead to more mosquito breeding grounds, such as neglected swimming pools or water-holding containers in the yard (Harrigan et al. 2010, e15437; Gibney et al. 2012, 895-901); the age of housing (Ozdenerol, Bialkowska-Jelinska, and Taff 2008, 1-12; Ruiz et al. 2004, 8-11; Ruiz et al. 2007, 10-11); and poor maintenance of homes such as basement flooding in apartments (Han et al. 1999, 230). Many of these factors can be included in a broader theme of urbanization, and indeed the process of urbanization has also been considered as a risk factor (Brown et al. 2008, 1539+). Another study grouped these characteristics into classes to evaluate the different combinations that may lead to higher risk of infection (Ruiz et al. 2004, 8-11; Ruiz et al. 2007, 10-11), and found that the class identified as the "Inner Suburb", characterized as having predominantly 1940-1960 era housing, moderate vegetation cover, and moderate population density, had the higher case rate.

The last portion of environmental factors is the social environment. This includes characteristics of the groups, relations, and societies to which people belong (Meade and Emch 2010), and are closely tied to population factors, but apply to the larger population instead of the individual. These factors have been shown to contribute to the incidence of West Nile Virus through population density, race, age, and socio-economic traits of the community (Ozdenerol, Bialkowska-Jelinska, and Taff 2008, 1-12; Ruiz et al. 2004, 8-11; Ruiz et al. 2007, 10-11; Harrigan et al. 2010, e15437). These studies have shown that higher percent of the population that is Black or African American and low income correlate positively with WNV infection rates, though this could be related indirectly due to correlated environmental factors or lifestyle (Ruiz et al. 2004, 8-11; Ruiz et al. 2007, 10-11).

Population factors

Population factors have also been shown to contribute to the incidence of West Nile Virus. These factors mainly involve the human host, such as genetic predisposition, underlying conditions, and age. In the West Nile Virus literature pool, the most significant risk factor is age. Though children have been shown to be 4.5 times more likely to be infected, they are 110 times less likely to contract the more severe West Nile Neuroinvasive Disease (WNND) (Mandalakas et al. 2005, 1774-1777), other studies have shown a more even distribution of seroprevalence (Mostashari et al. 2001, 261-264), but in general, advanced age correlates with a higher prevalence of severe disease (Nolan et al. 2012, 1-5; O'Leary et al. 2004, 61-70; Bode et al. 2005, 1174+; Bode et al. 2006, 1234-1240; HUHN et al. 2005, 768-776; Nolan, Schuermann, and Murray 2013, 137). Additionally, there appears to be slightly higher rates of infection in males, rather than females (O'Leary et al. 2004, 61-70; Nolan, Schuermann, and Murray 2013, 137). There are also several underlying conditions that affect the risk of a person contracting WNV. Studies have shown that diabetes and hypertension are associated with increased risk (Bode et al. 2005, 1174+; Bode et al. 2006, 1234-1240; Murray et al. 2006, 1325; Lindsey 2012, 179; 179-184; 184), while diabetes, hypertension, chronic renal disease, history of cancer, and history of alcohol abuse were associated with more severe illness (Murray et al. 2006, 1325; Lindsey 2012, 179; 179-184; 184).

Population factors are deeply involved in all three vertices of the human ecology of disease triangle, through the relationship between health and environmental factors, as well as how personal health impacts one’s behavior.

Behavioral factors

The third group of studies that comprise the human ecology of disease triangle is the group addressing behavioral factors. This vertex includes belief systems, technology, and perceptions. Behavior factors interact with the other vertices through the human creation of habitat, activities resulting in the exposure to hazards, human movement, and control of population characteristics through customs (Meade and Emch 2010). Because there is no cure for WNV, health officials emphasize the importance of prevention, and many of these suggestions are in the form of behavioral changes.

At the personal level, there are several methods of protective behavior in which an individual can engage. By instituting at least two methods of protective behavior, one can significantly lower the risk for infection (Loeb M FAU - Elliott, Susan,J. et al. 0518). These behaviors include avoiding mosquito problem areas, avoiding going outdoors (especially during evening hours), wearing long sleeves and pants, using mosquito repellent, and not working or attending school outside of the home (Mostashari et al. 2001, 261-264; Han et al. 1999, 230; Gibney et al. 2012, 895-901; Loeb M FAU - Elliott, Susan,J. et al. 0518). Behavior factors, tied closely to social environment factors, have also been explored through awareness and compliance studies (Averett et al. 2005, 1751+).

Adaptation of the Human Ecology of Disease Triangle for WNV

Based upon the literature review of contributing risk factors, the nature and geographic distribution of WNV incidence, and the availability of data in the study area, I have addressed the human ecology of disease triangle through the analysis of the following environmental factors: temperature, precipitation, land cover, elevation, presence of state parks, population density, proportion of population that is male and female, median age of population, proportion of population by race, income, housing density, and occupation of housing units.

One element not typically considered in the human ecology disease triangle is the impact of policies. Because of the nature of vector-borne diseases, a vertex labeled “Policies” would create a more complete representation of the factors affecting the rate of incidence. A policy vertex would also be considered interdependent with the other vertices, and might alter the interpretation to look as follows.

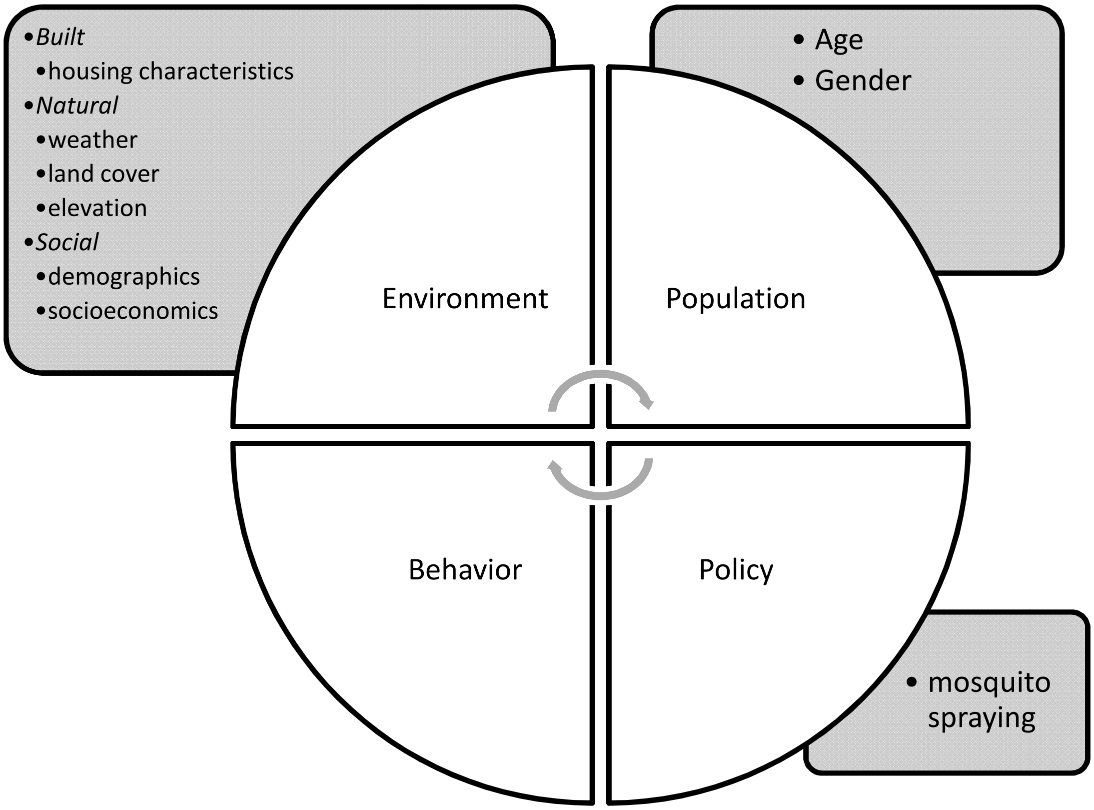

Figure 5. Human Ecology of Disease model, adapted for WNV

This data will be investigated through ordinary least squares regression, spatial lag models, and spatial error models to determine the relationship and effect of these identified risk factors on the incidence of WNV in Dallas and Houston. The details of data collection and processing, and the methodology of the analysis will be described in the next chapter.

The spatial analysis of West Nile Virus

While much of the literature surrounding West Nile Virus examines the factors that contribute to its presence, there is a growing field of literature about the study of West Nile Virus through spatial statistics and geographic information systems. Because of the difficulty of prediction in the future of West Nile Virus (Petersen and Fischer 2012, 1281-1284), it is generally agreed that better systems of tracking and addressing it as a public health concern are of vital importance. There have been several studies that examine the use of geographic information systems in order to visualize data in a timely way. The CDC collects this data for the entire country and presents it online for the public to access (Anonymous2010; Centers for 2005). The use of real-time data adds the benefit of sharing information immediately among public health officials and the public in order to better prepare for a potentially heightened exposure risk (Gosselin et al. 2005, 21-12). Using temporal data also appears to be critical in effectively predicting risk of infection, and has been organized into a Nearest Neighbor Distance Time Model (Ghosh, Manson, and McMaster 2010, 149+) and the Dynamic Continuous-Area Space-Time system (Theophilides et al. 2006, 103-115; Theophilides et al. 2003, 843-854), in order to better understand the risk of exposure in areas where host, vector, and reservoir are in close proximity. However, it is also acknowledged that no single spatial statistical model produces an adequate outcome with any consistency (Griffith 2005, 18-14), so these models must be constantly modified and updated with changing factors. Lastly, there are some drawbacks to using GIS to visualize the risk of WNV infection (Kitron 2000, 324-325). While GIS can provide a clear visualization of a large amount of data, it can also be used to mask a lack of data across areas with boundaries that may be arbitrary in determining actual risk.

Because of these concerns, it is important to apply the methods and models put forward in previous studies to data collected in the 2012 outbreak of West Nile Virus. The changing nature of the disease and the contributing factors provide an opportunity to increase our understanding of the risk of infection by investigating the patterns more deeply, and in the hopes of better addressing West Nile Virus as a public health concern.

III. Data and Methods

This chapter will review the data and methods used in the thesis. It will begin with a review of the study area, and the dependent and independent variables, and then discuss the process used to develop an appropriate model.

Study area

The study area for this research was defined as the metropolitan areas of Dallas and Houston, in the state of Texas. It is a topographically diverse state, ranging from nearly 9,000 feet in elevation in the west, to near- or below-sea level in the east along the Gulf Coast, and covers over 265,000 square miles. There are four principal physical regions of Texas, including the Gulf Coastal Plains, the Interior Lowlands, the Great Plains, and the Basin & Range Province. These regions range from hilly and wooded to arid and dry. The southern border of Texas is defined by the Rio Grande, while the northern border connects Texas with New Mexico, Oklahoma, Arkansas, and Louisiana (Texas State Historical Association).

As of 2012, the Texas Department of State Health Services projects a total population of over 26,000,000 residents, over 20,000,000 of which identify as either White or Hispanic, and 23,000,000 of which live in metropolitan areas. Houston and Dallas metropolitan areas are by far the largest.

Dallas and Tarrant counties have a combined population of over 4,600,000, making it slightly larger than Harris County, which includes the city of Houston, which has a population of 4,500,000. The two cities lie approximately 230 miles apart, in the eastern half of the state. While Houston has a larger population in the city limits, Dallas and Tarrant counties include the cities of Dallas, Fort Worth, and Arlington. The Dallas area covers a slightly larger area of land. Both areas fall in the Gulf Coastal Plains region of Texas.

