Temperature and Precipitation as Possible Environmental Indicators for Lyme Disease Incidence in the Northeastern United States, and Surveillance Data Limitations

A Senior Honors Thesis in Community Health

Aislinn Cook

Committee Chair: Dr. David Gute

Readers: Dr. Karen Claire Kosinski

Dr. Shalini Tendulkar

Tufts University, 2016

Acknowledgements

I would like to thank my committee, Dr. David Gute, Dr. Karen Claire Kosinski and Dr. Shalini Tendulkar for their unwavering support and guidance in writing this thesis. I would not have been able to overcome the bumps along the way without your availability and encouragement.

I am grateful to the Data Lab graduate assistants for answering my many questions and guiding me in my spatial and statistical analyses.

In addition, thank you to Dr. Jennifer Allen for supporting me two years ago when I first started studying Lyme disease and for encouraging me to pursue this topic for my senior honors thesis.

Table of Figures

Abstract

Lyme disease is the largest vector borne disease in the United States with approximately 30,000 cases reported each year. These cases are concentrated in the Northeast and Midwest U.S. According to several studies, these reported cases are a gross underestimation of the true Lyme disease burden, estimated at closer to 300,000 cases per year. Contributing to this underestimation are major flaws in the surveillance system in place. Given the large case burden, it would be beneficial to public health officials if yearly Lyme disease incidence could be predicted. Many different environmental variables contribute to the occurrence of Lyme disease. This study examines weather as one such variable, testing summer and winter temperature and precipitation for statistically significant associations to Lyme disease incidence with a two-year, one-year and same-year time lag. Weather data from the National Climatic Data Center were processed in ArcMap to extract temperature and precipitation data by county for 12 states in the Northeast U.S. Spatial regression models were used to determine whether 18 weather variables were associated with Lyme disease incidence. All results except winter temperature 2013 with Lyme disease incidence 2013 and 2014 were statistically significant at $p<0.05$. Based on the regression coefficients, summer temperature in all years was positively associated with the largest percentage point (approximately 4.0) increase in Lyme disease incidence. Winter temperature 2012 was also positively associated with an increase in Lyme disease incidence in 2013 and 2014. Summer and winter precipitation were positively associated with Lyme disease incidence; however, the regression coefficients indicate a less than one percentage point increase in Lyme disease incidence associated with these variables. The results of this study indicate that further investigation of summer and winter temperature as two of several environmental variables associated with Lyme disease is warranted.

Background

Lyme disease burden in the United States

Lyme disease is the most common vector-borne disease in the United States (Schulze et al., 2009). It is caused by the spirochete bacteria, *Borrelia burgdorferi*, and transmitted on the east coast by *Ixodes scapularis,* commonly known as the black-legged tick (Guerra et al., 2002). According to the U.S. Centers for Disease Control and Prevention (CDC), 14 states in the United States account for 96% of all reported Lyme disease cases ("Data and Statistics | Lyme Disease | CDC," 2015). These states are Wisconsin and Minnesota in the Midwest and Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont and Virginia on the east coast.

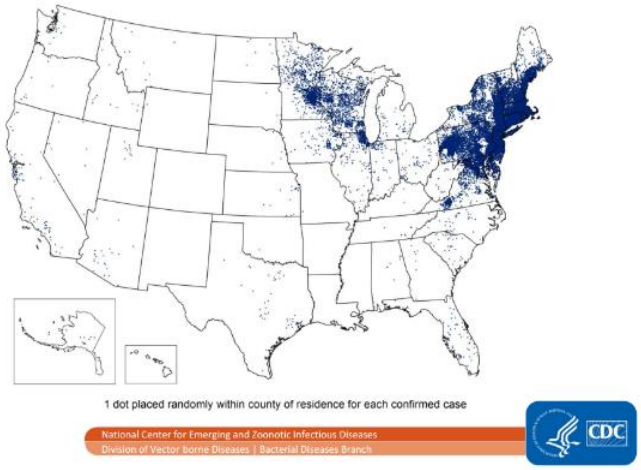

Figure 1. Reported Lyme disease cases 2013 (CDC).

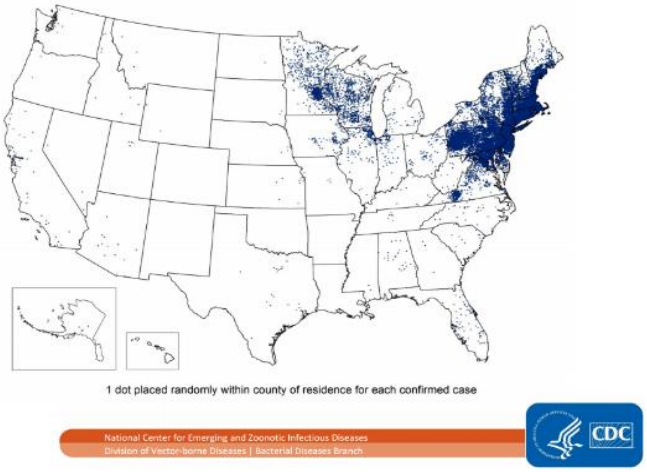

Figure 2. Reported Lyme disease cases 2014 (CDC).

Of these 14 states, the states in the Northeastern United States present a much higher burden of Lyme disease cases than those in the Midwest (Figure 1 and Figure 2). In 2013 Vermont, New Hampshire and Maine had the highest incidence rates per 100,000 people at

107.6, 100.0 and 84.8 respectively ("Lyme disease data tables | Lyme Disease | CDC," n.d.). Minnesota and Wisconsin had incidence rate of 26.4 and 25.2 respectively in 2013. Incidence rates for 2014 were lower than 2013, however Maine and Vermont still had the highest incidence rates of 87.9 and 70.5 respectively while Minnesota and Wisconsin had incidence rates of 16.4 and 17.2 respectively (Table 1) (CDC, Lyme Disease Incidence Rates by State, 2015).

Table 1. Lyme disease incidence rates by state organized by 2013 incidence from lowest to highest. The states summarized account for 96% of Lyme disease cases in the U.S. Incidence is the number of confirmed cases per 100,000 people ("Lyme disease incidence rates by state, 2005-2014 | Lyme Disease | CDC," 2015)

State	2013 Incidence	2014 Incidence
Virginia	11.2	11.7
Maryland	13.5	16.0
New York	17.9	14.4
Wisconsin	25.2	17.2
Minnesota	26.4	16.4
New Jersey	31.3	29.0
Pennsylvania	39.0	50.6
Rhode Island	42.2	54.0
Delaware	43.2	36.4
Massachusetts	57.0	54.1
Connecticut	58.7	47.8
Maine	84.8	87.9
New Hampshire	100.0	46.9
Vermont	107.6	70.5

The United States had 33,461 cases reported in 2014 (CDC, 2015) including confirmed and probable cases . Of the 33,461 cases reported in 2014, New England had 11,292, the Mid-

Atlantic had 14,509 in total and the South Atlantic had 3,678 cases and Wisconsin and Minnesota had 2,777 total cases combined. Overall, the South/Mid-Atlantic to Northeastern United States accounted for approximately 88% of the total reported cases. Case totals for the fourteen states that account for 96% of the Lyme disease burden in the United States are summarized in Table 2.

Table 2. Reported cases of Lyme disease in 2014 by region and state. The fourteen states summarized account for 96% of Lyme disease cases in the United States. ("Notice to Readers: Final 2014 Reports of Nationally Notifiable Infectious Diseases, 2014)

Region	State	Total Reported Cases (Confirmed and Probable)
New England	Vermont	599
	New Hampshire	724
	Rhode Island	904
	Maine	1401
	Connecticut	2360
	Massachusetts	5304
Mid-Atlantic	Delaware	417
	New Jersey	3286
	New York (Upstate and City combined)	3736
	Pennsylania	7487
South-Atlantic	Virginia	1346
	Maryland	1373
Midwest	Wisconsin	1361
	Minnesota	1416

Lyme Disease Surveillance

Lyme disease has been a nationally notifiable disease since 1991. The National Notifiable Diseases Surveillance System (NNDSS) is a national system that allows public health officials to report the occurrence of many infectious and noninfectious diseases to the CDC to centralize surveillance and control the spread of these conditions. States also have their own laws that mandate the reporting of certain conditions. Public health officials at the local level report notifiable conditions to the state based on state regulations and state health departments report nationally notifiable conditions to the CDC (National Notifiable Diseases Surveillance System | CDC, 2015).

Reporting procedures and case definitions are determined at the state level for Lyme disease based on recommendations from the Council of State and Territorial Epidemiologists (CSTE). The CSTE has updated its case definition of Lyme disease several times, with the most recent update in 2011. The changes have been minor but the CDC cautions that it could have an effect on reporting thus need to be considered when analyzing the data before the definition changed (Lyme disease surveillance and available data | CDC, 2015). The current CDC surveillance definition includes the criteria for physician and laboratory diagnosis which are used for reporting. A laboratory diagnosis can be confirmed multiple ways: a positive culture for the organism, *Borrelia burgdorferi* or using two-tiered serologic testing (Figure 3). The first tier is an enzyme-linked immunosorbent assay (ELISA). If the ELISA is positive, a Western Blot is conducted on both immunoglobulin M (IgM) and immunoglobulin G (IgG) assays. IgM antibody tests have higher rates of false-positivity than IgG and are therefore only meaningful during the

first four weeks of infection (MMWR, 1995) and positive results should be considered with results from an IgG assay. An IgG assay is used when a patient has been symptomatic longer than four weeks (MMWR, 1995).

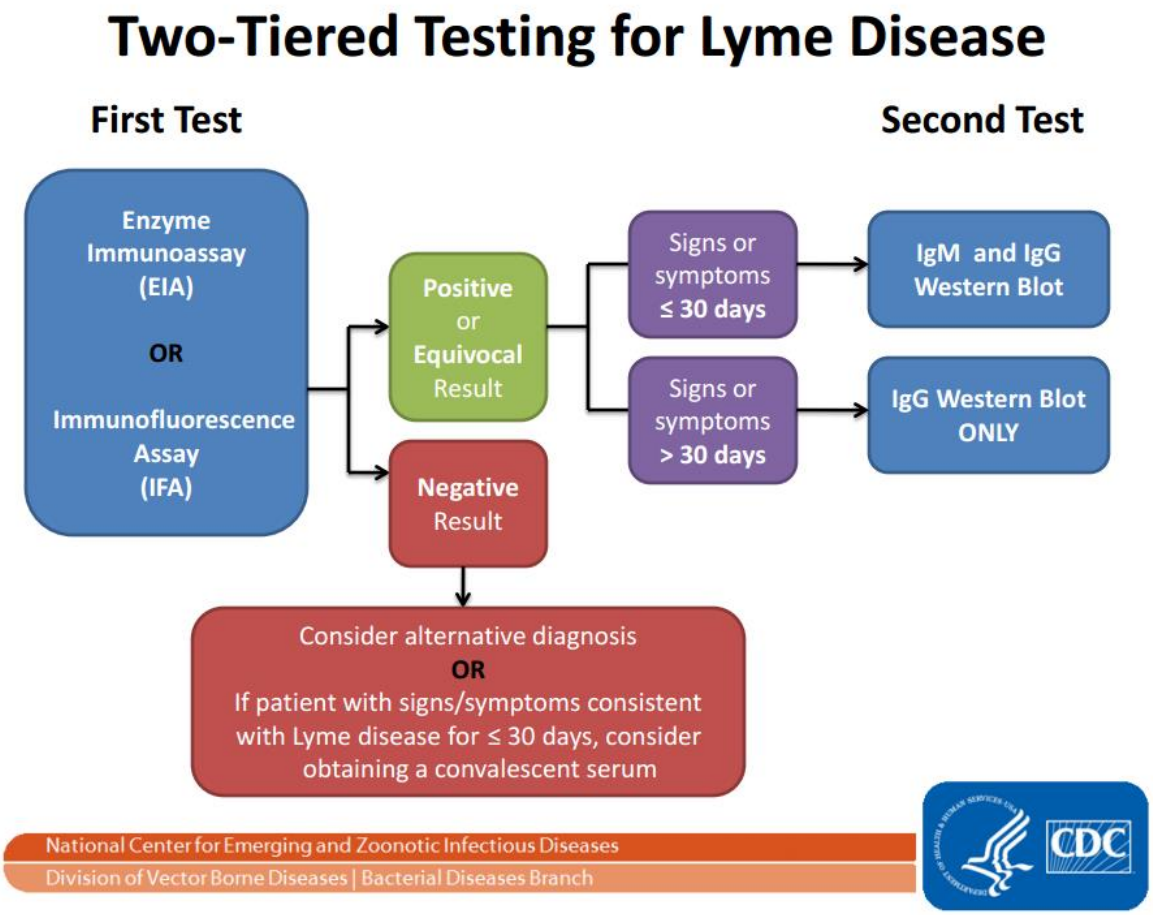

Figure 3. Diagram illustrating the recommended two-tiered testing for a laboratory diagnosis of Lyme disease.

An erythema migrans (EM) is the only clinical manifestation of Lyme disease that can be used as the sole diagnostic factor without a laboratory confirmation (Lyme Disease Case Definition 2011, CDC). The CDC classifies cases into three categories: suspected, probable and confirmed. Only probable and confirmed cases are reported in the NNDSS system. A confirmed case meets one of three criteria: a case with an erythema migrans and known exposure, a case

with an erythema migrans, laboratory confirmation and no known exposure, or a late stage manifestation with laboratory confirmation (CDC, Lyme disease 2011 Case Definition). A probable case is a case with any other physician-diagnosed Lyme disease manifestation with laboratory confirmation (CDC, Lyme disease 2011 Case Definition). A complete description of clinical manifestations of Lyme disease can be found in the "Signs and Symptoms" section. Exposure, in the event of a confirmed case, is defined as having been in a Lyme endemic county in potential tick habitat (woods, brush, grass) less than 30 days before an erythema migrans (CDC, Lyme disease 2011 Case Definition).

Lyme disease cases can be reported to the state health department by the physician or the laboratory. Many state public health departments have an electronic reporting system to automatically receive positive laboratory results. States determine their own surveillance criteria but many use the CDC case definition for consistent monitoring (Table 3).

Table 3. Lyme disease case definition and reporting process by state[1].

State Department name	Case Definition for Reporting	Reporting Process
Connecticut Department of Public Health	CDC/CTSE definition Note: Dept of Public Health notes that patients may be treated on a suspected case basis so data doesn't capture all diagnosed or treated cases	Not stated
Delaware Division of Public Health	CDC/CTSE definition	Delaware Electronic Reporting and Surveillance System – automatically receive positive laboratory results; physicians submit disease reports electronically
Maine Center for Disease Control and Prevention	CDC/CTSE definiton Note: Maine CDC notes that definition is for surveillance and may not be used perfectly in the clinical setting	Can submit by telephone, fax, electronic lab reporting, other written forms or using the standard Lyme disease case report form

[1] A complete list of references used to create this table can be found in Appendix 2.

Massachusetts Local Board of Health Massachusetts Department of Public Health (MDPH)	Any case diagnosed by a healthcare provider with or without laboratory testing is reported to MDPH MDPH uses CDC/CTSE definition to report to CDC	Physicians and laboratories report case to local board of health Local board of health completes Lyme disease case report form to submit to MDPH
Maryland	CDC/CTSE definition Note: report form notes that definition is for surveillance not clinical diagnosis/treatment	Physicians use Lyme disease case report form – can be faxed or mailed Laboratories use a different case report form. Note: both are obligated to report the same case.
New Hampshire Department of Health and Human Services	Case status uses CDC/CTSE definition	Physicians use Lyme disease case report form – can be faxed or mailed Laboratories fill out same form – fax or mail
New Jersey Department of Health	CDC definition	Communicable Disease Reporting and Surveillance System (CDRSS) Electronic reporting system from laboratories into CDRSS Cases reported to local health department – follow up investigation via Lyme disease case investigation form
New York Department of Health	Not stated	Mail case report form to local health department – no specific Lyme disease form Physicians and laboratory
Pennsylvania Department of Health	Not stated	Electronic disease reporting system – physicians and laboratories submit reports
Rhode Island Department of Health	Not stated	General case report form with Lyme disease subsection – can be mailed, faxed Laboratory reports can be submitted by mail, fax or electronic reports
Vermont	CDC/CTSE definition Note: not to be used in clinical diagnosis	Web-based case report form or paper case report form (mail or fax)
Virginia	Not stated	Reportable by physician and laboratory – general Virginia disease report form, CDC surveillance form, or other secure electronic report (at discretion of VDH)

Surveillance Limitations

Several limitations to the surveillance of Lyme disease must be considered when examining data. The biggest issue in any surveillance system, particularly with Lyme disease, is underreporting. Currently, around 30,000 cases are reported to the CDC each year; however, it is estimated that the actual number of patients with Lyme disease is closer to 300,000 each

year (Lyme disease surveillance, CDC 2015). Two CDC conducted studies provide the basis for this estimate: one examines the volume of Lyme disease laboratory tests, the other examines health insurance claims for physician-diagnosed Lyme disease cases. The laboratory testing study collected data from large commercial laboratories in the United States. Of the 7 laboratories that participated in this study, 3.4 million Lyme disease tests were conducted in 2008 on 2.4 million specimens (Hinckley et al., 2014). 62% of these tests were conducted using the two-tiered testing method. The researchers also examined the number of positive laboratory tests in 4 Lyme-endemic states. These 4 states reported 31% of the total tests conducted and 36% of the reported Lyme disease cases from 2007-2009 (Hinckley et al., 2014). Of the total tests conducted in these 4 states, 12% were true positives. This study used the 12% true positivity rate to estimate the national rate and multiplied it by the number of specimens tested resulting in an estimated 288,000 Lyme disease infections in 2008 (Hinckley et al., 2014). The second CDC study, examining health insurance claims from 2005-2010 found 45,430 Lyme disease case claims by physicians for 2005-2010 out of 17,309,054 people/year (Nelson et al., 2015). Standardizing these cases and accounting for claim code omissions, this study estimates 329,000 cases of Lyme disease a year for this time period (Nelson et al., 2015). This estimates a national incidence of 106.6 (per 100,000 people) compared to the reported incidence of 9.4 (per 100,000 people) for that time period (Nelson et al., 2015). These two CDC studies provide strong evidence that the actual occurrence of Lyme disease is nearly ten times greater than current surveillance data suggests.

