MODELING THE IMPACT OF PROJECTED LAND COVER ON LYME DISEASE EMERGENCE

Jayashree Surendrababu

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE
In
GEOGRAPHY

Korine N. Kolivras, Chair
James B. Campbell
Jie Li
Stephen Prisley

May 2, 2014

Blacksburg, Virginia

Keywords: GIS, Lyme disease, Land cover projections, IPCC scenarios

MODELING THE IMPACT OF PROJECTED LAND COVER ON LYME DISEASE EMERGENCE

Jayashree Surendrababu

ABSTRACT

Lyme disease is a common tick borne disease in the US. Lyme disease emerged from the Northeast and in the past decade, Virginia has been witnessing a rapidly increasing trend in incidence. This thesis uses land cover projection data as a basis to look at the potential future trend of Lyme disease incidence in Virginia for the IPCC (Intergovernmental Panel on Climate change) scenarios of A1B and A2, which indicate a global and regional focus respectively. This study is a continuation of previous work done by an NSF funded research team at Virginia Tech, in exploring the variables affecting Lyme disease in Virginia.

A Poisson point process is implemented in this thesis with land cover parameters (implemented land, water bodies, and edge metrics) and demographic parameters (population percentage and per capita income) as the spatial covariates. Lyme disease incidence data obtained from the Virginia Department of Health was used for model validation. The overall model was implemented using Python and its associated libraries while ArcGIS software was used for preliminary covariate analysis and data visualization.

This thesis generates risk maps for A1B and A2 scenarios for each decade from 2010 through 2060. Spatial occurrence of disease incidence has been generated by the Poisson point process and the risk level of each county in Virginia has been calculated based on the incidence count predicted for it. Population and area at risk under each scenario for each decade was calculated. Results show that in A1B scenario 22.1% and 42.9% of the total population of Virginia are under high risk and in the A2 scenario, 21% and 33% of the total population of Virginia are under high risk of Lyme disease in 2010 and 2060 respectively. In terms of the area, A1B scenario has 28% under high risk in 2010 and 66% of the total area under high risk in 2060, while A2 scenario has 22.4% under high risk of Lyme disease in 2010 62.7% of the total area in Virginia is under high risk in 2060.

ACKNOWLEDGEMENTS

I take this opportunity to express gratitude to my advisor, Dr. Korine N. Kolivras who has not only steered my thesis in the right direction but has also been my morale booster, mentor and a wonderful friend. It was because of her that I understood the importance of research for a graduate student and I will always cherish the experience of working with her the past two years. I also take this opportunity to thank my extremely supportive committee members Dr. James B. Campbell, Dr. Jie Li and Dr. Stephen Prisley for their valuable suggestions in helping me frame a methodology apt for this thesis and for their constant support throughout the course of this study in terms of their technical expertise and encouragement. I wish to collaborate with all my committee members sometime in the future as well!

I would like to thank Sarah Elizabeth Dymond who carried out the previous phase of this research for providing me the basic data and heads up on getting started with this research. I would also like to thank the Lyme disease research group of Virginia Tech for their excellent knowledge sharing sessions through the team meetings, which helped me understand spatial statistical processes.

I would like to thank the Center for Geospatial Information Technology (CGIT) and Network Dynamics and Simulation Science Lab (NDSSL) at Virginia Tech for funding my first and second year of graduate program at Virginia Tech respectively. I would also like to thank the faculty, staff and graduate students of the Department of Geography for making my thesis journey memorable and for their excellent knowledge sharing sessions.

I thank my parents Surendrababu Konian and Dr. P.T Uma Maheswari for their love and encouragement and for being there for me always. I would also like to thank my friend, Prasanna Kumar Ragavan from Virginia Tech, for all the help and encouragement during the course of my thesis journey.

ATTRIBUTION

The third chapter of this thesis has been prepared as a paper for a journal publication. Four colleagues have helped in framing the research methodology and writing for this chapter; their specific contributions are described below.

Chapter 3: Modeling the impact of projected land cover on Lyme disease emergence.

Chapter 3 has been prepared for submission to Geospatial Health.

Korine N. Kolivras, PhD is currently an Associate Professor at the Department of Geography, Virginia Tech. Dr. Kolivras is a co-author on this paper and has taken care of the overall management of the project and has also helped in organizing and writing this paper.

Stephen Prisley, PhD is currently an Associate Professor at the Department of Forest Resources and Environmental Conservation, Virginia Tech. Dr. Prisley is a co-author and has helped in framing the methodology for all the spatial analyses performed in this paper.

Jie Li, PhD is currently an Assistant Professor at the Department of Statistics, Virginia Tech. Dr. Li is also a co-author on this paper and has helped with the statistical analyses used in this paper.

James B Campbell, PhD is currently a Professor at the Department of Geography, Virginia Tech. Dr. Campbell is a co-author on this paper and has provided his expertise in making editorial comments and improvements.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION Page

CHAPTER 2: LITERATURE REVIEW

CHAPTER 3: MANUSCRIPT

LIST OF TABLES

CHAPTER 3: MANUSCRIPT Page

CHAPTER 1: INTRODUCTION

Lyme disease is caused by a spirochete, the bacterium *Borrelia burgdorferi*, which, in the eastern United States, is transmitted by the blacklegged tick *Ixodes scapularis* (CDC, 2011). Early symptoms of Lyme disease include fever, headache, fatigue, depression, and a characteristic circular skin rash called *erythema migrans,* and its seriousness results from the fact that its later symptoms may involve the joints, heart, and central nervous system (CDC, 2011). In most cases, antibiotics eliminate the infection and its symptoms, especially if the illness is treated early. Delayed or inadequate treatment can lead to the more serious symptoms, which can be disabling and difficult to treat (CDC, 2011).

According to Franke et al. (2013), Lyme disease is the most frequently reported vector-borne disease in the Northern Hemisphere, and also the most common tick-borne disease in the world. Even within the United States (US), Lyme disease is the most common tick-borne disease (Hu L. 2011). In the US, Lyme disease incidence is particularly high in the Northeast and Mid-Atlantic states from Maine to Virginia, the Midwest including Minnesota, Wisconsin, and Michigan, and in the West in northern California (Hu L, 2011). The emergence of *Lyme borreliosis* as a public health burden within the last two decades has stimulated renewed interest in tick-borne infections (Gubler et al. 2001). Lyme disease is the sixth most commonly notified disease at the national level. Figure 1 shows the extent of Lyme disease expansion across the previous decade, with the main concentrations in the Northeast and upper Midwest.

Figure 1: Expansion of Lyme disease in the US (Source: CDC, 2011)

1.1 LYME DISEASE IN THE UNITED STATES

According to the Centers for Disease Control and Prevention (CDC, 2011), the total number of confirmed cases of Lyme disease has increased from 19,804 in 2004 to 29,959 in 2009 in the US. Figure 2 shows the incidence of Lyme disease from 2002 through 2011 across the US (CDC, 2011). According to the CDC (2013), more than 300,000 Lyme disease cases are being reported every year and at the same time, approximately 10 times of the yearly reported number, are being underreported. Over 20 years ago, White et al. (1991) mentioned that the geographic range of Lyme disease has been increasing steadily as well, especially across the northeastern United States, and that pattern has continued.

While Lyme disease is expanding its geographic range within the US, it can be limited by unfavorable habitat conditions for both the tick and host including environmental elements that may reduce ambient humidity or the likelihood of interaction between vectors and hosts (Keirans et al., 1996). Specifically, the risk of human infection is closely associated at the landscape scale with woodland and edge habitats in the northeastern United States.

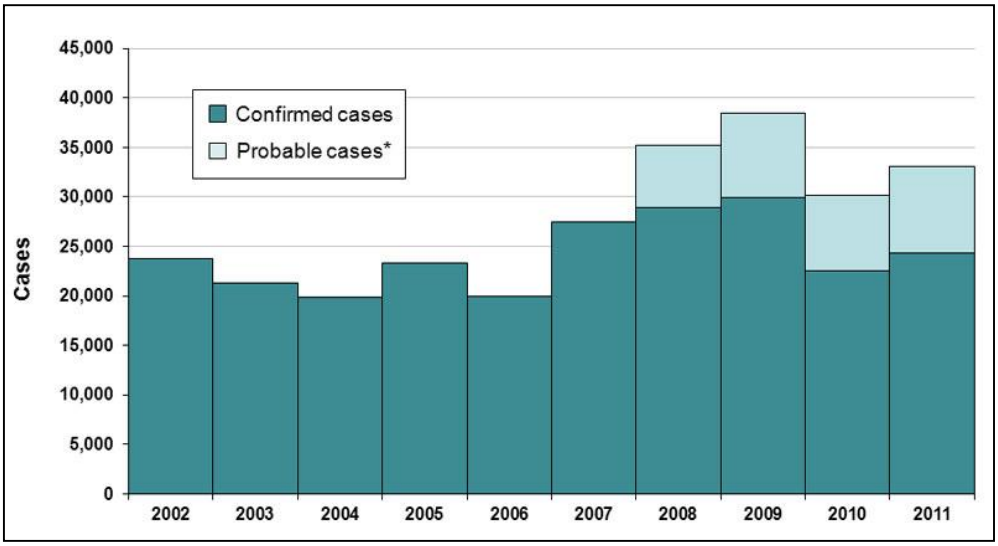

Figure 2: Reported Cases of Lyme disease by Year in the US 2002-2011 (Source: CDC, 2011)

Lyme disease occurs in humans when the enzootic cycle between ticks and disease reservoirs spills over into the human population. Small mammals, for instance, the white footed mouse, are both reservoirs and hosts for ticks in the northeastern US (Barbour et al., 1993). Deer are generally critical hosts for adult tick feeding and reproduction in these regions, but are not disease reservoirs as they are not infected with *B. burgdorferi* at levels high enough to continue the disease cycle. Once the disease gets transmitted to humans, the spirochetes cause illness giving rise to flulike symptoms and manifest into neurologic, cardiac, or musculoskeletal conditions, after the first couple of months (CDC, 2011).

1.2 LYME DISEASE AND TRENDS IN VIRGINIA

Virginia, in particular, has been witnessing an increasing trend in the annual number of newly

identified Lyme disease cases since 2000 (VDH, 2011). For instance, from 2003 to 2009, Lyme disease incidence increased in Virginia by five times. The Virginia Department of Health (VDH) has found that the state is facing an expansion in the range of Lyme disease; the spread of the disease is trending southwards and westwards since 2000. The current study therefore holds significance in choosing the study area as Virginia to examine Lyme disease emergence. Figure 3 shows that disease has become endemic in approximately 75% of Virginia's counties.