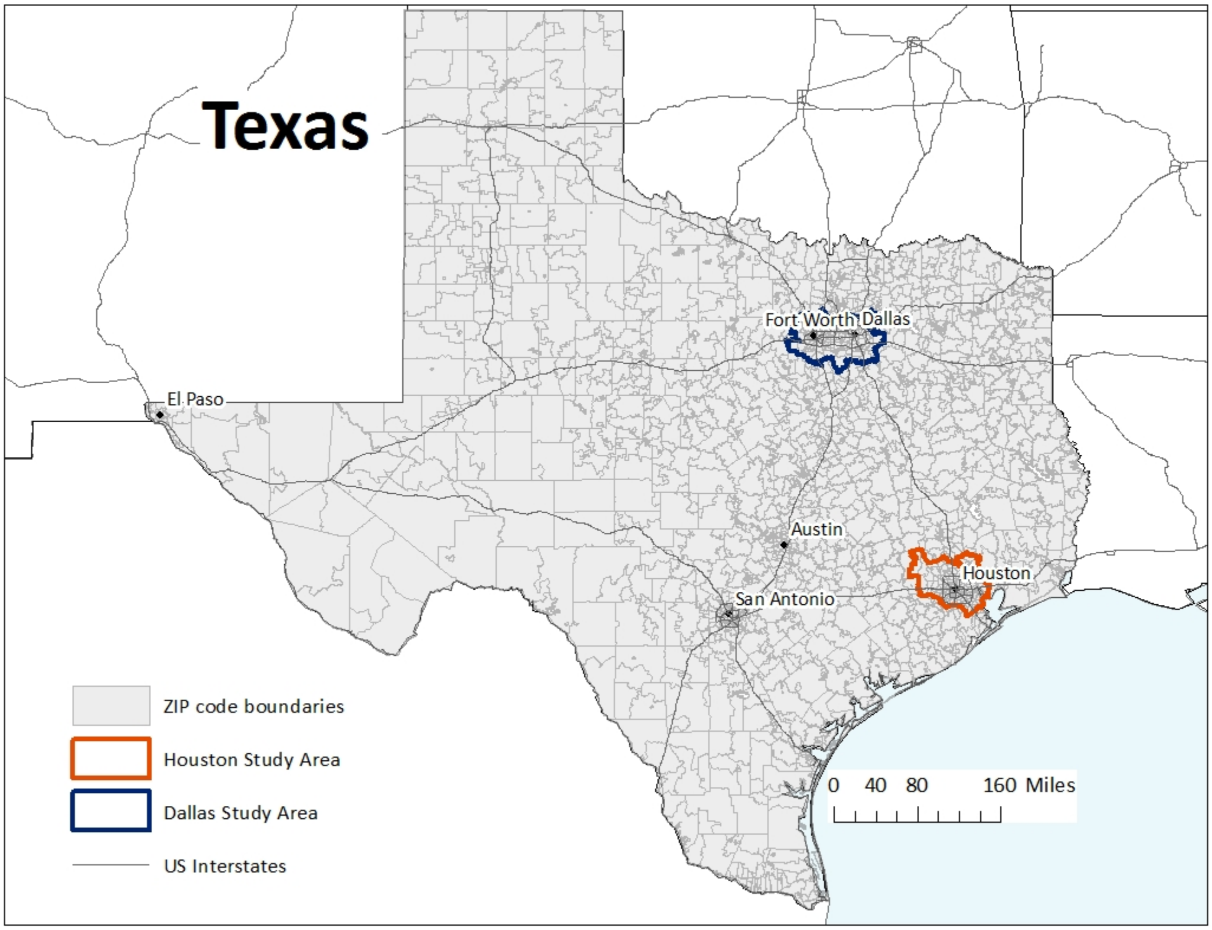

Figure 6. Map of Texas, with study areas highlighted

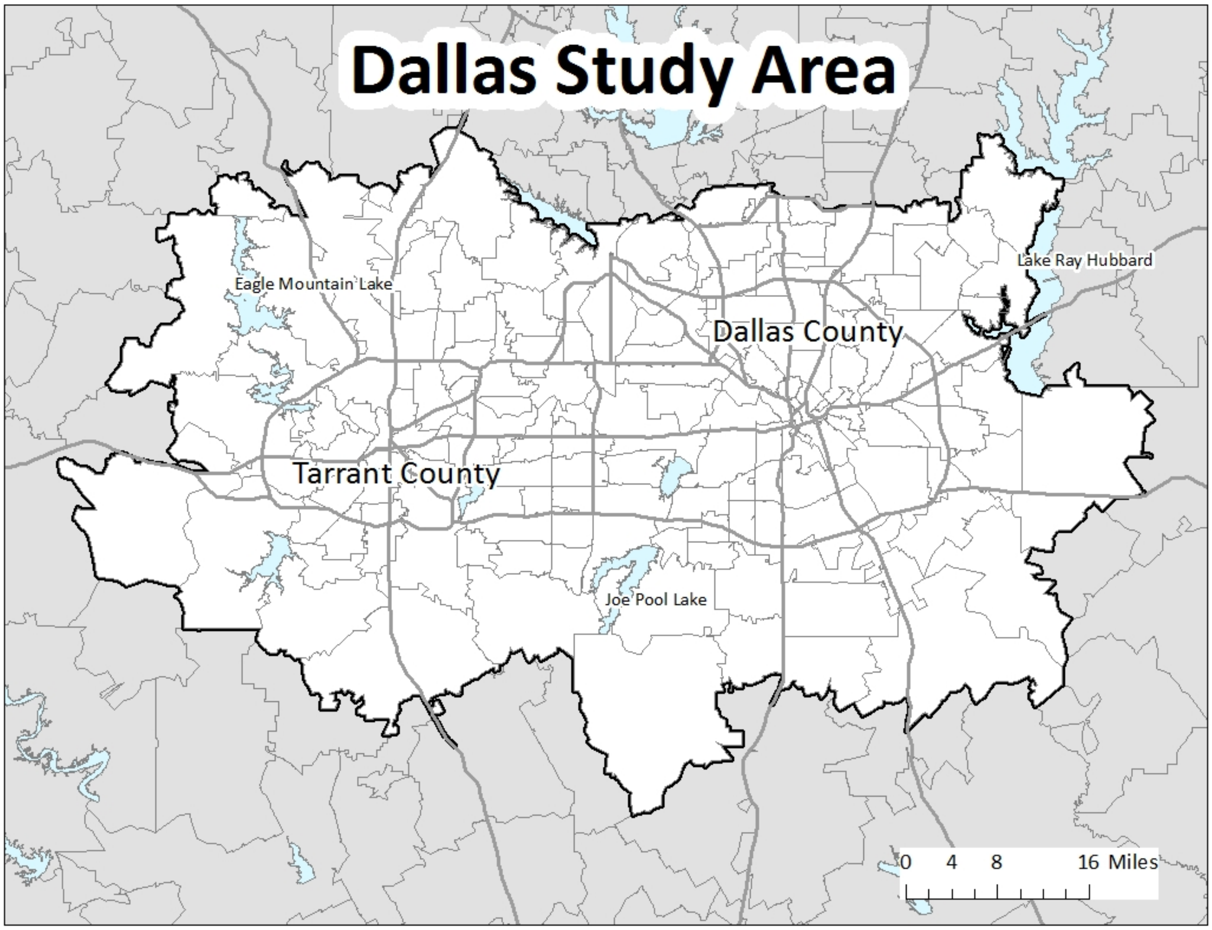

Figure 7. Dallas study area, including 161 ZIP codes included in study

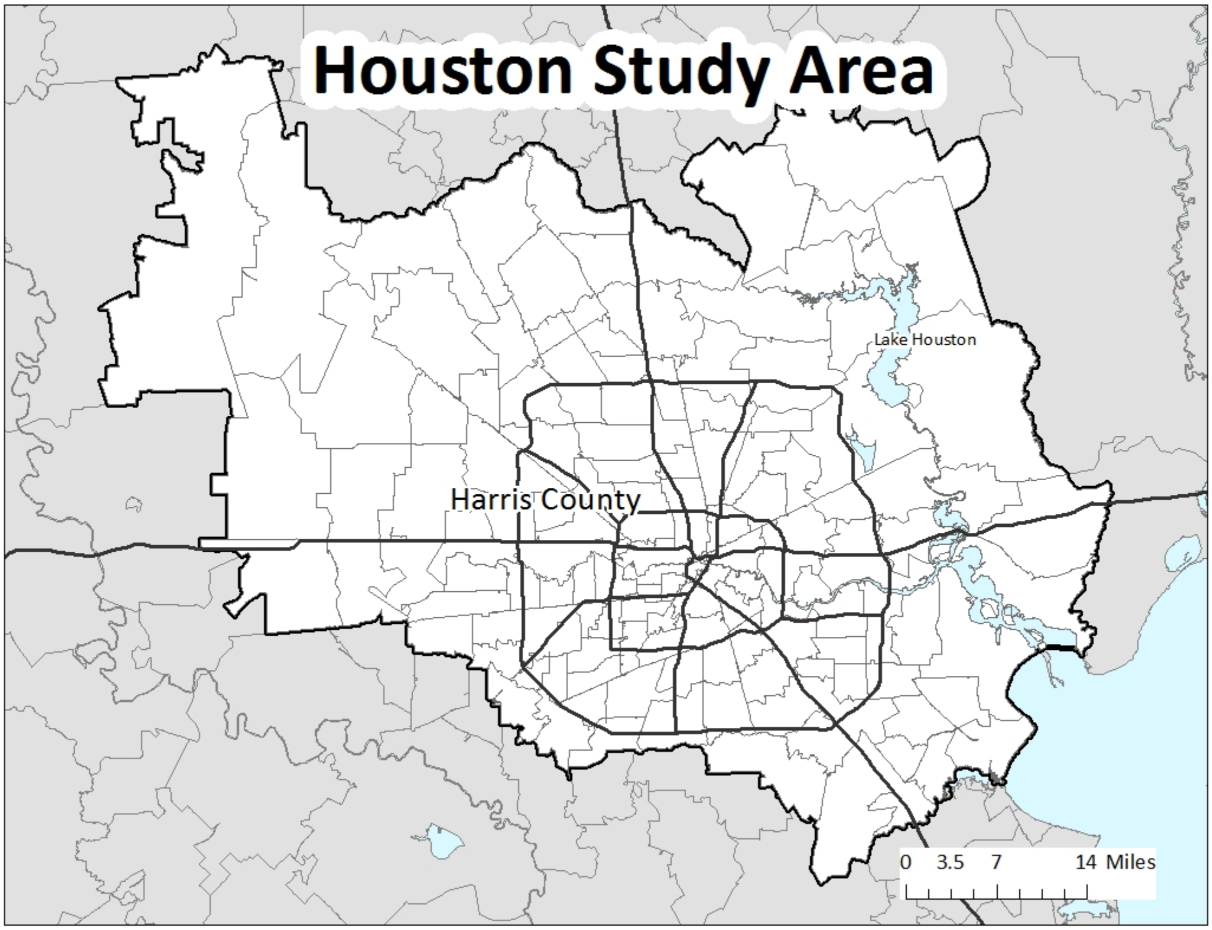

Figure 8. Houston study area, including 140 ZIP codes included in study

The unit of analysis used in this study is the United States Postal Service ZIP code. The Texas Department of State Health Services, in order to anonymize patient data, reports only the ZIP code in which a patient resides as the geographic location. The ZIP codes used were defined by ESRI Business Analyst, an extension of ESRI ArcGIS that combines demographic and business data based on the US decennial census data. There are 1,889 ZIP codes in the state of Texas. Because of the nature of ZIP codes, some are unique to a single high-volume building or mail recipient and are therefore small in spatial area, whereas ZIP codes in more rural areas can cover 5,000 square miles.

In urban areas however, there is a much smaller range of area. Within the Dallas area, which for this study includes Dallas and Tarrant counties, there are 161 ZIP codes intersecting the outer boundary, and the largest is 152 square miles. In the Houston area, which includes 140 ZIP codes intersecting the Harris county boundary, the largest is 203 square miles.

Dependent Variable

The dependent variable in this study is the incidence rate of West Nile Virus in a given ZIP code. Of over 5,000 cases of West Nile Virus reported nation-wide to the Center for Disease Control (CDC) for the year 2012, over 1,800 were located in Texas, and the majority of those were in the Dallas area (Petersen and Fischer 2012, 1281-1284). West Nile Virus is included in the National Notifiable Diseases Surveillance System, which is administered by the CDC, and supports the surveillance of infectious diseases. These numbers represent a significant leap over 2011, which had 712 cases confirmed in the United States, and only 27 in the state of Texas. Since WNV first appeared in Texas in 2002, yearly case numbers ranged from 27 in 2011 to 720 in 2003, when there was a significant nation-wide outbreak. This shows an alarming increase of West Nile Virus in the state of Texas, and warrants further investigation into the risk factors involved.

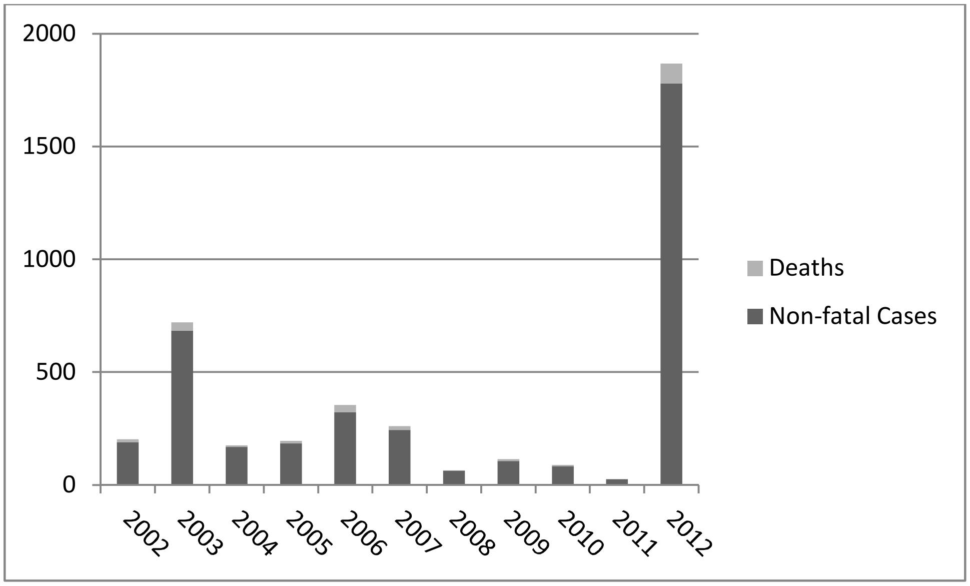

Figure 9. Annual cases of WNV in Texas, 2002-2012. Source: CDC

The original data was collected by the Texas Department of State Health Services in the form of a list of reported patients for the year 2012 as of February 4, 2013, and was estimated to be within 2-3% of the full total. This data is collected from regional and local health departments, and reported to the state and then the CDC. There were 1,846 cases included in the data, 1,819 of which had exact ZIP codes assigned to them and could therefore be used in this study. Each patient record included a unique identifier, the city, ZIP code and county in which the patient resided. Other information provided included health region, date of onset, age of patient at date of onset, birth date of patient, sex of patient, race of patient, and diagnosis. Of the confirmed cases of West Nile Virus in Texas with full information provided, 827 (45%) were female, 992 (55%) were male, and 1,515 (83%) were identified as White. Of the cases reported, 830 (46%) were diagnosed as West Nile Fever, and 989 (54%) were classified as West Nile Neuroinvasive Disease. Within each gender, the severity was split approximately evenly, as well as within patients identified as White. Of the 1,819 cases, over 1,400 were in patients above the age of 40, and there were 25 cases in patients under the age of 10. In patients below 70 years of age, most cases presented as West Nile Fever, whereas in patients above 70 years of age, most cases had progressed to West Nile Neuroinvasive Disease.

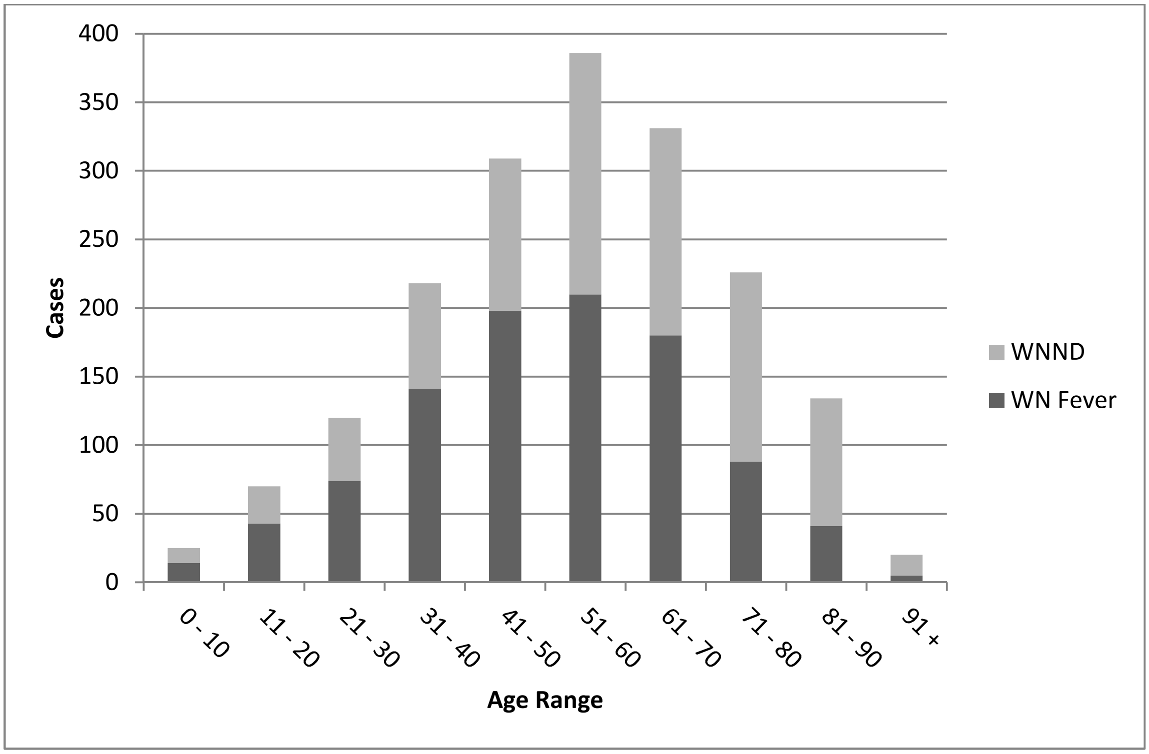

Figure 10. WNV cases in Texas, 2012, by age range

Patient numbers were aggregated by ZIP code for use in this research, and the incidence rate for each ZIP code was calculated by dividing the number of cases by the total population in the ZIP code and then multiplied by 100,000.

