Ticks and Tick-borne Diseases 12 (2021) 101822

Contents lists available at ScienceDirect

Ticks and Tick-borne Diseases

journal homepage: www.elsevier.com/locate/ttbdis

Spatio-temporal modeling for confirmed cases of lyme disease in Virginia

Naresh Neupane [a,*], Ari Goldbloom-Helzner [b], Ali Arab [c]

[a] Georgetown University, Department of Biology, Washington, DC 20057, USA
[b] Princeton University, Electoral Innovation Lab, NJ 08544, USA
[c] Georgetown University, Department of Mathematics and Statistics, Washington, DC 20057, USA

ARTICLE INFO

Key words:
Spatio-temporal modeling
Hierarchical Bayesian models
Negative binomial model
Hurdle model
Zero-inflation
Disease mapping
Epidemiological modeling
Vector-borne diseases

ABSTRACT

Epidemiological data often include characteristics such as spatial and/or temporal dependencies and excess zero counts, which pose modeling challenges. Excess zeros in such data may arise from imperfect detection and/or relative rareness of the disease in a given location. Here, we studied the spatio-temporal variation in annual Lyme disease cases in Virginia from 2001-2016 and modeled the disease with a spatio-temporal hierarchical Bayesian model. Using observed ecological and environmental covariates, we constructed a predictive model for the disease spread over space and time, including spatial and temporal random effects. We considered several different models and found that the negative binomial hurdle model performs the best for such epidemiological data. Among the various ecological predictors, the North-South (V component) of winds and relative humidity significantly contributed to predicting the Lyme cases. Our model results provide important insights on the spread of the disease in Virginia and the proposed modeling framework offers epidemiologists and health policymakers a useful tool for improving disease preparedness and control plans for the future.

1. Introduction

Spatio-temporal count data, such as cases of infectious and/or rare diseases, often contain a large number of zeros. In reality, these zeros could be true zero counts (also called structural zeros) or they could be generated due to the data collection process and imperfect detection of cases (also called sampling zeros). When modeling such data, therefore, these zeros should not be ignored. Previous studies have shown that standard probability models (e.g., Poisson, Binomial, and Negative-Binomial) do not accurately model such data (Arab, 2015; Fang et al., 2016). The class of zero-modified models designed to handle such cases of excess zeros include hurdle models (Cragg, 1971; Mullahy, 1986; Hilbe, 2014) and zero-inflated models (e.g., zero-inflated Poisson; Lambert, 1992; Welsh et al., 1996). The main difference between hurdle and zero-inflated models is in the process that generates zeros; a hurdle model assumes that all the zeros are generated from a zero-generating process and all remaining observations (non-zero counts) are generated from a counting process, while the zero-inflated model assumes a mixture of two processes where one generates zeros only and the other generates both zeros and non-zeros. More specifically, zero-inflated models consider zeros to be originated from two sources: structural zeros (generated due to structural reasons, e.g., absence of a disease in an area) and sampling zeros (may arise due to chance or perhaps as a function of sampling procedures such as inability to perfectly detect a phenomenon), while hurdle models only consider structural zeros and assume a structural difference between zero and non-zero values (e.g., zeros correspond to non-prevalence of Lyme disease in an area while non-zero values represent the count of Lyme cases where and when present). Considering both of these approaches to modeling data with excess zeros, here we develop a Bayesian hierarchical model to model the confirmed cases of Lyme disease.