Another major limitation of CDC surveillance data is that cases are reported by the diagnosing physician or laboratory, thus cases are geographically classified by county of

residence of the patient rather than county of infection (Lyme disease surveillance and available data | CDC, 2015). Surveillance is also limited by each state's capacity, human and financial, to identify all cases (Lyme disease surveillance and available data | CDC, 2015). Based on the multi-step and inconsistent processes each state has, there is the likelihood of high variability in the accuracy of surveillance data between states and in the data that makes it from the physician all the way up to the CDC.

Biology of *Borrelia burgdorferi*

Borrelia burgdorferi, an obligate parasite, is a gram-negative bacteria and a member of the *Spirochaetes* family. It exhibits a spiral shaped body with periplasmic flagella contained between the outer and inner membrane (Shapiro and Gerber, 2000), unlike most other bacterial species. The structure of the flagella allow the spirochetes to move through a gel-like medium such as connective tissue, skin and endothelial cells that would normally slow other bacteria (Li et al., 2000 and Sultan et al., 2012). *B. burgdorferi* moves in a flat wave rather than helical movement of other spirochetes. This motility has proven critical for the survival of *B. burgdorferi* in the tick gut as well as its infectivity of mammalian hosts (Sultan et al., 2013). *B. burgdorferi* survives in the tick and mammalian host, two very different environments, due to its ability to alter its gene expression depending on the environment (Anguita et al., 2003). Due to its small genome, *B. burgdorferi* doesn't have the ability to produce many proteins, thus relying on its host for most of its nutrients (Tilly et al., 2009).

Life Cycle of *Ixodes Scapularis*

I. scapularis has a two-year life cycle consisting of 3 different stages: larvae, nymph, and adult (Figure 4). Larvae, predominant during the summer months, feed almost exclusively on the white-footed mouse, *Peromyscus leucopus*, a natural reservoir of *B. burgdorferi* (Shapiro and Gerber, 2000). Ticks are most likely to become infected with *B. burgdorferi* during their larval stage possibly due to their near-exclusive feeding on *P. leucopus*, often considered the most important reservoir for *B. burgdorferi* (Tilly et al., 2009). It has been estimated that *P. leucopus* causes *B. burgdorferi* infection in 40 to 90% of feeding larvae (LoGiudice et al., 2003). After feeding, larvae enter a period of diapause through the winter (Duik-Wasser et al., 2012). Nymphs emerge the following spring and continue to feed throughout the summer on small mammals, and humans are accidental hosts. Nymphs can emerge infected with *B. burgdorferi* if they were infected as larvae and uninfected nymphs can also become infected with *B. burgdorferi* during feeding. Infected nymphs may transmit *B. burgdorferi* to small mammals, particularly the white footed mouse, the Eastern chipmunk (*Tamias striatus*), the masked shrew (*Sorex cinereus*) and the short-tailed shrew (*Sorex brevicauda*) (Levi et al., 2012) during this feeding rendering them new reservoirs for infection.

Humans are most likely to get Lyme disease from nymph ticks because of their small size, making it difficult to identify and remove the tick before it transmits *B. burgdorferi* (Tilly et al., 2009). Nymphs are also common during the spring and summer months, when humans spend the most time outside, thus increasing chances of tick-human interaction and the risk of contracting Lyme disease (Shapiro and Gerber, 2000).

Adult ticks emerge in the fall, where they seek they seek out a host to feed, mate and lay eggs. The white-tailed deer, *Odocoileus virginianus,* is the primary host of adult *I. scapularis*; however, they are not competent resevoirs for *B. burgdorferi.* Steere (2004) reports that white-tailed deer seem to be critical for adult tick survival. Reduced *I. scapularis* density has been found to be correlated with a reduced white-tailed deer abundance based on observation and experimental studies (Kilpatrick et al., 2014).

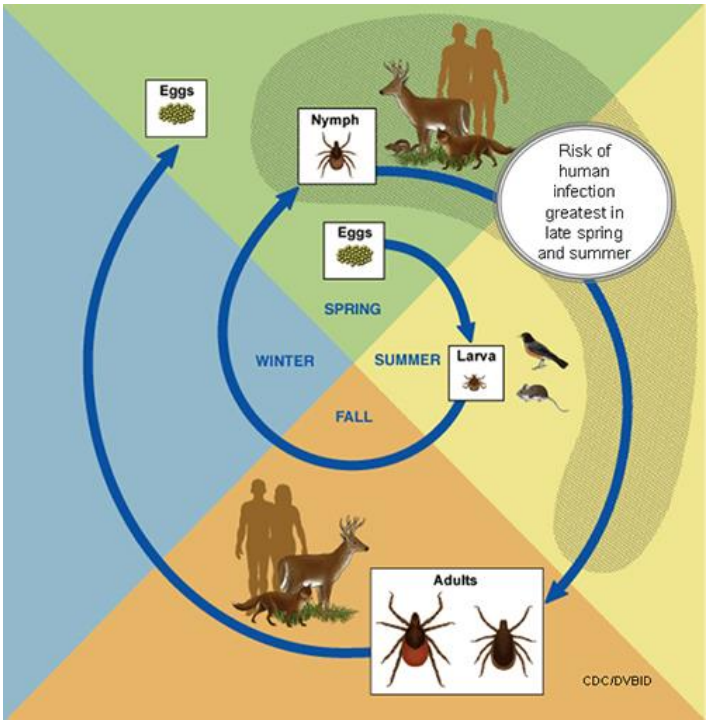

Figure 4. The life cycle of *Ixodes scapularis*, the black-legged tick, the vector for Lyme disease. The life cycle consists of 4 stages and takes approximately 2 years for completion. ("Life Cycle of Blacklegged Ticks | Lyme Disease | CDC", n.d.)

Transmission of *B. burgdorferi*

Lyme disease is transmitted to humans mostly during the tick's nymphal stage. Nymphs are approximately the size of a poppy seed, therefore making it difficult to identify and remove them in a timely manner (Shapiro and Gerber, 2000). When a tick bites, it takes several hours for it to implant its mouthparts; during this time, it secretes enzymes that prevent bites from being painful. Ticks must be attached for more than 36 hours to transmit Lyme disease (Wormser et al., 2006); however, risk of infection greatly increases after 72 hours of feeding (Hojgaard et al., 2008, Piesman et al., 1986). The feeding time needed for transmission is directly related to the engorgement of the tick, in that it must be engorged with blood before the spirochete can be transmitted (Shapiro and Gerber, 2000). During feeding, *B. burgdorferi*, residing in the midgut of *I. scapularis*, multiplies, migrates to the tick salivary glands and is transmitted into the host (Tilly et al., 2009, Dai et al., 2010).

Signs and Symptoms

Lyme disease is a systemic infection with clinical manifestations that are categorized into three distinct stages: early localized stage, early disseminated stage and late stage disease (Wright et al., 2012). If not diagnosed and treated early, Lyme disease can have serious clinical manifestations (Wright et al., 2012).

Early Localized Stage

The early localized stage occurs 3 to 30 days after the tick bite, averaging between 7 and 10 days (Shapiro and Gerber, 2000). During this stage, an erythema migrans (EM), a round,

growing lesion at the site of the tick bite occurs in approximately 90% of patients (Shapiro and Gerber). The average size of an EM lesion is 16cm (Wright et al., 2012) and can be entirely rosy or can have a bulls-eye like appearance (Wormser et al., 2006). This lesion can also be accompanied by non-specific virus-like symptoms such as fever, fatigue, and chills.

Early-Disseminated Stage

Early-disseminated stage occurs 3-5 weeks after the initial tick bite (Shapiro and Gerber, 2000). Multiple EM lesions are a common symptom of early-disseminated Lyme disease. The disseminated lesions tend to be smaller than the initial localized lesion (Shapiro and Gerber, 2000). Early-disseminated Lyme disease can also be classified by neurological manifestations, and sometimes, cardiac manifestations. The most common neurological manifestation is a cranial neuropathy, seventh nerve palsy, which is a form of facial paralysis (Wormser et al., 2006). One in 4 patients with seventh nerve palsy may be diagnosed with Lyme disease (Wright et al., 2012). Lyme meningitis is another neurological manifestation of early-disseminated Lyme disease, occurring a month or so after initial infection (range 3 weeks to 3 months) (Pachner, 1995). Patients with Lyme meningitis usually (Wright et al., 2012).

Although rare, cardiac involvement sometimes occurs in patients with Lyme disease. It can occur between one week and seven months after infection, on average one to two months after infection (Wright et al., 2012). Lyme carditis occurs in approximately 4 to 10% of patients (Wormser et al., 2012). It usually presents as atrioventricular heart block[2], with the block

[2]Atrioventricular heart block is when the electrical signal traveling from the atria to the ventricles is blocked from reaching the ventricle thus affecting the ability of the ventricles to contract (Cleveland Clinic, 2015).

occurring above the bundle of His[3] (Wormser et al., 2006). Patients with Lyme carditis usually present with dyspnea[4], syncope[5], chest pain, and palpitations (Shapiro and Gerber, 2000) and are kept under continuous monitoring in a hospital setting given the life-threatening nature of the condition (Wormser et al., 2006).

Late Stage

Late Lyme disease commonly presents as arthritis, either monoarticular or oligoarticular and affecting large joints such as the knees. Late Lyme disease generally presents several months to years after the initial infection. Lyme arthritis, in contrast to non-Lyme arthritis, is usually characterized by swelling much larger than the degree of pain (Shapiro and Gerber, 2000) and with intermittent inflammation and spontaneous resolution every few weeks to months (Wormser et al., 2006).

Late stage Lyme disease can also present with neurological symptoms: encephalomyelitis[6], peripheral neuropathy[7] or encephalopathy[8] (Wormser et al., 2006). Perhaps due to increased awareness of early stage Lyme disease and more timely treatment, neurologic manifestations of late-stage Lyme disease are increasingly rare (Wormser et al., 2006). Lyme encephalomyelitis can be misdiagnosed as multiple sclerosis; however, serologic testing can differentiate the two (Wormser et al., 2006). Lyme-associated peripheral

---

[3]The bundle of His is a group of heart cells that play a key role in electrical conduction from the AV node (point of electrical impulses between atria and ventricles) to the ventricles (University of Minnesota, 2014).
[4] Dyspnea = shortness of breath
[5]Syncope = fainting
[6] Encephalomyelitis = inflammation of the brain and spinal cord (broad term)
[7] Peripheral neuropathy = results from damage to peripheral nerves; characterized by weakness, numbness, pain in hands and feet (Mayo Clinic, 2014)
[8]Encephalopathy = general term for brain damage or abnormal brain function such as memory loss, difficulty concentrating, personality change, and seizures

neuropathy patients present with intermittent “pins and needles” sensation in their limbs and some radiating pain (Wormser et al., 2006). Lyme-associated encephalopathy is characterized by non-specific symptoms such as memory and cognitive function abnormalities (Wormser et al., 2006).

Environmental Influences on *Ixodes scapularis* and the Life Cycle of *I. scapularis*

Given the multiple vector, reservoir, and host interactions needed for Lyme disease to occur, researchers investigated possible climatic explanations for variability in Lyme disease both between regions and from year to year. Given the complicated life cycle of *I. scapularis*, weather variables could influence tick survival and abundance both directly and indirectly. Tick distribution, survival, and host-seeking behavior could be directly impacted by weather (Schulze et al., 2009). Weather can also affect the food supply and abundance of reservoirs and hosts such as the white-footed mouse, which can indirectly impact tick abundance, survival and infection rates (Ostfeld et al., 2006).

Tick-Density and Lyme Disease

Several studies have examined the risk of infection with *B. burgdorferi* in association with the density of infected and uninfected *I. scapularis.* In general, human risk of vector-borne diseases tends to increase with greater vector abundance and high infection rates of vectors (Ostfeld, et al., 2006). A small scale study in Rhode Island found a positive correlation between

entomological risk, the product of nymph abundance and infection rate of nymphs with *B. burgdorferi*, and Lyme disease incidence (Mather et al., 1996). The relationship between nymph abundance and Lyme disease incidence was linearly strong with an r^2 of 0.978. A more recent large-scale study examining the geographic variation between Lyme disease incidence and tick density found that both the density of nymphs and the density of infected nymphs were statistically significantly associated with Lyme disease incidence (Pepin et al., 2012).

Temperature and *Ixodes scapularis*

Temperature is commonly investigated as a variable to explain *I. scapularis* abundance and possible Lyme disease infection rates among humans. Temperature can affect tick survival due to their susceptibility of desiccation while seeking a host (Schauber et al., 2005). Several studies have examined temperature as a part of an investigation of the effects of climate change and the range of *I. scapularis* (Ogden et al., 2014, Ostfeld et al., 2006, Subak, 2003). Other studies have examined air temperature or temperature at the soil surface as possibly influencing the life cycle of *I. scapularis*. The ticks seem to require moderate temperatures for survival, meaning high summer temperatures and very low winter temperatures both can affect *I. scapularis* survival (Schulze et al., 2009). Larval survival has found to be drastically decreased at temperatures above 27°C (Jones and Kitron, 2000).

A study by Diuk-Wasser et al. (2012) examined the density of infected nymphs along with several environmental variables, such as elevation, mean vapor pressure deficit, maximum and minimum temperature. The investigators sampled 304 sites in the northeastern United States

for nymphs and infected nymphs and found that the distribution of infected nymphs was similar to that of overall nymph density. When examining environmental variables, the authors found that seasonal temperature changes had an influence on the likelihood of finding nymphs in a certain area. High summer temperatures can negatively impact tick survival, as can very low winter temperatures. Extreme temperatures can also impact host-seeking behavior exhibited by the tick as well the ability to find an appropriate host (Diuk-Wasser et al., 2012).

A laboratory study of *I. scapularis* found decreased survival rates in high temperatures (Needham, 1991). Ostfeld et al. (2006) conducted a multivariate risk analysis for Lyme disease, examining several climate factors, host abundance and host food as possible indicators. They found that increased temperature, measured in growing degree days, from one year prior increased the density of nymphs and the density of infected nymphs. A prior study by Schauber et al. (2005) also found that summer temperature from one year prior was also significantly associated with Lyme disease incidence in New York. A study conducted in Illinois found a negative correlation between prior year temperatures and larval abundance of *I. scapularis* found on trapped mice (Jones and Kitron, 2000).

In a different species of tick, *Ixodes pacificus*, the vector responsible for Lyme disease transmission on the west coast, temperature accounted for 31-66% of the variation in tick abundance (Loye and Lane, 1988). This study also found that host-seeking activity by *I. pacificus* was negatively correlated with temperature (Loye and Lane, 1988).

Vail and Smith (1998) examined the effect of temperature at the microclimate level (soil surface), where ticks spend 95% of their lives (Ostfeld et al., 2006) and found that *I. scapularis* nymph density decreased as summer temperature (degrees Celsius) increased. In this case,

temperature explained 33% of the variation in nymph abundance. In the larger context of climate, a Canadian study modeling the effect of climate change on *I. scapularis* reproduction using mean monthly temperature and the number of days above 0°C showed that increasing temperatures are associated with an expanding range of *I. scapularis* into Canada (Ogden et al., 2014).