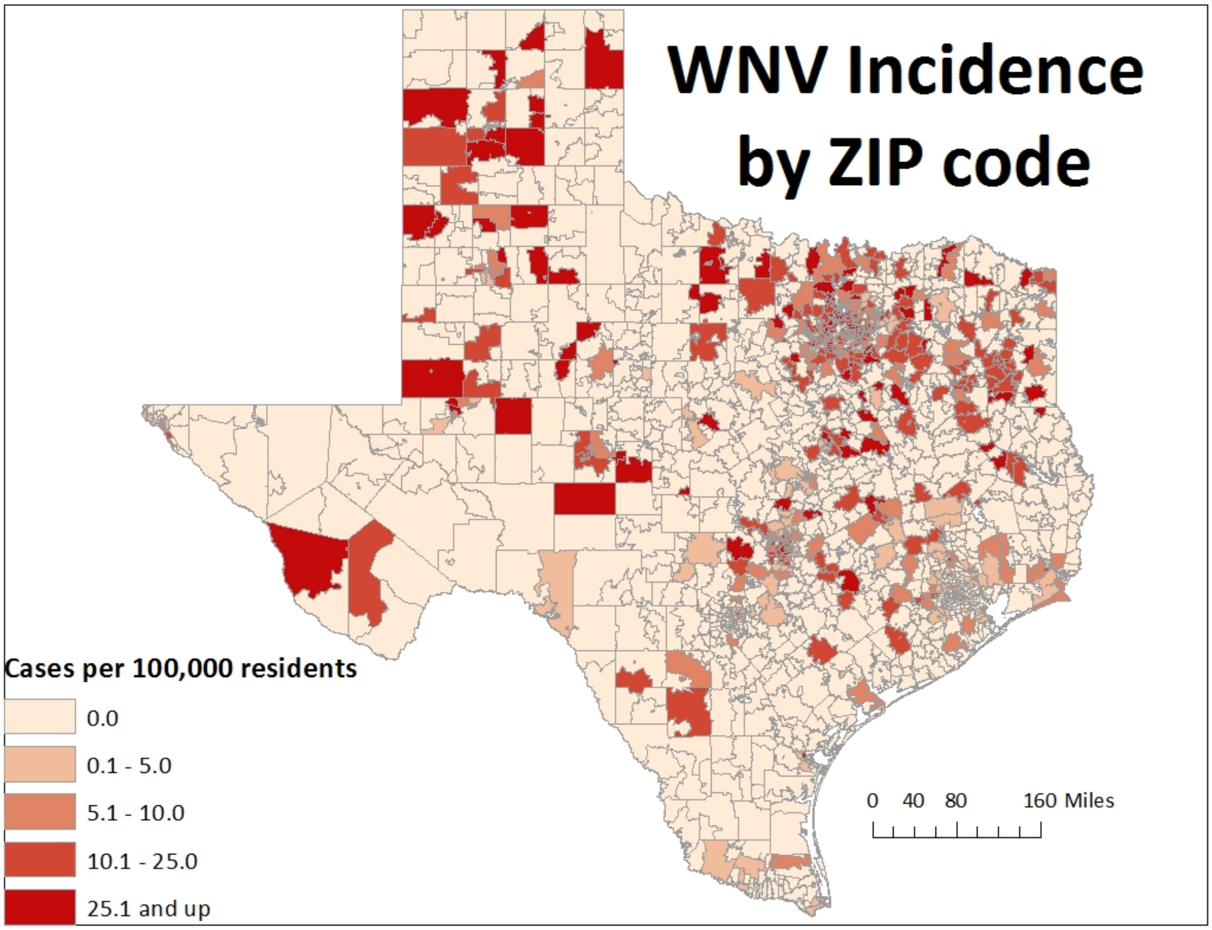

Figure 11. WNV incidence rate in Texas, by ZIP code

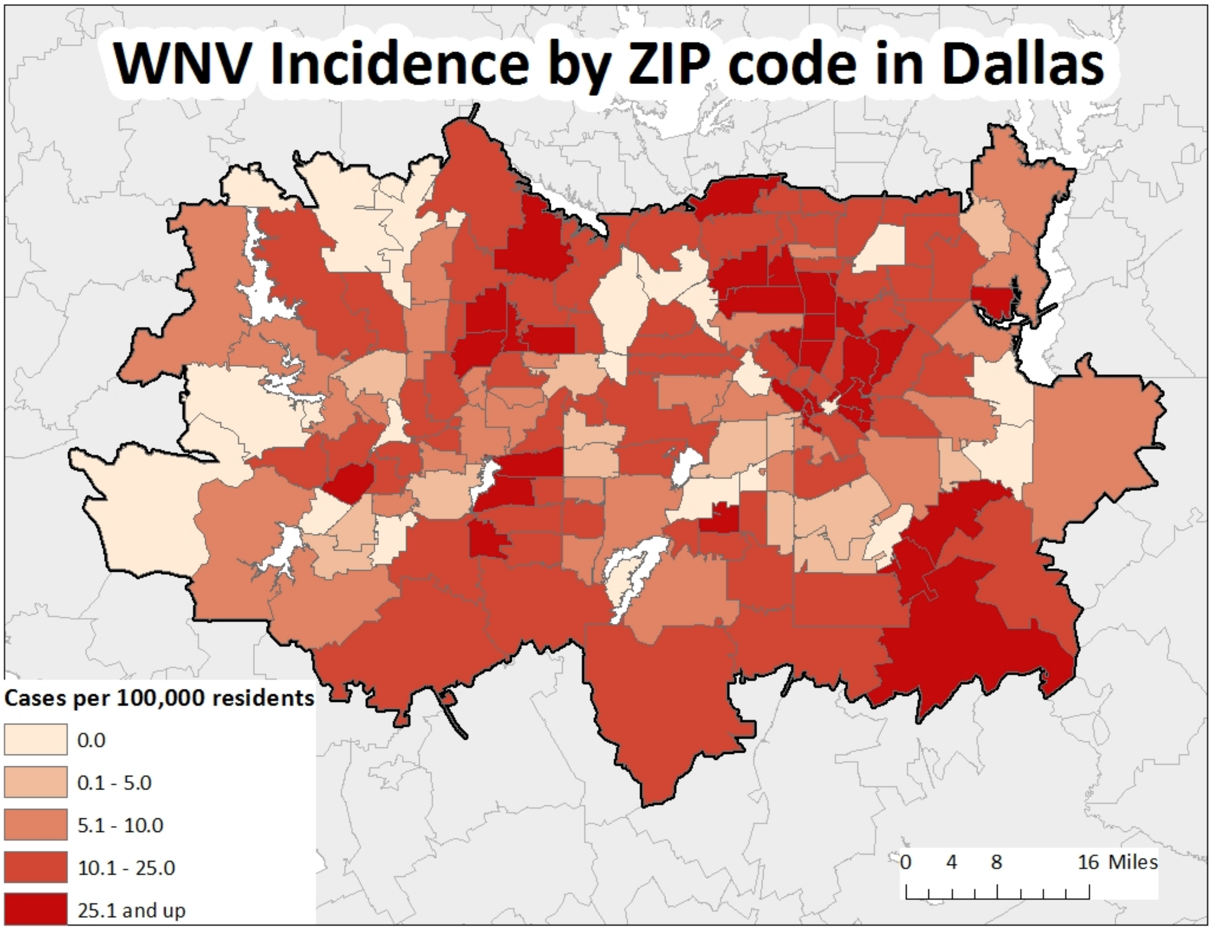

Figure 12. Incidence rates for Dallas study area, by ZIP code

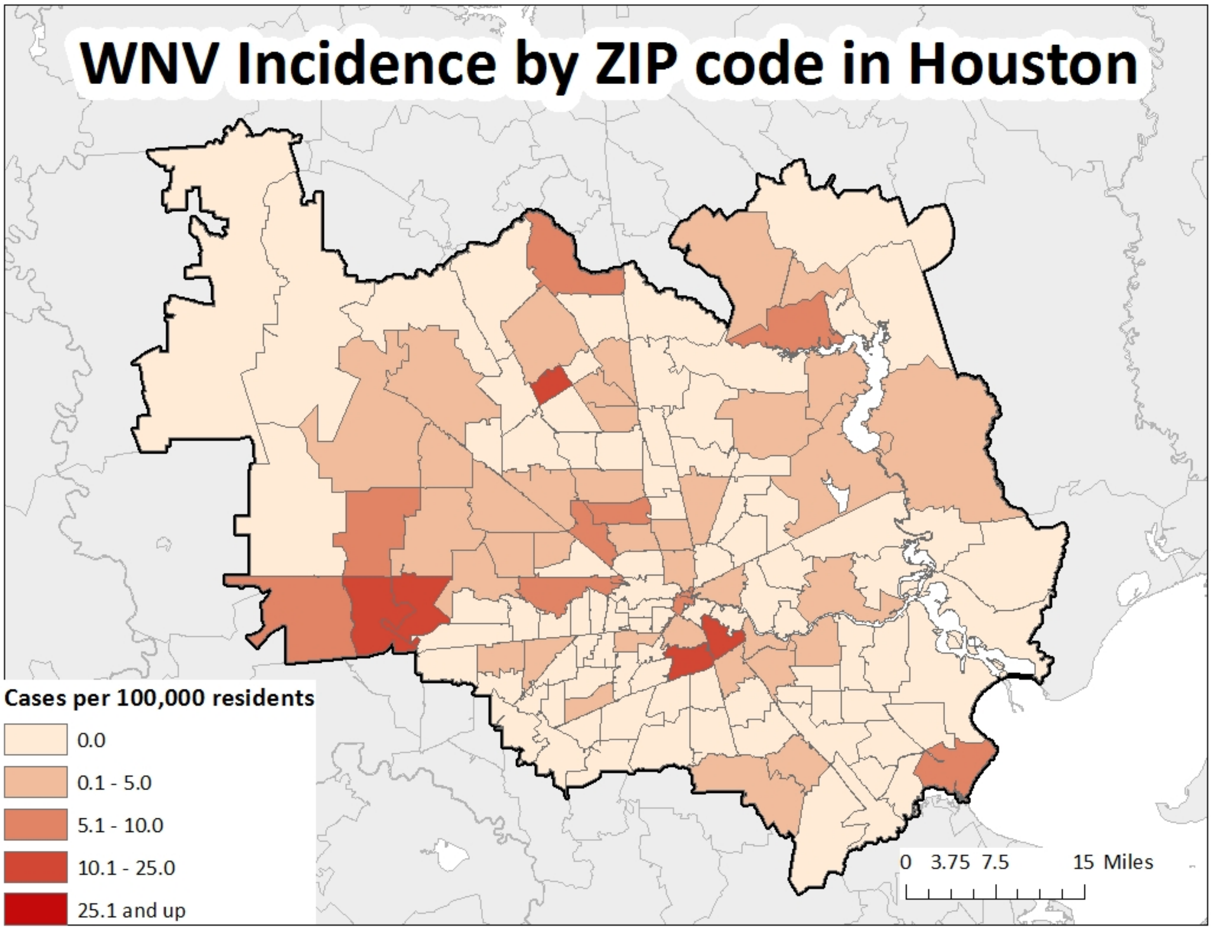

Figure 13. Incidence rates in Houston study area, by ZIP code

Independent Variables

The risk factors included in this analysis are encompassed by the environmental vertex of the human ecology of disease triangle. The variables chosen to represent the natural environment included weather data, land cover, elevation, and number of state parks. Monthly weather variables were then reported through precipitation, average temperature, maximum and minimum temperature, and heating and cooling degree days. The social environment was represented by population density, proportion of population that is male and female, median age of population, proportion of population by race, and income. The built environment variables included housing density and occupation of housing units.

Category	Risk Factor	Data Format	Source	Expected Relationship
Environment - Social	Total population	Vector	ESRI Business Analyst	positive
Environment - Social	Population Density	Vector	ESRI Business Analyst	positive
Environment - Social	Median age of Population	Vector	ESRI Business Analyst	positive
Environment - Social	Male (% of population)	Vector	ESRI Business Analyst	positive
Environment - Social	Median age of Male Population	Vector	ESRI Business Analyst	positive
Environment - Social	Female (% of population)	Vector	ESRI Business Analyst	negative
Environment - Social	Median age of Female Population	Vector	ESRI Business Analyst	positive
Environment - Social	White (% of population)	Vector	ESRI Business Analyst	positive
Environment - Social	Black (% of population)	Vector	ESRI Business Analyst	negative
Environment - Social	American Indian (% of population)	Vector	ESRI Business Analyst	negative
Environment - Social	Asian (% of population)	Vector	ESRI Business Analyst	negative
Environment - Social	Pacific Islander (% of population)	Vector	ESRI Business Analyst	negative
Environment - Social	Other race (% of population)	Vector	ESRI Business Analyst	negative
Environment - Social	Hispanic (% of population)	Vector	ESRI Business Analyst	positive
Environment - Social	Median Household Income	Vector	ESRI Business Analyst	negative
Environment - Social	Average Household Income	Vector	ESRI Business Analyst	negative
Environment - Social	Per Capita Income	Vector	ESRI Business Analyst	negative

Figure 14. Social Environment variables

Category	Risk Factor	Data Format	Source	Expected Relationship
Environment - Built	Total housing units	Vector	ESRI Business Analyst	positive
Environment - Built	Housing occupied by owner (% of total housing units)	Vector	ESRI Business Analyst	negative
Environment - Built	Housing occupied by renter (% of total housing units)	Vector	ESRI Business Analyst	positive
Environment - Built	Vacant housing (% of total housing units)	Vector	ESRI Business Analyst	positive

Figure 15. Built Environment variables

Category	Risk Factor	Data Format	Source	Expected Relationship
Environment - Natural	Water (% of total area)	Raster	National Land Cover Dataset (2006)	positive
Environment - Natural	Developed land (% of total area)	Raster	National Land Cover Dataset (2006)	positive
Environment - Natural	Barren land (% of total area)	Raster	National Land Cover Dataset (2006)	negative
Environment - Natural	Forest (% of total area)	Raster	National Land Cover Dataset (2006)	positive
Environment - Natural	Shrub (% of total area)	Raster	National Land Cover Dataset (2006)	negative
Environment - Natural	Herbaceous (% of total area)	Raster	National Land Cover Dataset (2006)	negative
Environment - Natural	Cultivated land (% of total area)	Raster	National Land Cover Dataset (2006)	positive
Environment - Natural	Wetlands (% of total area)	Raster	National Land Cover Dataset (2006)	positive
Environment - Natural	Number of state parks in ZIP code	Raster	Texas Parks and Wildlife Department	positive
Environment - Natural	Average elevation of ZIP code	Raster	Texas Natural Resources Information System	negative
Environment - Natural	Range of elevation of ZIP code	Raster	Texas Natural Resources Information System	negative
Environment - Natural	Total inches of precipitation in each	Raster	NOAA	varies by month

	month			
Environment - Natural	Average temperature in each month	Raster	NOAA	varies by month
Environment - Natural	Maximum temperature in each month	Raster	NOAA	varies by month
Environment - Natural	Minimum temperature in each month	Raster	NOAA	varies by month
Environment - Natural	Number of heating degree days in each month	Raster	NOAA	varies by month
Environment - Natural	Number of cooling degree days in each month	Raster	NOAA	varies by month

Figure 16. Natural Environment variables

Category	Risk Factor	Data Format	Source	Expected Relationship
Policy	Aerial spraying (Y/N)	Vector	DallasCityHall.com, Harris County Public Health & Environmental Services	negative

Figure 17. Policy variable

The weather data was obtained from land-based weather stations as reported to the National Climate Data Center (NCDC) at the National Oceanic and Atmospheric Administration (NOAA). The data used were part of the Quality Controlled Local Climatological Dataset (QCLCD), and included 93 land-based weather stations in Texas, though not all stations had complete data for all six weather variables considered in this research. For each weather station, a monthly value was recorded for the water equivalent of precipitation in inches, average temperature, maximum temperature, minimum temperature, and heating and cooling degree days. After all values were collected, the 93 weather stations were plotted by using latitude and longitude coordinates, then a raster surface was interpolated using the Inverse Distance Weighted technique. The ZIP code shapefile provided by ESRI Business Analyst was used in the Tabulate Area tool, in order to find the mean value of each variable for each ZIP code. All interpolations, raster processing, and area tabulations were done in ESRI ArcGIS 10.1.

The following table shows the average value for each weather variable across all 161 ZIP codes in Dallas and 140 ZIP codes in Houston for each month of 2012.

Variable	Dallas	Houston
January precipitation	6.1	4.4
February precipitation	2.0	4.8
March precipitation	5.7	5.3
April precipitation	2.4	2.4
May precipitation	1.9	4.9
June precipitation	1.7	3.6
July precipitation	0.8	6.1
August precipitation	3.8	3.3
September precipitation	2.8	3.5
October precipitation	0.9	0.9
November precipitation	0.1	0.7
December precipitation	1.8	3.6
January average temperature	50.0	58.5
February average temperature	52.0	59.9
March average temperature	64.2	68.7
April average temperature	70.0	73.2
May average temperature	77.4	78.3
June average temperature	85.3	83.3
July average temperature	87.4	82.7
August average temperature	86.0	85.4
September average temperature	79.4	79.5
October average temperature	66.2	70.7
November average temperature	58.8	63.2
December average temperature	48.5	59.1
January maximum temperature	79.3	79.7
February maximum temperature	81.7	82.1

March maximum temperature	84.0	85.2
April maximum temperature	89.2	87.0
May maximum temperature	95.8	93.3
June maximum temperature	105.2	102.3
July maximum temperature	107.4	95.9
August maximum temperature	107.2	98.8
September maximum temperature	104.1	96.6
October maximum temperature	88.6	89.4
November maximum temperature	87.4	87.2
December maximum temperature	82.3	83.7
January minimum temperature	25.7	29.2
February minimum temperature	26.5	35.0
March minimum temperature	36.8	40.8
April minimum temperature	45.9	50.9
May minimum temperature	56.6	62.7
June minimum temperature	59.3	68.3
July minimum temperature	70.5	71.3
August minimum temperature	63.8	72.3
September minimum temperature	54.9	58.9
October minimum temperature	34.4	41.2
November minimum temperature	31.2	39.0
December minimum temperature	20.4	29.9
January heating days	460.4	216.0
February heating days	375.0	192.2
March heating days	109.4	42.7
April heating days	13.8	0.7
May heating days	0.0	0.0
June heating days	1.3	0.0
July heating days	0.2	0.0
August heating days	2.8	0.0
September heating days	0.1	0.0
October heating days	110.3	42.7
November heating days	216.0	120.9
December heating days	451.6	243.4
January cooling days	2.0	26.2
February cooling days	6.8	49.9

March cooling days	90.2	164.7
April cooling days	171.0	255.2
May cooling days	391.5	412.2
June cooling days	571.4	555.5
July cooling days	701.5	555.9
August cooling days	654.1	636.4
September cooling days	437.5	440.4
October cooling days	152.7	219.0
November cooling days	37.0	79.7
December cooling days	26.4	74.5

Figure 18. Mean value for each weather variable, by month 2012

Heating days, as reported in the NCDC QCLCD, are defined as the sum of negative differences between the mean daily temperature (calculated by averaging the daily maximum and minimum temperatures) and a base of 65 degrees Fahrenheit. This is a measure of energy demand. The following table shows the average number of heating days per month across all 1,889 ZIP codes in Texas, the lowest value, and the highest value for each month of 2012, and the same data for Dallas and Houston.

Cooling days, being the inverse of heating days, are the sum of positive differences between the mean daily temperature and a base of 65 degrees Fahrenheit. The following table shows the average number of cooling days per month across all 1,889 ZIP codes in Texas, the lowest value, and the highest value for each month of 2012, and the same data for Dallas and Houston.

The weather variables were collected from the following number of stations.

Variable	Number of stations
Precipitation	86
Average temperature	61

Maximum temperature	93
Minimum temperature	93
Heating Days	63
Cooling Days	63

Figure 19. Number of stations each weather variable was collected from

The natural environment is also represented in this research through a measure of land cover. The data used for this variable was obtained from the National Land Cover Database (NLCD) 2006. This is the most current land cover data available for the study region. The NLCD 2006 uses 16 classes applied at a resolution of 30 meters. The 16 classes can be grouped into the following eight groups, which were used in this study: water, including open water and perennial ice or snow; developed, including developed, open space and developed, low-, medium-, and high-density; barren; forest, including deciduous forest, evergreen forest, and mixed forest; shrubland; herbaceous, including grassland, sedge, lichens, and moss; planted/cultivated, including pasture and hay and cultivated crops; and wetlands, including woody wetlands and emergent herbaceous wetlands (US Department of the Interior and US Geological Survey 2013).

The original raster data sets were downloaded to cover the study area and then reclassified from 16 classes to eight, as laid out above. The ESRI Business Analyst ZIP code shapefile was used to calculate Zonal Statistics, giving an output table of percentage of each ZIP code covered by each land cover class. The table below shows the average percentages calculated for each 161 ZIP codes in Dallas and 140 ZIP codes in Houston.

Variable	Dallas	Houston
Water	0.50	1.15
Developed Land	11.91	62.56
Barren Land	0.39	0.19

Forest	9.95	7.67
Shrub	23.55	0.88
Herbaceous Land	18.19	15.54
Cultivated Land	31.29	9.97
Wetland	4.23	2.04

Figure 20. Average percent of land covered by different land cover types

The last natural environmental factor used in this research was elevation. Both average elevation and elevation range for each ZIP code were included. This data was collected from the Texas Natural Resource Information System, part of the Texas Water Development Board. This dataset was originally part of the National Elevation Data (NED), but already cropped for the boundary of Texas. The NED is a raster data set created by the U.S. Geological Survey aimed at providing accurate elevation data at a resolution of one arc-second, or approximately 30 meters (US Department of the Interior and US Geological Survey 2006). For this research, the Zonal Statistics tool was used again in order to determine the average, the maximum and minimum, and therefore the range of elevation for each of the 161 ZIP codes in Dallas and 140 ZIP codes in Houston.

Variable	**Dallas**	**Houston**
Average Elevation	577.30	70.20
Elevation Range	125.08	33.94

Figure 21. Mean elevation values

The social and built environments are also represented in this research through several variables. In order to consider the effect of the built environment, housing data from ESRI Business Analyst was included. For each ZIP code, housing density, the percentage of homes occupied by owner, the percentage occupied by renters, and the percentage left vacant were included as possible contributing factors to the incidence rate of West Nile Virus. These characteristics can contribute to how the built environment is cared for and to factors of neighborhood maintenance. The social environment factors include population density; the gender breakdown of the population; the median age of the population; the percentage of the population identified as white, black, and Hispanic; and the median household income for each ZIP code. The data included in ESRI Business Analyst were tested in a blind study against other data vendors and found to be the most accurate information available (ESRI 2012). It is based on projections from the 2010 US decennial census.

Variable	Dallas	Houston
Total Population	29179.54	32374.15
Population Density	3108.47	3668.21
Median Age	33.85	33.19
Percent Male	49.60	50.04
Median Male Age	33.04	32.38
Percent Female	50.40	49.96
Median Female Age	34.56	33.97
Percent White	60.93	56.89
Percent Black	18.79	19.88
Percent American Indian	0.68	0.65
Percent Asia	4.72	5.64
Percent Pacific Islander	0.10	0.06
Percent Other Race	11.95	13.86
Percent Hispanic	30.61	39.47
Median Household Income	55538.55	51925.27
Average Household Income	72358.65	68475.43
Per Capita Income	28126.02	25773.31
Total Housing Units	11593.32	12556.52
Percent Owner-occupied housing	52.59	51.89
Percent Renter-occupied housing	38.53	37.70
Percent Vacant housing	8.88	10.41

Figure 22. Mean values for demographic and socioeconomic variables

Methods

The following methods were used in order to develop an appropriate model for explaining the difference in West Nile incidence rates among ZIP codes in Dallas and Houston.