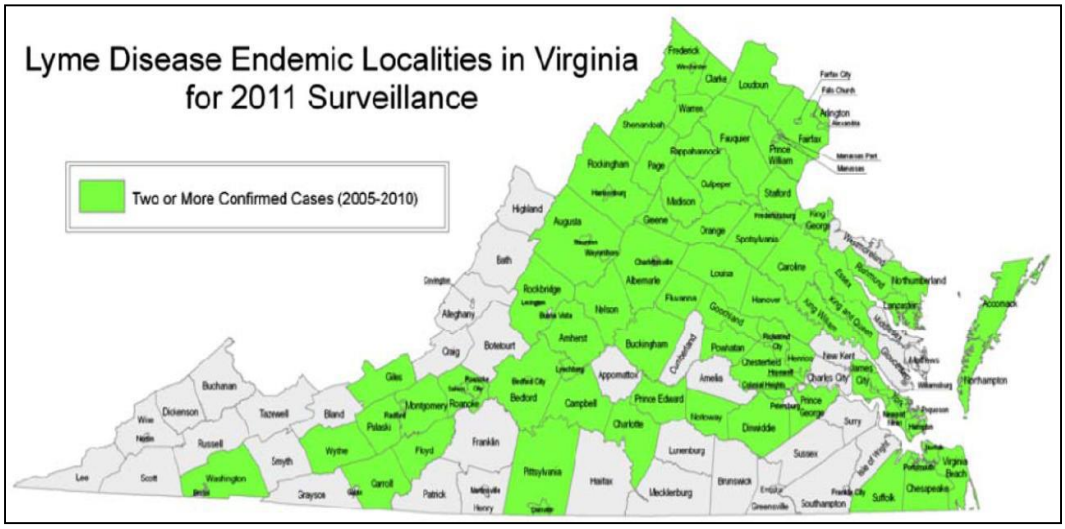

Figure 3: Endemic localities of Lyme disease in Virginia (Source: VDH, 2011)

In Virginia, Lyme disease leads the rate of increase of cases when compared to other tick borne diseases present within the state. Figure 4 depicts the increasing trend in Lyme disease amongst the population of Virginia, when compared to Rocky Mountain spotted fever and ehrlichiosis.

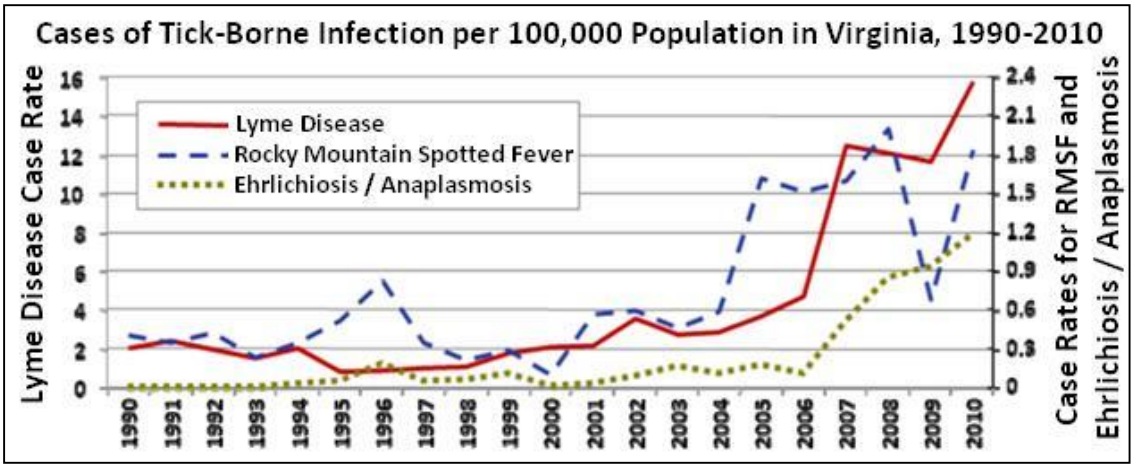

Figure 4: Graph showing the Tick-Borne infection cases in Virginia (Source: CDC, 2011)

1.3 PROBLEM STATEMENT

VDH (2011) mentions that the increase in the number of Lyme disease cases seen in figure 3 is not because of reporting changes, but it is due to the fact that Virginia is undergoing an expansion in the range of Lyme disease. With the emergence of Lyme disease in the US, especially across the northeastern US, it is important to predict the risk factors of the disease to plan ahead for mitigation measures. Thus, the research focuses on studying the potential future distribution of Lyme disease emergence under different land cover conditions using projected land cover datasets for Virginia. The thesis is a continuation of an on-going research project on the emergence of Lyme disease in Virginia.

The objective of this thesis is to predict future Lyme disease distributions in Virginia based on

different land cover scenarios. Studies that have attempted to predict future Lyme disease distributions have used variables such as environmental data (Guerra et al., 2002), climatic parameters (Brownstein et al., 2005), and host density (Schauber et al., 2005), but none has used projected land cover data to predict human incidence of the disease. Guerra et al. (2002) studied environmental factors such as soil, topography and vegetation that cause development of ticks, while Brownstein et al. (2005) studied effects of monthly vapor pressure, and minimum, maximum, and mean monthly temperature to generate an auto logistic model of Lyme disease risk in Canada. Subak et al. (2003), reported variability of Lyme disease incidence in seven northeastern US states, that was an outcome of weather variability in the summer drought indices. Along similar lines, Schauber et al. (2005) used mouse population data along with weather parameters to predict Lyme disease risk and found that the population of mouse was more strongly correlated to the probability of disease incidence occurrence than acorn production. Given the emergence of tick-borne diseases in Virginia (Figure 3), the objectives of the thesis are significant because they will help us understand as to how predicted land cover changes may impact the diffusion of the disease.

1.4 RESEARCH QUESTIONS

This thesis addresses both basic and applied research questions. The basic research question focuses on the gap in the literature that this thesis aims to address, specifically prediction of future Lyme disease incidence in Virginia using projected land cover change. The applied question addresses the needs of the end user, VDH.

Thus, the basic research question of the thesis is**:** How does future Lyme disease risk vary with respect to projected land cover change in Virginia? In order to answer the basic research question, this thesis aims to achieve the following objectives: 1) Model the spatio-temporal pattern of Lyme disease incidence against changes in the future land cover in Virginia; and 2) Determine high, medium and low risk areas of Lyme disease incidence in Virginia under A1B and A2 land cover projections.

Results generated to solve the basic research question will be used as input to address the following applied research question: How can VDH use the outcomes of the thesis to carry out preventive measures and public awareness? The applied research question will be addressed through the following objective: 1) Implement an interactive risk analysis tool, to help VDH view disease risk maps for each census tract in Virginia up to 2060.

REFERENCES:

Barbour AG, Fish D, 1993. The biological and social phenomenon of Lyme disease. Science 260, 1610–1616.

Brownstein, JS, Theodore RH, and Durland F, 2005. Effect of climate change on Lyme disease risk in North America. EcoHealth 2.1, 38-46.

CDC, 2011. Reported Lyme disease cases by state 2002-2011. Available at: http://www.cdc.gov/lyme/stats/chartstables/casesbyyear.html. Access date: November 7, 2012.

CDC, 2013. CDC provides estimates of Americans diagnosed with Lyme disease each year. Available at: http://www.cdc.gov/media/releases/2013/p0819-lyme-disease.html. Access date: September 2, 2013.

Franke J, Anke H, and Wolfram D, 2013. Exploring gaps in our knowledge on Lyme borreliosis spirochaetes–updates on complex heterogeneity, ecology, and pathogenicity. Ticks and tick-borne diseases 4.1, 11-25.

Guerra M, Walker E, Jones C, Paskewitz, S, Cortinas, MR, Stancil A, Kitron U, 2002. Predicting the Risk of Lyme Disease: Habitat Suitability for *Ixodes scapularis* in the North Central United States.

Gubler DJ, Reiter P, Ebi KL, Yap W, Nasci R, Patz JA, 2001. Climate variability and change in the United States: Potential impacts on vector- and rodent-borne diseases. Environ Health Perspect 109 (Suppl 2), 223–233.

Hu L, 2011. Patient information: Lyme disease symptoms and diagnosis (Beyond the Basics). Available at: http://www.uptodate.com/contents/lyme-disease-symptoms-and-diagnosis-beyond-the-basics#H1. Access date: November 7, 2012.

Keirans JE, Hutcheson HJ, Durden LA, Klompen JSH, 1996. *Ixodes scapularis* (Acari: Ixodidae): redescription of all active stages, distribution, hosts, geographical variation, and medical and veterinary importance. Journal of Medical Entomology 33, 297–318.

Schauber EM, Ostfeld RS, Andrew SE, 2005.What is the best predictor of annual Lyme disease incidence: Weather, mice, or acorns? Ecological Applications 15.2, 575-586.

Subak S, 2003. Effects of climate on variability in Lyme disease incidence in the northeastern United States. American Journal of Epidemiology 157.6, 531-538.

VDH, 2011. Lyme Disease. Available at: http://www.vdh.virginia.gov/epidemiology/DEE/Vectorborne/factsheets/lyme.html. Access date: November 7, 2012.

White DJ, Chang H-G, Benach JL, Bosler EM, Meldrum SC, Means RG, Debbie JG, Birkhead GS, Morse DL, 1991. The geographic spread and temporal increase of the Lyme disease epidemic. JAMA 266, 1230–1236.

CHAPTER 2: LITERATURE REVIEW

2.1 MEDICAL GEOGRAPHY AND LANDSCAPE EPIDEMIOLOGY

Medical geography, a specific branch of geography that is used to study disease epidemiology (Woodward, 2013) is one of the fastest growing fields of geography. Physical, environmental and economic parameters are responsible for introducing variations in disease distribution. Geographic analysis as a tool becomes useful to analyze disease patterns (Brown et al., 2012). Brown et al. (2002) identifies location, landscape and multilevel analyses as the three most important details to consider in a medical geography analysis. According to Woodward (2013), epidemiology is the study of disease distribution. Thus, geography plays an important role in epidemiology. This research is a part of the medical geography domain which uses Geographical Information System (GIS) and statistics to analyze the risk of Lyme disease projections for future IPCC scenarios. According to Mead and Emch (2010), medical geography is the branch of geography which combines environmental sciences and spatial analyses to study health problems and disease patterns. This research involves, medical geography, landscape epidemiology and GIS concepts to analyze and predict Lyme disease risk for future land cover projections.

Landscape epidemiology, a subset of medical geography, involves the identification of geographical areas where disease is transmitted. According to Meade and Emch (2010), the key environmental factors which influence the presence, development, activity, and longevity of pathogens, vectors, zoonotic reservoirs of infection, and their interactions with humans are elevation, temperature, rainfall, and humidity. Meade and Emch (2010) illustrate the significance

that landscape epidemiology holds in medical geography in tracing and understanding the patterns or possible patterns (for forecast modeling) of disease emergence and highlights the impact of landscape modification in the epidemiological research. More specifically, a central theme in landscape epidemiology, that "Landscape modification can create, or be used to prevent, the establishment of disease cycles" (Meade and Emch, 2010, p.111), is particularly relevant to this research as we consider how future landscape modifications may cause expansion or contraction of the Lyme disease risk area.

Spatial analysis of disease patterns combined with landscape epidemiology can be an effective tool to generate risk maps of diseases. Landscape epidemiology, therefore, helps in determining the geographical extent of disease emergence and will be applied as a theoretical framework in this study. This thesis is a continuation of the research study at Virginia Tech on the examination of the environmental factors affecting the emergence of Lyme disease. Thus, the current thesis would focus on predicting Lyme disease incidence risk in Virginia based on projected land cover data.