In recent decades, Lyme disease has been on the rise in the U.S. (Piesman, 2006; Hoen et al., 2009; Khatchikian et al., 2015), and Virginia has not been an exception. In 2001, there were a total of 100 Lyme cases in Loudoun and Fairfax counties (two highly populated counties in Northern Virginia), whereas the number of cases increased to 294 in 2016. Fig. 1 displays maps of the confirmed cases of Lyme disease in 2001 (Fig. 1a) and 2016 (Fig. 1b). This disease has also become prevalent in parts of Virginia where it used to be rare (Lantos et al., 2015). The number of reported cases in southwestern regions of the state (as in Wythe, Pulaski, Montgomery, Carroll, and Floyd counties) increased from zero in 2001 to 200 in 2016. Given the complex nature of this data, we consider the zero-modified modeling framework to model the spatial and temporal expansion of the disease. We test both hurdle and

* Corresponding author.
E-mail address: Naresh.Neupane@georgetown.edu (N. Neupane).

https://doi.org/10.1016/j.ttbdis.2021.101822
Received 6 October 2020; Received in revised form 18 March 2021; Accepted 9 August 2021
Available online 4 September 2021

zero-inflated models also because the increases in the disease cases may be due to improvements in diagnosis or reporting of the disease (Borgermans et al., 2015; Waddell et al., 2016; Sadilek et al., 2020).

Chen et al., 2006 modeled the distribution of Lyme disease in New York from 1990-2000 using a Bayesian hierarchical framework. They compared the fit of various probability models (Poisson, zero-inflated Poisson, and negative binomial). The negative binomial model provided the best . Patterns of the disease distribution were found not to be random but clustered. Likewise, several other studies arrived at similar conclusions for disease distributions in Virginia (Li et al., 2014 studied data for 1998-2011; Lantos et al., 2015a and b studied data for 2000-2014).

Arab (2015) developed a spatio-temporal model for the confirmed cases of Lyme in Illinois from 2007-2011. The study used Poisson and negative binomial models as well as their zero-modified versions and concluded that the zero-Inflated Poisson and Poisson hurdle with a regression model for the zero-inflation probability provided the best fit for the data. Here, following the hierarchical Bayesian modeling approach proposed by Arab (2015), we develop a spatio-temporal model for the confirmed cases of Lyme in Virginia. We focus on the choice of the probability distribution that best describes the data. Finally, we conduct a model comparison among various possible models to find the "best" model.

Lyme disease is caused by the bite of nymphal or adult ticks (genus *Ixodes*) that carry bacteria, *Borrelia burgdorferi* sensu lato (Tugwell et al., 1997). Green vegetation is considered the primary habitat of such ticks. As a proxy to the habitat availability, researchers often use the Normalized Difference Vegetation Index (NDVI, Estrada-Peña, 2002; Brownstein et al., 2005; Kalluri et al., 2007; Barrios et al., 2012). Following these findings, we included the effect of green vegetation in our model as a potential predictor that may explain the disease variability. In addition to NDVI, we included several environmental variables, described in the next section, that are often related to the disease.

2. Data

We used the annual confirmed cases of Lyme disease in Virginia counties and cities from 2001-2016. This data is available at the county/city level from the Centers for Disease Control and Prevention (CDC, 2018). We eliminated four counties (Arlington, Bedford, Bland, and Highland) that did not have enough data related to one of our predictors (the deer population index). We considered 3 cities (Chesapeake, Suffolk, and Virginia Beach) and 91 counties, with a total of 1504 observations. The data reported by the CDC only includes confirmed cases of Lyme (there is no data available for unreported and unconfirmed cases). Consequently, the results of our analysis should be interpreted accordingly and, therefore, it could be inherently impacted by the data collection mechanism including any issues of diseases diagnosis and reporting (i.e., access to healthcare, socio-economic factors, variability in the quality of care and diagnostics, among others).

Across all 94 cities/counties, there are 781 observations (out of a total of 1504) with non-zero confirmed cases of Lyme diseases from 2001-2016. Fig. 2 (a, b) displays the frequency and map of the disease distributions over the study period. The excessive zero counts contained in the data are evident in Fig. 2. There were 723 observations with zero cases of Lyme, and this may contain the absence of the disease and possibly unreported/unconfirmed cases, too. Likewise, there were 682 cases between 1 and 25 counts and frequency abruptly tapers down to one for intervals between 275 and 300. A few counties/cities display relatively high cases. Some northern counties (Loudoun, Fairfax, and Prince William) display a total of up to 3000 cases during the study period. These counties are also the top three most densely populated counties in Virginia. Similarly, there is another cluster of counties in the southwestern region that includes Montgomery, Pulaski, and Floyd counties, which showed up to 1000 cases over the period. In general, counties with a large number of cases are located in the northern regions. Southern counties have relatively low or zero counts (e.g., Lee and Scott counties in the southwest, and Southampton and Greenville counties in the southeast).