Analyzing the incidence of West Nile Virus

In order to see the spatial distribution of incidence rates across the state of Texas, a hot spot analysis was run in ArcGIS, which uses the Global Moran's I statistic. The specification of the hot spot analysis technique is as follows. First, because of the nature of ZIP code shapes and sizes in the state, the zone of indifference method was used to conceptualize the spatial relationships. In this method, neighboring features inside the specified distance threshold receive a weight of 1, while past this distance the influence of neighboring features diminishes with distance. To determine an appropriate threshold, the Spatial Autocorrelation tool was run at 5,000 meter intervals starting at 5,000 meters across the state of Texas. There was drop in the z-score at a threshold of 95,000 meters, which was chosen as the critical distance threshold. The hot spot analysis tool uses the Getis-Ord Gi* statistic in order to identify statistically significant hot and cold spots, or clusters of high and low values, respectively. This analysis identified a significantly hot spot in the Dallas area. In contrast, the Houston area was identified as a significantly cold spot. This led to the two areas being the focus of this research.

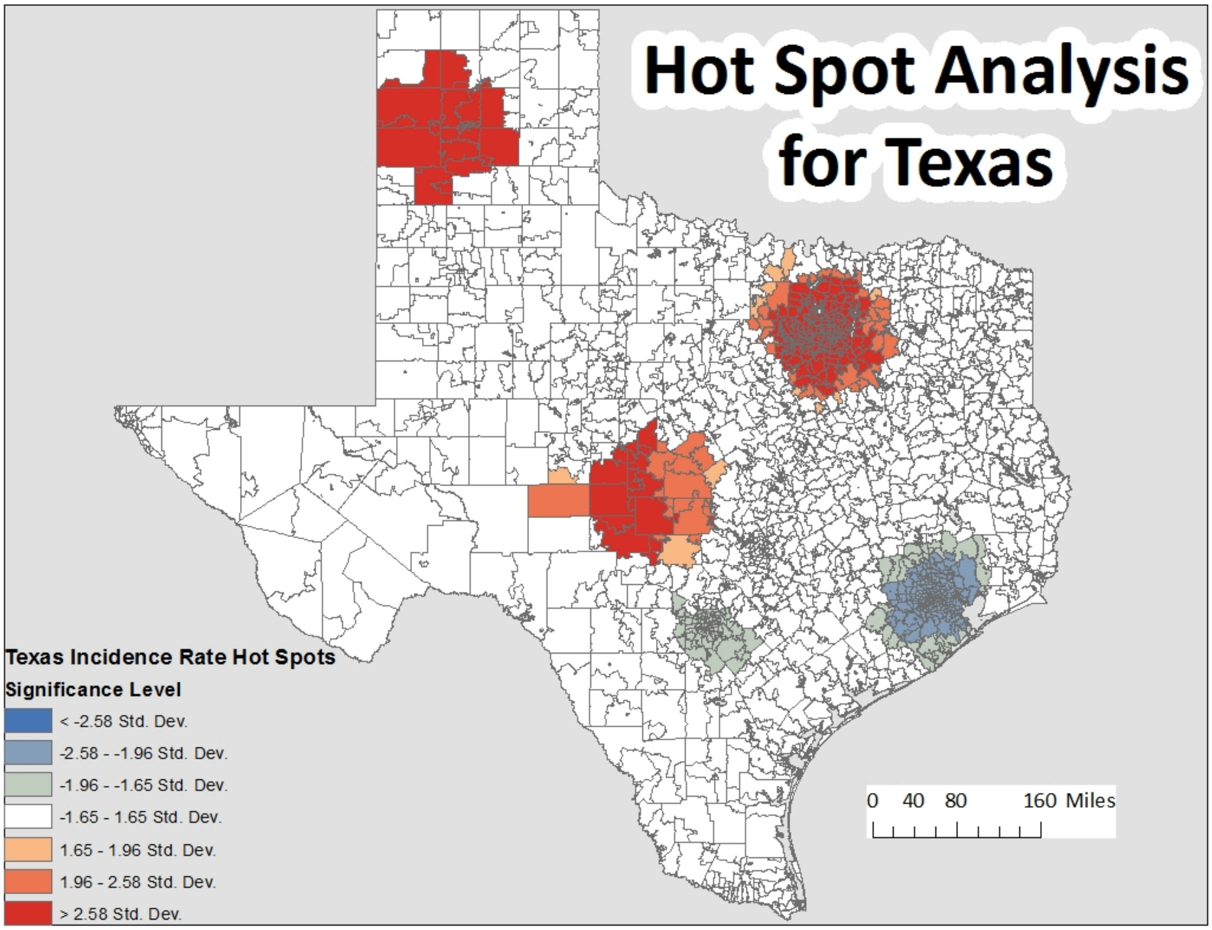

Figure 23. Hot Spot Analysis for Texas, 2012

A similar process was run on both the smaller study areas to explore the spatial distribution of incidence rates in each area. The Spatial Autocorrelation tool was run at 500 meter intervals, starting at 3,000 meters for both Dallas and Houston, until an appropriate threshold was determined. For the Dallas ZIP codes, a threshold of 8,500 meters was chosen. In Houston, the threshold chosen was 6,500 meters. These same thresholds were used with the zone of indifference method in the Hot Spot Analysis.

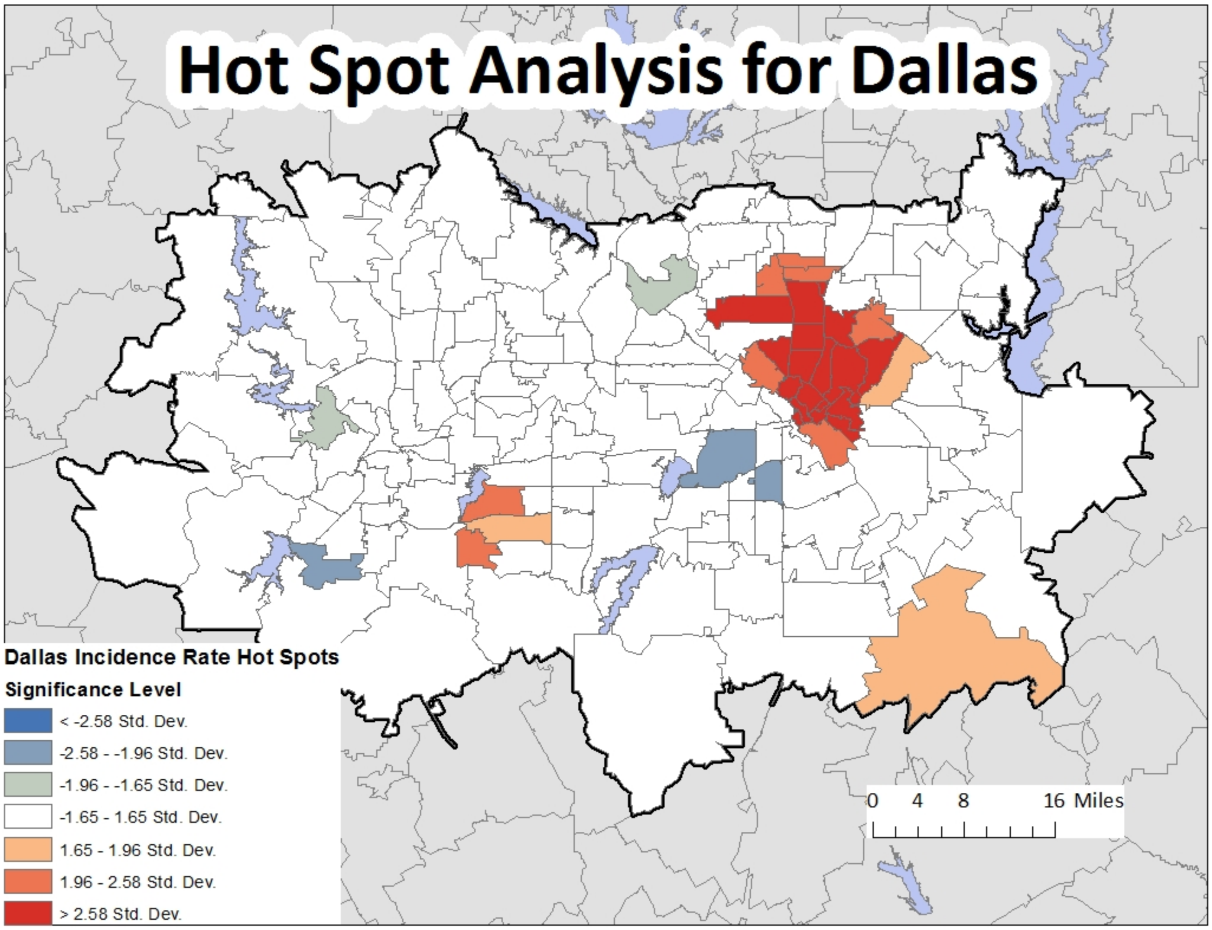

Figure 24. Hot Spot Analysis for Dallas study area

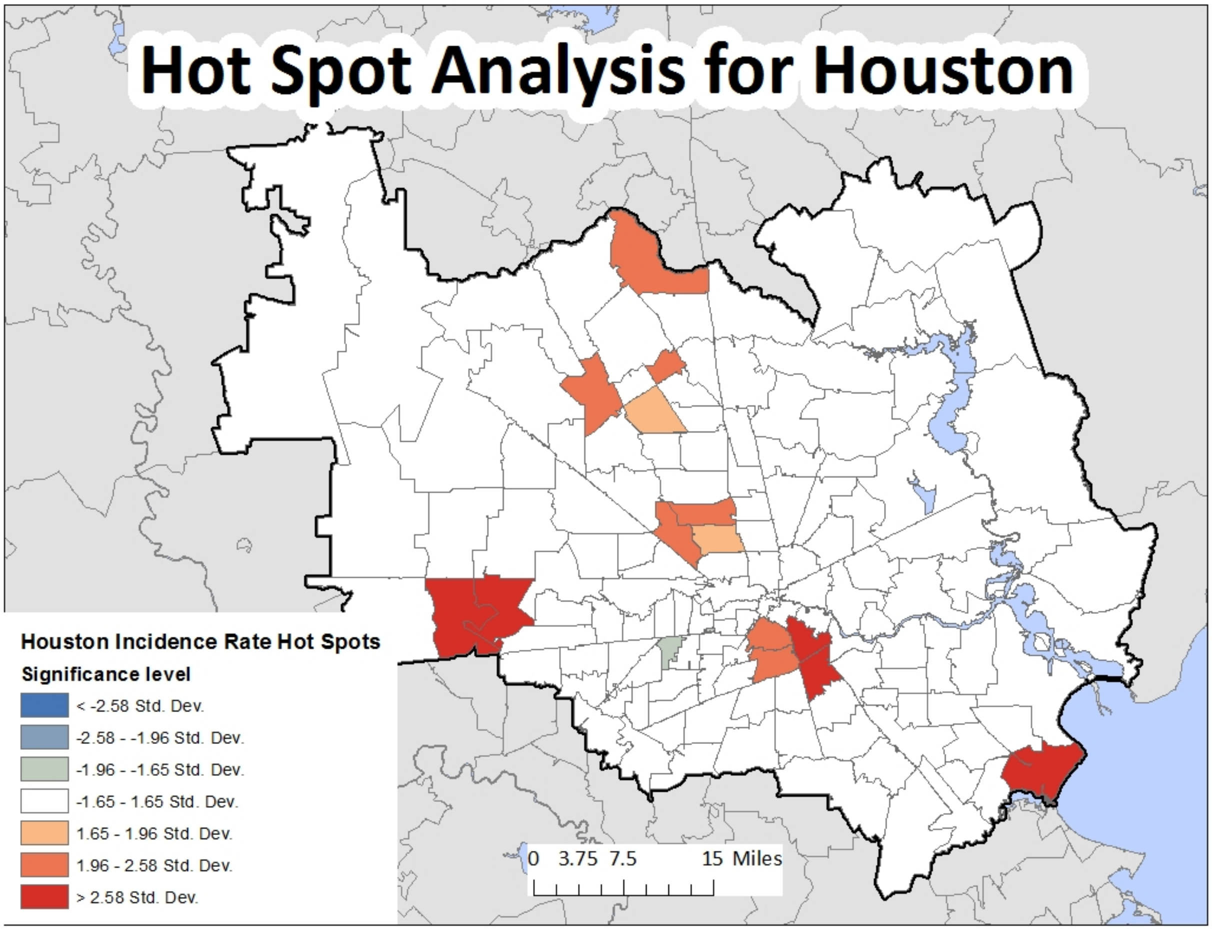

Figure 25. Hot Spot Analysis for Houston study area

Following that, further explorations of the incidence rate were run in SPSS.

The incidence rates of West Nile Virus across ZIP codes in Dallas showed a Pearson skewness coefficient of 1.50, and there were 21 ZIP codes with zero cases of WNV. In Houston, the Pearson skewness coefficient was 2.198, and there were 89 ZIP codes with zero cases of WNV. Because of this positive skew, the log of each incidence rate was taken, creating a more normal distribution to use in the model.

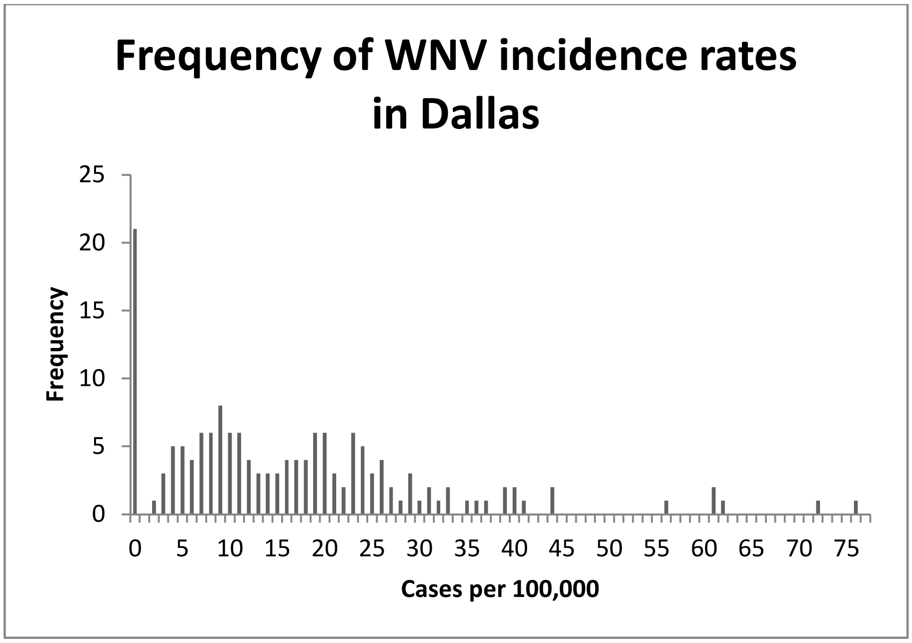

Figure 26. Dallas incidence frequency

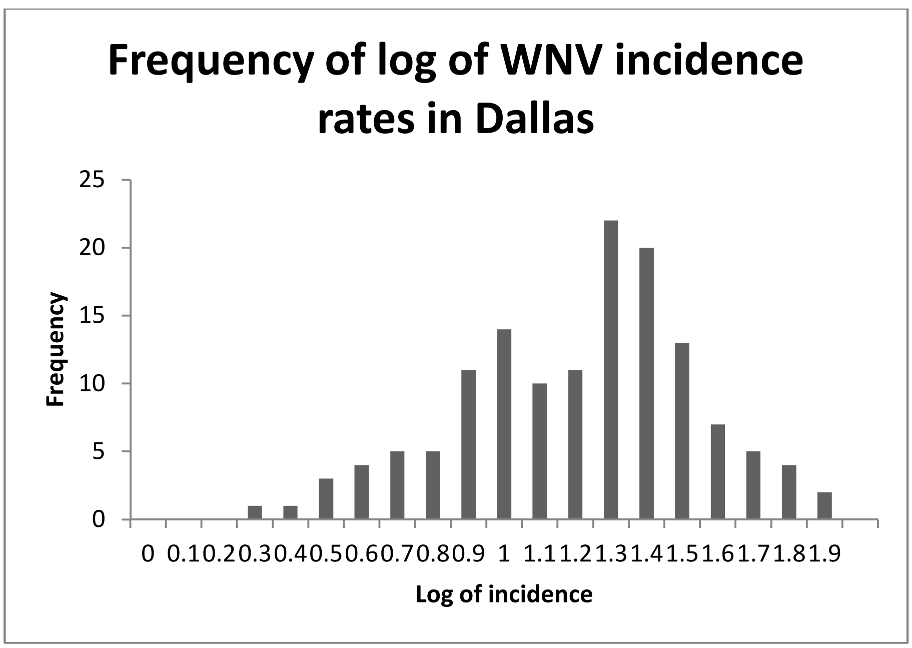

Figure 27. Dallas log of incidence frequency

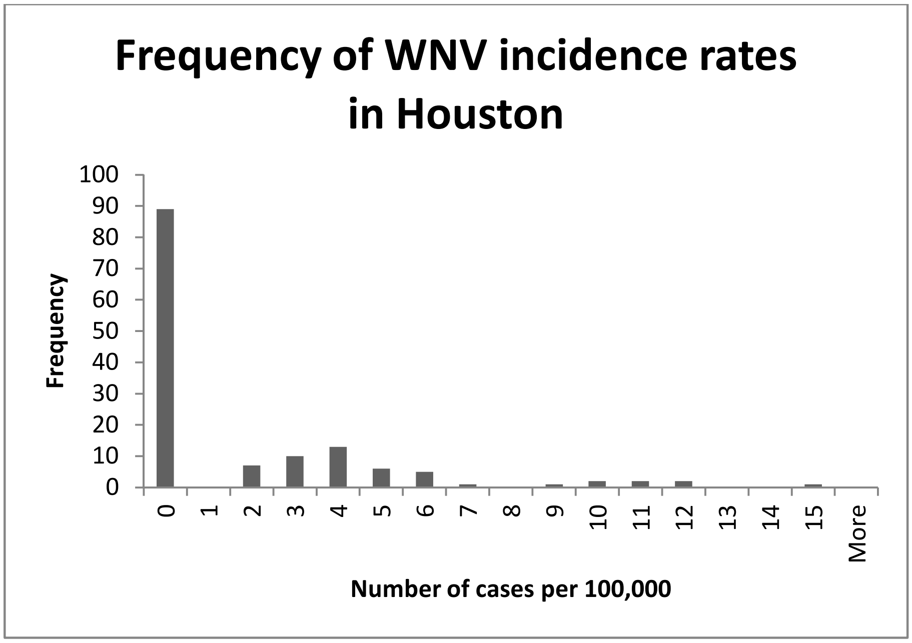

Figure 28. Houston incidence frequency

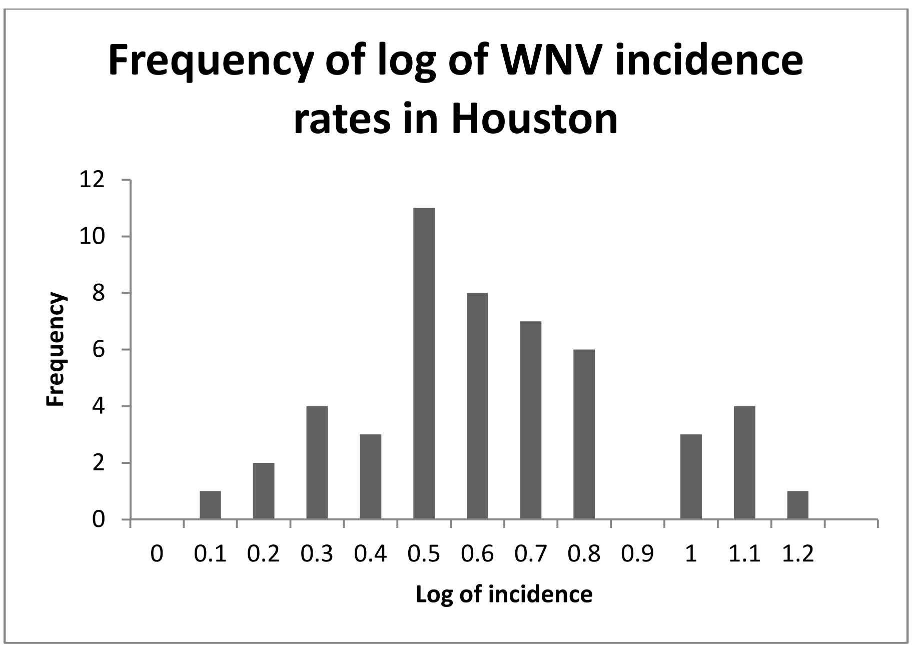

Figure 29. Houston log of incidence frequency

Model Selection

Once all the independent variables were processed as listed above, the following steps were taken to develop an appropriate model for explaining the difference in West Nile Virus incidence rates among ZIP codes in Dallas and Houston. The same process was used for each of the study areas.