2.2 SPATIAL ANALYSIS AND DISEASE INCIDENCE

The World Health Organization (WHO, 2013) refers to disease incidence as the count of newly diagnosed cases in a geographical unit. Spatial analysis of disease incidence is especially important in medical geography when it comes to making projected models because environmental variables impact the occurrence of diseases and these environmental variables follow geographic patterns. Eisen et al. (2008) discuss the need for understanding the significance of spatial modeling of

human risk to vector-borne disease incidence with analyses of West Nile virus, tularemia and Lyme disease. The authors mention that it is important to estimate human risk of exposure to vectors and their related pathogens for effective surveillance and control measures such as vaccination, drug administration or education campaigns, and the use of sentinel sites to monitor vector reservoirs and in the identification of areas for the most effective use of pesticides.

Past research has predicted disease incidence using spatial analysis. Authors of Tran et al. (2002) study the relationship between land cover classes and malaria incidence. In this paper, land cover classes surrounding the villages were classified using the Maximum Likelihood technique and correlated using a univariate Poisson regression model. More specifically, Tran et al. (2002) compares population densities and disease incidence rates obtained from census data with those obtained from remote sensing (satellite) data. Disease risk was estimated as a function of the disease density per pixel of land cover data, wherein each pixel represented an urban area indicating the presence of people.

There have also been studies that generate spatio-temporal risk maps as opposed to simple risk maps to indicate how the predicted risk of diseases varies across time. Kulldorff (1997) developed software to determine the statistical significance of space-time disease clusters of high cancer rates. Kulldorff (1997) used zip codes as the unit of analysis within New York. Modules were developed in this software to predict incidence of cancer and used mortality rates, socio-demographic and lifestyle covariates for each zip-code level, as input variables. Wimberly et al. (2012) used satellite-derived land cover data to determine the environmental indices of the causative factors of malaria

from 2000 through 2011. These environmental variables were used as input to develop an ecological forecasting model of mosquito borne disease risks, in R statistical software. This thesis implements a similarly dynamic Lyme disease risk model which generates prediction maps for future projections for the years 2010 to 2060.

2.3 RELATIONSHIP BETWEEN ENVIRONMENTAL VARIABLES AND LYME DISEASE EMERGENCE

According to Khatichikian et al. (2012), anthropogenic changes to the environment are often reasons for increases in emergence of vector borne diseases in human population. Changes in the environmental factors increase the geographic spread of the disease causing pathogens and their vectors (Ogden et al., 2009). Spatial variability of risk and incidence rates of vector-borne and zoonotic diseases are becoming primary foci of ecological and epidemiological research (Ostfeld et al., 2005). Geographic Information Systems (GIS) are being used to study the spatial distributions of epidemiologic and environmental variables affecting diseases (Chrisman, 1989). GIS can used to study role of hosts and environment in determining distributions of ticks. For instance, in Kitron et al. (1991), GIS and spatial statistics were used to examine how white-tailed deer become infected by ticks with presence of woods and sandy soil in the state of Illinois. Kitron et al. (1992), found *I. Scapularis* to be less spread compared to its host, the white tailed deer. This paper compared spatial distributions of deer in tick-infested areas and non-infested areas to find that infestations were significantly clustered around tick sources despite the fact that the host distribution was several times wider than the tick distribution.

Remote sensing is an another emerging technique which provides a means to collect data at different spatial scales for the landscape features which may determine the distribution of hosts and ticks (Washino et al., 1994). Remote sensing, as used in medical geography and epidemiology, is based on the development of a sequence of steps which relates measures of radiation made by a sensor to measures of a disease and its vectors (Crombie et al., 1999). These steps include the collection of remotely sensed data to provide information on land cover and thus habitats of the causative factors of the disease (Innes et al., 1998), the geographical pattern of a disease and its vectors. Hence, remote sensing can be used to provide information on the spatial distribution of the vector-borne diseases.

Barbour et.al (1993) mention that the emergence of Lyme disease in the US is due to ecological changes in the northeastern and midwestern parts of the country in the late nineteenth century and throughout the twentieth century, like reforestation of previously abandoned farmlands. Reforestation resulted in the repopulation of the white-tailed deer (*Odocoileus virginianus*), the primary host for the adult stage of the tick *Ixodes scapularis*.

Ticks have a complicated lifecycle that relates to the enzootic cycle of Lyme disease, and subsequent human infection (Figure 1, Steere et al. 2004). *Ixodes scapularis* ticks need blood meals at several stages of their life cycle (which is two years in length generally) and feed once during each of the three stages of their life cycle. Larval ticks take one blood meal in late summer, nymphs feed in the following late spring and early summer, and adults feed during the fall, after which female tick lays eggs that hatch the next summer. In both larval and the nymphal stages, it is

important for ticks to feed on the same kind of host, a mammal, reptile or bird for disease transmission to occur (Steere et al, 2004), especially given that the life cycle of the bacteria depends on horizontal transmission from infected nymphs to rodents in spring and from infected rodents to larvae in the late summer. Disease gets transmitted through the bacteria to the human population when the ticks initially feed on an infected host in the larval stage, followed by feeding on a human in the nymphal or adult stage

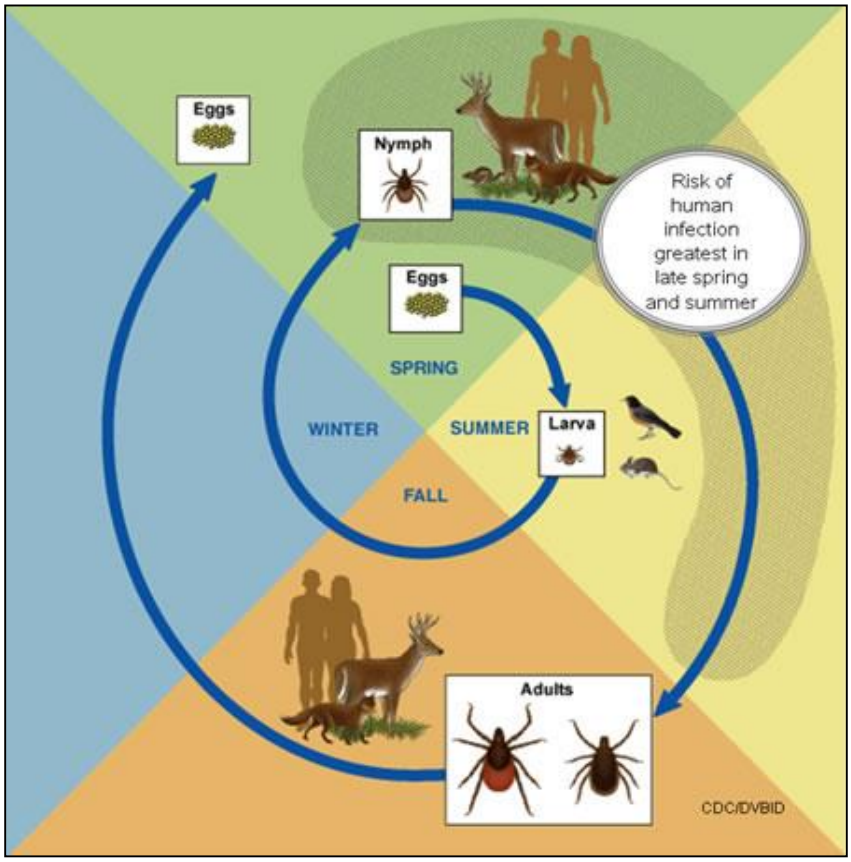

Figure 1: Lifecycle of Lyme disease (CDC, 2011)

The white footed mouse and white tailed deer are the most important hosts for ticks in the

northeastern part of the US (Jackson et al. 2006). The white tailed deer provides blood meal for the vector and also carries ticks from one place to another, even though it is not a reservoir for the bacteria. Barbour et al., (1993) attributes recent increase in tick population to increasing densities of their hosts, especially deer. However, it is difficult to decide if the tick population could be controlled by artificially reducing the host population. Van Buskirk et al. (1995) performed a computer simulation to test if human incidence of Lyme disease could be reduced by reducing populations of the ticks. The authors proved that Lyme disease risk was a function of the tick densities and the proportion of ticks carrying the spirochetes; results of their simulation model suggested that Lyme disease incidence could be reduced without affecting tick densities but by just encouraging diverse communities of the host vertebrates in land cover which involve high interaction of humans with the vertebrates. Although ticks could attach to humans in all three stages, it is the tick in the nymph stage that transmits the infection. With the presence of these vectors and hosts, Lyme disease incidence is related to the woodland and edge habitats of the landscape (Cromley et al. 1998). Thus, by studying the habitat and environmental factors which help the tick, mouse and deer interact, it is possible to identify places where humans could get infected by Lyme disease. This way, we can trace the factors that cause the emergence of human Lyme cases.

Morse (1995) mentions that the most important factors attributed to the emergence of Lyme disease are movement of tick carriers like humans and deer, land cover changes, urbanization and climate variability. Suburbanization, a specific type of land cover change, is associated with

emergence of Lyme disease (Jackson, 2006). The subsequent rise in human incidence of Lyme disease in the past decade in the Northeast is due to land cover changes through high developmental activities in the suburbs (Barbour et al., 1993). Maximum interaction between humans, vectors and hosts takes place in an individual residential unit. However, these units do not reflect the landscape epidemiology of Lyme disease when it comes to broader scales (Jackson, 2006). Therefore, Jackson et al. (2006) focused on endemic regions of Maryland and analyzed risk of Lyme disease in forested areas adjacent to major roads. It was found in this paper that fragmented forested areas were those associated with increased Lyme disease risk.

In a similar study, Glass et al. (1995) used GIS and a spatial epidemiological approach to generate Lyme disease risk maps and found a greater risk of Lyme disease along forest boundaries; areas of high development on the other hand had a lower risk. Land cover is an important variable in defining the human exposure level to Lyme disease because the density of infected ticks on a residential property is positively correlated with the proportion of vegetation cover on the property (Frank et.al 1998).

While most past work focus on Lyme disease risk or emergence at the regional scale, Brownstein et al. (2005) applied a slightly different approach. They focused on implementing an autologistic regression model to generate a risk map for the entire US by estimating the range of the distribution of *Ixodes scapularis* up to the year 2080. Figure 2 shows the county based distribution of *Ixodes scapularis* from 2005 to 2080 (Brownstein et al., 2005). The Brownstein et al. (2005) model uses climate change data and considers the effects of both greenhouse gas and

sulfate aerosols. Guerra et al. (2002) focused on micro and meso level of environmental data to generate risk and no-risk grids in Wisconsin and Illinois by developing a regression equation to determine the probability of occurrence of *Ixodes Scapularis*. The authors found that tick habitat was positively correlated with deciduous forest and barren lands while land cover types such as grasslands, coniferous forests and wetlands were found to be negatively correlated. Rizzoli et al. (2002) found that ticks were abundant in limestone type of substratum, vegetation cover with deciduous forests and high densities of the host. Tick abundance was also found at elevation below 1300 meters above sea level (Rizzoli et al., 2002).