Diuk-Wasser et al. (2012) also discuss the effect that temperature can have on the timing of the *I. scapularis* life cycle. Cooler fall temperatures may limit the ability of larvae to feed before diapause, whereas earlier and warmer spring hasten the emergence of nymphs and increase the season during which humans are at risk of Lyme disease (Diuk-Wasser et al., 2010, Steere et al., 2004).

Subak (2003) examined the effect of winter temperature in the same year and one year prior on Lyme disease incidence. She found that winter temperature one year prior was positively associated with Lyme disease incidence in several northeastern states. Several studies have found that *I. scapularis* winter activity is affected by a threshold temperature of between 4°C and 9°C and activity linearly increases as temperatures increase above this threshold (Vail and Smith, 1998).

Temperature may indirectly impact *I. scapularis* through decreased abundance of the white-footed mouse during cold temperatures. A decrease in the mouse population decreases the number of mice carrying *B. burgdorferi* and the likelihood that as many larvae or nymphs will become infected through feeding. Increased populations of *P. leucopus* during larval feeding have also been associated with an increased abundance of infected nymphs the following year (Subak, 2003).

Precipitation and *Ixodes scapularis*

I. scapularis requires a relatively humid microclimate for survival, as shown by laboratory studies that found high death rates at low humidity (Vail and Smith, 1998). A study conducted on *I. pacificus* found that humidity was able to explain 40-66% of variation in activity (Vail and Smith, 1998). Dry conditions have been found to negatively impact tick survival (Schulze et al., 2009). Humidity is influenced in many places by the amount of precipitation (McCabe and Bunnell, 2004) where increased precipitation increases humidity.

Ostfeld et al. (2006) found that the density of nymphs and the density of infected nymphs were affected by same-year precipitation. However, they found that tick abundance was not positively correlated in a linear fashion with precipitation; rather, moderate levels of precipitation were associated with the highest abundance of nymphs. Low survival rates with high precipitation could be explained through drowning (Ostfeld et al., 2006).

A study examining the effect of drought conditions and moisture on Lyme disease incidence found that the summer Palmer Hydrological Drought Index (PHDI), a moisture index, from two years prior was positively associated with Lyme disease incidence (Subak, 2003). Another study examining the effect of precipitation on *I. scapularis* found larval density on trapped mice to be significantly lower, approximately 20% lower than the average density, following two drought years (Jones and Kitron, 2000). A decrease in *P. leucopus* population was also observed during a drought year and in the year following a drought year at approximately 57% less and between 30% and 46% less than the average number of mice trapped respectively (Jones and Kitron, 2000). Another study found a significant association between precipitation in

all seasons in the US and tick abundance. Soil that was too wet was negatively correlated with the abundance of *I. scapularis* (Guerra et al., 2002). McCabe and Bunnell (2004) found that Lyme disease and precipitation in early summer were positively correlated. They also found that precipitation and Lyme disease cases 2 years later were associated (McCabe and Bunnell, 2004). This evidence indicates that precipitation can affect tick survival and abundance. Precipitation can influence soil conditions, such as humidity, which *I. scapularis* requires for survival, as well as directly impact tick survival through flood-death or desiccation. Several of these studies found a time lag between tick abundance or Lyme disease and precipitation, particularly with regards to humidity and precipitation two years prior (Subak, 2003 and McCabe and Bunnell, 2004).

Environmental Conditions and Humans

Aside from directly affecting tick and host survival, environmental conditions can affect human exposure to *I. scapularis*. Human activity increases during the summer months, the same time that nymphs are in highest abundance and the most commonly reported months of onset for Lyme disease (MMWR, 2008). Particularly hot summer temperatures or times of high precipitation could cause humans to spend less time outdoors, reducing their risk of Lyme disease through reduced exposure to the vector (Subak, 2003). The most commonly reported months of Lyme disease onset are June, July and August

Significance

Lyme disease is the most common vector-borne disease in the United States. Given it's non-specific symptoms, many cases go untreated, which can result in serious health problems. *I. scapularis* ticks have a very complex life cycle with many vector, reservoir, host and environmental interactions needed for Lyme disease to persist. Being able to predict years in which high incidence is expected would be of great use to public health officials and physicians. Environmental conditions can impact survival and abundance of *I. scapularis*, but there are no definitive conclusions about the effect of climatic conditions on tick abundance and Lyme disease incidence. Temperature and precipitation likely influence the abundance of *I. scapularis*, and thus human infection of Lyme disease, but the temporal effect is unclear and there is huge variability in the indicators used to measure temperature and precipitation. This makes it challenging to compare across studies and to determine which indicator is most appropriate for a predictive model.

Research Question

This study is testing for statistically significant associations of mean summer and winter temperature and cumulative summer and winter precipitation with Lyme disease incidence. Associations are examined with a same-year, one-year and two-year time lag across all four variables. This study has four specific aims:

1. Test for statistically significant associations between summer precipitation in 2012, 2013, 2014 and Lyme disease incidence in 2013 and 2014.

2. Test for statistically significant associations between summer temperature in 2012, 2013, 2014 and Lyme disease incidence in 2013 and 2014.
3. Test for statistically significant associations between winter precipitation 2012 and 2013 and Lyme disease incidence in 2013 and 2014.
4. Test for statistically significant associations between winter temperature 2012 and 2013 and Lyme disease incidence in 2013 and 2014.

Methods

Study Design

The large burden of Lyme disease found in the Mid-Atlantic and Northeastern United States provided the basis for the selection of this region as the study area. Minnesota and Wisconsin were excluded from this analysis because they are not a part of this continuous region. Temperature and precipitation were chosen as study variables based on evidence from prior studies indicating an association between temperature and precipitation and Lyme disease. A same-year, one-year and two-year time lag are being examined based on the two-year life cycle of *I. scapularis* as well as findings from prior studies that indicate a possible lag effect of weather on tick abundance and Lyme disease.

Data Sources:

Lyme Disease Data

Lyme disease data was downloaded from the Centers of Disease Control and Prevention (CDC) website. The data includes the number of cases reported to the CDC between 2000 and 2014, which were stratified at the county level for all US states. The cases were reported to the CDC from state health departments using the most up-to-date CDC case definition.

Population Data

Population data were downloaded from the United States Census Bureau. The bureau produces annual population estimates every year, stratified at different levels. County-level population estimates for 2013 and 2014 were used for this study. The population estimates are for July 1, 2013 and July 1, 2014. The Census bureau calculates the estimates using birth, death, and migration variables (U.S. Census, 2014). The 2010 Census data provides the population base for the estimates (U.S. Census). Methods for calculating population estimates can be found on the U.S. Census website under "Population Estimates."

Temperature and Precipitation Data

Temperature and Precipitation data were downloaded from the National Climatic Data Center (NCDC) associated with the National Oceanic and Atmospheric Association (NOAA). To assess summer and winter temperature effects, monthly summaries for December 2011, January 2012, February 2012, June 2012, July 2012, August 2012, December 2012, January 2013, February 2013, June 2014, July 2014 and August 2014 were used. The data were searched

using these date queries by state in order to capture all weather stations in the state for the selected time frame. Downloaded data included station identification code, station name, geographical location (latitude, longitude and elevation), date, total precipitation (TPCP) and monthly mean temperature (MNTM). Latitude was reported in decimated degrees where northern hemisphere values >0, longitude was reported in decimated degrees with western hemisphere values <0, and elevation was reported in thousandths of meters above sea level. Total precipitation was reported in tenths of millimeter and temperature was reported in tenths of degrees Celsius. A guide to the weather data produced by the National Climatic Data Center is available from NOAA (Global Historical Climatology Network (GHCN) Documentation, 2016)

Data Processing

Lyme Disease Incidence

Lyme disease case counts in 2013 and 2014 were extracted from the CDC dataset for all counties in Connecticut, Delaware, Washington D.C., Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont and Virginia. The 2013 and 2014 population estimates from the U.S. Census Bureau were added into their respective Microsoft Excel sheet (one for 2013 and one for 2014), matched by state and county name. Bedford City, Virginia, an independent city, became a part of Bedford County, Virginia in the middle of 2013; therefore, population estimates for Bedford City were within Bedford County. However, the CDC did not make this change in their dataset so Bedford City and Bedford County were still two separate entities in their dataset. There was only one case of Lyme

disease in Bedford City in 2013 so to account for the census change, this one case was added to the case count for Bedford County.

Once cases counts and populations were matched appropriately, Federal Information Processing Standards (FIPS) codes, unique identifiers for every county in the United States, were added to the sheet, matched with their corresponding county. Lyme disease incidence rates per 100,000 people for 2013 and 2014 were calculated in Microsoft Excel.

Temperature and Precipitation Data

Temperature and precipitation data were processed in Microsoft Excel using pivot tables to sort monthly summaries into the correct season. Variables for this study are categorized as seasonal variables, weather-related variable and temporal variables, resulting in 10 variables (Table 4). Summer is defined as June, July and August. Winter is defined as December of the preceding year, January and February. For example, winter 2012 is defined as December 2011, January 2012 and February 2012.

Table 4. Categories for analysis of weather variables.

Season	**Weather**	**Time**
Summer	Precipitation Temperature	2012 2013 2014
Winter	Precipitation Temperature	2012 2013

Temperature (MNTM) (to tenths of degrees, Celsius) and precipitation (TPCP) data (in tenths of millimeters). In order to render these values easier to use, they were converted into degrees Celsius and millimeters.

Once sorted by season, time and weather category, the data were put into separate Excel sheets for each of the 10 variables for future analysis. Every row in each sheet was a different month during the respective season at different weather stations (Table 5). Any monthly entries that had missing temperature or precipitation data were removed. The seasonal measurement is an average of the months for which data was available at each station. For example, January and February 2013 did not have any precipitation data at the Barton 3.0 ENE VT US weather station in Vermont so these entries was deleted leaving December 2012 precipitation data to comprise "Winter Precipitation 2013" from this weather station. Some stations did not have any temperature or precipitation data for any of the months during a respective season, other stations had data for one or two months but not all three, and some stations had data for all three months of the respective season. The total number of stations with data was different for each of the 10 variables. The existing monthly values for each station were averaged, a process repeated across the 10 categories.

Table 5. Total number of weather stations by state organized from fewest to most.

State	Weather Station Count
Delaware	75
Rhode Island	84
Connecticut	174
Vermont	264
New Hampshire	373
Maine	428

Maryland	442
Massachusetts	442
New Jersey	558
Virginia	868
Pennsylvania	1192
New York	1361

Spatial Analysis

Temperature and Precipitation Raster Map Building

Given that temperature and precipitation measurements are taken from fixed points, but the analysis evaluates county-level correlations, the temperature and precipitation data were interpolated to form a smooth surface with measurements in every county. This was achieved using geographic information systems (GIS) to process the data across the twelve states. A base map of the United States was selected in ArcMap (ESRI, 2014) with a coordinate system of GCS NAD 1983 (meters). Using the clip tool, the area from Virginia to Maine was selected as the total map area (Figure 5). This area was then projected into USA Contiguous Albers Equal Area Conic coordinate system to more accurately portray the selected states spatially.

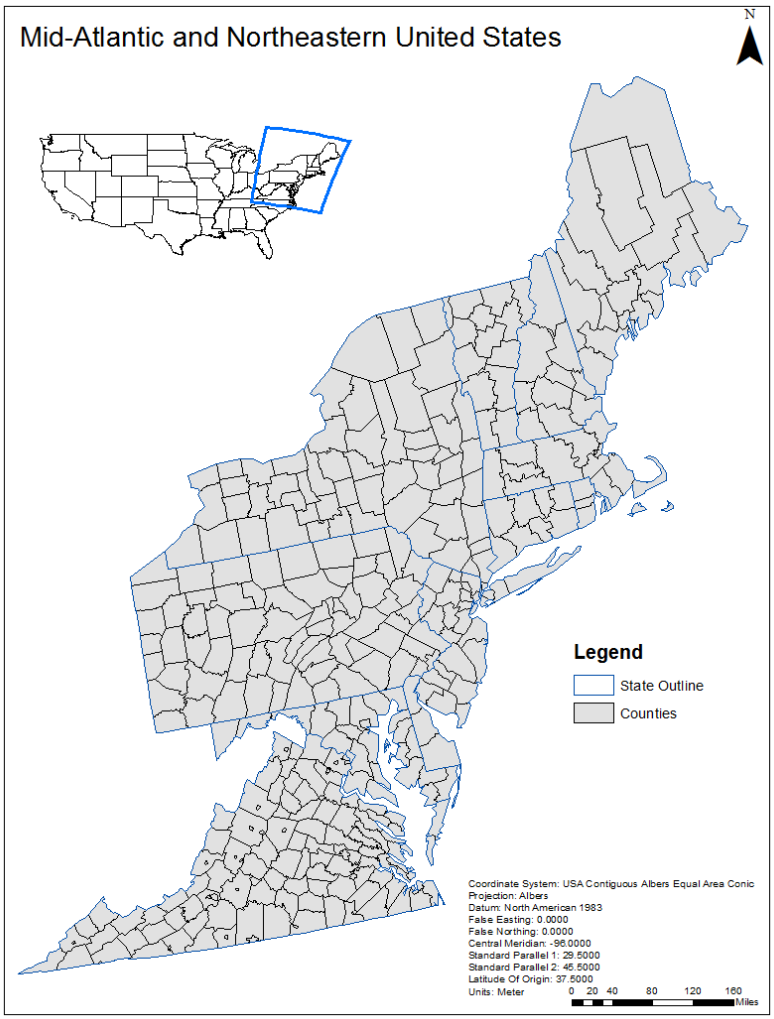

Figure 5. Basemap of the Mid-Atlantic and Northeastern United States. States included are Connecticut, Delaware, District of Columbia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York. Pennsylvania, Rhode Island, Vermont and Virginia.

The Microsoft Excel sheets containing study variables were imported into ArcMap (ESRI, 2014) as tables. Using the "display XY" feature, the weather stations were plotted as points on the base map. This layer was exported and saved as a shape file and then projected into USA Contiguous Albers Equal Area Conic in order to match the base map. As a shape file, every station point now had precipitation or temperature measurements associated with that geographic location.

In order to have a continuous area of precipitation or temperature across the entire region using just the plotted points, the data was further manipulated. Using the IDW (Inverse Distance Weighted) interpolation tool and the weather stations as the input points, the data were interpolated into a raster. Given that weather does not have large amounts of variation at a local level, IDW interpolation was used because it assigns values weighted by the distance of nearby neighbors (ESRI, Raster Interpolation Toolset). The processing extent was set to the same as the states and the mask was also set to the states layer in order to keep the interpolation within the boundaries of interest. After the interpolation was run, the raster was added to the base map and displayed in a color gradient where each pixel had a specific precipitation or temperature measurement.

Zonal Statistics

Since a raster is in a different form than the data for county boundaries (vector polygon) and the Lyme disease data (vector polygon), the weather variables were formatted for statistical analysis and comparison with Lyme disease counts. This formatting was achieved using the "Zonal Statistics" function in ArcMap. Zonal statistics allows raster data to be summarized using the boundary of an existing shape file. In this case, the "zones" were the county boundaries, as specified by Federal Information Processing Standards (FIPS) codes, unique identifiers for every county in the United States. The raster input was the weather variable (i.e. winter precipitation 2012) across the entire map area. The zonal statistic extracted was the "Mean", which calculated the mean of all of the raster pixel values within each county.

Under the environment settings, in the raster analysis sub-category, cell size was set to 60 in order to extract weather data from the very small counties in Virginia.

Temperature and Precipitation Map-Building with Zonal Statistics

The calculated zonal statistics were exported into an Excel spreadsheet from ArcMap and matched by FIPS code with 2014 Lyme disease incidence. In order to make the visual display of the weather data clear, this table was imported back into ArcMap and joined to the existing states layer, matching by FIPS codes, to provide the data in the table spatial locations. Once all of the zonal statistics (mean "variable") were matched with their corresponding spatial location, the data was exported as a shape file for ease of processing in ArcMap. This map-building process from weather stations to temperature and precipitation gradients by county was repeated for each of the 10 weather variables. This resulted in 10 separate maps, one for each weather variable (Appendix 2).

Lyme Disease Map

In order to visually display Lyme disease incidence trends across the northeast region for 2013 and 2014, two maps were developed using the same base map of Virginia to Maine that was used for the temperature and precipitation variables. The Microsoft Excel sheets for Lyme disease incidence 2013 and Lyme disease incidence 2014 were imported into ArcMap as tables. These tables were joined to the counties base map using FIPS codes to ensure county incidence was correctly located. Once joined, the layer was exported and added to the map document as its own shape file.