There were a total of 105 variables for each study area. Initially, the Pearson correlation coefficient was calculated for each pair of independent variables using SPSS. Pairs which had a coefficient of 0.8 or higher were identified and one variable removed from each pair. This addressed concerns regarding multicollinearity in the independent variables.

Next, an Ordinary Least Squares regression was run on the subset of uncorrelated variables. However, this produced low adjusted R-squared values. Given the nature of the independent variables, which could be spatially autocorrelated, other regression models were tested to check for improved performance. GeoDa, an open-source software designed for geospatial analysis, was used to test a Spatial Lag Model and a Spatial Error Model. GeoDa uses ESRI's shapefile structure, so the same Business Analyst data was used in this software. For spatial analysis in GeoDa, an appropriate weights file must be created. For this study, a queen weights matrix was used, so that all surrounding ZIP codes, including those connected at a vertex, are used to define a location's neighbor. The Spatial Lag Model incorporates a spatially lagged dependent variable, which can assess the existence and strength of spatial interaction among neighbors (Anselin 1999). The Spatial Error Model assumes that the errors of a model are spatially correlated, so that the model's shortcomings in one location are similar to its shortcomings in the neighbors of that location (Anselin 1999).

However, both of these tests showed low R-squared values and very low spatial autocorrelation values, indicating other regression models would provide better results.

Because of the relatively low number of ZIP codes, compared to the number of independent variables used in the model, the low R-squared value potentially resulted from an inappropriate feature or independent variable selection. The statistical software R was used to perform a leaps function on the independent variables. By using an efficient branch-and-bound algorithm, the function searches for the best subsets of the independent variables for predicting the dependent variable (Venables, Smith, and the R Core Team 2013). This method produced the following subsets of variables to explain the differences in the log of WNV incidence in each metropolitan area.

Dallas
Population Density
Median Age
Percent Black
Percent Hispanic
Percent Vacant
Percent Water
Percent Developed
Average Elevation
March min temp
May max temp
May heating days
June max temp
June heating days
July precipitation

Houston
Population Density
Median Age
Percent Black
Percent Hispanic
Percent Vacant
Percent Water
Percent Wetland
Park count
January maximum temp
February precipitation
March precipitation
April precipitation
June maximum temp
August precipitation
August maximum temp

Figure 30. Final set of variables for each study area

Model Interpretation

The results of this spatial analysis will also be compared to the spatial patterns found in known health disparity indicators. Many of the known risk factors of West Nile Virus are also considered health disparity indicators, such as age, race, and income. A hot spot analysis will be run with the same parameters as the incidence rates on each of the health disparity indicators to identify any similarities.

Based upon the data, processing, and methodology described in this chapter, the next chapter will describe the results in detail.

IV. Results

Based on the methodology laid out in the previous chapter, this chapter describes the results in detail.

Hot Spot Analysis

After identifying Dallas and Houston as significantly hot and cold spots, respectively, of incidence of West Nile Virus in Texas, the same analysis was run within each study area to identify any potential clusters within those areas. Using the zone of indifference method, and the thresholds determined by the Spatial Autocorrelation tool (Dallas: 8500 meters, Houston: 6500 meters), significant hot spots were found in both study areas, and significant cold spots in the Dallas study area only.

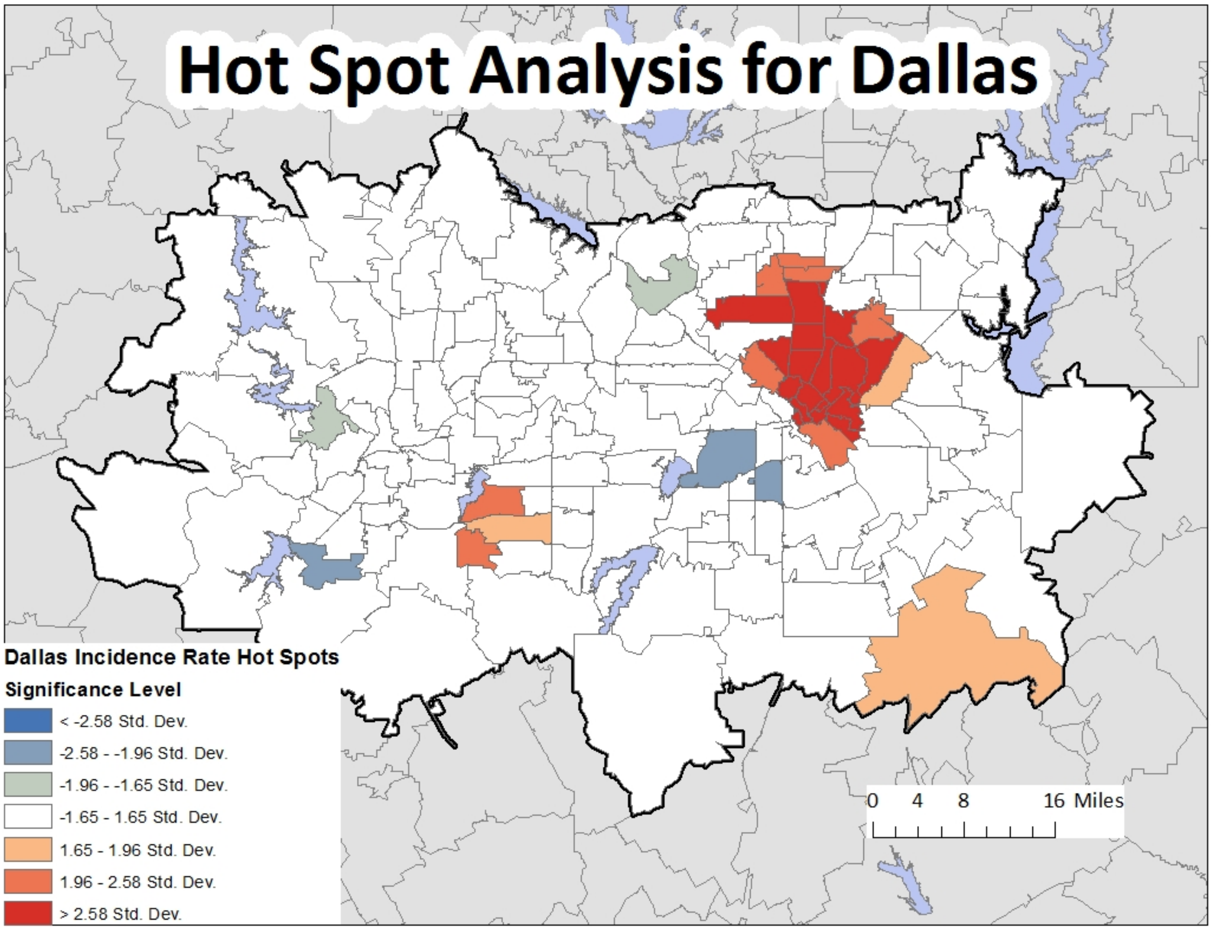

Figure 31. Hot Spot Analysis for Dallas study area

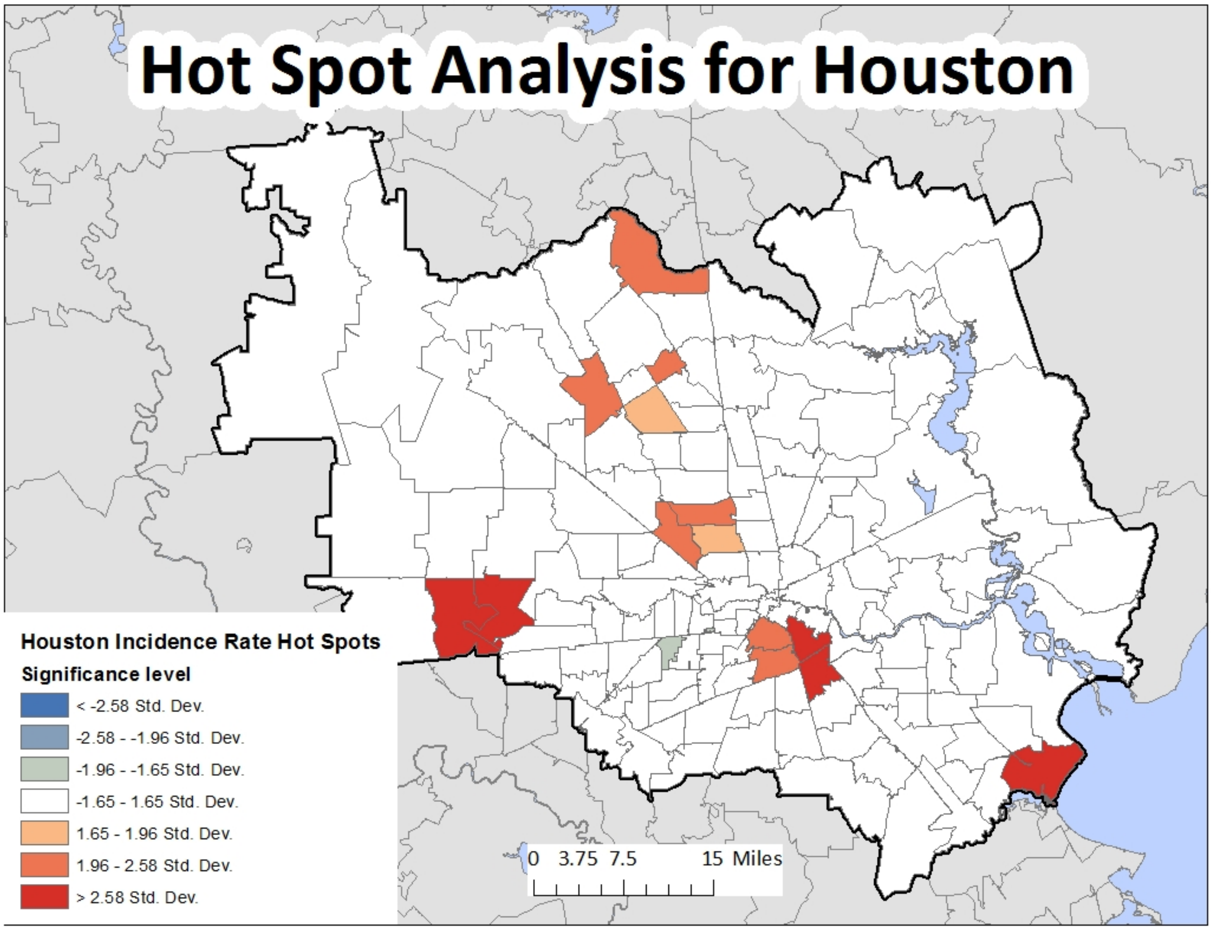

Figure 32. Hot Spot Analysis for Houston study area

In the Dallas study area, there is a distinct cluster of high incidence values in the northeast quadrant of Dallas County. This area includes Dallas city limits, as well as the areas of Highland Park and University Park. There are 24 ZIP codes in this area. There are also two other ZIP codes with significantly high values in the Dallas study area. They are located in Tarrant County, and include the cities of Kennedale and Dalworthington Gardens. Of the 26 ZIP codes included in the Dallas hot spots, the highest incidence rate of 75.8 in ZIP code 75223. Interestingly, a neighboring ZIP code, 75226, had zero cases of WNV. These were located in the middle of Dallas. The average incidence rate for these 26 ZIP codes was 34.3 cases per 100,000 residents.

In the Houston study area, there was no large cluster of high incidence values, but rather seven groups of one to four ZIP codes with significantly high incidence rates. There were a total of 12 ZIP codes with high incidence rates. Two ZIP codes (77070 and 77068) had no cases of WNV, but were still considered hot spots. The highest incidence rate in the Houston area hot spots was 14.5 cases per 100,000 residents. This was in the 77021 ZIP code, which is right in the middle of the study area, and contiguous with three other ZIP codes considered hot spots. The average incidence rate for these 12 ZIP codes was 7.2 cases per 100,000 residents.

Correlations and Model Testing

Using SPSS, the Pearson correlation coefficient was determined for each pair of variables. Because of high levels of multicollinearity in many of the independent variables, this created a smaller subset of variables to be used to explain the difference in the log of WNV. In the Dallas study area, this subset consisted of 39 independent variables. In Houston, a subset of 30 uncorrelated variables was chosen. For example, there was significant correlation between the median age of the entire population, and the median age of each gender; there was significant correlation between the percentage of homes occupied by the owner, and those occupied by renters or that were vacant; and there was significant correlation between many of the climate variables.

These subsets were then used in an Ordinary Least Squares (OLS) regression, using SPSS. The resulting adjusted R-squared value for the Dallas area was 0.353 with a standard error of 11.68. In the Houston area, the adjusted R-squared value was 0.0144 and the standard error was 2.832.

The results showed few variables of significance, as well as high levels of multicollinearity.

The next models considered were the Spatial Lag Model and the Spatial Error Model. First, the Spatial Lag Model was run on both study areas. The Spatial Lag Model for the Dallas study area returned a nearly identical R-squared value as the OLS, and the Akaike Information Criterion (AIC) value actually increased. It showed no spatial autocorrelation and high multicollinearity among the 39 variables used in the subset. The Spatial Lag Model for the Houston study area showed a slightly improved R-squared value, but an identical AIC value. It also showed no spatial autocorrelation, and high multicollinearity among the 30 variables used in the subset. These results indicate that the spatial lag model did not improve the model performance and was not a good fit for the given data.

Next, the Spatial Error Model was run on both study areas. The Spatial Error Model for the Dallas study area also had disadvantages, resulting in a slightly improved R-squared value over the OLS model, but a nearly-identical AIC value. Again, there was no spatial autocorrelation, and high multicollinearity among the variables. The Spatial Error Model for the Houston study area showed a drop in the AIC value, but no improvement in the R-squared value. The model diagnostics showed no spatial autocorrelation and high multicollinearity among the variables.

These results indicated that the Ordinary Least Squares regression model could provide the best fit, but that the original list of variables was too large and could potentially be improved by removing more collinear variables and identify a more appropriate set of independent variables. To find a better subset of variables, the leaps function from the R statistical software package was used.

For the Dallas study area, a subset of 14 variables was returned. For the Houston study area, there were 15 variables in the final subset. Both subsets showed comparable or improved adjusted R-squared values, and statistical significance in the independent variables. Hence these models were selected as the best fit models.

Final Model

The following models were developed for each study area.

Dallas

For the Dallas study area, the adjusted R-squared value of this model was 0.4734 and a standard error of 0.5885. There was statistical significance, and low multicollinearity among the independent variables. The model for the Dallas area which had the best fit is as follows:

$$\log_{10} y = \beta_0 + \beta_1 x_1 + \cdots + \beta_{14} x_{14} + e$$

where y equals the incidence of West Nile Virus in a ZIP code, β_0 is the intercept, $\beta_1 \ldots \beta_{14}$ represent the coefficients associated with each independent variable $x_1 \ldots x_{14}$, and e is the error.