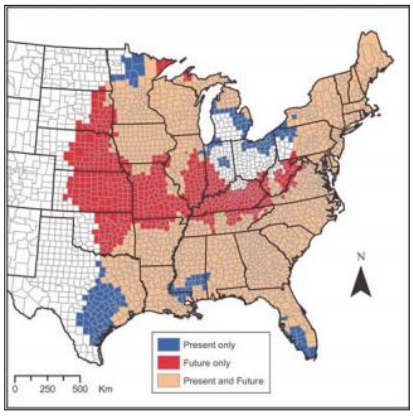

Figure 2: Change in *Ixodes scapularis* distribution from 2005 to 2080. (Brownstein et al., 2005)

Schauber et al. (2005) also examined climate variables and Lyme disease risk, and found a positive correlation between mean summer temperature and the rate of Lyme disease in the northeastern US. Schauber et al. (2005) also suggest a similar tie between lagged precipitations as

measured by the Palmer Hydrologic Drought Index (a dryness measurement index based on recent precipitation and temperature) and Lyme disease incidence in the Northeast.

In conclusion, previous research depicts land cover variables like suburbanization, forest type, and forest fragmentation as factors involved in the expansion of the host and tick populations leading to human Lyme disease incidence. These works also highlight the significance of using different scales of analysis in understanding the Lyme disease risk. However, none of the previous works uses projected land cover data, a noted important variable for current Lyme disease risk, to predict human incidence of Lyme disease risk. Thus, the current thesis aims to fill this gap and use projected land cover data to model Lyme disease risk in the state of Virginia.

REFERENCES:

Barbour AG, Fish D, 1993. The biological and social phenomenon of Lyme disease. Science 260, 1610–1616.

Brown T, Graham M, 2012. Geography and global health. The Geographical Journal 178.1, 13-17.

Brownstein JS, Holford TR, Fish D, 2005. Effect of climate change on Lyme disease risk in North America. EcoHealth 2.1, 38-46.

CDC, 2013. CDC provides estimates of Americans diagnosed with Lyme disease each year. Available at: http://www.cdc.gov/media/releases/2013/p0819-lyme-disease.html. Access date: September 2, 2013.

Chrisman NR, Cowen DJ, Fisher PF, Goodchild MF, Mark DM, 1989. Geographic information systems. Geography in America, 353-375.

Crombie MK, Gillies RR, Arvidson RE, Brookmeyer P, Weil GJ, Sultan M, Harb AM, 1999. An application of remotely derived climatological fields for risk assessment of vector-borne diseases: A spatial study of filariasis prevalence in the Nile Delta, Egypt. Photogrammetric engineering and remote sensing 65.12, 1401-1409.

Cromley EK, Cartter ML, Mrozinski RD, Ertel SH, 1998. Residential setting as a risk factor for Lyme disease in a hyperendemic region. American Journal of Epidemiology 147, 472-477.

Eisen, RJ, Eisen L, 2008. Spatial modeling of human risk of exposure to vector-borne pathogens based on epidemiological versus arthropod vector data. Journal of Medical Entomology 45.2, 181-192.

Frank DH, Fish D, Moy FH, 1998. Landscape features associated with Lyme disease risk in a suburban residential environment. Landscape Ecology 13, 27-36.

Franke J, Hildebrandt A, Dorn W, 2013. Exploring gaps in our knowledge on Lyme borreliosis spirochaetes–updates on complex heterogeneity, ecology, and pathogenicity. Ticks and tick-borne diseases 4.1, 11-25.

Glass GB, Schwartz J, Morgan III JM, Johnson DT, Noy P, Israel E, 1995. Environmental risk factors for Lyme disease identified with Geographic Information Systems. American Journal of Public Health 85.7, 944-948.

Gubler DJ, Reiter P, Ebi KL, Yap W, Nasci R, Patz JA, 2001. Climate variability and change in the United States: Potential impacts on vector- and rodent-borne diseases. Environ Health Perspect 109.2, 223–233.

Guerra M, Walker E, Jones C, Paskewitz, S, Cortinas, MR, Stancil A, Kitron U, 2002. Predicting the Risk of Lyme Disease: Habitat Suitability for *Ixodes scapularis* in the North Central United States.

Innes JL, Koch B, 1998. Forest biodiversity and its assewnent by remote sensing. Global Ecology and Biogeography Letters 7.397, 419.

Jackson LE, Hilborn ED, Thomas JC, 2006. Towards landscape design guidelines for reducing Lyme disease risk. International Journal of Epidemiology 35.2, 315-322.

Jackson LE, Levine JF, Hilborn ED, 2006. A comparison of analysis units for associating Lyme disease with forest-edge habitat. Community Ecology 7.2, 189-197.

Khatchikian CE, Prusinski M, Stone M, Backenson PB, Wang IN, Levy MZ, Brisson D, 2012. Geographical and environmental factors driving the increase in the Lyme disease vector *Ixodes scapularis*. Ecosphere 3.10, art85.

Kitron U, Bouseman JK, Jones CJ, 1991. Use of the ARC/INFO GIS to study the distribution of Lyme disease ticks in Illinois. Preventive Veterinary Medicine 11, 243-8.

Kitron U, Jones CJ, Bouseman JK, Nelson JA, Baumgartner DL, 1992. Spatial analysis of the distribution of *Ixodes dammini* (Acari: Ixodidae) on whitetailed deer in Ogle County, Illinois. Journal of Medical Entomology 29, 259–266.

Kulldorff M, 1997. A spatial scan statistic. Communications in Statistics: Theory and Methods 26,1481-1496.

Mausner JS, Mausner KS, Epidemiology B, 1985. An introductory text. Phila-delphia, PA: WB Saunders Co.

Mayer JD, 1983. The role of spatial analysis and geographic data in the detection of disease causation. Social Science & Medicine 17.16, 1213-1221.

Meade MS, Emch M, 2010. Medical Geography, 3rd edition. New York: The Guilford Press.

Morse SS, 1995. Factors in the emergence of infectious diseases. Emerging Infectious Diseases 1, 7-15.

Ogden, N. H., L. R. Lindsay, M. Morshed, P. N. Sockett, and H. Artsob. 2009. The emergence of Lyme disease in Canada. Canadian Medical Association Journal 180:1221–1224.

Ostfeld RS, Glass GE, Keesing F, 2005. Spatial epidemiology: an emerging (or re-emerging) discipline. Trends in Ecology & Evolution 20, 328–336.

Rizzoli A, Merler S, Furlanello C, Genchi C, 2002. Geographical information systems and bootstrap aggregation (bagging) of tree-based classifiers for Lyme disease risk prediction in Trentino, Italian Alps. Journal of medical entomology 39.3, 485-492.

Schauber EM, Ostfeld RS, Andrew SE, 2005.What is the best predictor of annual Lyme disease incidence: Weather, mice, or acorns? Ecological Applications 15.2, 575-586.

Steere AC, Coburn J, Glickstein L, 2004. The emergence of Lyme disease. Journal of Clinical Investigation 113.8, 1093-1101.

Tran A, Gardon J, Weber S, Polidori L, 2002. Mapping disease incidence in suburban areas using remotely sensed data. American journal of epidemiology 156.7, 662-668.

Van Buskirk J, Ostfeld RS, 1995. Controlling Lyme disease by modifying the density and species composition of tick hosts. Ecological Applications 5.4, 1133-1140.

Washino RK, Wood BL, 1994. Application of remote sensing to vector arthropod surveillance and control. The American journal of tropical medicine and hygiene 50, 34-44.

Woodward M, 2013. Epidemiology: study design and data analysis. CRC Press.

WHO, 2013. Immunization surveillance, assessment and monitoring, 2013. Available at: http://www.who.int/immunization_monitoring/data/data_subject/en/. Access date: December 21, 2013.

Wimberly MC, Chuang TW, Henebry GM, Liu Y, Midekisa A, Semuniguse P, Senay G, 2012. A computer system for forecasting malaria epidemic risk using remotely-sensed environmental data. Proceedings of the Sixth Biennial Conference of the International Environmental Modelling and Software Society.

MODELING THE IMPACT OF PROJECTED LAND COVER ON LYME DISEASE EMERGENCE

This manuscript has been prepared for submission to *Geospatial Health*

Jayashree Surendrababu[**] *(sjayas7@vt.edu)*, Korine N. Kolivras[1] *(korine @vt.edu)*, Stephen Prisley[2] *(prisley@vt.edu)*, Jie Li[3] (*jieli@vt.edu*), James B. Campbell[1] *(jayhawk@vt.edu)*

1. Department of Geography, Virginia Tech, 115 Major Williams Hall, Blacksburg, VA - 24060
2. Department of Forest Resources and Environmental Conservation,Virginia Tech, 313 Cheatham Hall, Blacksburg, VA, 24060
3. Department of Statistics, Virginia Tech, 406-A Hutcheson Hall, Blacksburg, VA, 24060

**Corresponding Author:
Jayashree Surendrababu
115 Major Williams Hall,
Virginia Polytechnic Institute and State University,
Blacksburg, VA 24061
Phone 540-449-9150
Email: sjayas7@vt.edu

Abstract

Lyme disease in the northeastern United States and, in particular the state of Virginia, has experienced a rapid increase in incidence and continued emergence spatially over the past decade. Previous research has identified land cover as significantly contributing to Lyme disease's current emergence pattern, but few studies have attempted to predict future Lyme disease distributions. This research predicts risk of Lyme disease in Virginia using projected land cover data created under two scenarios (A1B and A2) developed by the Intergovernmental Panel on Climate Change (IPCC) from 2010 to 2060. In this study, a Poisson point process is implemented with land cover parameters and demographic parameters as the spatial covariates. Lyme disease incidence data, obtained from the Virginia Department of Health, was used for model validation. Risk maps for both IPCC scenarios were generated from 2010 through 2060 by estimating incidence in each county in Virginia, and the land area and population affected were calculated. Results show that the A1B scenario has a 22.34% higher incidence estimate for the year 2060 than the A2 scenario. In the A1B scenario, 28% of the total area of Virginia is under high risk in 2010 which increases to 66% in 2060; in the A2 scenario, 22.4% is under high risk in 2010 leading to an increase in 2060 to 62.7%. In terms of population, 22% is in the high risk zone for A1B scenario in 2010 which increases to 42.9% in 2060; in the A2 scenario, 21% of the population is under high risk in 2010 which increases to 33% in 2060.

Keywords

GIS, Lyme disease, Land cover projections, IPCC scenarios

Introduction

Lyme disease, caused by the bacterium *Borrelia burgdorferi*, is the most frequently reported vector borne disease in the United States (Franke et al. 2013). The bacterium, which circulates in different reservoir populations such as the white-footed mouse *Peromyscus leucopus*, is transmitted by the blacklegged tick *Ixodes scapularis* (CDC, 2011) in the northeastern US, which sees the highest number of cases in the country. Within the US, Lyme disease cases are high in the Northeast and Mid-Atlantic States, the Midwest, and in northern California (Hu L, 2011), thereby making the disease a national public health

In 2012, with over 22,000 reported cases, Lyme disease was the seventh most common nationally notifiable disease (CDC, 2013). The Centers for Disease Control and Prevention (CDC 2013) reports that there are likely actually more than 300,000 Lyme disease cases annually when unreported cases are included in the total.