In order to develop a useful and ecologically meaningful modeling framework, we considered the effect of several different environmental variables (summarized in Table 1). Previous studies have shown that the disease-causing bacteria's activity increases with the increase in atmospheric humidity (Vail and Smith, 2002; Perret et al., 2003; Moore et al., 2014; Bennet et al., 2006). Following this, we incorporated relative humidity in our model. Studies have also shown that various species of deer serve as hosts of the disease-carrying ticks. The incidence of Lyme was relatively higher in regions with higher deer population (Shapiro et al., 1992; Goldstein et al., 2001; Yabsley et al., 2005; El Khoury et al., 2012; Levi et al., 2012). While deer population density is an important predictor of Lyme incidence (Shapiro et al., 1992), a deer population index is not available at an appropriate scale in Virginia. Therefore, we used data on antlered buck kill per square mile of estimated deer habitat provided by Virginia's Department of Game & Inland Fisheries as a proxy for the deer population. Several other studies have shown that the background wind circulation plays a major role in transporting vector-borne diseases (Olsen et al., 1995; Scott and Durden, 2009). Taking this into consideration, we incorporated wind as a predictor in our model. In addition, we also included precipitation and temperature as predictor variables.

We extracted environmental predictor variables from available observations or reanalysis based on the following procedure: first, we obtained the centroid of each county. Next, we created averaging regions centered at the county centroid. Rather than extracting a value that corresponds to the county centroid, we chose to average the values that lie in the enclosed region between centroid longitude $\pm 0.1^{\circ}$ and latitude $\pm 0.1^{\circ}$. With an approximation of 1-degree latitude/longitude to equal about 100 km within the latitudes that span Virginia (~ around 38°N,

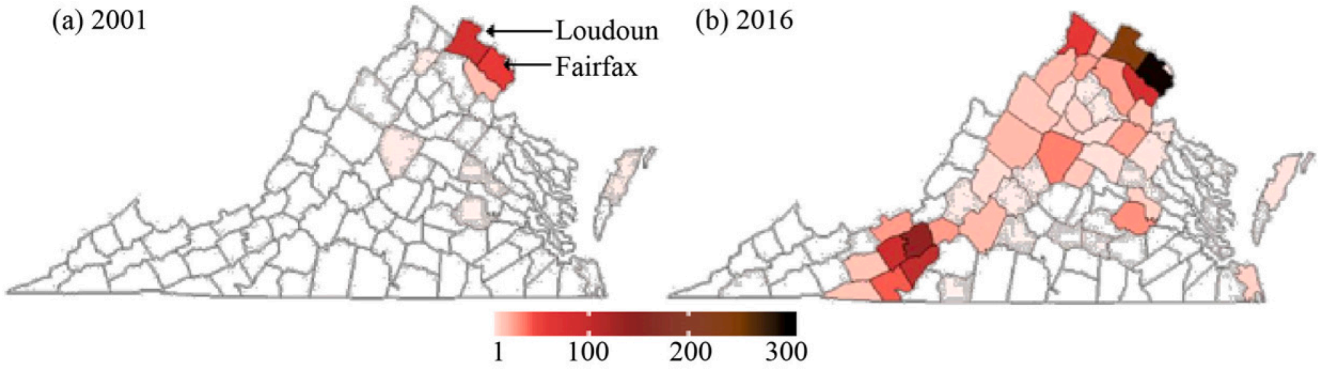

Fig. 1. Confirmed cases of Lyme disease in Virginia in (a) 2001 and (b) 2016.