Map Formatting

For use in a spatial statistics program, Lyme disease incidence for 2013 and 2014 were added to the individual shape files for each weather variable. For example, the attribute table of the shape file of winter precipitation 2012 had a column for incidence 2013 and incidence 2014 alongside its precipitation data.

Statistical Analysis:

GeoDa

GeoDa is a spatial statistics program from the GeoDa Center for Geospatial Analysis and Computation at Arizona State University. Data can be imported in the form of ESRI shape files from ArcMap. For use in GeoDa, a shape file containing every weather variable and Lyme disease incidence for 2013 and 2014 was created in ArcMap using the joining process described above and imported into GeoDa. All statistical analyses were conducted in GeoDa in order to account for spatial correlation of the variables.

First, a univariate local Moran's I test was conducted to examine the spatial clustering of Lyme disease incidence in 2013 and 2014 respectively. This was conducted on untransformed Lyme disease incidence data. In order to conduct the local Moran's I test, a weights file was created that was distance-based and used Euclidean distance with the threshold distance, the distance to be considered neighbors, set to the default. The default threshold ensures that each area has at least one neighbor (GeoDa). Euclidean distance is used for projected data, such as USA Contiguous Albers Equal Area Conic, the format of these shape files.

Secondly, an ordinary least squares (OLS) regression was conducted for each combination of variables to assess the association, if any, between the different weather variables and Lyme disease incidence in 2013 or 2014. In this case, a regression model was used rather than a correlation because the research question is examining how temperature and precipitation act on Lyme disease incidence, the dependent variable. With a regression model, there are several assumptions. The first is that the data are normally distributed. Histograms of each weather variable, Lyme disease incidence 2013 and Lyme disease incidence 2014 were created to look for normality (Figure 6 and Figure 7). The histograms of the weather variables showed approximate normality, confirmed with normality (Q-Q) plots in SPSS (Chicago, Version 22). Lyme disease incidence 2013 and 2014 were not normally distributed in the histograms. Due to the nature of counting disease cases, the distribution was also truncated at 0. In order to correct for this distribution, the data for Lyme disease incidence 2013 and Lyme disease incidence 2014 were transformed logarithmically (Figure 8 and Figure 9).

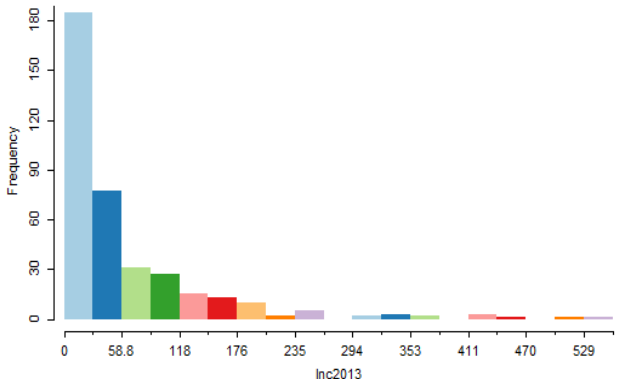

Figure 6. Distribution of Lyme disease incidence 2013. The x-axis is Lyme disease incidence 2013

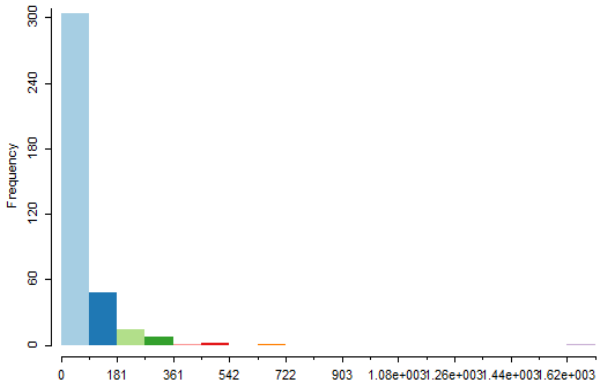

Figure 7. Distribution of Lyme disease incidence 2014. The x-axis is Lyme disease incidence 2013.

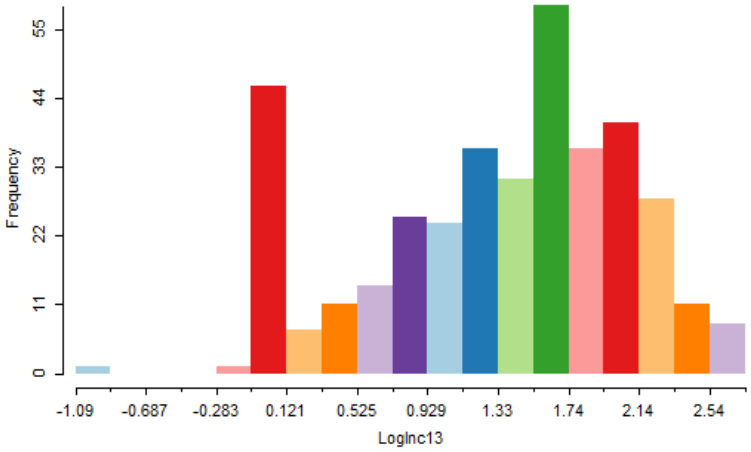

Figure 8. Distribution of Lyme disease incidence 2013 when logarithmically (log_{10}) transformed. The x-axis is the log_{10} of Lyme disease incidence 2013.

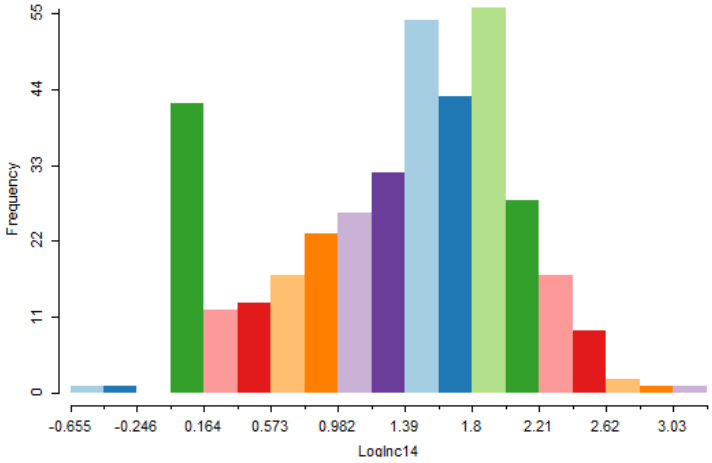

Figure 9. Distribution of Lyme disease incidence 2014 when logarithmically (log_{10}) transformed. The x-axis is the log_{10} of Lyme disease incidence 2014.

The second assumption of a regression is that the variables are independent of one another. Given the nature of the temperature, precipitation and Lyme disease incidence, there are several issues with this assumption including seasonality, temporal variance and spatial autocorrelation. GeoDa was used for this analysis because of its ability to correct for spatial autocorrelation in its calculations. Due to the statistical complexity of correcting for seasonal and temporal correlation between the variables, these issues could not be adjusted for in this model.

Taking into account the benefits and limitations of using a regression, 18 different bivariate OLS regressions were run with the various weather variables as the independent variable and the transformed Lyme disease data as the dependent variable. In order to account for spatial autocorrelation, a weights file was created for this dataset using Euclidian distance with the threshold distance set to default as above and used with the regression.

Using the output from the OLS regression as a diagnostic to determine which type of spatial regression model, spatial lag or spatial error, will be better for the data to account for spatial autocorrelation. The best indicator are the Lagrange Multiplier (LM) statistics. The first one considered is the standard LM-lag. If the p-value for this is significant, the standard LM-error is examined. If both p-values are significant, the robust LM-lag and robust LM-error are examined. If both robust values are statistically significant the more significant value is selected; however, if both are significant at the same level, the higher of the two robust LM test statistic values indicates which regression model to use (Anselin, 2005). This selection process (Figure 10) was repeated 18 times for the different variable combinations. Once the correct

spatial regression models were chosen, they were run on each weather variable/Lyme disease incidence pair. As with the diagnostic OLS regression, the weather variable was on the x-axis and the logarithmic transformation of Lyme disease incidence was on the y-axis.

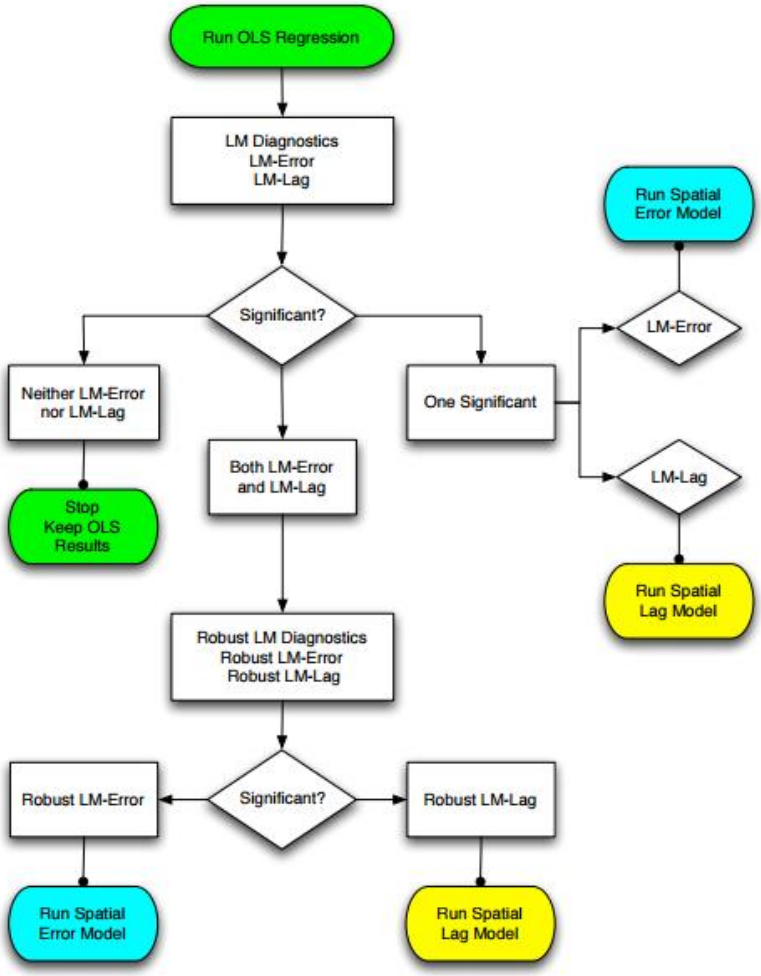

Figure 10. OLS regression diagnostics and selection process of a spatial regression model (Anselin, 2005).

Results

ArcMap

The map building process in ArcMap for each temperature and precipitation variable resulted in 10 separate maps (Appendix B). Summer precipitation did not exhibit any clear visual patterns (Figures 17, 18, 19). Summer temperature maps exhibited a temperature gradient with southern states being warmer and the northern most states being the coolest (Figures 20, 21, 22). Winter precipitation did not exhibit any obvious visual gradients (Figures 23, 24). Winter temperature exhibited a gradient as well, with the southernmost counties being the warmest and the northern-most states the coldest (Figures 25, 26)

The map building process for Lyme disease incidence resulted in two maps: incidence 2013 (Figure 11) and incidence 2014 (Figure 12) by county. Incidence rates are displayed using a red color gradient with lighter shades indicating lower incidence and darker shades indicating higher incidence. Counties in central Pennsylvania, southern Vermont, on the southern coast of Maine and southeast Connecticut and Massachusetts have the highest incidence rates in both 2013 and 2014.

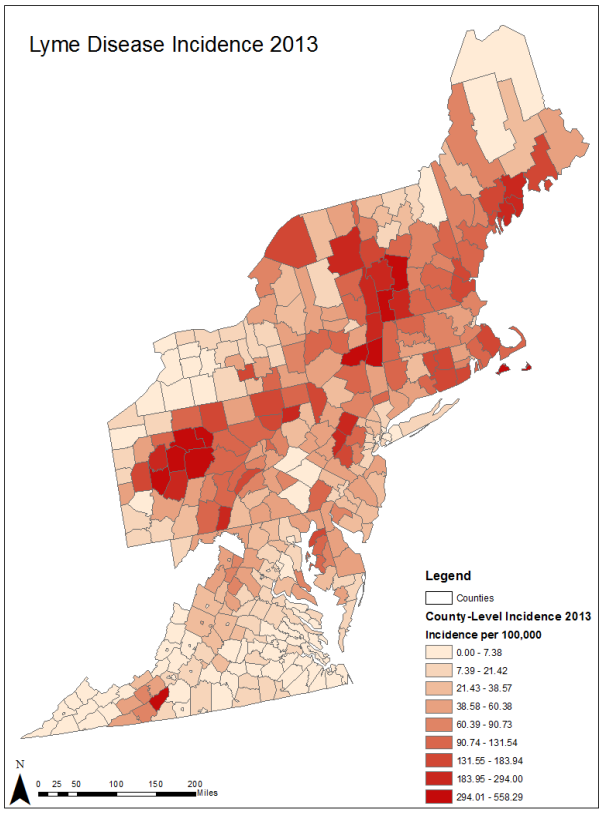

Figure 11. Lyme disease incidence 2013 by county.

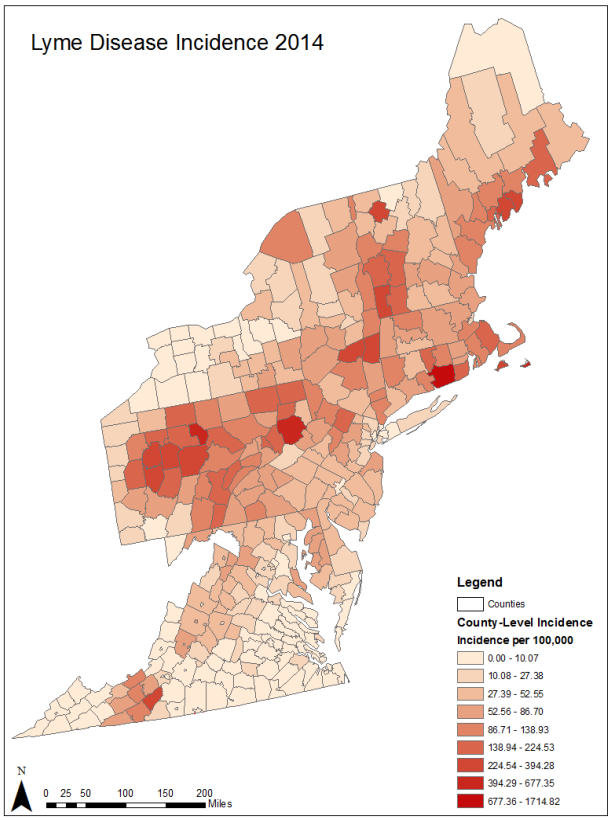

Figure 12. Lyme disease incidence 2014 by county.

GeoDa

Lyme Disease Cluster Maps

The output from the univariate local Moran's I test for Lyme disease incidence 2013 and Lyme disease 2014 produced a cluster map, a significance map and Moran's I scatter plot. The cluster map shows 4 different categories of autocorrelation resulting in two cluster categories and two outlier categories (Table 6). These categories also correspond to the four quadrants on the Moran's scatterplot. The significance map shows which areas of the cluster map are significant and the significance level. Significance is based on a pseudo-p value that can change depending on the number of random permutations run (99 in the case of these variables) (GeoDa, Glossary of Key Terms). The cluster map (Figure 13) and significance map (Figure 14) for Lyme disease incidence 2013 and the cluster map (Figure 15) and significance map (Figure 16) for Lyme disease incidence 2014 indicate spatial clustering and spatial outliers for the distribution of Lyme disease.