Those afflicted with Lyme disease can experience severe symptoms that degrade quality of life and can incur high costs for treatment. The most prominent initial symptoms are erythema migrans (bull's eye rash), fever, and muscle and joint aches. Later stages of untreated Lyme disease are marked by more severe symptoms such as neurological and cardiac problems as well as arthritis (VDH, 2011). Delayed diagnosis also adds to the disease burden by increasing the treatment costs.

Zhang et al. (2006) report that costs incurred in the year 2002 were around $203 million in the US for treating Lyme disease. Considerable research has examined links between environmental characteristics and Lyme disease incidence, given that biotic and abiotic factors impact geographic ranges of reservoirs and vectors (Estrada-Peña et al., 2014). Specifically, previous research indicates the importance of land cover in understanding the spatial variability of Lyme disease. Forest edges and woodlands are associated with increased probability of Lyme disease incidence (Cromley et al., 1998). Cromley et al., (1998) also found that highly developed areas have lower susceptibility for human infection than suburbanized and less developed areas. Frank et al. (1998) found that scrub lands were negatively correlated with tick density, while forests were positively correlated with tick density. In Allan et al. (2003), forest fragmentation was compared to tick infection to determine the relationship between ticks and forest fragment size, and it was found that the nymphal infection rate decreased in a significantly linear manner with an increase in forest fragmentation. Jackson et al. (2006) compared forest fragmentation to disease incidence in Maryland and found that regions with high levels of forest fragmentation had high incidence. The authors also found that low-density residential and suburban areas are correlated with high Lyme disease incidence.

Suburban areas enhance the environment for the white tailed deer (VDH, 2011). VDH (VDH, 2011) mentions that forested areas that are adjacent to residential settings, bring human population into greater contact with the tick vector's natural habitat. Glavanakov et al. (2001) found that

expansion in forest fragmentation and suburban regions increases the population of ticks and their hosts. Lastly, the results of Seukep et al. (under review) were quite unlike the studies by Jackson et al. (2006) and Allan et al. (2003) indicating that both forest fragments (<2ha) and the combined perimeter of the small forest fragments (<2 ha) within the tracts were insignificant.

Despite this body of research examining environmental characteristics and Lyme disease, only a few studies have attempted to predict future Lyme disease incidence based on projected environmental conditions. Anthropogenic changes to the environment affect the population and distribution of hosts, disease reservoirs, and vectors (Ogden et al., 2009), and several studies have predicted Lyme disease risk in different regions in North America. At the northern end of Lyme disease's range, Ogden et al. (2008) generated Lyme disease risk maps for Canada based on climatic projections using the Canadian Coupled Global Climate Model for the years 2020, 2050, and 2080. This study found that a warmer climate changed bird migration patterns, allowing for greater tick dispersion. Within the US, Brownstein et al. (2005) used climate change scenarios to generate prediction models for habit suitability of ticks. This study revealed that there was a significant expansion of infected ticks, with an increase in suitable habitat of 213% by the year 2080. Such attempts to predict disease risk could be extremely useful in disease surveillance and control measures.

Given the strength of the association between land cover and Lyme disease noted in previous studies, and the lack of research considering the potential future risk of disease, this research models future Lyme disease incidence based on projected land cover. In addition to addressing our

lack of understanding of Lyme disease's potential future distribution, this study has applied significance as well. An improved understanding of the disease's potential future emergence can assist public health officials and land managers in decision making. Thus, the objective of this study is to predict potential future Lyme disease distributions in Virginia, at the southern end of Lyme disease's current range, based on future land cover scenarios.

Materials and Methods

Study Area

Virginia is on the southern edge of Lyme disease's range expansion and has experienced a rapid increase in the number of cases over the past decade (Figure 1, VDH 2011). According to the Virginia Department of Health (VDH, 2012) a total of 1110 cases were reported in 2012, which represented an 8% increase from the cases reported in 2011 and a 9% increase from the five-year average from 2005 to 2009 of 1013.6 cases yearly. According to the Virginia Department of Health (VDH, 2011), the increase in Lyme disease incidence is not because of disease reporting changes; rather, it is because Virginia is on the front line of a southwestward expansion of Lyme disease's range. Previous work has examined links between environmental variability and Lyme disease's emergence within the state (Seukep et al., under review), and therefore, it is a natural extension to project the disease's distribution into the future under different land cover scenarios. Additionally, changes in land cover and land use patterns in Virginia make it an apt study area for the current research. In particular, over the past several decades, Virginia has experienced rapid suburbanization, a primary factor that increases risk of human Lyme disease.

Data

Three datasets were used to carry out this study: projected land cover data and projected demographic (population density and per capita income) data were applied to an existing model to predict future Lyme disease incidence, and Lyme disease incidence data of 2001 and 2006 forVirginia were used to validate the model. Projected land cover data, at a resolution of 250m, were obtained from the Land Carbon Project of the United States Geological Survey (Sleeter et al., 2012), a national assessment project aimed at understanding impacts of greenhouse gas fluxes on future land cover. Climate change scenarios defined by the Intergovernmental Panel on Climate Change (IPCC), based on demographic and technological changes, as well as changes in economic development, were incorporated to develop projected land cover data (Sleeter et al., 2012). The specific scenarios used to develop the datasets included the A1B, A2, B1 and B2 scenarios (Sohl et al., 2012).

Two specific scenarios, A1B and A2 (Figure 2), which represent contrasting scenarios in terms of economic development, technological advances, and demographic change (Sohl et al., 2012), were selected for this study in order to understand the possible risk zones of Lyme disease under different potential future situations. . The A1B scenario focuses on global economic and technological development. Higher end technologies increase the demand for energy thereby increasing the use of biofuels. A2 scenario also focuses on economic growth, but with a regional economic and technological development perspective. This scenario is marked by a higher population that eventually increases the use of natural resources and fossil fuels. Urban sprawl also

increases significantly in this scenario to support the population increase. Figures 3 and 4 illustrate the land cover changes in each scenario from 2010 to 2060 for developed, herbaceous, water and forest cover. Of all the input land cover variables included in this study, the developed category has the maximum change with an increase from 20.7% of total area of Virginia in 2010 to approximately 26% in 2060 in the A1B scenario. In the A2 scenario, this trend changes from about 26% in 2010 to 29.47% in 2060. Population will increase under both scenarios but the A2 scenario has a higher population growth compared to the A1B scenario. Thus, population growth is the driving force behind a higher growth rate in the developed land cover category. Projected demographic data, specifically projected personal income and population for each Virginia county for the A1B and A2 scenarios, has been obtained from the Forest 2010 RPA Assessment (Zarnoch et al., 2010).

Lyme disease incidence data by county for 2001 and 2006 were obtained from the VDH (VDH, 2011) and used for model validation. Since the basic geographic unit used for this model is a census tract, county wise population was downscaled to census tracts. For this purpose, census data for 2010 was obtained from the census bureau (US Census Bureau, 2013), then the percentage of population within all census tracts of a county were calculated. These percentages were then applied to downscaled population from county to census tract for projected demographic data from 2010 to 2060.

Methodology

In order to examine the potential distribution of Lyme disease in Virginia under projected land cover scenarios, we applied a previously developed model relating land cover and current incidence to the projected land cover data (Seukep et al., under review). This model was implemented using a spatial Poisson regression model to test the relationship between land cover and demographic variables and 2006-2010 Lyme disease incidence at the census tract level in Virginia. Ten variables, listed in Table 1, were statistically significantly associated with Lyme disease in Virginia, and it is this model that we now apply to projected land cover in order to predict future Lyme disease distributions. In particular, our model includes three land cover variables, percent developed, percent herbaceous cover, and percent water; four landscape metrics that address the boundary or edge between forested and herbaceous land, and herbaceous and developed land; and three demographic variables, population density, age, and income.

Several steps were required to prepare the input data for analysis and model application. The land cover data was reclassified into 4 broad categories - developed, forest, herbaceous and water using *ArcGIS Desktop 10* (ESRI, 2012). The edge metrics, Total Edge Contract Index (TECI) and Contrast Weighted Edge Index (CWED), were calculated in ArcGIS using the Patch Analyst extension in *ArcGIS Desktop 10* for the reclassified NLCD data. Finally, *Python 2.7* was used to automate the Poisson modeling processes and the geoprocessing steps.

Poisson point process

The significant variables (Table 1) were used to apply and implement a Poisson point process model (Kingman, 1992) for each decade from 2010 to 2060 for both IPCC scenarios. A Poisson point process follows a Poisson curve is used when the input is count data and when the model varies at a specified time interval, in this case every 10 years. Since the interval is constant at 10 years, this method is called a homogenous Poisson point process. A homogenous Poisson point process is given by the equation shown below:

$$\rho = P\,[(N(t+\tau) - N(t)) = i] = (e^{-\lambda\tau}(\lambda\tau)^i)/i! \qquad i = 0, 1, \ldots,$$

where $N(t + \tau) - N(t) = i$ = Total number of risk events in the given time interval $(t, t + \tau)$

Parameter λ is estimated with Maximum Likelihood Estimation through Python 2.7 using NLCD land cover data and Lyme case data of 1998 and 2010. *Statsmodel* and *Pandas* are the libraries that were used to implement the statistical functions and the data dictionaries in Python. The final ρ value was used to determine the number of cases per census tract which in turn was aggregated for each county boundary. The process was repeated for both scenarios for the decades 2010 to 2060, and individual risk maps for Virginia were generated. For the final visualization, estimated cases were aggregated for each county to generate county wise risk values which in turn were standardized for 100,000 people to determine the disease incidence per county.

The implemented model was applied to the National Landcover Dataset (NLCD) data for 2001,

2006, and 2010 for model validation. Calculated case data were compared to actual case data for 2001 and 2006, and the Root Mean Squared Error (RMSE) was calculated.

Results

Count Estimates

From 2010 to 2060, 2790 average annual cases were estimated in the A1B scenario and 2298 cases were estimated for the A2 scenario as shown in figure 7. Overall, the A1B scenario witnessed an 332% increase in cases from 2010 to 2060 along with a 57% population increase, while the A2 scenario was marked by a 227% increase in cases and a 74% increase in population.

Overall area and number of people affected in each risk category and their respective average counts under each scenario is depicted in Table 2. Twenty-eight percent of the total area in Virginia is under high risk of Lyme disease under the A1B scenario for the year 2010, while under the A2 scenario it is 22.4%; this trend increases to 66% and 62.7% of area under high risk for 2060 for A1B and A2 scenarios respectively.

In terms of population, for A1B scenario 1,725,235 people live within the high risk category in 2010, accounting for 22% of the total population, while 5,268,365 live within the high risk area in 2060. For A2 scenario, the population falling within the high risk zone in 2010 is 1,709,006 people, constituting approximately 21% of the total population, while around 4,674,738 of them

are under high risk in 2060, constituting 33% of the projected population. Total population for the year 2010 is different between the A1B and A2 scenarios at 7806495 and 8138124 respectively; this discrepancy occurs because projected population data was used to calculate population estimates for 2010 instead of the actual census data collected in that year.