Table 1. Dallas OLS model results

Variable	Beta value	Standard Error	t-statistic	Significance
Constant	12.075	8.573	1.409	0.162
Population Density	-0.000023	0.000	-1.627	0.106
Median Age	***0.24***	***0.008***	***3.064***	***0.003***
Percent Black	***-0.008***	***0.002***	***-5.019***	***0.000***
Percent Hispanic	-0.004	0.002	-2.416	0.017
Percent Vacant	***0.021***	***0.007***	***2.884***	***0.005***
Percent Water	0.051	0.031	1.647	0.102
Percent Developed	-0.001	0.001	-1.176	0.242
Average Elevation	0.000	0.000	-0.962	0.338
March min temp	***0.98***	***0.033***	***2.978***	***0.003***
May max temp	***-0.176***	***0.084***	***-2.103***	***0.037***
May heating days	***-4.495***	***1.316***	***-3.415***	***0.001***
June max temp	0.023	0.016	1.685	0.169
June heating days	***0.148***	***0.054***	***2.733***	***0.007***
July precipitation	***-0.843***	***0.237***	***-3.556***	***0.001***

Houston

For the Houston study area, the adjusted R-squared value of this model was 0.5956 and a standard error of 0.1699. There was statistical significance, though still had multicollinearity among some independent variables. The model for the Houston area which had the best fit is as follows:

$$\log_{10} y = \beta_0 + \beta_1 x_1 + \cdots + \beta_{15} x_{15} + e$$

where y equals the incidence of West Nile Virus in a ZIP code, β_0 is the intercept, $\beta_1 \ldots \beta_{15}$ represent the coefficients associated with each independent variable $x_1 \ldots x_{15}$, and e is the error.

Table 2. Houston OLS model results

Variable	Beta value	Standard Error	t-statistic	Significance
Constant	-56.646	56.762	-1.051	0.301
Population Density	***-0.000048***	***0.000***	***-3.142***	***0.003***
Median Age	***0.040***	***0.011***	***3.552***	***0.001***
Percent Black	0.004	0.003	1.330	0.192
Percent Hispanic	0.002	0.002	1.075	0.290
Percent Vacant	***0.021***	***0.009***	***2.431***	***0.020***
Percent Water	***0.017***	***0.010***	***1.680***	***0.102***
Percent Wetland	***0.010***	***0.003***	***3.075***	***0.004***
Park count	***-0.625***	***0.223***	***-2.807***	***0.008***
January max temp	***-1.860***	***0.676***	***-2.752***	***0.009***
February precipitation	***-0.800***	***0.281***	***-2.847***	***0.007***

March precipitation	***2.067***	***0.837***	***2.471***	***0.019***
April precipitation	***-2.328***	***0.815***	***-2.857***	***0.007***
June max temp	-0.022	0.021	-1.054	0.299
August precipitation	***1.128***	***0.384***	***2.936***	***0.006***
August max temp	***2.067***	***0.654***	***3.162***	***0.003***

A discussion of these results will follow in the next chapter.

V. Discussion and Conclusion

This chapter will review and discuss the following research questions presented in the introduction to this thesis and address each answer.

1) In 2012, what was the spatial distribution of West Nile Virus in Texas?
2) Were areas of high or low incidence clustered with known health disparity indicators?
3) What natural, built, and social environmental or policy factors affect the incidence rate of WNV in Dallas and Houston?

This is followed by a brief discussion of challenges to WNV prevention and control, and a conclusion including limitations and future research.

Spatial Distribution of West Nile Virus and Health Disparity Indicators

Research questions 1 and 2 were addressed in the following manner. During the course of this research, as noted, one method of exploring the spatial distribution of WNV incidence rates was through the presence of hot spots, or clusters of high incidence rates. As noted in the previous chapter, the spatial distribution of West Nile Virus showed a cluster of high incidence rates in the Dallas area, and a cluster of low incidence rates in the Houston area. The distribution within each study area was also investigated, and then compared with higher proportions of other known health disparity indicators. These included age, race, and socioeconomic status (Adler and Newman 2002, 60-76; Braveman et al. 2010, - S186; Carter-Pokras and Baquet 2002, 426; Krieger et al. 2003, 1655). These disparities can be defined as any difference in health outcome due to these underlying factors. Socioeconomic status has been shown to be a determinant of health care, environmental exposure, and health behavior (Adler and Newman 2002, 60-76), tying it closely to the human ecology of disease triangle and therefore also to the incidence of West Nile Virus. For each study area, the hot spot analysis was run on the percentage of the population above 70 years old, the percentage of the population under 10, the percentage of the population that is female, the percentage of the population identified as black, the percentage of the population identified as Hispanic, the percentage of homes occupied by renters, and the median income of a given ZIP code. While the data is the same as what was included in the model, this was another way of interpreting the relationships.

In the Dallas study area, there is a distinct cluster of high incidence values in the northeast quadrant of Dallas County. Of the health disparity indicators, there is some overlap with a cluster of ZIP codes that have a high percentage of the population over 70 years old, but there is another cluster of an older population in Tarrant County that did not have a significant clustering of high West Nile Virus incidence rates. The age cut-off of 70 years old was because over the age of 70 was when the majority of reported cases were neuroinvasive disease, rather than West Nile Fever. The only other factor that showed a visual correlation with the cluster of high West Nile Virus incidence rates was the percentage of renter-occupied homes. There is a large area of high rates of renters in the northern half of Dallas County. This aligns with the literature on urban dwelling patterns, and could be associated with the correlation between home ownership and home maintenance (Dietz and Haurin 2003, 401-450).

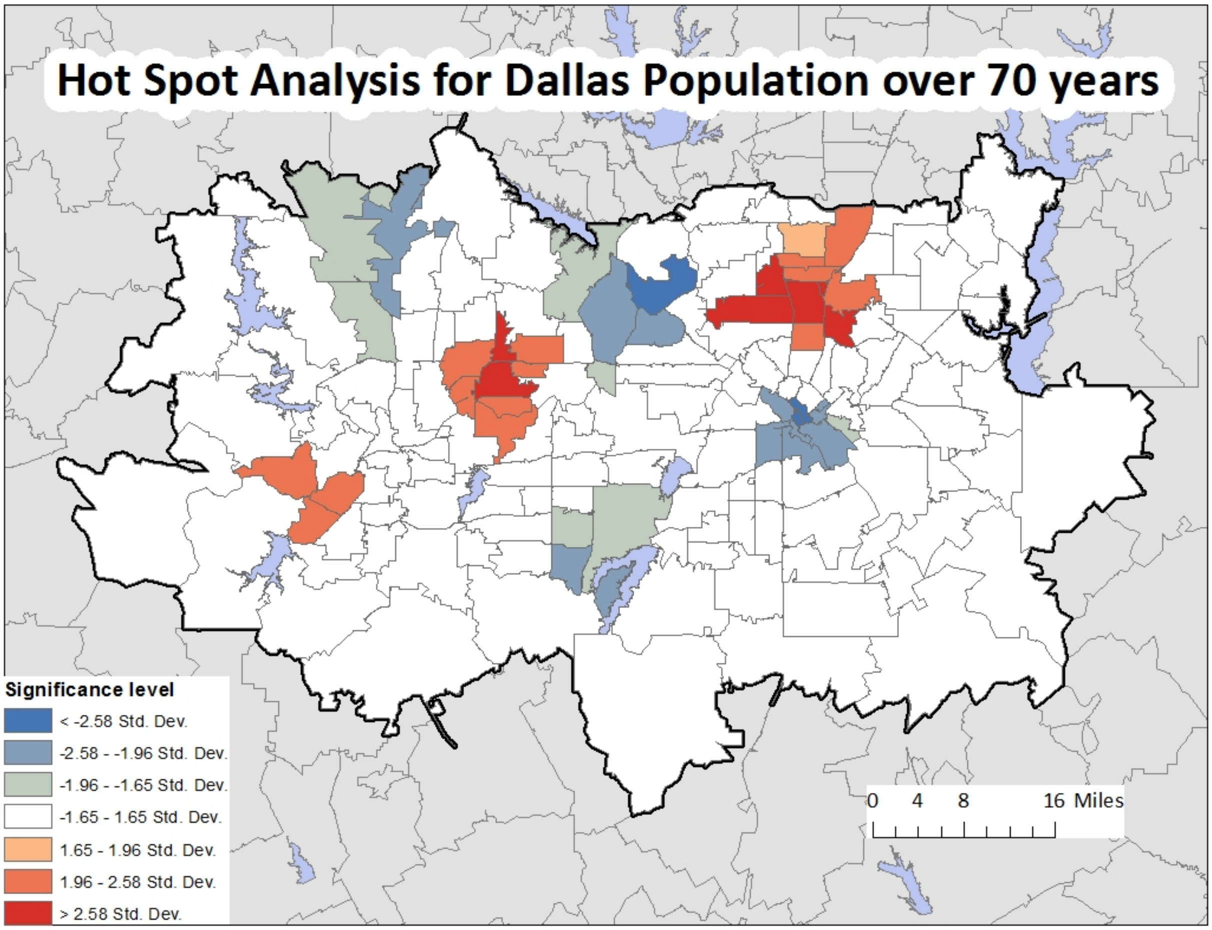

Figure 33. Hot Spot Analysis for Dallas study area - population over 70 years old

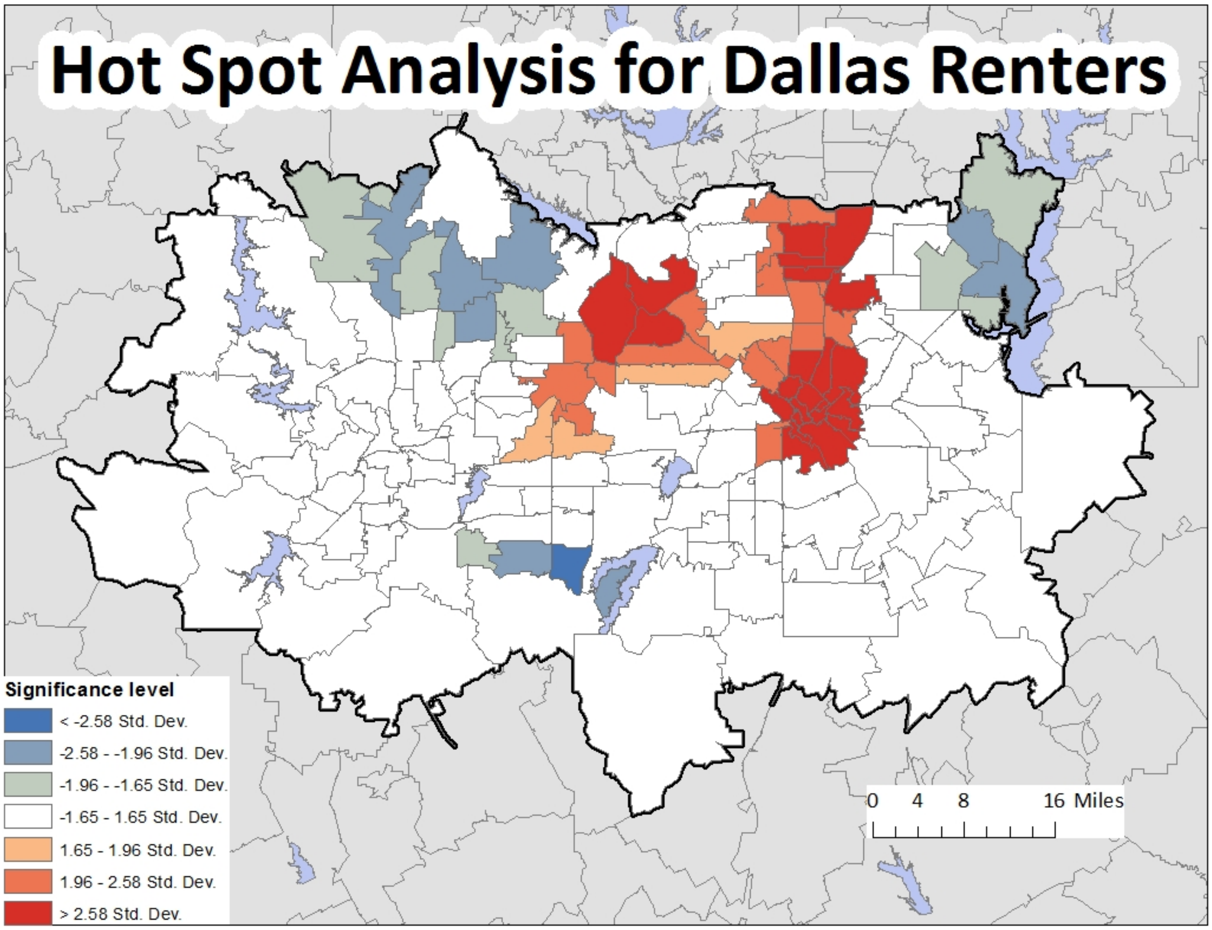

Figure 34. Hot Spot Analysis for Dallas study area - renter-occupied homes

In the Houston area, the hot spots of high West Nile Virus incidence rates are much more dispersed, rather than one contiguous cluster. The indicators of health disparities showed expected patterns, but none correlated visually with the WNV hot spot areas. Interestingly, there were no statistical "cold" spots, or clusters of low values, in Houston, implying a more even distribution in general.

Significant Factors in Determining WNV Incidence

The following subsections will address research question 3 by discussing the factors which were found to be significant in determining the incidence rate of WNV in each of the study areas.

Dallas

Based on the model developed for the Dallas study area through the previously-described research methods, there were several explanatory variables that were significant at the level of p = 0.005. However, some of these had a very weak influence on the incidence of West Nile Virus, or in other words, their coefficient values were small. This included the percent of the population identified as black and the percent of homes that were vacant. Other variables had a stronger effect on the incidence rate of West Nile Virus. The factor with the strongest influence was the number of May heating degree days, which was negatively correlated with incidence. This means that as the number of May heating degree days increased, the incidence of WNV decreased. This could imply that a cooler late spring led to lower rates of WNV, perhaps because of the impact on mosquito populations. Another factor that had a significant impact on the incidence of WNV was the average amount of precipitation in July. It was also negatively correlated with incidence, though not as strongly. Two variables that were positively correlated at a significant level with WNV incidence rates were the median age of the population and the average minimum temperature in March. These also align with the literature in that as the median age of a community increases, so does the incidence rate of WNV, and a warmer average minimum temperature in early spring could lead to higher numbers of mosquitoes and therefore higher rates of incidence of WNV. Of these variables, age, race, and home ownership are considered indicators of health disparities.

Houston

The model for the Houston study area had many of the same variables in it with some key differences to the model for the Dallas study area. In terms of significant variables, median age was also a factor in Houston, but the effect of the median age, though positive, was miniscule. The population density and the percentage of the ZIP code classified as “wetland” were also significant, but with very little influence on the incidence rate. The only variable that was statistically significant with a noticeable impact was the average maximum temperature recorded for August. This had a positive correlation with WNV incidence rates, so as the maximum temperature increased, so did the incidence rate. By looking at variables that were significant at the 0.01 level, there were four others that had a large impact on the incidence rate of WNV. The average minimum temperature in January was negatively correlated with WNV incidence rates, which was unexpected. The average amount of precipitation in April was also negatively correlated with WNV incidence rates, which means a drier April leads to higher rates of WNV incidence. The average amount of precipitation in August is positively correlated with WNV incidence. Both of these agree with what the literature has put forward.

Temporal Effects of Weather

The temporal aspect of an outbreak of West Nile Virus is another critical consideration in determining contributing factors of high incidence rates. As previously discussed, the literature has addressed some of the spatio-temporal effects of weather on dead birds and human cases. While this study did not directly include a temporal aspect, it is important to acknowledge its potential effect, seeing as precipitation during a specific month was noted as a significant factor in each study area. Therefore, the timing of the cases may also be impacted by that factor. Specifically in 2012, the drought that affected much of the United States, including Texas, may have affected the mosquito population or the rate of WNV incidence. The difference between 2012 and the 30-year average precipitation amounts may have contributed to the dramatic increase in the 2012 incidence rates over previous years. The relationship between the incidence rate and the month or previous month's precipitation amounts may also be of note.