Table 6. Guide to cluster map interpretation. Adapted from BioMedware SpaceStat Help (Interpreting univariate Local Moran Statistics)

Color	Category	Interpretation
Blue	Low-low	Spatial Cluster – lower than average and neighbors are low
Red	High-high	Spatial Cluster – higher than average and neighbors are high
Pink	High-low	Spatial Outlier – higher than average but neighbors are low
Purple	Low-high	Spatial Outlier – lower than average but neighbors are high

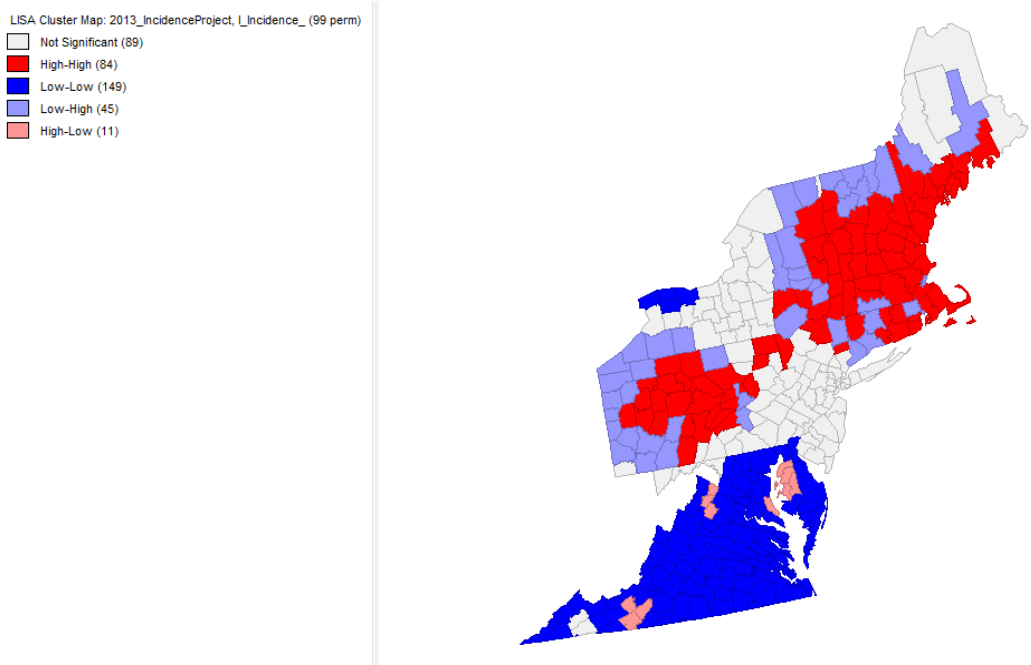

Figure 13. Local Indicators of Spatial Association (LISA) cluster map for Lyme disease incidence 2013.

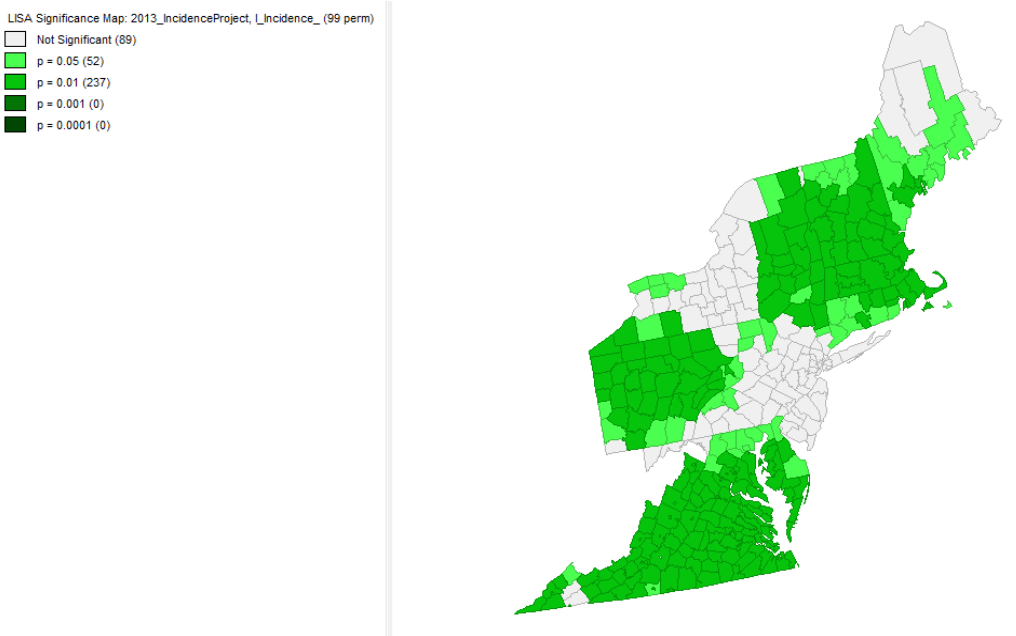

Figure 14. Local Indicators of Spatial Association significance map for Lyme disease incidence 2013. Indicates significance of spatial clusters and outliers in Figure 11.

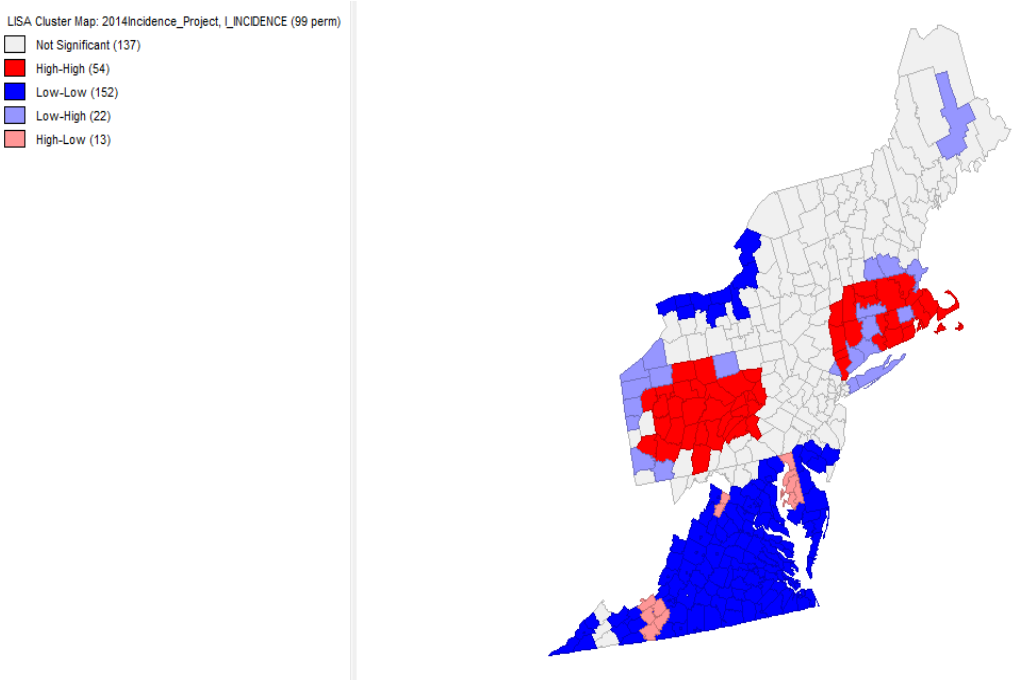

Figure 15. Local Indicators of Spatial Association cluster map for Lyme disease incidence 2014.

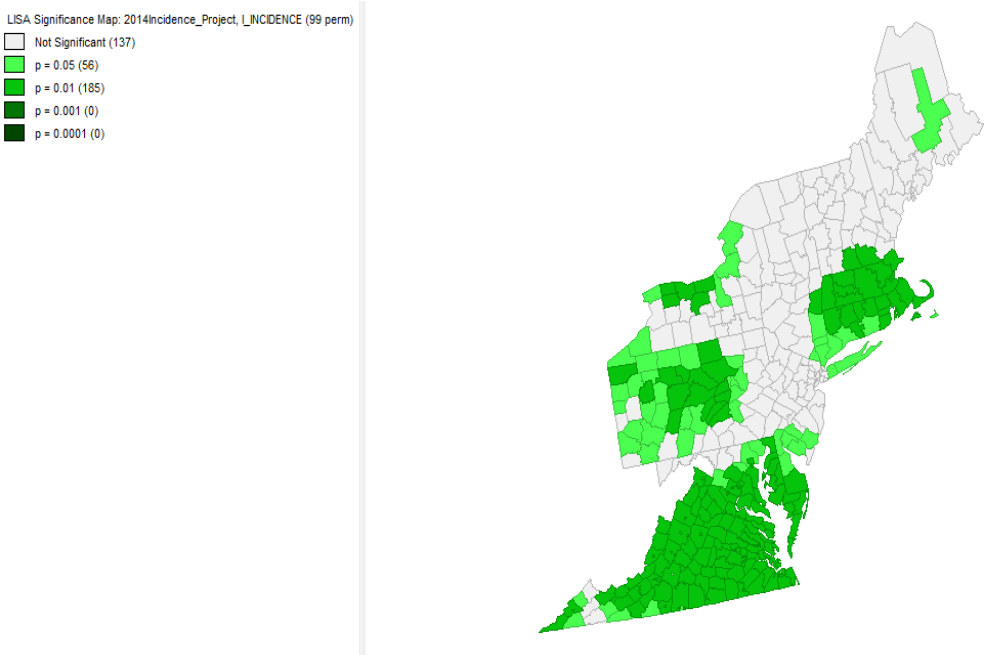

Figure 16. Local Indicators of Spatial Association significance map for Lyme disease incidence 2014. Indicates significance of spatial clusters in Figure 15.

The Lyme disease cluster maps indicate which areas have higher than average incidence rates and which areas have lower than average incidence rates while also indicating the counties that do not follow the expected spatial clustering of Lyme disease. New England and parts of Pennsylvania have more areas of high-high autocorrelation, where a county has a high incidence and its neighbors do as well whereas the Mid-Atlantic area including Maryland, Virginia, Delaware, Washington D.C. and parts of New Jersey have more areas of low-low autocorrelation where most areas have low incidence rates as do their neighbors. Based on those patterns, it can be observed that New England and Pennsylvania also have more areas of low-high outliers where those counties have low incidence rates while their neighbors have high incidence rates. Southern states in this analysis tended to have a few areas of high-low autocorrelation indicating counties with high incidence while their neighbors have low incidence rates. In the 2014 cluster map, one county in Maine is indicated as a low-high spatial outlier. This means that the neighboring counties have high incidence even if they are not a high-high spatial cluster and this county has low incidence. Of note, 2013 Lyme disease incidence had many more areas of spatial association than 2014 incidence particularly in regards to high-high clusters in New England and low-high clusters in New England and Pennsylvania.

The significance maps for both 2013 and 2014 indicate significance to the .01 level in the low-low cluster of Virginia and the high-high cluster of Massachusetts, Rhode Island and Connecticut and the high-high cluster of central Pennsylvania. Northwest Pennsylvania had low-high outliers in both 2013 and 2014; however, the 2013 outliers were more statistically

significant than the 2014 outliers. Eastern Maryland had high-low outliers in both years which were statistically significant at the .01 level.

Regressions

Eighteen different diagnostic ordinary least squares (OLS) regressions were run to determine which spatial regression model, spatial lag or spatial error, would work better for each bivariate relationship examined. The OLS regressions for all of the bivariate relationships between weather variables and logarithmic Lyme disease incidence indicated that a spatial regression model was the better model to use. All of the variables indicated the use of a spatial error model except for winter temperature 2012 and 2013. Diagnostics of the OLS models for winter temperature 2013 for both Lyme disease years indicated a spatial lag model as the better fit. For winter temperature 2012 with logarithmic Lyme disease incidence 2013, both standard Lagrange Multiplier (LM) values (lag and error) returned with statistically significant values. In this case, it was necessary to examine the robust LM values for lag and error. Both robust LM values were statistically significant, with $p < 0.00000$ for the robust LM-lag and $p = .02926$ for the robust LM-error. Given the large difference in statistical significance, the spatial regression model with the most significant result is used. In this case, a spatial lag model was indicated. For winter temperature 2012 with logarithmic Lyme disease incidence 2014, both standard LM values were statistically significant, when examining the robust LM values, $p<0.00000$ for LM-lag and $p=.00745$ for LM-error. As above, the large statistical difference indicates that a spatial lag model is the better fit for this data. The selection process for winter temperature 2013 regression models was slightly different. For both logarithmic Lyme disease

variables (2013 and 2014), the standard LM values were statistically significant at p<0.00000. In examining the robust LM values, both were also statistically significant at p<0.00000. This indicated that the test statistic values needed to be considered to decide which model is a better fit. In the case of winter temperature 2013 with logarithmic Lyme disease incidence 2013, the test statistic for robust LM-lag was 69.8894 and the test statistic for robust LM-error was 22.9034 indicating that a spatial lag model is a better fit in this case. For winter temperature 2013 with logarithmic Lyme disease 2014, the test statistic for robust LM-lag was 68.3402 and the test statistic for robust LM-error was 29.2416 indicating that a spatial lag model is a better fit. For all of the other weather variables, both standard and robust LM statistics were significant at the p<0.00000 level so the higher of the robust test statistic values were used to determine that the spatial error model was the best fit. The results of the OLS diagnostics are summarized in Table 7.

Table 7. Results of Ordinary Least Squares regressions diagnostics for spatial dependence. 18 different bivariate relationships between weather and Lyme disease incidence were examined. The results of the robust-LM test statistics determine which spatial regression model to use.

Variables	Robust LM-lag test statistic	Robust LM-error test statistic	Model Decision
Summer Precip 2012 x Lyme Incidence 2013	42.3885	230.9995	Spatial error
Summer Precip 2013 x Lyme Incidence 2013	47.1495	313.3447	Spatial error
Summer Temp 2012 x Lyme Incidence 2013	31.5924	408.5317	Spatial error
Summer Temp 2013 x Lyme Incidence 2013	32.8773	394.8091	Spatial error
Winter Precip 2012 x Lyme Incidence 2013	67.3939	160.2678	Spatial error
Winter Precip 2013 x Lyme Incidence 2013	64.9288	236.9139	Spatial error

Winter Temp 2012 x Lyme Incidence 2013	78.6176 (p<0.00000)	4.7525 (p=0.02926)	Spatial lag
Winter Temp 2013 x Lyme Incidence 2013	69.8894	22.9034	Spatial lag
Summer Precip 2012 x Lyme Incidence 2014	41.2142	223.1509	Spatial error
Summer Precip 2013 x Lyme Incidence 2014	27.2033	233.7433	Spatial error
Summer Precip 2014 x Lyme Incidence 2014	35.0870	317.8627	Spatial error
Summer Temp 2012 x Lyme Incidence 2014	36.7550	358.9703	Spatial error
Summer Temp 2013 x Lyme Incidence 2014	37.5047	349.4401	Spatial error
Summer Temp 2014 x Lyme Incidence 2014	40.2398	330.1594	Spatial error
Winter Precip 2012 x Lyme Incidence 2014	72.9831	184.8660	Spatial error
Winter Precip 2013 x Lyme Incidence 2014	71.7937	238.3330	Spatial error
Winter Temp 2012 x Lyme Incidence 2014	78.6904 (p=0.00745)	7.1623 (p<0.00000)	Spatial lag
Winter Temp 2013 x Lyme Incidence 2014	68.3402	29.2416	Spatial lag

Once the type of spatial regression was determined, the 18 bivariate regressions were run between the logarithmic Lyme disease incidence of 2013 or 2014 and the 10 weather variables. Significance was assessed at the p=0.05 level. All of the results were significant at p<0.00000 except winter temperature 2012 with logarithmic Lyme disease incidence 2013 was significant at p=0.00312 and winter temperature 2012 with logarithmic Lyme disease incidence 2014 was significant at p=0.00464. Winter temperature 2013 with logarithmic Lyme disease incidence 2013 (p=0.06311), winter temperature 2013 with logarithmic Lyme disease incidence

2014 (p=0.07397) were not significant. All correlation coefficients were positive. The magnitude of the regression coefficients ranged from 0.0053 (summer precipitation 2013 x log Lyme incidence 2014) to 0.0418 (summer temperature 2014 x log Lyme incidence 2014). The complete results of the spatial regressions are summarized in Table 8.

Table 8. Results of spatial regressions. All models were spatial error models except those for winter temperature regressions which were spatial lag models.