Risk maps were generated for A1B and A2 scenarios based on the incidence per county across Virginia as shown in figure 8 from 2010 to 2060. In both the scenarios, the progression of higher incidence rates are from the northern counties towards the southern counties along the western boundary of the state. However, A1B scenarios have more counties with incidence rates from 25 to 50 compared to A2 scenario which have more counties with higher incidence rates from 50 to 245. Although the total count estimated for A1B is higher than that of A2 (Figure 8), A1B has fewer counties with the highest incidence rates of 200 to 250. In the A2 scenario, in the year 2010, high incidence rates are mostly found in the northern parts of Virginia where there are lower levels of developed, most of the area being herbaceous cover.

Model Validation

The point process model was implemented on NLCD datasets for the years 2001 and 2006 and the counts estimated were compared against the actual count data obtained from the VDH for those two years. The model has Root Mean Square Deviation (RMSE) values of 0.99 and 1.44 number of cases respectively for the years 2001 and 2006 when implemented for the NLCD dataset as shown in table 3, and figures 5 and 6.

Discussion

Risk scenarios

This study has predicted potential future distributions of Lyme disease in Virginia based on projected land cover under the A1B and A2 scenarios from 2010 to 2060. As illustrated in the data section, figures 3 and 4 compare the land cover variables used in this study with the predicted Lyme disease rates from 2010 to 2060. It can be observed that in the A1B scenario, disease incidence rate follows the rate of change of edge interspersion of Herbaceous-Developed and Forest-Herbaceous. However, in A2 scenario the increasing rate aligns more with the developed land cover variable. Although, developed area was found to be negatively correlated with disease incidence in past studies, the A2 scenario witnesses an increasing rate in its forest cover and forest-herbaceous interspersion that could be attributed to the increase in disease rates as well.

Count estimated for A1B scenario is higher than that of A2 but A1B has fewer counties with the highest incidence rates of 200 to 250 because A1B is more environmentally stable than A2 scenario. Besides, highly populated counties in the northern part of the state in 2010 are classified with low disease incidence which is in accordance with Seukep's findings of the population being negatively correlated to Lyme disease incidence (Seukep et al., under review). Apart from human population being negatively correlated to Lyme disease incidence, the other main reason that these regions fall under high risk zones could be because these regions are marked by higher levels of interspersions between forest and herbaceous areas. In the A2 scenario, higher incidence zones spread toward the southwestern part of the state by 2060. This region of the state is marked by

lower levels of population growth compared to other parts of Virginia from 2010 to 2060.

Limitations and directions for future research

The main limitations of this study relate to the resolution of the land cover data and the generalizability of the results. Data used to develop the initial model was 30m land cover (NLCD) data; however the model had to be rescaled because the projected land cover data was only available at 250m resolution. Higher resolution land cover data would help us obtain higher accuracies while calculating the area occupied by each land cover dataset. Thus, future research can focus on higher resolution land cover data for fine scale analysis at either the county level or census tract level.

Another limitation with the land cover data used in this study is that the land cover classification could be missing certain important true risk causative factors such as suburbanization and the wild land (which includes urban forested lands) and urban interface. Since the model implemented uses projected demographic data, extrapolating this study and the model to other regions could be relatively weak; however, the general methods could be applied to other areas after the model is rebuilt and implemented separately in a new study area. Also, projected population estimates used for modeling in this study adds additional uncertainty to the implemented model. Besides, this study gives us an idea of how the disease could be expected to spread under the A1B and A2 climate change scenarios but it does not account for prior disease in the study area for a given year. Thus, for public health and surveillance measures, a more robust way of monitoring the disease is needed.

Conclusion

Lyme disease is a vector borne disease spreading across the northern hemisphere and identifying its risk through this study could increase disease control measures and help public health officials prepare for continued emergence. Despite Lyme disease's recognition as the most commonly reported vector borne disease in the United States and clear links between land cover and the disease's continued emergence, no other study has attempted to project future Lyme disease distributions based on potential land cover scenarios. This study contributes to the overall body of research on spatio-temporal disease prediction and more specifically, that of Lyme disease. The final risk maps show counties where Lyme disease incidence increase could potentially occur under two possible future IPCC scenarios. This model can be used as a base by public health officials to carry out improved surveillance and public education, and potentially enact control measures of ticks and hosts in the risk areas determined through this research. Additionally, public health officials can allocate optimal resources for carrying out such activities in each county depending on the risk classification of generated through this study.

References

Allan BF, Keesing F, Ostfeld RS, 2003. Effect of forest fragmentation on Lyme disease risk. Conservation Biology 17.1, 267-272.

Bacon RM, Kugeler, KJ, Mead PS, 2008. Surveillance for Lyme disease--United States, 1992-2006. MMWR Surveill Summ 57, 1-9.

Brownstein JS, Holford TR, Fish D, 2005. Effect of climate change on Lyme disease risk in North America. EcoHealth 2.1, 38-46.

CDC, 2013. CDC provides estimates of Americans diagnosed with Lyme disease each year. Available at: http://www.cdc.gov/media/releases/2013/p0819-lyme-disease.html. Access date: September 2, 2013.

Cromley EK, Cartter ML, Mrozinski RD, Ertel SH, 1998. Residential setting as a risk factor for Lyme disease in a hyperendemic region. American Journal of Epidemiology 147, 472-477.

Seukep LD, Kolivras KN, Hong Y, Li J, Prisley S, Campbell JB, Dymond RL, 2013. Lyme Disease Emergence in Virginia: An Examination of the Demographic and Environmental Variables Correlated to the Spatial Pattern of Disease Incidence. Manuscript submitted for publication to the International Journal of Health Geographics.

ESRI, 2012. ArcGIS. Available at: http://www.esri.com/software/arcgis. Access date: June 25, 2013.

Estrada-Peña A, Farkas R, Jaenson TG, Koenen F , Madder M, Pascucci I, Jongejan F, 2013. Association of environmental traits with the geographic ranges of ticks (Acari: Ixodidae) of medical and veterinary importance in the western Palearctic. A digital data set. Experimental and Applied Acarology 59.3, 351-366.

Estrada-Peña A, Ostfeld RS, Peterson AT, Poulin R, De la Fuente J, 2014. Effects of environmental change on zoonotic disease risk: an ecological primer. Trends in parasitology 30.4, 205-214.

Frank DH, Fish D, Moy FH, 1998. Landscape features associated with Lyme disease risk in a suburban residential environment. Landscape Ecology 13, 27-36.

Franke J, Anke H, Wolfram D, 2013. Exploring gaps in our knowledge on Lyme borreliosis spirochaetes–updates on complex heterogeneity, ecology, and pathogenicity. Ticks and tick-borne diseases 4.1, 11-25.

Glavanakov S, White DJ, Caraco T, Lapenis A, Robinson GR, Szymanski BK, Maniatty WA, 2001. Lyme disease in New York State: Spatial pattern at a regional scale. American Journal of Tropical Medicine and Hygiene 65.5, 538-545.

Hu L, 2011. Patient information: Lyme disease symptoms and diagnosis (Beyond the Basics). Available at: http://www.uptodate.com/contents/lyme-disease-symptoms-and-diagnosis-beyond-the-basics#H1. Access date: November 7, 2012.

Jackson LE, Hilborn ED, Thomas JC, 2006. Towards landscape design guidelines for reducing Lyme disease risk. International Journal of Epidemiology 35.2, 315-322.

Jones CJ, Kitron UD, 2000. Populations of *Ixodes scapularis* (Acari: Ixodidae) are modulated by drought at a Lyme disease focus in Illinois. Journal of Medical Entomology 37**,** 408-15.

Keirans JE, Hutcheson HJ, Durden LA, Klompen JSH, 1996. *Ixodes scapularis* (Acari: Ixodidae): redescription of all active stages, distribution, hosts, geographical variation, and medical and veterinary importance. Journal of Medical Entomology 33, 297–318.

Killilea ME, Swei A, Lane RS, Briggs CJ, Ostfeld RS, 2008. Spatial dynamics of Lyme disease: A review. EcoHealth 5.2, 167-195.

Kingman JFC, 1992. Poisson processes Vol 3. Oxford university press.

McGarigal K, Cushman SA, Ene E, 2012. FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps.

Meade MS, Emch M, 2012. Medical geography. Guilford Press.

Ogden NH, Lindsay LR, Beauchamp G, Charron D, Maarouf A, O'Callaghan CJ, Waltner-Toews D, Barker IK, 2004. Investigation of relationships between temperature and developmental rates of tick *Ixodes scapularis* (Acari: Ixodidae) in the laboratory and field. Journal of Medical Entomology 41.4, 622-633.

Ogden NH, St-Onge L, Barker IK, Brazeau S, Bigras-Poulin M, Charron DF, Francis CM, Heagy A, Lindsay LR, Maarouf A, Michel P, Milord F, O'Callaghan CJ, Trudel L, Thompson RA, 2008. Risk maps for range expansion of the Lyme disease vector, *Ixodes scapularis*, in Canada now and with climate change. International Journal of Health Geographics 7**,** 24.

Ogden NH, Lindsay LR, Morshed M, Sockett PN, Artsob H, 2009. The emergence of Lyme disease in Canada. Canadian Medical Association Journal 180, 1221–1224.

Schauber EM, Ostfeld RS, Evans Jr AS, 2005. What is the best predictor of annual Lyme disease incidence: Weather, mice, or acorns? Ecological Applications 15.2, 575-586.

Sleeter BM., Sohl TL, Bouchard MA, Reker RR, Soulard CE, Acevedo W, Griffith GE, Sleeter RR, Auch RF, Sayler KL, Prisley S, Zhu Z, 2012. Scenarios of land use and land cover change in the conterminous United States: utilizing the special report on emission scenarios at ecoregional scales. Global Environmental Change 22. 4, 896-914.

Sohl TL, Sleeter BM, Zhu Z, Sayler KL, Bennett S, Bouchard M, Reker R, Hawbaker T, Wein A, Liu S, Kanengieter R, Acevedo W, 2012. A land-use and land-cover modeling strategy to support a national assessment. Applied Geography 34, 111-124.

Subak S, 2003. Effects of climate on variability in Lyme disease incidence in the northeastern United States. American Journal of Epidemiology 157.6, 531-538.

US Census Bureau, 2013. 2010 Census Data Products: United States. Available at: http://www.census.gov/population/www/cen2010/glance/. Access Date: July, 2013.

VDH, 2011. Lyme Disease. Available at: http://www.vdh.virginia.gov/Epidemiology/Surveillance/SurveillanceData/AnnualReports/Reports/Diseases%202011/Lyme_Disease2011.pdf Access date: November 7, 2012.

VDH, 2012. Lyme Disease Case Reporting. Available at: http://www.vdh.virginia.gov/Epidemiology/Surveillance/SurveillanceData/AnnualReports/Reports/Diseases%202012/Lyme_Disease2012.pdf Access date: April 1, 2014.