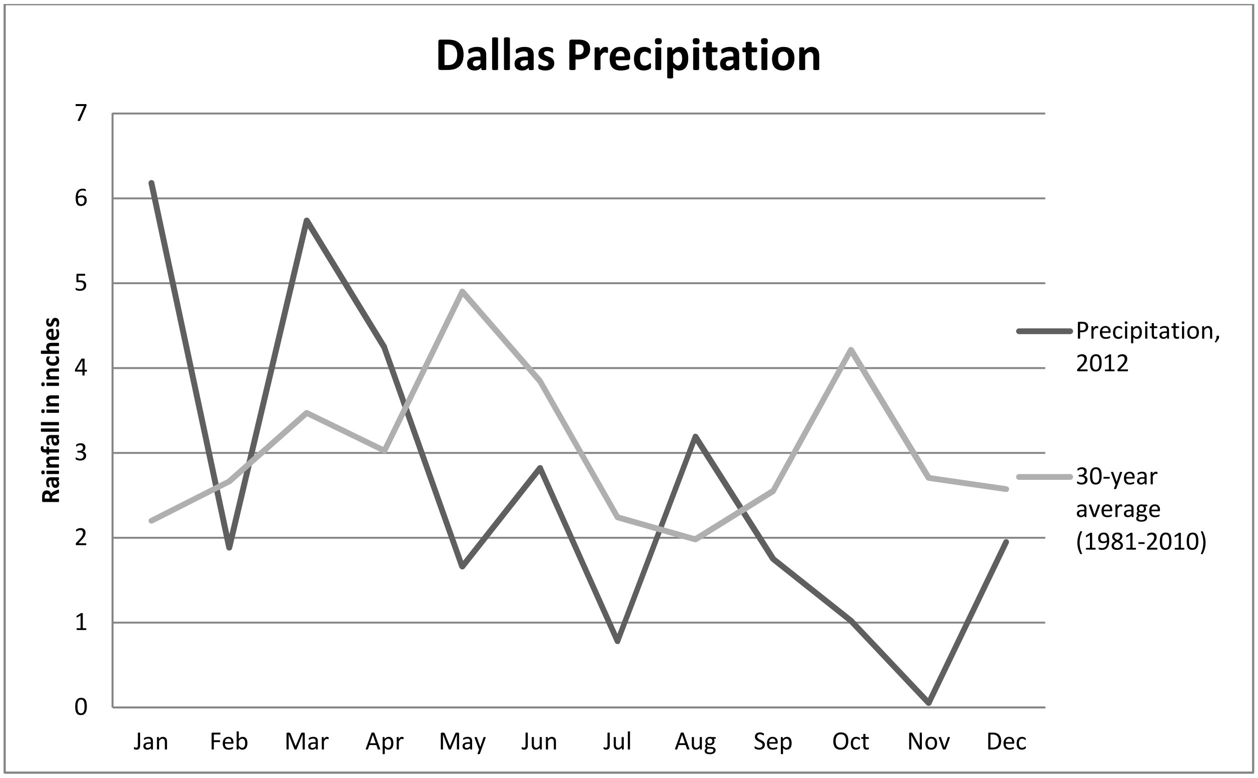

Figure 35. Dallas precipitation, 2012 and 30-year average

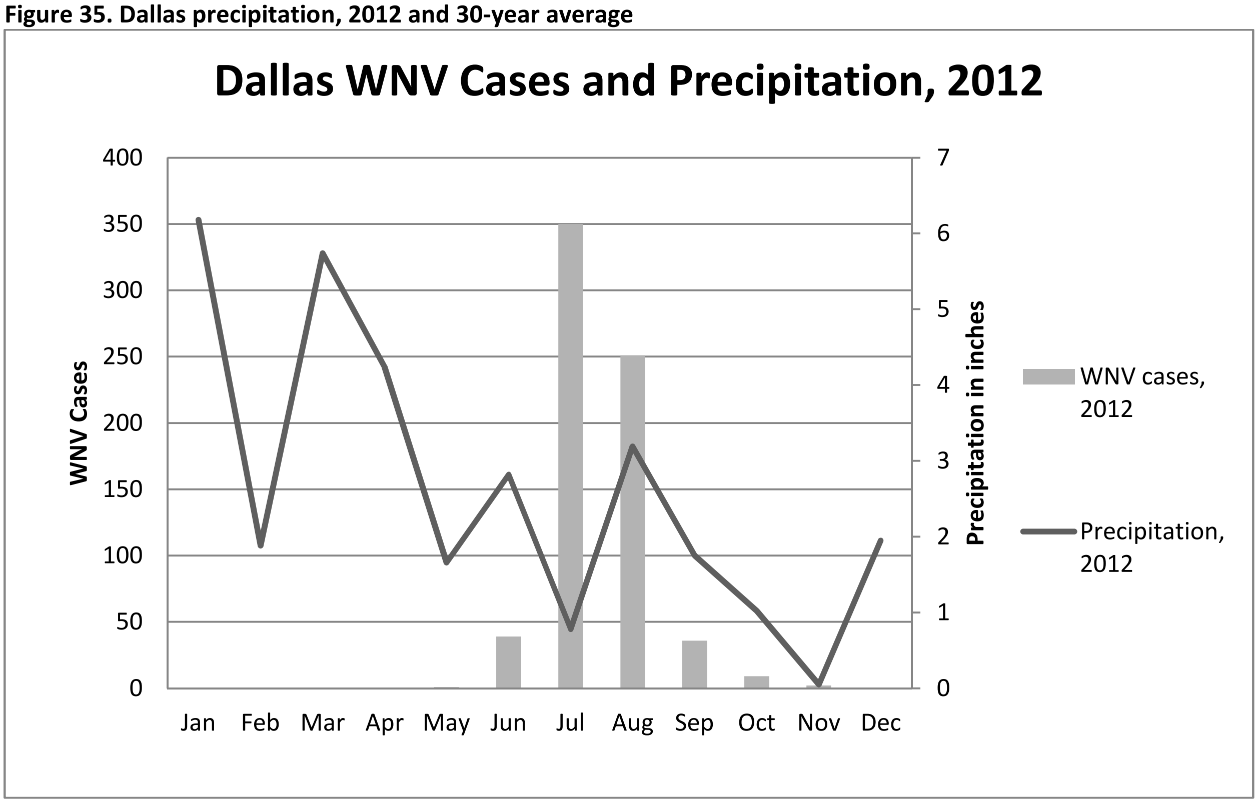

Figure 36. Dallas WNV Cases and Precipitation, 2012

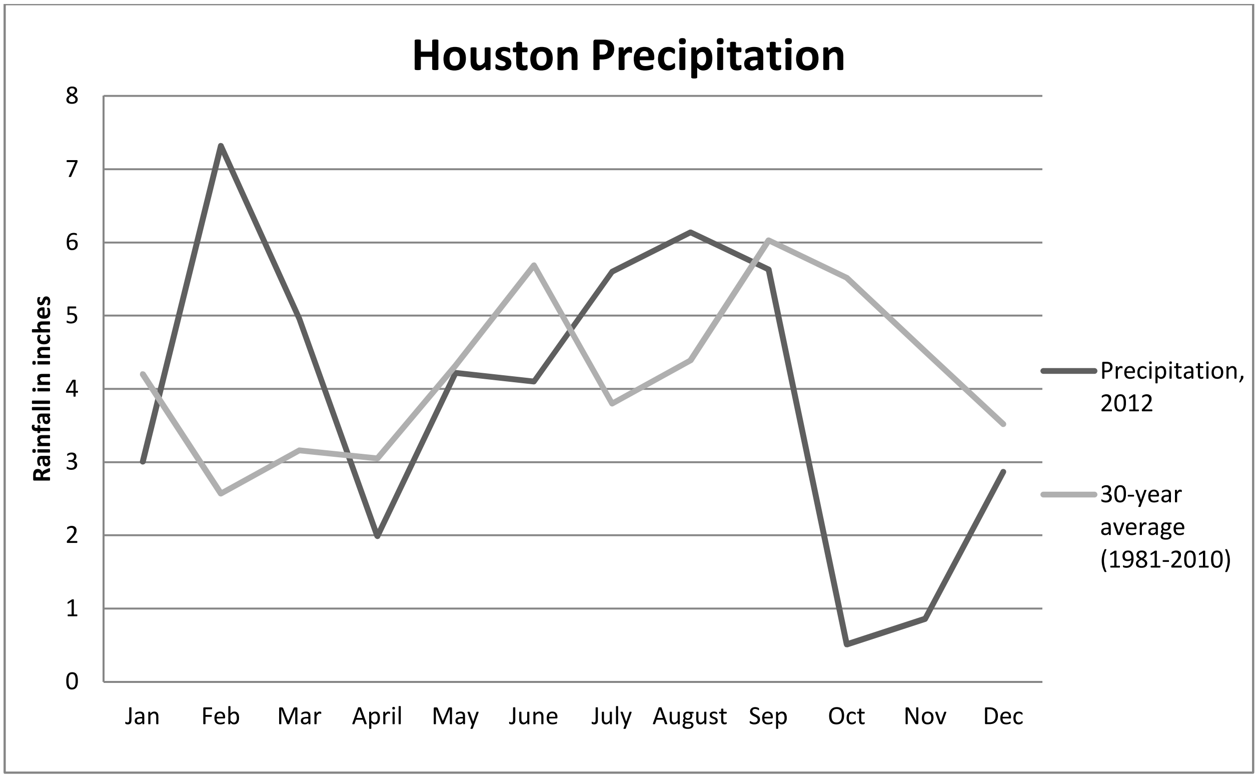

Figure 37. Houston precipitation, 2012 and 30-year average

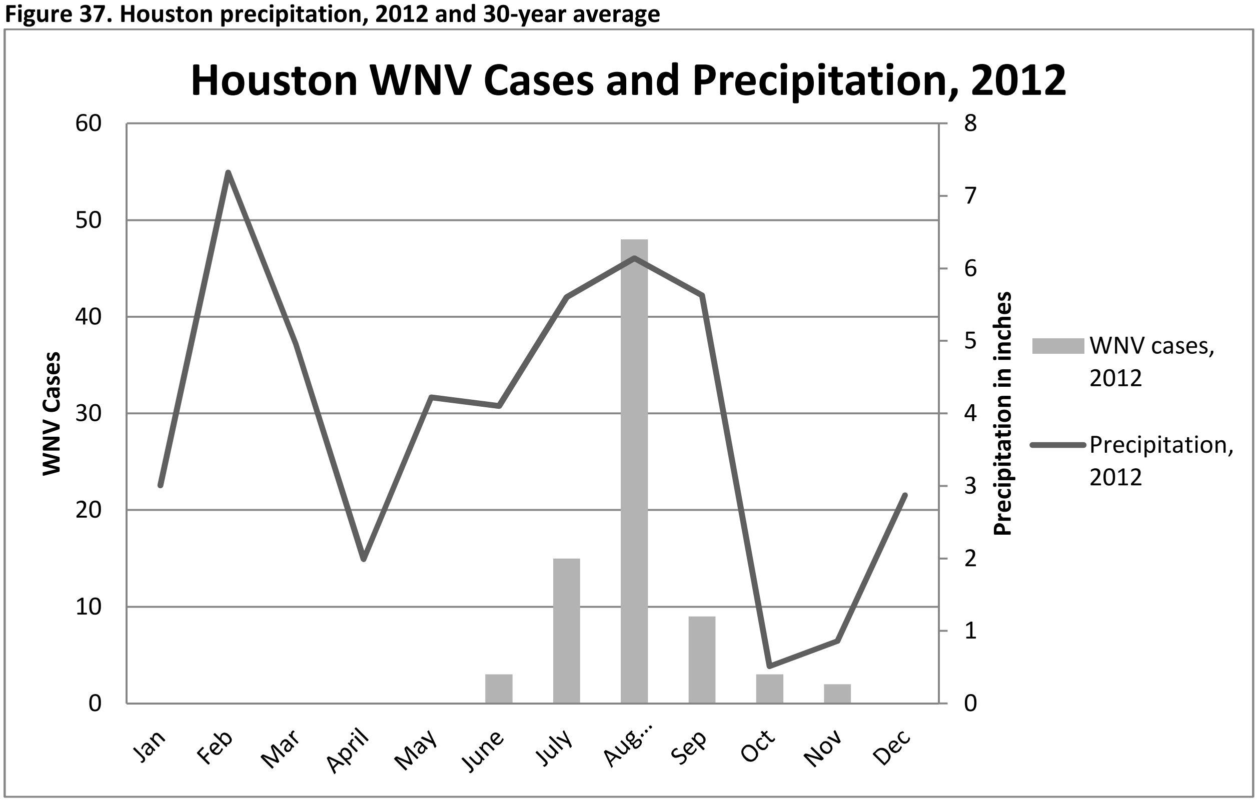

Figure 38. Houston WNV cases and Precipitation, 2012

Challenges to Prevention and Control

There are several challenges facing policy-makers when addressing the prevention and control of West Nile Virus. These include reporting and awareness of the disease, appropriate vector-control policies and concerns regarding these methods, and even the structure of the health services.

While considering the effect of health disparities on the incidence of West Nile Virus, other factors that may be also be affected by common health disparities were noted, and recognized as potentially interdependent. First, because of the prevalence of asymptomatic presentation of West Nile Virus, it is important to consider the number of cases that are not reported. Second, this can also be affected by awareness. Public health campaigns which promote awareness of prevention methods may be amplifying the differences in incidence among different groups. This plays interdependently with the social environment factors and behavioral factors.

Another factor that is important to examine is the use of mosquito abatement methods. While ground spraying was used ubiquitously in both study areas during the summer of 2012, aerial spraying was used more sparingly, typically because of cost. Both Dallas and Houston employed aerial spraying in small areas. These areas showed no significant difference in incidence rates, but because the temporal analysis of cases was not included in this research, there could be unknown effects. This spraying is generally carried out by small aircraft flying low at night that are usually undetectable. There is some concern regarding the effectiveness and safety of the insecticides used in this prevention method (Elnaiem et al. 2008, 751-757; Hemingway et al. 2006, 308-312; Knickerbocker 2012, N.PAG), which contributes to its minimal usage. Due to the lack of evidence of a significant reduction in WNV rates, further research regarding effective targeting and timing of spraying is necessary.

Lastly, in Texas, the Department of State Health Services is broken into eleven different health service regions (Texas Department of State Health Services 2013). Because of the uneven distribution of West Nile Virus cases across the state, some regions carry a much heavier burden of disease control than others. In 2012, the region containing the Dallas study area had over 1,000 cases (or one-fifth of the total) reported, while Region 5, just east of the Dallas region, had none. There are also localities within Texas where the local health department provides services, rather than the regional, all of which have separate budgets and concerns to address.

Conclusion

The 2012 outbreak of West Nile Virus in Texas was far above what has been seen previously in the state. Because there is no vaccination or cure for the disease, understanding the potential causes for higher incidence rates can better prepare health officials to address the disease through effective prevention measures which targeting high risk areas and populations. Because there are so many potentially influential factors, coming from all three vertices of the traditional human ecology of disease triangle, the activities involved in controlling West Nile Virus need to consider all possible angles and adjust for the factors that may have a more significant impact on the local incidence rates of WNV.

As this study has shown, even comparable cities within a few hundred miles of each other can still have different driving factors of incidence rates of West Nile Virus, and may need to take different approaches to controlling future outbreaks. While there are some factors that cannot be controlled, such as climate, or even nominally controlled, such as elevation, these factors can be used to help identify potential areas of high incidence rates. Other factors, such as the racial, age, or socioeconomic characteristics of a community can contribute to developing effective awareness and treatment campaigns.

While this research attempted to minimize the effect of limitations, there are several that could still have impacted the thoroughness of the results. First, within the data, only the ZIP code of residence is reported. This does not account for where a patient was exposed to West Nile Virus, which could potentially better correlate with factors associated with mosquito populations. Also within the data, ZIP codes are not a perfect spatial unit, and while consistency was maintained throughout by using the ESRI Business Analyst boundaries and socioeconomic data, it is possible other data did not line up as neatly. Lastly, the considerations of population and behavioral characteristics of each patient potentially had a strong impact on the risk of contracting WNV, but were not included in this study because of limitations in the data. While data was collected at the individual level, because of confidentiality concerns, it was aggregated to the ZIP code level. Future research could include this individual level of analysis.

These considerations may lead to productive future research possibilities. By looking more closely at the programs currently in place in each study area, a more complete picture could be drawn of effective control techniques. Additionally, as this research addressed the environmental factors in the human ecology of disease triangle, though interconnected with the other vertices, a survey of current trends in population and behavioral factors could also provide a more thorough picture of the West Nile Virus risk in Texas.

References

CDC: West nile virus ArboNET. 2010 [cited 04/28 2012]. Available from http://www.cdc.gov/ncidod/dvbid/westnile/usgs_frame.html.

Adler, Nancy E., and Katherine Newman. 2002. Socioeconomic disparities in health: Pathways and policies. *Health Affairs* 21 (2) (March 01): 60-76.

Anselin, Luc. 1999. Spatial econometrics.

Artsob, H., D. J. Gubler, D. A. Enria, M. A. Morales, M. Pupo, M. L. Bunning, and J. P. Dudley. 2009. *West Nile Virus in the New World: Trends in the Spread and Proliferation of West Nile Virus in the Western Hemisphere*. Vol. 56 Wiley-Blackwell.

Averett, Ellen, Michael H. Fox, Gail Hansen, and John S. Neuberger. 2005. Evaluation of West Nile Virus education campaign. *Emerging Infectious Diseases* 11 (11; 2013/4): 1751+.

Bernkopf, H., S. Levine, and R. Nerson. 1953. Isolation of West Nile Virus in Israel. *The Journal of Infectious Diseases* 93 (3) (Nov. - Dec.): 207-18.

Bode, Amy V., Grant L. Campbell, Edward B. Hayes, Robert S. Lanciotti, James J. Sejvar, and Sherif R. Zaki. 2005. Virology, pathology, and clinical manifestations of West Nile Virus disease. *Emerging Infectious Diseases* 11 (08; 2013/4): 1174+.

Bode, Amy V., James J. Sejvar, W. J. Pape, Grant L. Campbell, and Anthony A. Marfin. 2006. West Nile Virus Disease: A descriptive study of 228 patients hospitalized in a 4-county region of Colorado in 2003. *Clinical Infectious Diseases* 42 (9) (5): 1234-40.

Braveman, Paula A., Catherine Cubbin, Susan Egerter, David R. Williams, and Elsie Pamuk. 2010. *Socioeconomic Disparities in Health in the United States: What the Patterns Tell us*. - American Journal of Public Health. 100th ed. American Public Health Association.

Brown, Heidi E., James E. Childs, Maria Diuk-Wasser, and Durland Fish. 2008. Ecological Factors Associated with West Nile Virus Transmission, Northeastern United States. *Emerging Infectious Diseases* 14 (10; 2013/4): 1539+.

Brownstein, John S., Hilary Rosen, Dianne Purdy, James R. Miller, Mario Merlino, Farzad Mostashari, and Durland Fish. 2002. Spatial Analysis of West Nile Virus: Rapid risk assessment of an introduced vector-borne zoonosis. *Vector-Borne and Zoonotic Diseases* 2 (3) (09/01; 2013/03): 157-64.

Campbell, Grant L., Edward B. Hayes, Nicholas Komar, Susan P. Montgomery, Roger S. Nasci, and Daniel R. O'Leary. 2005. Epidemiology and Transmission Dynamics of West Nile Virus Disease. *Emerging Infectious Diseases* 11 (08; 2013/4): 1167+.

Campbell, Grant L., Anthony A. Marfin, Robert S. Lanciotti, and Duane J. Gubler. 2002. West Nile Virus. *The Lancet Infectious Diseases* 2 (9) (9): 519-29.

Carter-Pokras, O., and C. Baquet. 2002. What is a "health disparity"? *Public Health Reports* 117 (5): 426.

Center for Disease Control and Prevention. West Nile Virus. 2013 [cited 5/1 2013]. Available from http://www.cdc.gov/westnile/index.html.

Centers for, Disease Control. 2005. West Nile Virus Statistics, Surveillance and Control.

Chuang, Ting-Wu, Michael B. Hildreth, Denise L. Vanroekel, and Michael C. Wimberly. 2011. Weather and land cover influences on mosquito populations in Sioux Falls, South Dakota. *Journal of Medical Entomology* 48 (3) (05/01; 2013/05): 669-79.

Cooke, William H., Katarzyna Grala, and Robert Wallis. 2006. Avian GIS models signal human risk for West Nile Virus in Mississippi. *International Journal of Health Geographics* 5 (01): 36-54.

Crowder, David W., Elizabeth A. Dykstra, Jo Marie Brauner, Anne Duffy, Caitlin Reed, Emily Martin, Wade Peterson, Yves CarriÃ¨re, Pierre Dutilleul, and Jeb P. Owen. 2013. West Nile Virus prevalence across landscapes is mediated by local effects of agriculture on vector and host communities. *PLoS ONE* 8 (1) (01/30): e55006.