Variables	Regression Coefficient	p-value
Summer Precip 2012 x Lyme Incidence 2013	0.00627	p<0.00000
Summer Precip 2013 x Lyme Incidence 2013	0.00623	p<0.00000
Summer Temp 2012 x Lyme Incidence 2013	0.04119	p<0.00000
Summer Temp 2013 x Lyme Incidence 2013	0.04181	p<0.00000
Winter Precip 2012 x Lyme Incidence 2013	0.01005	p<0.00000
Winter Precip 2013 x Lyme Incidence 2013	0.00863	p<0.00000
Winter Temp 2012 x Lyme Incidence 2013	0.02761	p=0.00312
Winter Temp 2013 x Lyme Incidence 2013	0.01621	p=0.06311
Summer Precip 2012 x Lyme Incidence 2014	0.00591	p<0.00000
Summer Precip 2013 x Lyme Incidence 2014	0.00533	p<0.00000
Summer Precip 2014 x Lyme Incidence 2014	0.00773	p<0.00000
Summer Temp 2012 x Lyme Incidence 2014	0.04014	p<0.00000
Summer Temp 2013 x Lyme Incidence 2014	0.04015	p<0.00000
Summer Temp 2014 x Lyme Incidence 2014	0.04182	p<0.00000
Winter Precip 2012 x Lyme Incidence 2014	0.00994	p<0.00000

Winter Precip 2013 x Lyme Incidence 2014	0.00793	p<0.00000
Winter Temp 2012 x Lyme Incidence 2014	0.02457	p=0.00464
Winter Temp 2013 x Lyme Incidence 2014	0.01448	p=0.07397

In addressing aim 1, summer precipitation 2012 and 2013 were statistically significantly associated with the transformed Lyme disease incidence data for 2013. The regression coefficient for summer precipitation 2012 x Lyme disease 2013 indicates that a 0.627 percentage point increase in Lyme disease incidence 2013 was associated with a one unit increase in precipitation (mm) in summer 2012. The regression coefficient for summer precipitation 2013 x Lyme disease 2013 indicates a 0.623 percentage point increase in Lyme disease incidence 2013 associated with a one unit increase in precipitation (mm) in summer 2013. Summer precipitation 2012, 2013 and 2014 were also statistically significantly associated with the transformed Lyme disease incidence data for 2014. The regression coefficient for summer precipitation 2012 x Lyme disease 2014 indicates a 0.591 percentage point increase in Lyme disease incidence in 2014 associated with a one unit increase in precipitation (mm) in summer 2012. The regression coefficient for summer precipitation 2013 x Lyme disease incidence 2014 indicates a 0.533 percentage point increase in Lyme disease incidence 2014 associated with a one unit increase in precipitation (mm) in summer 2013. The regression coefficient for summer precipitation 2014 x Lyme disease incidence 2014 indicates a 0.773 percentage point increase in Lyme disease incidence 2014 associated with a one unit increase in precipitation in summer 2014.

In addressing aim 2, summer temperature 2012 and 2013 were statistically significantly associated with the transformed Lyme disease data for 2013. The regression coefficient for summer temperature 2012 x Lyme disease 2013 indicates a 4.119 percentage point increase in Lyme disease incidence 2013 associated with a one unit increase in temperature (°C) in summer 2012. The regression coefficient for summer temperature 2013 indicates a 4.181 percentage point increase in Lyme disease incidence associated with a one degree (°C) increase in temperature. Summer temperature 2012, 2013, and 2014 were also statistically significantly associated with the transformed Lyme disease data for 2014. The regression coefficient for summer temperature 2012 indicates a 4.014 percentage point increase in Lyme disease incidence 2014 associated with a one degree (°C) increase in temperature in summer 2012. The regression coefficient for summer temperature 2013 indicates a 4.015 percentage point increase in Lyme disease incidence 2014 associated with a one degree (°C) increase in temperature in summer 2013. The regression coefficient for summer temperature 2014 indicates a 4.182 percentage point increase in Lyme disease incidence 2014 associated with a one degree (°C) increase in temperature in summer 2014.

In addressing aim 3, winter precipitation 2012 and 2013 were statistically significantly associated with the transformed Lyme disease incidence data for 2013. The regression coefficients indicate a 1.005 and 0.863 percentage point increase in Lyme disease incidence 2013 associated with a one unit increase in winter precipitation (mm) for 2012 and 2013 respectively. Winter precipitation 2012 and 2013 were also statistically significantly associated with the transformed Lyme disease incidence data for 2014. The regression coefficients indicate

a 0.994 and 0.793 percentage point increase in Lyme disease incidence 2014 associated with a one unit increase in winter precipitation (mm) for 2012 and 2013 respectively.

In addressing aim 4, winter temperature 2012 was statistically significantly associated with the transformed Lyme disease incidence data for 2013. The regression coefficient indicates a 2.761 percentage point increase in Lyme disease incidence 2013 associated with a one degree (°C) increase in winter temperature in 2012. Winter temperature 2012 was also statistically significantly associated with the transformed Lyme disease data for 2014 and the regression coefficient indicates a 2.457 percentage point increase in Lyme disease incidence 2014 associated with a one degree (°C) increase in winter temperature 2012. Winter temperature 2013 was not statistically significantly associated with Lyme disease incidence 2013 or 2014 thus the regression coefficients were not considered.

Discussion

Lyme Disease Maps

The Lyme disease incidence maps created in ArcMap provide a visual representation of Lyme disease distribution by county for 2013 and 2014. There are counties in several states that exhibit high incidence across both years including counties in central Pennsylvania, southwestern Vermont, southeast Connecticut and Massachusetts and the coast of Maine. Knowing the many factors that influence Lyme disease such as land cover (Guerra et al., 2002), animal reservoirs and tick-human interaction points, it is expected that areas of high Lyme

disease incidence one-year would also be areas of high Lyme disease incidence in the following year just as low incidence areas one-year are expected to also be low the following year.

The Lyme disease LISA cluster maps provide one further step of interpretation of the patterns seen in the basic Lyme disease incidence maps. These maps indicate areas of spatial clustering and spatial outliers. First considering the spatial clusters, based on the possible environmental factors and the range of animal reservoirs/hosts, it can be assumed that counties in a similar spatial location would have similar incidence rates of Lyme disease. The cluster maps show the counties in red that have high incidence and are surrounded by areas that also have high incidence. The significance map associated with the cluster map shows which areas of high-high incidence are significant. In the case of 2013, central Pennsylvania and a large portion of New England had high-high clusters of which both clusters showed significance at the p=0.01 level. Due to the significant high-high cluster in these areas, one can inference that several factors in the environmental conditions, animal reservoirs, and/or human-tick interaction sphere are particularly conducive to Lyme disease. The high-high clusters for 2014 are much smaller than those for 2013, although they are still focused in the same areas: central Pennsylvania and southeastern New England including all of Massachusetts, eastern Connecticut and Rhode Island. The high-high clusters in both of these cases are surrounded by far fewer counties of low-high outliers than 2013 which begs the question of what is happening in those counties that they are experiencing such a difference in Lyme disease compared to their neighbors. Aside from possible differences in tick-human interaction, or other environmental factors that are limiting vector or reservoir activity, there is the possibility that physicians and labs are not reporting Lyme disease cases in these areas.

The cluster map for 2013 and 2014 Lyme disease incidence also shows an area of significant low-low clustering in Virginia. This cluster is also expected and indicates that the conditions for Lyme disease to occur are not ideal across that area. However, Virginia does have several counties with significant high-low spatial outliers, indicating cases of Lyme disease in an unexpected area. Limitations in reporting practices, in that cases are reported based on county of residence, not county of exposure could indicate that the cases reported in this area are imported cases (diagnosed but not acquired in that county).

Regressions

All of the combinations of weather variables and Lyme disease incidence were significant except for winter temperature 2013 when examined with Lyme disease incidence 2013 and with Lyme disease incidence 2014.

Summer precipitation 2012 and 2013 were significantly associated with Lyme disease incidence in 2013 and 2014 and summer precipitation 2014 was significantly associated with Lyme disease incidence in 2014. Although the relationship was statistically significant, the regression coefficients were quite small, ranging from 0.00533 (summer precip 2013 x Lyme 2014) to 0.00773 (summer precip 2014 x Lyme 2014) indicating a range of 0.533 to 0.773 percentage point increase in Lyme disease associated with an increase in precipitation (mm). Although some studies have found a relationship between summer precipitation and Lyme disease incidence in the same year and two years later (McCabe and Bunnell, 2004 and Subak, 2003), other studies have found that cumulative precipitation from one year prior was not associated with nymph abundance (Schulze et al., 2009). Ostfeld et al. (2006) found an

association with precipitation in the same year; however, this association was weak and indicated moderate precipitation as the highest association with nymph abundance. The findings from the regressions conducted in this study are consistent with prior findings in which a weakly significant relationship exists between summer temperature and Lyme disease incidence. Similar to the findings of McCabe and Bunnell (2004) and Ostfeld et al. (2006), the strongest association between summer precipitation and Lyme disease incidence was in the same year (summer precipitation 2014 and Lyme disease incidence 2014) indicated by the largest regression coefficient of this series.

Summer temperature 2012 and 2013 were significantly associated with Lyme disease incidence in 2013 and 2014 and summer temperature 2014 was significantly associated with Lyme disease incidence in 2014. All of the relationships, including the one and two-yeartime lag, had regression coefficients around 0.04 meaning that a 4 percentage point increase in Lyme disease incidence in the respective year was associated with a one degree (°C) increase in summer temperature. The one-year lag, from summer 2012 temperature to Lyme disease incidence 2013 and summer temperature 2013 to Lyme disease incidence 2014 was also found by Schauber et al. (2005) in New York. Ostfeld et al. (2006) also found that increased temperature from one year prior also increased the density of infected nymphs, which could indirectly suggest an increased Lyme disease risk. On the other hand, Schulze et al. (2009) did not find a significant association between summer temperature one year prior and *I. scapularis*. All of the associations found in this regression analysis were positive, which is the opposite finding of Vail and Smith (1998) who found a negative correlation between increasing temperature and nymph abundance. Of note on the positive associations found in this

regression analysis is that the average maximum temperature in summer 2012 was 27.25°C, 26.47°C in 2013 and 25.96°C in 2014 which are at or below the temperature threshold of 27°C found to decrease larval survival (Jones and Kitron, 2000). This means that the relationship between summer temperature and Lyme disease might have a different trajectory at temperatures above this threshold and thus caution should be taken in using this model to interpret the relationship above 27°C. These results suggest that southern states would have a higher incidence of Lyme disease given their warmer temperatures; however, there are other conditions necessary for Lyme disease to occur such as vector, deer and mouse abundance. The presence or lack of these other conditions in southern states can explain why Lyme disease incidence isn't high despite the higher temperatures.

Winter precipitation 2012 and 2013 were significantly associated with Lyme disease incidence in 2013 and 2014; however, the regression coefficients for all relationships are quite small ranging from 0.00793 (winter precip 2013 x Lyme disease incidence 2014 to 0.01005 (winter precip 2012 x Lyme disease incidence 2013). Based on a search of the literature, no prior studies have examined winter precipitation on its own in association with Lyme disease incidence. Although this relationship was significant, the small regression coefficients indicate a very small percentage point increase in Lyme disease incidence associated with winter precipitation in all years. A possible explanation for the positive association between winter precipitation and Lyme disease incidence is that snow pack, generally much of winter precipitation accumulation in the northeast United States, perhaps protects ticks from harsh winter temperatures. *I. scapularis* larvae, nymphs and adults enter a period of dormancy

through the winter at the soil surface in leaf litter thus increased snow cover could provide extra insulation even as winter temperatures drop (Lindsay et al., 1998).

Winter temperature 2013 was not significantly associated with Lyme disease incidence in 2013 or 2014; however winter temperature 2012 was found to have a significant association with both years of Lyme disease. A positive association between winter temperature and Lyme disease incidence makes sense with the current literature about temperature and *I. scapularis* survival. It is known that *I. scapularis* winter activity increases above an average temperature threshold of 6.5°C (Vail and Smith, 1998). It has been found that *I. scapularis* can survive short term temperatures as cold as -10°C (Lindsay et al., 1998). Under snow cover, winter temperatures at the soil surface were rarely below -4°C (Lindsay et al., 1998) even when air temperatures reached -18°C (Lindsay et al., 1998) and *I. scapularis* survival was quite high at slightly more than 80% (Lindsay et al., 1998). The regression coefficients in this study indicate that a 2.76 and 2.46 percentage point increase in Lyme disease incidence 2013 and 2014 respectively is associated with a one degree (°C) temperature increase in winter 2012. It is unclear why the one-year time lag between winter temperature 2012 and Lyme disease incidence 2013 was significant but the one-year time lag between winter temperature 2013 and Lyme disease incidence 2014 was not. A significant one-year lag can be explained by the life cycle of the tick where a milder winter in 2012 allowed more adult ticks to survive the winter and lay eggs, thus more larvae emerging in summer 2012, resulting in more nymphs the following summer when humans are susceptible. Subak (2003) also found a significant relationship between winter temperature one year prior and Lyme disease incidence; however, there was multicollinearity in this model between winter temperature and moisture therefore

the actual influence of winter temperature on Lyme disease incidence remains unclear. The two-year time lag, from winter 2012 to Lyme disease incidence 2014 could be related to a similar life cycle phenomenon as the one-year time lag. In this case, milder winter temperatures could allow more larvae to survive through the winter of 2012 (Diuk-Wasser et al., 2012), resulting in more nymphs emerging summer 2012, adults through winter 2013, larvae summer 2013 and nymphs emerging summer 2014 again. Although this is a possible explanation due to the life cycle of *I. scapularis*, there are many parts of the life cycle that rely on other external factors thus winter temperature two years prior is a very weak connection to Lyme disease incidence 2014. This weak connection is also exhibited in the regression coefficient for winter temperature 2012 x Lyme disease 2014 where a 2.457 percentage point increase in Lyme disease in 2014 is associated with a one degree (°C) increase in winter temperature.

Based on the regression analysis and existing literature, the association between temperature and precipitation and Lyme disease incidence is still unclear. The magnitude of the regression coefficients indicate that summer temperature in two years prior, one year prior and the same year could be factors to continue investigating in association with Lyme disease incidence. Existing literature supports the recommendation that summer temperature could be an indicator of Lyme disease incidence (Schulze et al., 2009, Diuk-Wasser et al., 2012, Ostfeld et al., 2006, Schauber et al., 2005) based on its effect on *I. scapularis* survival. One thing to consider when examining summer temperature is that tick survival seems to have a threshold temperature of 27°C above which, tick survival decreases (Jones and Kitron, 2000). Prior studies have found that humidity might be a better variable to consider than cumulative precipitation

(Subak, 2003, McCabe and Bunnell, 2004, Vail and Smith, 1998) when examining associations between weather and Lyme disease incidence.

Implications of Surveillance Limitations

The regression models conducted in this study were conducted on the existing Lyme disease data for 2013 and 2014; however, based on flaws in the surveillance system, this dataset was incomplete rendering this regression study limited in its conclusions. The surveillance of Lyme disease is a flawed process. The CDC estimates that only 10% of cases are reported, meaning the largest vector borne disease in the United States actually has ten times more occurring than are reported. This underreporting is due in part to the complicated reporting process and also possibly due to misdiagnosis, lack of detection or a patient not seeking care. Although under reporting does not change the geographic range of Lyme disease, the county and state level variability could be hugely different if the 90% of missing cases were reported. This large gap in available data and estimated case burden makes it difficult for researchers to study the factors influencing Lyme disease directly, instead researchers must use *I. scapularis* abundance and infected nymph density as proxies to project human risk of Lyme disease. While these methods are effective and one piece of necessary research, having complete (or more complete) Lyme disease data is a necessity for public health officials to know where to focus their resources for prevention education campaigns.

In order to improve Lyme disease estimates, physicians and laboratories need to increase their timely reporting. Adding another form for already-busy physicians to fill out can be a barrier to reporting. A large-scale study of physicians in Massachusetts found that 33.3% of

physicians who had diagnosed a Lyme disease case in the past year did not report any cases, and only 45.7% reported all of their cases. Those that didn't report all of their cases were asked why and the top three responses were that the lab reports it so they don't have to, they didn't know they had to and that they don't have time to fill out the form (JSI, 2008). Table 3 of this study highlights the complicated and inconsistent reporting process between states. The information about reporting Lyme disease was not readily available on the Department of Health websites for the states highlighted in Table 3, thus providing another disincentive for busy physicians to report cases.

Another major flaw in the current surveillance system is that Lyme disease cases are reported based on county of residence of the patient not the county the case was acquired in. This means that county-level case variability is not necessarily accurate. In figures 11 and 12, the Lyme disease incidence maps for 2013 and 2014 we can see huge variability in incidence with spatial clustering as well as counties of high incidence surrounded by counties of low incidence. The reporting flaw means that some of the counties with very high incidence might include imported cases, thus skewing their actual burden and risk of Lyme disease. It is also possible that spatial outliers, counties with low incidence next to counties of high incidence might not actually be that low depending on where people live. Most cases are acquired in June, July and August, when people tend to spend a lot of time outside, children are out of school, and people also tend to take vacations. All of these factors can increase the amount of human-tick interaction as well as the possibility that cases are being acquired outside of county of residence.