Wormser GP, Dattwyler RJ, Shapiro ED, Halperin JJ, Steere AC, Klempner MS, Krause PJ, Bakken JS, Strle F, Stanek G, Bockenstedt L, Fish D, Dumler JS, Nadelman RB, 2006. The clinical assessment, treatment and prevention of Lyme disease, human granulocytic anaplasmosis and babesiosis: Clinical practice guidelines. Clinical Infectious Diseases, 1089-1134.

Cordell HK, Betz CJ, Langner L, 2010. Projecting county-level populations under three future scenarios: a technical document supporting the Forest Service 2010 RPA Assessment. US Department of Agriculture Forest Service, Southern Research Station, 8.

Zhang X, Meltzer MI, Peña CA, Hopkins AB, 2006. Economic Impact of Lyme disease. Emerging infectious diseases 12, 653-660.

Table 1: List of Significant variables (Seukep et al., under review)

LANDCOVER/LANDSCAPE METRIC VARIABLE	ESTIMATE	95% CI LOWER	95% CI UPPER
Percentage Developed	**-0.115**	**-0.144**	**-0.086**
Percentage Herbaceous/scrub	**0.076**	**0.017**	**0.137**
Percentage Water	**0.137**	**0.053**	**0.211**
CWED[1] Forest-Scrub	**0.048**	**0.025**	**0.073**
TECI[2] Forest-Scrub	**0.053**	**0.033**	**0.072**
CWED Scrub-Developed	**-0.048**	**-0.078**	**-0.015**
TECI Scrub-Developed	**-0.047**	**-0.079**	**-0.013**
Population Density	**-0.086**	**-0.161**	**-0.017**
Mean Income	**0.184**	**0.108**	**0.26**

Table 2: Total Population and Area of Virginia under Lyme disease risk in A1B and A2 scenario from 2010 to 2060

	A1B						A2				
	Risk Category	Population	Population (%)	Total Population	Area (km^2)	Area (%)	Population	Population (%)	Total Population Affected	Area (km^2)	Area (%)
2010	High	1725235.4	22.1	7806495	31019.8	28	1709006	21	8138124	24815.84	22.4
	Med	1600331.5	20.5	7806495	21270.72	19.2	1928735	23.7	8138124	22157	20
	Low	2591756.5	33.2	7806495	40547.31	36.6	3182006	39.1	8138124	41101.235	37.1
	No risk	1889171.9	24.2	7806495	17947.17	16.2	1318376	16.2	8138124	22710.925	20.5
2020	High	1524435.3	17.4	8761123	17947.17	16.2	1928007	20.9	9224915	27917.82	25.2
	Med	1209034.9	13.8	8761123	15953.04	14.4	1162339	12.6	9224915	22932.495	20.7
	Low	3670910.4	41.9	8761123	53730.725	48.5	2075606	22.5	9224915	33457.07	30.2
	No risk	2356742	26.9	8761123	23154.065	20.9	4058963	44	9224915	26477.615	23.9
2030	High	2247264.7	23.1	9728419	30133.52	27.2	3699498	35.7	10362738	32681.575	29.5
	Med	1021484	10.5	9728419	13294.2	12	1160627	11.2	10362738	17947.17	16.2
	Low	3755169.6	38.6	9728419	52068.95	47	3305713	31.9	10362738	37445.33	33.8
	No risk	2704500.4	27.8	9728419	15288.33	13.8	2196900	21.2	10362738	22710.925	20.5
2040	High	1193731.5	11.2	10658317	29801.165	26.9	2883869	25	11535475	33678.64	30.4
	Med	3122887	29.3	10658317	13737.34	12.4	4625725	40.1	11535475	21159.935	19.1
	Low	1385581.3	13	10658317	44424.785	40.1	2514733	21.8	11535475	31795.295	28.7
	No risk	4956117.6	46.5	10658317	22821.71	20.6	1511147	13.1	11535475	24151.13	21.8
2050	High	2404835.4	20.9	11506389	48413.045	43.7	1930244	15.1	12783073	56721.92	51.2
	Med	1208170.9	10.5	11506389	13404.985	12.1	2173122	17	12783073	19276.59	17.4
	Low	5292939.1	46	11506389	27142.325	24.5	3655959	28.6	12783073	15509.9	14
	No risk	2600444	22.6	11506389	21824.645	19.7	5023748	39.3	12783073	19276.59	17.4
2060	High	5268365.2	42.9	12280572	73228.885	66.1	4674738	33	14165873	69462.195	62.7
	Med	2161380.6	17.6	12280572	14734.405	13.3	2124881	15	14165873	14180.48	12.8
	Low	2456114.3	20	12280572	13848.125	12.5	5666349	40	14165873	15731.47	14.2
	No risk	2394711.4	19.5	12280572	8973.585	8.1	1699905	12	14165873	11410.855	10.3

Table 3: Model validation for 2001, 2006 data

Year	R^2	RMSE
2001	0.974	0.99
2006	0.991	1.44

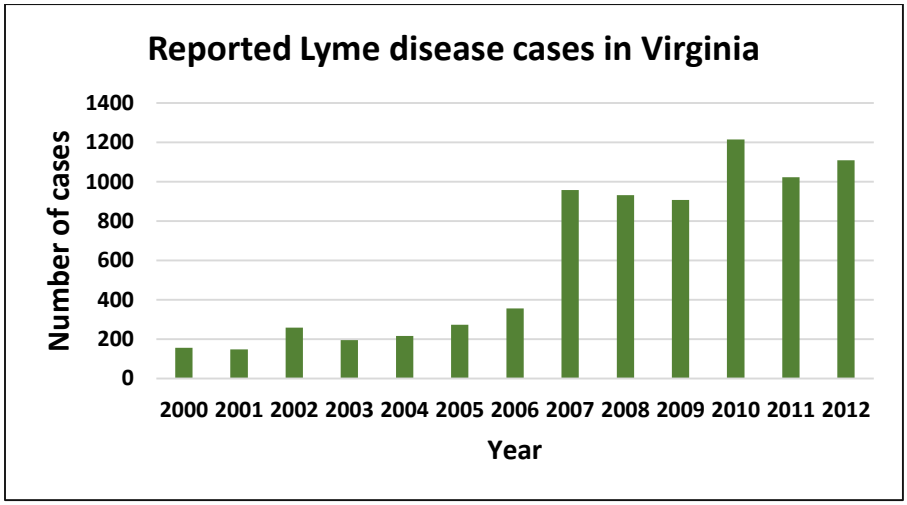

Figure 1: Lyme disease Cases per Year in Virginia (VDH, 2012)