DeGroote, John P., Ramanathan Sugumaran, Sarah M. Brend, Brad J. Tucker, and Lyric C. Bartholomay. 2008. Landscape, demographic, entomological, and climatic associations with human disease incidence of West Nile Virus in the state of Iowa, USA. *International Journal of Health Geographics* 7 (01): 1-16.

Dietz, Robert D., and Donald R. Haurin. 2003. The social and private micro-level consequences of homeownership. *Journal of Urban Economics* 54 (3) (11): 401-50.

Dohm, David J., Monica L. Oâ€™Guinn, and Michael J. Turell. 2002. Effect of environmental temperature on the ability of culex pipiens (diptera: Culicidae) to transmit west nile virus. *Journal of Medical Entomology* 39 (1) (01/01; 2013/03): 221-5.

Eidson, Millicent, Nicholas Komar, Faye Sorhage, Randall Nelson, Tom Talbot, Farzad Mostashari, and Robert McLean. 2001. Crow deaths as a sentinel surveillance system for west nile virus in the northeastern united states, 1999. *Emerging Infectious Diseases* 7 (4) (Jul): 615.

Eidson, Millicent, Jim Miller, Laura Kramer, Bryan Cherry, and Yoichiro Hagiwara. 2001. Dead crow densities and human cases of West Nile Virus, New York State, 2000. *Emerging Infectious Diseases* 7 (4) (Jul): 662.

Elnaiem, Dia-Eldin A., Kara Kelley, Stan Wright, Rhonda Laffey, Glenn Yoshimura, Marcia Reed, Gary Goodman, et al. 2008. Impact of aerial spraying of pyrethrin insecticide on culex pipiens and culex tarsalis (diptera: Culicidae) abundance and West Nile Virus infection rates in an Urban/Suburban area of Sacramento County, California. *Journal of Medical Entomology* 45 (4) (07/01; 2013/04): 751-7.

Epstein, Paul R. 2001. West Nile Virus and the Climate. *Journal of Urban Health* 78 (2) (06/01): 367-71.

ESRI. 2012. Vendor Accuracy Study.

Ghosh, Debarchana, Steven M. Manson, and Robert B. McMaster. 2010. Delineating West Nile Virus transmission cycles at various scales: The nearest neighbor distance-time model. *Cartography and Geographic Information Science* 37 (04; 2013/3): 149+.

Gibbs, Samantha E. J., Michael C. Wimberly, Marguerite Madden, Janna Masour, Michael J. Yabsley, and David E. Stallknecht. 2006. Factors affecting the geographic distribution of West Nile Virus in Georgia, USA: 2002-2004. *Vector-Borne and Zoonotic Diseases* 6 (1) (03/01; 2013/03): 73-82.

Gibney, Katherine B., James Colborn, Steven Baty, Andrean M. Bunko Patterson, Tammy Sylvester, Graham Briggs, Tasha Stewart, et al. 2012. Modifiable risk factors for West Nile Virus infection during an Outbreak—Arizona, 2010. *The American Journal of Tropical Medicine and Hygiene* 86 (5) (May 01): 895-901.

Godsey MS Jr FAU - Blackmore, Mark,S., Blackmore MS FAU, Panella NA FAU, K. FAU Burkhalter, Gottfried K FAU - Halsey, Lawrence,A., Halsey LA FAU, Rutledge R FAU - Langevin, Stanley,A., et al. 0607. *West Nile Virus Epizootiology in the Southeastern United States, 2001*.

Gosselin, Pierre, Germain Lebel, Sonia Rivest, and Monique Douville-Fradet. 2005. The integrated system for public health monitoring of West Nile Virus (ISPHM-WNV): A real-time GIS for surveillance and decision-making. *International Journal of Health Geographics* 4 (01): 21-12.

Griffith, Daniel A. 2005. A comparison of six analytical disease mapping techniques as applied to West Nile Virus in the coterminous United States. *International Journal of Health Geographics* 4 (01): 18-4.

Gu, Weidong, Richard Lampman, Nina Krasavin, Robert Berry, and Robert Novak. 2006. Spatio-temporal analyses of West Nile Virus transmission in culex mosquitoes in Northern Illinois, USA, 2004. *Vector-Borne and Zoonotic Diseases* 6 (1) (03/01; 2013/03): 91-8.

Gubler, Duane J. 2007. The continuing spread of West Nile Virus in the western hemisphere. *Clinical Infectious Diseases* 45 (8) (10/15): 1039-46.

Han, Linda L., Florin Popovici, James P. Alexander Jr., Velea Laurentia, Leslie A. Tengelsen, Costin Cernescu, Howard E. Gary Jr., Nicolae Ion-Nedelcu, Grant L. Campbell, and Theodore F. Tsai. 1999. Risk factors for west nile virus infection and meningoencephalitis, romania, 1996. *Journal of Infectious Diseases* 179 (1) (1): 230.

Harrigan, Ryan J., Henri A. Thomassen, Wolfgang Buermann, Robert F. Cummings, Matthew E. Kahn, and Thomas B. Smith. 2010. Economic conditions predict prevalence of West Nile Virus. *PLoS ONE* 5 (11) (11/12): e15437.

HAYES, CURTIS G. 2001. West Nile Virus: Uganda, 1937, to New York City, 1999. *Annals of the New York Academy of Sciences* 951 (1): 25-37.

Hemingway, Janet, Barry J. Beaty, Mark Rowland, Thomas W. Scott, and Brian L. Sharp. 2006. The innovative vector control consortium: Improved control of mosquito-borne diseases. *Trends in Parasitology* 22 (7) (7): 308-12.

Huhn, Gregory D., Connie Austin, Carl Langkop, Kate Kelly, Robert Lucht, Richard Lampman, Robert Novak, et al., 2005. The Emergence of West Nile Virus During a Large Outbreak in Illinois, IN, 2002. *The American Journal of Tropical Medicine and Hygiene* 72 (6) (June 01): 768-76.

Kilpatrick, A. M., Laura D. Kramer, Matthew J. Jones, Peter P. Marra, and Peter Daszak. 2006. West Nile Virus Epidemics in North America are driven by shifts in mosquito feeding behavior. *PLoS Biology* 4 (4) (04): e82.

Kitron, U. 2000. Risk maps: Transmission and Burden of Vector-Borne Diseases. *Parasitology Today* 16 (8) (8/1): 324-5.

Knickerbocker, Brad. 2012. Dallas launches air war against West Nile mosquitoes: Is it safe? *Christian Science Monitor* (08/17): N.PAG.

Komar, N. FAU, S. FAU Langevin, S. FAU Hinten, N. FAU Nemeth, E. FAU Edwards, D. FAU Hettler, B. FAU Davis, R. FAU Bowen, and M. Bunning. 0513. *Experimental Infection of North American Birds with the New York 1999 Strain of West Nile Virus*.

Krieger, Nancy, Jarvis T. Chen, Pamela D. Waterman, David H. Rehkoph, and S. V. Subramanian. 2003. Race/Ethnicity, Gender, and Monitoring Socioeconomic Gradients in Health: A Comparison of area-based socioeconomic Measures—The

Public Health Disparities Geocoding Project. *American Journal of Public Health* 93 (10): 1655.

Kulasekera, Varuni L., Bryan Cherry, Carla Glaser, James R. Miller, Laura Kramer, Roger S. Nasci, Farzad Mostashari, and Susan C. Trock. 2001. West Nile Virus Infection in mosquitoes, birds, horses, and humans, Staten Island, New York, 2000. *Emerging Infectious Diseases* 7 (4) (Jul): 722.

Kunkel, Kenneth E., Robert J. Novak, Richard L. Lampman, and Wiedong Gu. 2006. Modeling the Impact of Variable Climate Factors on the Crossover of Culex Restraints and Culex Pipiens (Diptera: Culcidae), Vectors of West Nile Virus in Illinois.*The American Journal of Tropical Medicine and Hygiene* 74 (1) (January 01): 168-73.

Landesman, William J., Brian F. Allan, R. B. Langerhans, Tiffany M. Knight, and Jonathan M. Chase. 2007. Inter-annual Associations between Precipitation and Human Incidence of West Nile Virus in the United States. *Vector-Borne and Zoonotic Diseases* 7 (3) (09/01; 2013/03): 337-43.

Lindsey, N. P. 2012. Medical Risk Factors for Severe West Nile Virus Disease, United States, 2008-2010. *The American Journal of Tropical Medicine and Hygiene* 87 (1) (-07): 179; 179,184; 184.

Lindsey, N. P., Stephanie Kuhn, Grant L. Campbell, and Edward B. Hayes. 2008. West Nile Virus Neuroinvasive Disease Incidence in the United States, 2002-2006. *Vector-Borne and Zoonotic Diseases* 8 (1): 35.

Liu, Ann, Vivian Lee, Deron Galusha, Martin D. Slade, Maria Diuk-Wasser, Theodore Andreadis, Matthew Scotch, and Peter M. Rabinowitz. 2009. Risk factors for human infection with West Nile Virus in Connecticut: A Multi-Year Analysis. *International Journal of Health Geographics* 8 (01): 67-76.

Loeb M., Susan J. Elliott, B.Gibson, M. Fearon, R. Nosal, M. Drebot, C. D'Cuhna, et al. 0518. *Protective Behavior and West Nile Virus Risk*.

Mandalakas, Anna M., Christopher Kippes, Joseph Sedransk, Jeffery R. Kile, Asha Garg, John McLeod, Richard L. Berry, and Anthony A. Marfin. 2005. West Nile Virus Epidemic, Northeast Ohio, 2002. *Emerging Infectious Diseases* 11 (11) (11): 1774-7.

Meade, Melinda S., and Michael Emch. 2010. *Medical Geography*. 3rd ed. New York, New York: The Guilford Press.

Melnick, Joseph L., John R. Paul, John T. Riordan, VoHammie H. Barnett, Natan Goldblum, and Eva Zabin. 1951. Isolation from Human Sera in Egypt of a virus apparently identical to West Nile Virus. *Proceedings of the Society for Experimental*

Biology and Medicine.Society for Experimental Biology and Medicine (New York, N.Y.) 77 (4) (August 01): 661-5.

Mostashari, Farzad, Michel L. Bunning, Paul T. Kitsutani, Daniel A. Singer, Denis Nash, Michael J. Cooper, Naomi Katz, et al. 2001. Epidemic West Nile Encephalitis, New York, 1999: Results of a household-based seroepidemiological survey. *The Lancet* 358 (9278) (7/28): 261-4.

Murgue, B. S. Nurri, H. Triki, V. Deubel, and H. G. Zeller, 2001. West nile in the mediterranean basin: 1950-2000. *Annals of the New York Academy of Sciences* 951 (1): 117-26.

Murray, K., S. Baraniuk, M. Resnick, R. Arafat, C. Kilborn, K. Cain, R. Shallengerger, et al. 2006. Risk factors for Encephalitis and Death from West Nile Virus Infection. *Epidemiology & Infection* 134 (06): 1325.

Nasci, Roger S., Harry M. Savage, Dennis J. White, James R. Miller, Bruce C. Cropp, Marvin S. Godsey, Amy J. Kerst, Paul Bennett, Kristy Gottfried, and Robert S. Lanciotti. 2001. West Nile Virus in overwintering Culex Mosquitoes, New York City, 2000. *Emerging Infectious Diseases* 7 (4) (Jul): 742.

Nash, Denis, Farzad Mostashari, Annie Fine, James Miller, Daniel O'Leary, Kristy Murray, Ada Huang, et al. 2001. The Outbreak of West Nile Virus Infection in the New York City Area in 1999. *N Engl J Med* 344 (24) (06/14; 2013/03): 1807-14.

Nolan, Melissa S., Amber S. Podoll, Anne M. Hause, Katherine M. Akers, Kevin W. Finkel, and Kristy O. Murray. 2012. Prevalence of chronic kidney disease and progression of disease over time among patients enrolled in the Houston West Nile Virus Cohort. *PLoS ONE* 7 (7) (07): 1-5.

Nolan, Melissa S., J. Schuermann, and Kristy O. Murray. 2013. West Nile Virus Infection among Humans, Texas, USA 2002-2011. *Emerging Infectious Diseases* 19 (1): 137.

O'Leary, Daniel R., Anthony A. Marfin, Susan P. Montgomery, Aaron M. Kipp, Jennifer A. Lehman, Brad J. Biggerstaff, Veronica L. Elko, Peggy D. Collins, John E. Jones, and Grant L. Campbell. 2004. The Epidemic of West Nile Virus in the United States, 2002. *Vector-Borne and Zoonotic Diseases* 4 (1) (03/01; 2013/03): 61-70.

Ozdenerol, Esra, Elzbieta Bialkowska-Jelinska, and Gregory N. Taff. 2008. *Locating Suitable Habitats for West Nile Virus-infected Mosquitoes through Association of Environmental Characteristics with Infected Mosquito Locations: A Case Study in Shelby County, Tennessee*. Vol. 7 BioMed Central.

Petersen, Lyle R., and Marc Fischer. 2012. Unpredictable and difficult to control — the adolescence of West Nile Virus. *N Engl J Med* 367 (14) (10/04; 2013/03): 1281-4.

Petersen, Lyle R., and Anthony A. Marfin. 2002. West Nile Virus: A primer for the clinician. *Annals of Internal Medicine* 137 (3) (August 6): 173-9.

Philip, C. B., and J. E. Smadel. 1943. Transmission of West Nile Virus by infected aedes albopictus. *Proceedings of the Society for Experimental Biology and Medicine.Society for Experimental Biology and Medicine (New York, N.Y.)* 53 (1) (May 01): 49-50.

Reisen, W. K. 0218. *Landscape epidemiology of vector-borne diseases*.

Reisen, William K., Ying Fang, Hugh D. Lothrop, Vincent M. Martinez, Jennifer Wilson, Paul Connor, Ryan Carney, Barbara Cahoon-Young, Marzieh Shafii, and Aaron C. Brault. 2006. Overwintering of West Nile Virus in Southern California. *Journal of Medical Entomology* 43 (2) (03/01; 2013/03): 344-55.

Reisen, William K., Ying Fang, and Vincent M. Martinez. 2006. Effects of Temperature on the Transmission of West Nile Virus by culex tarsalis (diptera: Culicidae). *Journal of Medical Entomology* 43 (2) (03/01; 2013/03): 309-17.

Ruiz, Marilyn O., Carmen Tedesco, Thomas J. McTighe, Connie Austin, and Uriel Kitron. 2004. Environmental and social determinants of human risk during a West Nile Virus outbreak in the greater Chicago area, 2002. *International Journal of Health Geographics* 3 (01): 8-11.

Ruiz, Marilyn O., Edward D. Walker, Erik S. Foster, Linn D. Haramis, and Uriel D. Kitron. 2007. Association of West Nile Virus illness and urban landscapes in Chicago and Detroit. *International Journal of Health Geographics* 6 (01): 10-1.

Sejvar, James J. 2003. West Nile Virus: An historical overview. *The Ochsner Journal* 5 (3): 6.

Shaman, Jeffrey, Jonathan F. Day, and Marc Stieglitz. 2005. Drought-induced amplification and epidemic transmission of West Nile Virus in southern Florida. *Journal of Medical Entomology* 42 (2) (03/01; 2013/05): 134-41.

Smith, David L., Jonathan Dushoff, and F. E. Mckenzie. 2004. The risk of a mosquito-borne infection in a heterogeneous environment. *PLoS Biology* 2 (11) (11): 1957-64.

Smithburn, K. C., T. P. Hughes, A. W. Burke, and J. H. Paul. 1940. A neurotropic virus isolated from the blood of a native of uganda. *The American Journal of Tropical Medicine and Hygiene* s1-20 (4) (July 01): 471-92.

Su, T., Webb, James P., Meyer, R.P., and M. S. Mulla. 0805. *Spatial and temporal distribution of mosquitoes in underground storm drain systems in Orange County, California*.

Tachiiri, Kaoru, Brian Klinkenberg, Sunny Mak, and Jamil Kazmi. 2006. Predicting outbreaks: A spatial risk assessment of West Nile Virus in British Columbia. *International Journal of Health Geographics* 5 (01): 21-.

Taylor, R. M. 1956. A study of the ecology of West Nile Virus in Egypt. *The American Journal of Tropical Medicine and Hygiene* 5 (4): 579; 579,620 .; 620 .

Texas Department of State Health Services. Health Service Regions. 2013 [cited 04/02 2013]. Available from http://www.dshs.state.tx.us/regions/.

Texas State Historical Association. Texas Almanac. [cited 4/28 2013]. Available from https://www.texasalmanac.com/.

Theophilides, C. N., S. C. Ahearn, E. S. Binkowski, W. S. Paul, and K. Gibbs. 2006. First evidence of West Nile Virus amplification and relationship to human infections. *International Journal of Geographical Information Science* 20 (1) (01): 103-15.

Theophilides, C. N., S. C. Ahearn, S. Grady, and M. Merlino. 2003. Identifying West Nile Virus risk areas: The dynamic continuous-area space-time system. *American Journal of Epidemiology* 157 (9): 843-54.

US Department of the Interior, and US Geological Survey. National land cover database (2006) product legend. 2013 [cited 3/28 2013]. Available from http://www.mrlc.gov/nlcd06_leg.php.

———. National elevation dataset. 2006 [cited 3/28 2013]. Available from http://ned.usgs.gov/about.asp.

Venables, W. N., D. M. Smith, and the R Core Team. 2013. An introduction to R.

Wang, G., Minnis, Richard, B., Belant, J.L., and C. L. Wax. 0510. *Dry weather induces outbreaks of human West Nile Virus infections*.

Winters, Anna M., Rebecca J. Eisen, Saul Lozano-Fuentes, Chester G. Moore, W. John Pape, and Lars Eisen. 2008. Predictive spatial models for risk of West Nile Virus exposure in eastern and western Colorado. *The American Journal of Tropical Medicine and Hygiene* 79 (4) (October 01): 581-90.

Zou, Li, Scott N. Miller, and Edward T. Schmidtmann. 2006. Mosquito larval habitat mapping using remote sensing and GIS: Implications of coalbed methane development and West Nile Virus. *Journal of Medical Entomology* 43 (5) (09/01; 2013/03): 1034-41.