Study Limitations

This study had several major limitations. The first limitations are in regards to the raw data quality. Due to the above surveillance issues, Lyme disease case counts are not the most accurate at the county level. The temperature and precipitation was extracted from weather stations and some stations did not take measurements for all (or any) months of a given season. The weather stations are also not equally spaced across the region therefore the interpolated weather data for counties further away from a weather station are less accurate. In regards to statistical analysis, there are also several limitations. The first is that temperature and precipitation data are not independent of Lyme disease across several different years so the issue of multicollinearity needs to be considered when interpreting the results. Using a spatial regression model in GeoDa, the spatial autocorrelation of the data was able to be corrected for; however the issues of seasonality and temporal autocorrelation were not able to be corrected for due to limitations in statistical abilities. The temperature and precipitation data was normally distributed and a log transformation corrected the normality of the Lyme disease data; however, the other assumption of a regression is linearity and this data was not truly linear, exhibited by the statistically significant Breusch-Pagan test for heteroscedasticity in the spatial regression outputs. The model was used knowing linearity was an issue thus the test statistics need to be considered knowing this assumption was violated. There is also the issue of an increased risk of type 1 errors, in which the null hypothesis is rejected but it is actually true. The significance level helps to correct for this and the level of 0.05 was used for this study leaving a 5% probability of a type 1 error. However, the more regressions that are run across multiple variables, the risk of a type 1 error increases due to the fact that at least one of the

results will be significant just by chance. Given flaws in data and issues with autocorrelation, this study was the best possible estimate of the association between several different weather variables and Lyme disease incidence.

Study Implications

Being able to predict Lyme disease incidence would be hugely beneficial to public health officials. Predictive models could be beneficial to understand both the timing of high Lyme disease years and the spatial clusters of high incidence. If limited financial and human resources are an issue for a public health department, knowing where and when to concentrate resources could allow public education campaigns regarding Lyme disease to be most effective. Public health officials could also focus on physician education for signs and symptoms and treatment in high incidence areas to reduce the number of cases with severe clinical manifestations.

There are many environmental variables that can influence Lyme disease incidence both directly and indirectly, through human-tick interaction, reservoirs, hosts, host food and tick survival and abundance. Many studies have found connections between weather and climate variables and these environmental variables, however there is no consensus on the most important factors that could be used for a predictive model. A lot more work needs to be done on Lyme disease given its large public health burden and the benefits a predictive model could have.

References

2014-natstcopr-meth.pdf. (n.d.). Retrieved from https://www.census.gov/popest/methodology/2014-natstcopr-meth.pdf

An overview of the Raster Interpolation toolset. (n.d.). Retrieved April 8, 2016, from http://resources.esri.com/help/9.3/arcgisengine/java/gp_toolref/geoprocessing_with_3d_analyst/an_overview_of_the_raster_interpolation_toolset.htm

Anguita, J., Hedrick, M., Fikrig, E. (2003). Adaptation of *Borrelia Burgdorferi* in the tick and the mammalian host. *Federation of European Microbiological Societies Microbiology Reviews*. 27:493-504.

Anselin, L.(2001) Spatial Econometrics. *A companion to Theoretical Econometrics*. 310-330. Retrieved from http://web.pdx.edu/~crkl/WISE/SEAUG/papers/anselin01_CTE14.pdf

BioMedware SpaceStat Help - Interpreting univariate Local Moran statistics. (n.d.). Retrieved April 8, 2016, from https://www.biomedware.com/files/documentation/spacestat/Statistics/LM/Results/Interpreting_univariate_Local_Moran_statistics.htm

CDC - Transmission - Lyme Disease. (n.d.). Retrieved April 8, 2016, from http://www.cdc.gov/lyme/transmission/blacklegged.html

CDC.(1995). Notice to Readers Recommendations for Test Performance and Interpretation from the Second National Conference on Serologic Diagnosis of Lyme Disease. MMWR Weekly. 44(31); 590-591.

Cleveland Clinic. Heart Block. (2015). Retrieved from: http://my.clevelandclinic.org/services/heart/disorders/arrhythmia/heart-block

Conduction System Tutorial. (n.d.). Retrieved April 22, 2016, from http://www.vhlab.umn.edu/atlas/conduction-system-tutorial/atrioventricular-node-and-bundle-of-his.shtml

Data and Statistics | Lyme Disease | CDC. (n.d.). Retrieved April 8, 2016, from http://www.cdc.gov/lyme/stats/index.html

Diuk-Wasser, M. A., Hoen, A. G., Cislo, P., Brinkerhoff, R., Hamer, S. A., Rowland, M., … Fish, D. (2012). Human risk of infection with Borrelia burgdorferi, the Lyme disease agent, in eastern United States. *The American Journal of Tropical Medicine and Hygiene*, *86*(2), 320–327. http://doi.org/10.4269/ajtmh.2012.11-0395

E. M. Schauber, R. S. O. (2005). What is the best predictor of annual lyme disease incidence: Weather, mice, or acorns? *Ecological Applications*, *15*(2), 575–586.

GeodaTutorial093_7.pdf. (n.d.). Retrieved from http://www.uam.es/personal_pdi/economicas/coro/docencia/geomark/GeodaTutorial093_7.pdf

geodaworkbook.pdf. (n.d.). Retrieved from https://geodacenter.asu.edu/system/files/geodaworkbook.pdf

GHCNDMS_documentation.pdf. (n.d.). Retrieved from http://www1.ncdc.noaa.gov/pub/data/cdo/documentation/GHCNDMS_documentation.pdf

Glossary of Key Terms | GeoDa Center. (n.d.). Retrieved April 8, 2016, from https://geodacenter.asu.edu/node/390#lisa2

Guerra, M., Walker, E., Jones, C., Paskewitz, S., Cortinas, M. R., Stancil, A., … Kitron, U. (2002). Predicting the risk of Lyme disease: habitat suitability for Ixodes scapularis in the north central United States. *Emerging Infectious Diseases*, *8*(3), 289–297. http://doi.org/10.3201/eid0803.010166

Hinckley, A., Connally, N., Meek, J., Johnson, B., Kemperman, M., Feldman, K., White, J., Mead, P. (2014). Lyme disease testing by large commercial laboratories in the United States. *Clinical Infectious Diseases*. 59(5):676-681.

hood.pdf. (n.d.). Retrieved from https://geodacenter.asu.edu/pdf/hood.pdf

Interpreting Spatial Lag Models. (n.d.). Retrieved April 29, 2016, from http://rstudio-pubs-static.s3.amazonaws.com/5027_52298866e7924b18b54e5c9a0a21b450.html

Jones, C. J., & Kitron, U. D. (2000). Populations of Ixodes scapularis (Acari: Ixodidae) are modulated by drought at a Lyme disease focus in Illinois. *Journal of Medical Entomology*, *37*(3), 408–415. http://doi.org/10.1603/0022-2585(2000)037[0408:POISAI]2.0.CO;2

Lindsay, R., Barker, I., Surgeoner, G., McEwen, S., Gillespie, T., Addison, E. (1998). Survival and Development of the Different Life Stages of *Ixodes scapularis* (Acari: Ixodidae) Held within Four Habitats on Long Point, Ontario, Canada. *The Journal of Medical Entomology*, 35(3), 189-199.

Local Spatial Autocorrelation | GeoDa Center. (n.d.). Retrieved April 8, 2016, from https://geodacenter.asu.edu/node/393

LoGiudice, K., Ostfeld, R. S., Schmidt, K. A., & Keesing, F. (2003). The ecology of infectious disease: effects of host diversity and community composition on Lyme disease risk. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(2), 567–571. http://doi.org/10.1073/pnas.0233733100

Loye, J. E., & Lane, R. S. (1988). Questing behavior of Ixodes pacificus (Acari: Ixodidae) in relation to meteorological and seasonal factors. *Journal of Medical Entomology*, *25*(5), 391–398.

Lyme disease data tables | Lyme Disease | CDC. (n.d.). Retrieved April 8, 2016, from http://www.cdc.gov/lyme/stats/tables.html

Lyme disease maps | Lyme Disease | CDC. (n.d.). Retrieved April 8, 2016, from http://www.cdc.gov/lyme/stats/maps.html

Lyme disease surveillance and available data | Lyme Disease | CDC. (n.d.). Retrieved April 8, 2016, from http://www.cdc.gov/lyme/stats/survfaq.html

Mather, T. N., Nicholson, M. C., Donnelly, E. F., & Matyas, B. T. (1996). Entomologic Index for Human Risk of Lyme Disease. *American Journal of Epidemiology*, *144*(11), 1066–1069.

McCabe, G. J., & Bunnell, J. E. (2004). Precipitation and the occurrence of lyme disease in the northeastern United States. *Vector Borne and Zoonotic Diseases (Larchmont, N.Y.)*, *4*(2), 143–148. http://doi.org/10.1089/1530366041210765

McKinney, A., Day, J., Schneider, K., Robert, L., Carper, J., Cuénoud, D. (2008). Massachusetts Lyme Disease Needs Assessment: Final Report. John Snow, Inc.

Monthly Summaries of the Global Historical Climatology Network - Daily (GHCN-D) | Dataset Overview | National Climatic Data Center (NCDC). (n.d.). Retrieved April 8, 2016, from https://gis.ncdc.noaa.gov/geoportal/catalog/search/resource/details.page?id=gov.noaa.ncdc:C00841

Nelson, C., Saha, S., Kugeler, K., Delorey, M., Shankar, M., Hinckley, A. Mead, P. (2015). Incidence of Clinician-Diagnosed Lyme Disease, United States, 2005-2010. *Emerging Infectious Diseases*. 21(9):1625-1631.

Notice to Readers: Final 2014 Reports of Nationally Notifiable Infectious Diseases. (n.d.). Retrieved April 8, 2016, from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6436a8.htm

Ogden, N. H., Radojević, M., Wu, X., Duvvuri, V. R., Leighton, P. A., & Wu, J. (2014). Estimated Effects of Projected Climate Change on the Basic Reproductive Number of the Lyme Disease Vector Ixodes scapularis. *Environmental Health Perspectives*. http://doi.org/10.1289/ehp.1307799

Ostfeld, R. S., Canham, C. D., Oggenfuss, K., Winchcombe, R. J., & Keesing, F. (2006). Climate, Deer, Rodents, and Acorns as Determinants of Variation in Lyme-Disease Risk. *PLOS Biol*, *4*(6), e145. http://doi.org/10.1371/journal.pbio.0040145

Pepin, K., Eisen, R., Mead, P., Piesman, J., Fish, D., Hoen, A., Barbour, A., Hamer, S., Diuk-Wasser, M. (2012). Geographic Variation in the Relationship between Human Lyme Disease Incidence and Density of Infected Host-Seeking *Ixodes scapularis* Nymphs in the Eastern United States. *American Journal of Tropical Medicine and Hygiene.* 86(6). 1062-1071.

Peripheral neuropathy - Mayo Clinic. (n.d.). Retrieved April 22, 2016, from http://www.mayoclinic.org/diseases-conditions/peripheral-neuropathy/basics/definition/con-20019948

References | GeoDa Center. (n.d.). Retrieved April 29, 2016, from https://geodacenter.asu.edu/node/396#reg

Schulze, T. L., Jordan, R. A., Schulze, C. J., & Hung, R. W. (2009). Precipitation and temperature as predictors of the local abundance of Ixodes scapularis (Acari: Ixodidae) nymphs. *Journal of Medical Entomology*, *46*(5), 1025–1029.

Shapiro, E. D., & Gerber, M. A. (2000). Lyme Disease. *Clinical Infectious Diseases*, *31*(2), 533–542. http://doi.org/10.1086/313982

SPREG_Presentation.pdf. (n.d.). Retrieved from https://csde.washington.edu/services/gis/workshops/Resources/SPREG_Presentation.pdf

Steere, A. C., Coburn, J., & Glickstein, L. (2004). The emergence of Lyme disease. *Journal of Clinical Investigation*, *113*(8), 1093–1101. http://doi.org/10.1172/JCI21681

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. http://doi.org/10.1093/aje/kwg014

Sultan, S., Manne, A., Stewart, P., Bestor, P., Rosa, P., Charon, N., Motaleb, M. (2013). Motility is crucial for the infectious life cycle of *Borrelia burgdorferi*. *Infection and Immunity*. 81(6):2012-2021.

Tilly, K., Rosa, P. A., & Stewart, P. E. (2008). Biology of infection with Borrelia burgdorferi. *Infectious Disease Clinics of North America*, *22*(2), 217–234, v. http://doi.org/10.1016/j.idc.2007.12.013

University of Minnesota. Conduction System Tutorial: The Atrioventricular Node and Bundle of His: Specific Features (2014). Retrieved from: http://www.vhlab.umn.edu/atlas/conduction-system-tutorial/atrioventricular-node-and-bundle-of-his.shtml

US Census Bureau, D. I. D. (n.d.-a). Population Estimates. Retrieved April 8, 2016, from https://www.census.gov/popest/data/counties/totals/2014/CO-EST2014-01.html

US Census Bureau, D. I. D. (n.d.-b). Population Estimates. Retrieved April 8, 2016, from https://www.census.gov/popest/methodology/

Vail, S. G., & Smith, G. (1998). Air temperature and relative humidity effects on behavioral activity of blacklegged tick (Acari: Ixodidae) nymphs in New Jersey. *Journal of Medical Entomology*, *35*(6), 1025–1028.

Wormser, G., Dattwyler, R., Shapiro, E., Halperin, J., Steere, A., Klempner, M., Krause, P., Bakken, J., Strle, F., Stanek, G., Bockenstedt, L., Fish, D., Dumler, J.S., Nadelman, R. (2006). The Clinical Assessment, Treatment and Prevention of Lyme Disease, Human Granulocytic Anaplasmosis and Babesiosis: Clinical Practice Guidelines by the Infectious Diseases Society of America. *Clinical Infectious Diseases*. 43:1089-1134.

Wright, W., riedel, D., Talwani, R., Gilliam, B. (2012). Diagnosis and Management of Lyme Disease. *American Family Physician*. 85(11). 1086-1093.

Appendix A.

References used to develop surveillance and reporting table (Table 3)

7434.pdf. (n.d.). Retrieved from http://sos.ri.gov/documents/archives/regdocs/released/pdf/DOH/7434.pdf

case_report_form.pdf. (n.d.). Retrieved from http://healthvermont.gov/prevent/lyme/documents/case_report_form.pdf

CDRSS Information Guide, Communicable Disease Service. (n.d.). Retrieved April 29, 2016, from http://nj.gov/health/cd/cdrss.shtml#faq

cds-14.pdf. (n.d.). Retrieved from http://www.state.nj.us/health/forms/cds-14.pdf

CommunicableDiseases.pdf. (n.d.). Retrieved from http://www.health.ri.gov/forms/reporting/cases/CommunicableDiseases.pdf

Disease Reporting Forms | Bureau of Infectious Disease Control | NH Department of Health and Human Services. (n.d.). Retrieved April 29, 2016, from http://www.dhhs.nh.gov/dphs/cdcs/forms.htm

Disease Reporting Requirements/ Regulations, Communicable Disease Service. (n.d.). Retrieved April 29, 2016, from http://nj.gov/health/cd/reporting.shtml

doh-389_instructions.pdf. (n.d.). Retrieved from https://www.health.ny.gov/forms/instructions/doh-389_instructions.pdf

DPH: Lyme Disease Statistics. (n.d.). Retrieved April 29, 2016, from http://www.ct.gov/dph/cwp/view.asp?a=3136&q=399694

Epi1.pdf. (n.d.). Retrieved from http://www.vdh.virginia.gov/epidemiology/documents/pdf/Epi1.pdf

Infectious Disease Reporting: Department of Health. (n.d.). Retrieved April 29, 2016, from http://www.health.ri.gov/diseases/about/reporting/

ldcasereportform.pdf. (n.d.). Retrieved from http://www.dhss.delaware.gov/dph/epi/files/ldcasereportform.pdf

List of Reportable Diseases: Department of Health. (n.d.). Retrieved April 29, 2016, from http://www.health.ri.gov/diseases/infectious/resultsreportable.php

Lyme_CRF_2014.pdf. (n.d.). Retrieved from http://www.maine.gov/dhhs/mecdc/infectious-disease/epi/vector-borne/lyme/documents/Lyme_CRF_2014.pdf

Lyme_Disease_case_report_form.pdf. (n.d.). Retrieved from http://phpa.dhmh.maryland.gov/IDEHASharedDocuments/case_reports/Lyme_Disease_case_report_form.pdf

Lyme Disease - Infectious Disease Epidemiology Program - MeCDC; DHHS Maine. (n.d.). Retrieved April 29, 2016, from http://www.maine.gov/dhhs/mecdc/infectious-disease/epi/vector-borne/lyme/index.shtml#resourcephysicians

Lyme Disease Information. (n.d.). Retrieved April 29, 2016, from http://www.dhss.delaware.gov/dph/epi/lyme.html

Lyme Disease - Information, Communicable Disease Service. (n.d.). Retrieved April 29, 2016, from http://nj.gov/health/cd/lyme/techinfo.shtml

Lyme Disease in Vermont - Vermont Department of Health. (n.d.). Retrieved April 29, 2016, from http://healthvermont.gov/prevent/lyme/lyme_disease.aspx

lymediseasereport.pdf. (n.d.). Retrieved from http://www.dhhs.nh.gov/dphs/cdcs/documents/lymediseasereport.pdf

Maine Environmental Public Health Tracking (EPHT) Network - Public Data Portal - IBM Cognos Viewer. (n.d.). Retrieved April 29, 2016, from https://gateway.maine.gov/cognos/cgi-bin/cognosisapi.dll?b_action=cognosViewer&ui.action=run&ui.object=%2fcontent%2ffolder[%40name%3d%27CDC%20EOHP%20EPHT%20AVR%27]%2freportView[%40name%3d%27Maine%20Environmental%20Public%20Health

%20Tracking%20%28EPHT%29%20Network%20-%20Public%20Data%20Portal%27]&cv.header=false&cv.toolbar=false

notifiable-conditions.pdf. (n.d.). Retrieved from http://www.maine.gov/dhhs/mecdc/infectious-disease/epi/disease-reporting/documents/notifiable-conditions.pdf

PA Bulletin, Doc. No. 03-941. (n.d.). Retrieved April 29, 2016, from http://www.pabulletin.com/secure/data/vol33/33-20/941.html

Pennsylvania Department of Health. (n.d.). Retrieved April 29, 2016, from http://www.portal.state.pa.us/portal/server.pt?open=514&objID=557245&mode=2

reportable_disease_brochure.pdf. (n.d.). Retrieved from http://nj.gov/health/cd/documents/reportable_disease_brochure.pdf

ReportableDisease_HCP.pdf. (n.d.). Retrieved from http://phpa.dhmh.maryland.gov/IDEHASharedDocuments/ReportableDisease_HCP.pdf

reportable_disease_list.pdf. (n.d.). Retrieved from http://www.vdh.virginia.gov/epidemiology/documents/pdf/reportable_disease_list.pdf

TwoTieredTesting.pdf. (n.d.). Retrieved from http://www.maine.gov/dhhs/mecdc/infectious-disease/epi/vector-borne/lyme/documents/TwoTieredTesting.pdf

what-to-report. (n.d.). Retrieved April 29, 2016, from http://phpa.dhmh.maryland.gov/Pages/what-to-report.aspx

Appendix B.