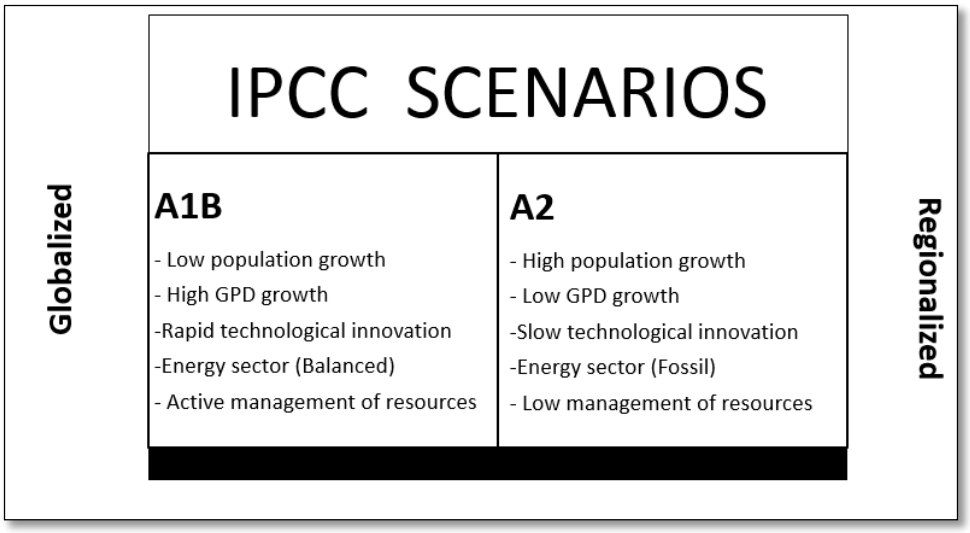

Figure 2: Comparison of A1B and A2 scenarios

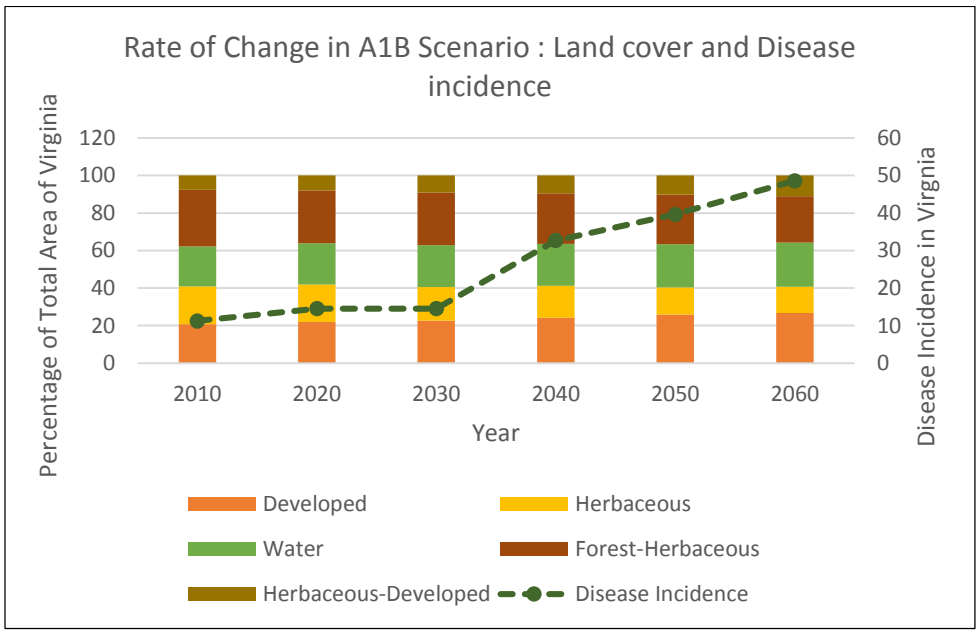

Figure 3: Rate of Change in A1B Scenario: Land cover and Disease incidence

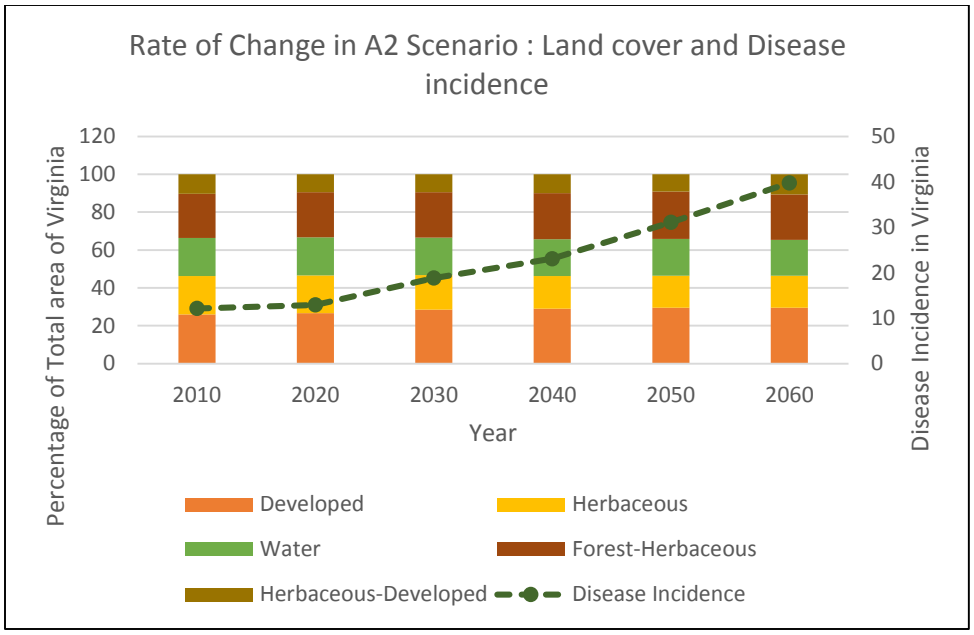

Figure 4: Rate of Change in A2 Scenario: Land cover and Disease incidence

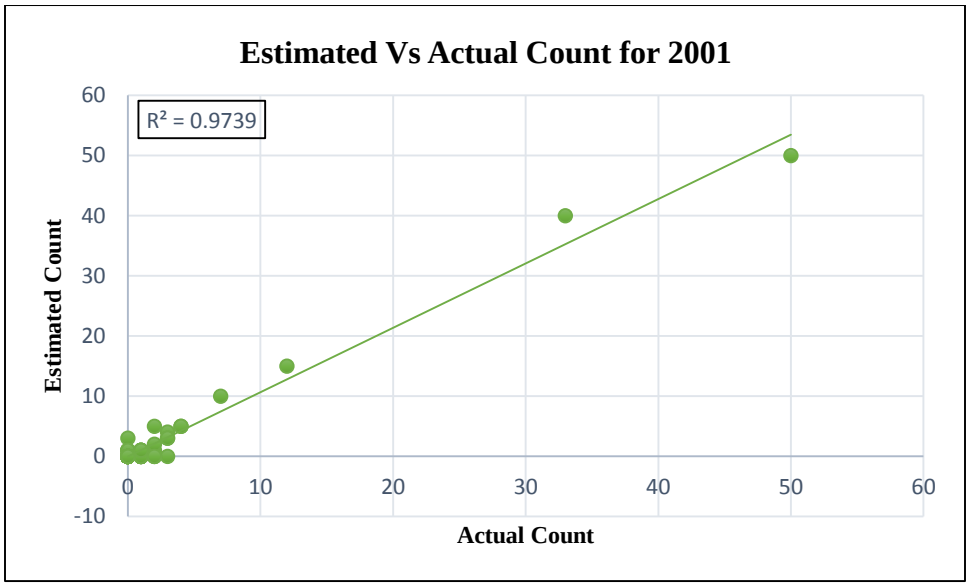

Figure 5: Scatter Plots of actual and estimated count data for 2001

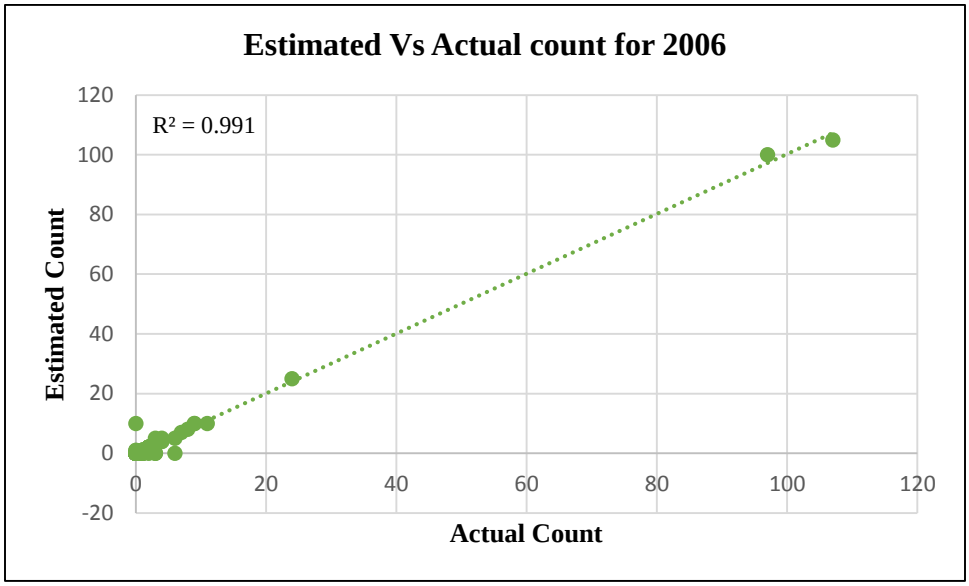

Figure 6: Scatter Plots of actual and estimated count data for 2006

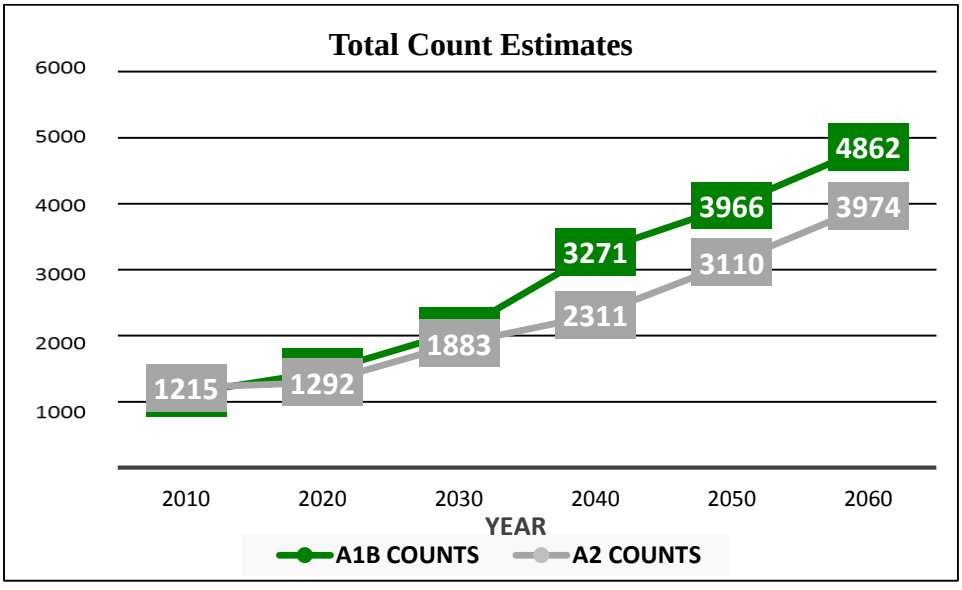

Figure 7: Year wise counts estimated for both the IPCC scenarios

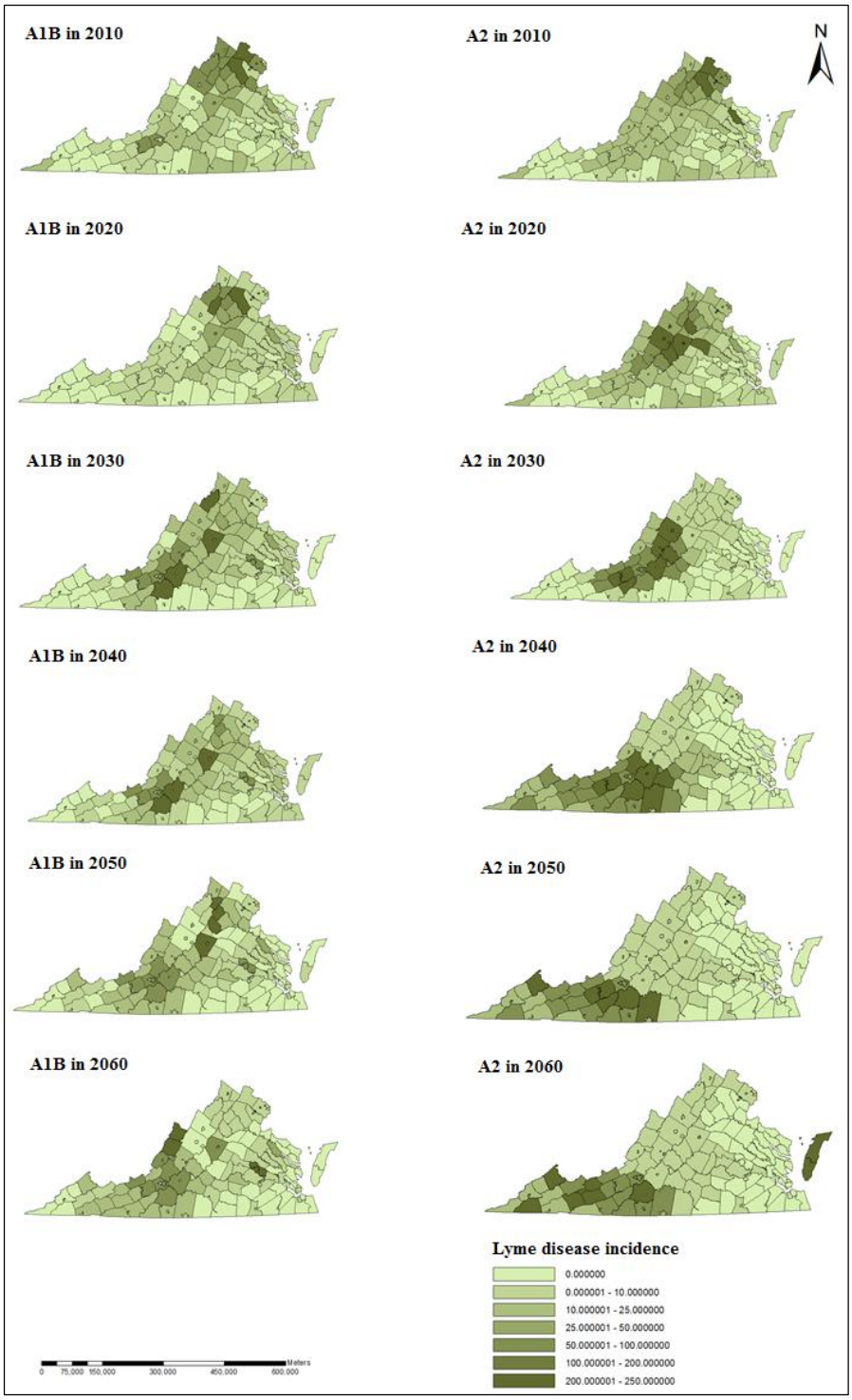

Figure 8: County wise aggregated Risk Maps for A1B and A2 scenario from 2010 to 2060