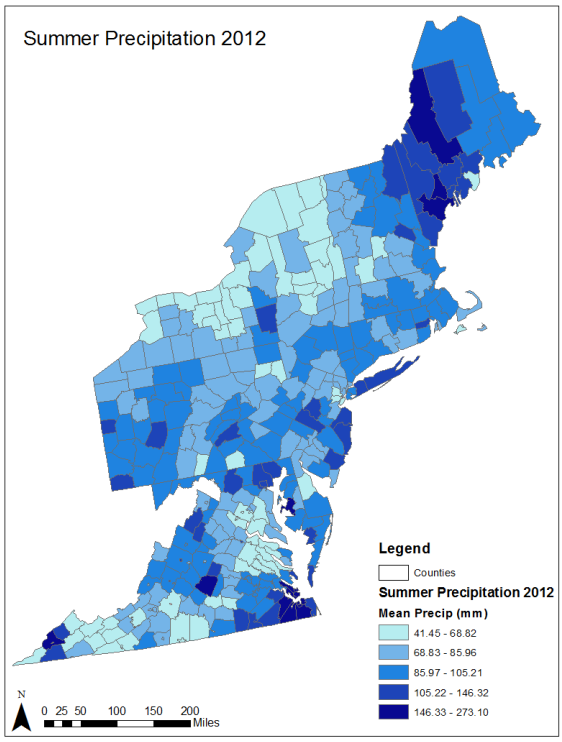

Figure 17. Summer Precipitation 2012.

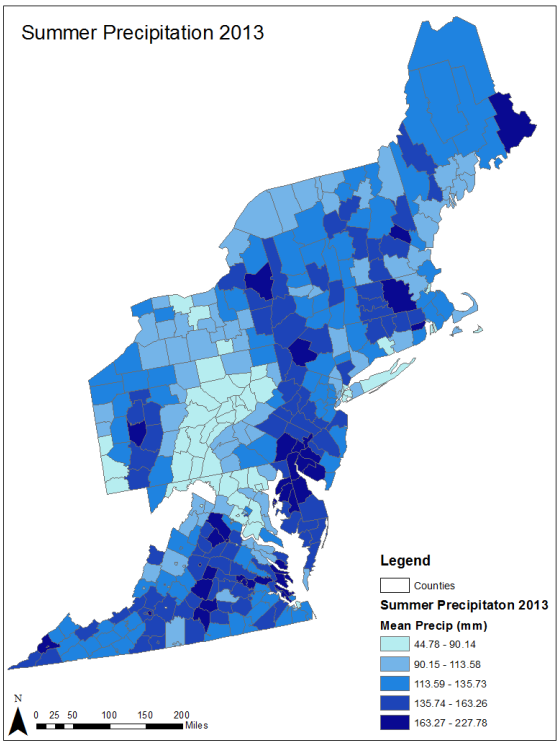

Figure 18. Summer Precipitation 2013.

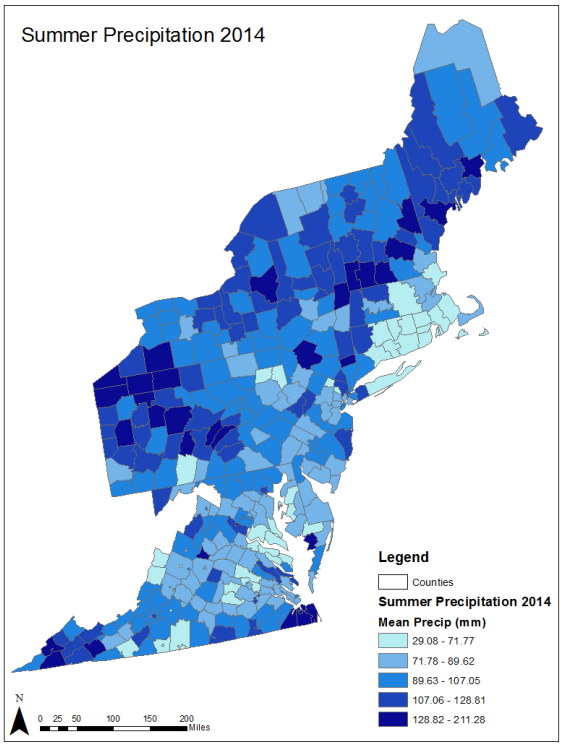

Figure 19. Summer Precipitation 2014.

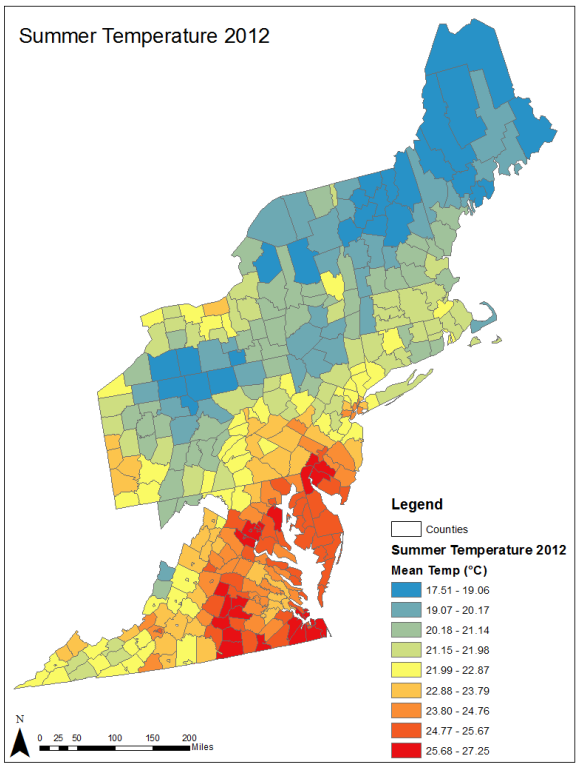

Figure 20. Summer Temperature 2012

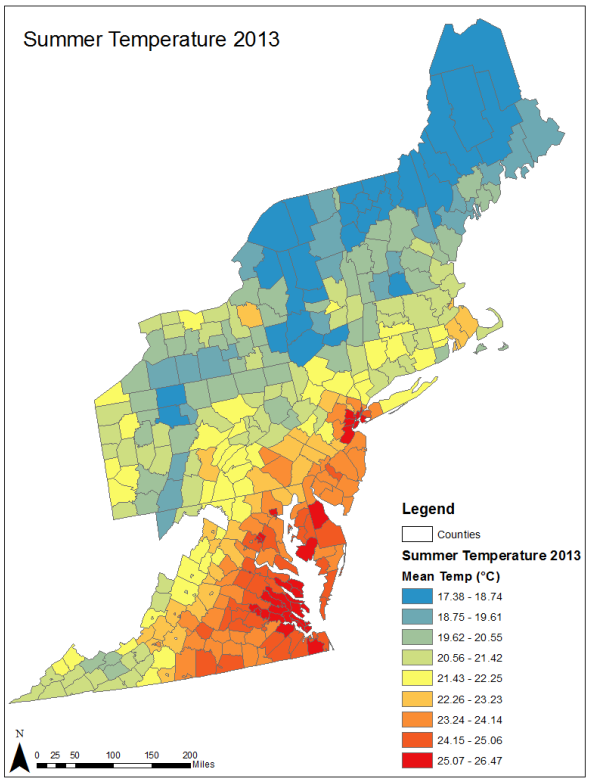

Figure 21. Summer Temperature 2013

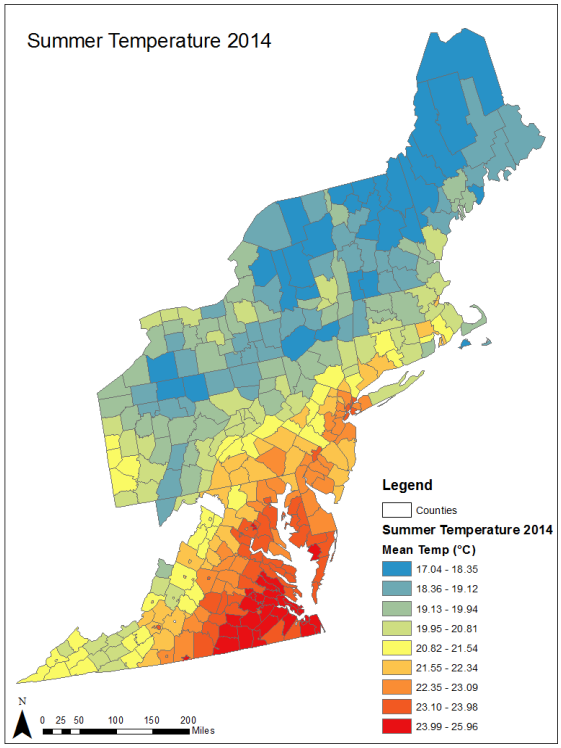

Figure 22. Summer Temperature 2014

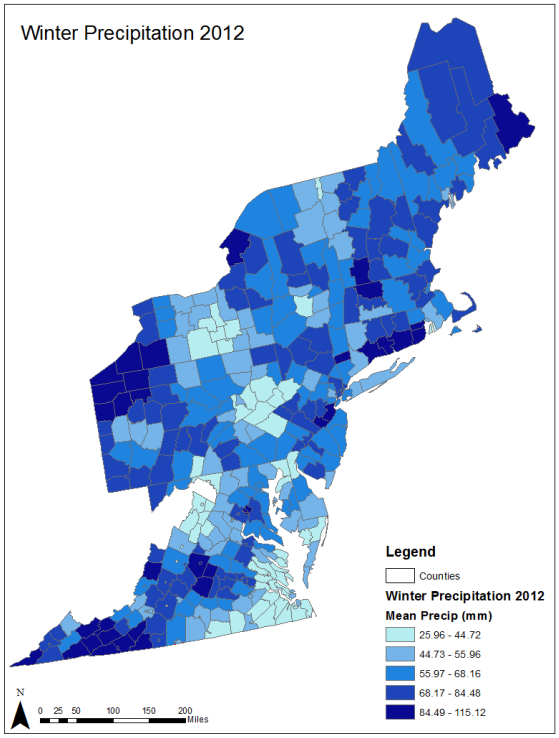

Figure 23. Winter Precipitation 2012.

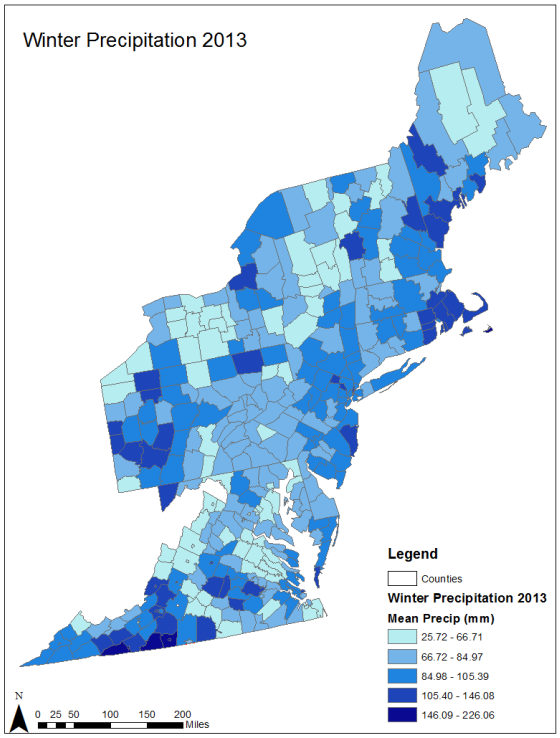

Figure 24. Winter Precipitation 2013.

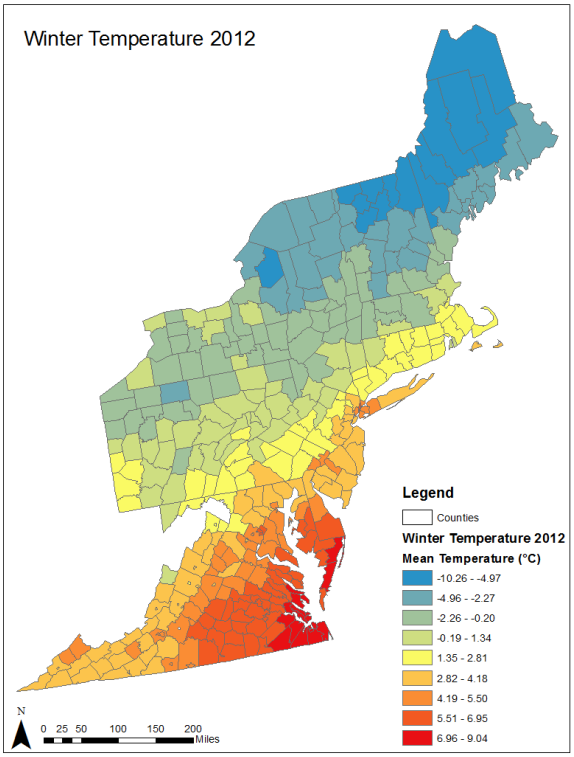

Figure 25. Winter Temperature 2012.

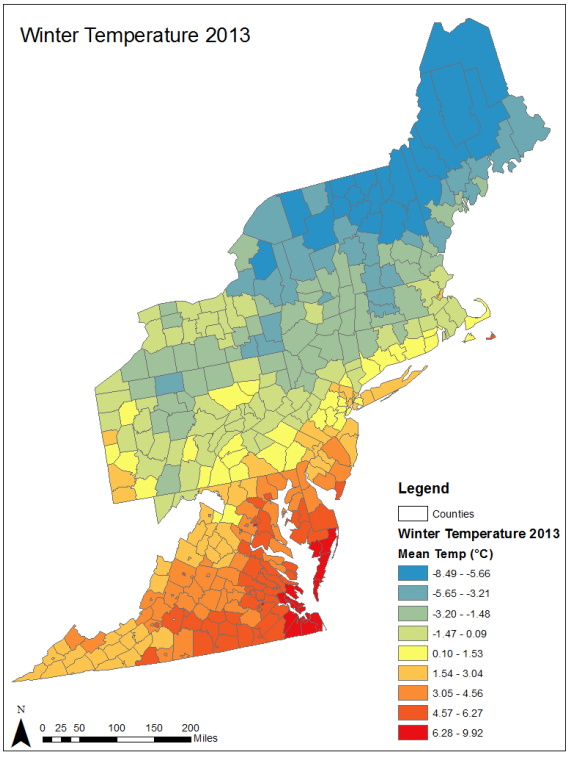

Figure 26. Winter Temperature 2013.