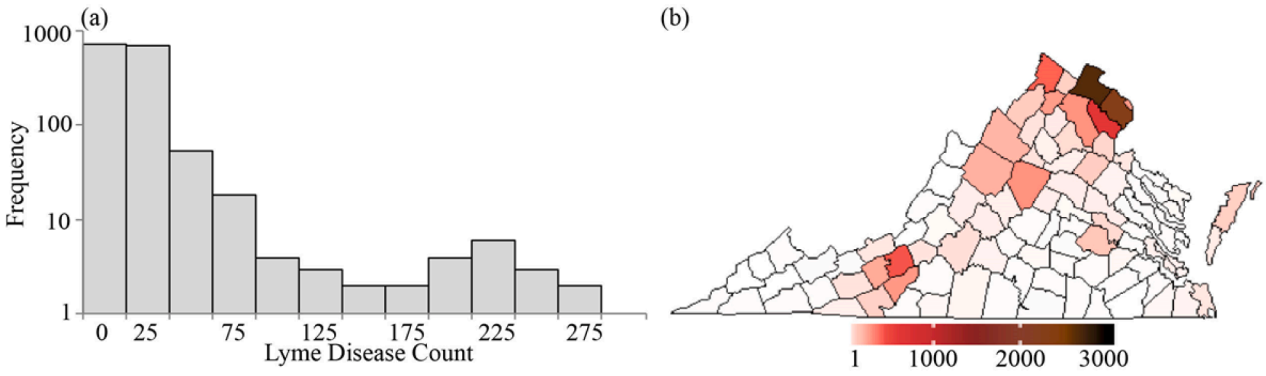

Fig. 2. (a) Confirmed cases of Lyme disease in Virginia across all years (2001 – 2016). (b) Map of total number of confirmed Lyme cases in Virginia from 2001-2016.

Table 1
Summary of predictor variables.

Variables	Source	Spatio-temporal scale
Precipitation	NARR*	Monthly, Centroid box
Temperature	"	"
Relative Humidity	"	"
East-West wind(U)	"	"
North-South wind (V)	"	"
Greenness index (NDVI)	MODIS**	"
Deer population index	CDC***	Monthly, County/city average

* NARR, North American Regional Reanalysis.
** MODIS, Moderate Resolution Imaging Spectroradiometer.
*** Centers for Disease Control .

https://www.usgs.gov/faqs), 0.2 ° equals to about 20 km (We selected this area as a county representative because most counties have land areas greater than 400 km^2; these regions were well within county geographies and avoid the problem of overlapping especially for the smaller counties/ cities). Then, for these regions, we averaged and extracted the environmental covariates on a monthly timescale.

Surface temperatures (T, °C), relative humidity (RH, %), precipitation (P, mm/day), and winds were obtained from the North American Regional Reanalysis (NARR, Mesinger et al., 2006). NARR is available from 1979 to present at 32 km spatial and 3-hourly temporal resolutions, and 3-hourly values are interpolated to monthly mean. The reanalysis precipitation values are assimilated using observations. We included both Meridional (V-component, wind component along latitude, m/s) and zonal (U-component, wind component along longitude, m/s) wind components in the model to separately model the potential effects of the latitudinal and longitudinal components of wind. These 5 types of environmental variables were averaged for each month, providing 60 predictor variables.

We used MODIS (Moderate Resolution Imaging Spectroradiometer) Aqua (Instrument: MOD13Q1)/ Terra (Instrument: MYD13Q1 product) satellite Normalized Difference Vegetation Index (NDVI) to estimate the monthly surface greenness index (Huete et al., 1999). NDVI is available at 250-m spatial and 16-day temporal resolutions from 2000 to present. This index ranges from 0 (no vegetation) to 1 (green vegetation); values are atmospherically corrected, cloud-free, continuous, and calculated from surface reflectance values. We extracted this index for each of the county/city centroids, or the nearest land point in cases where the centroid is over water. Monthly average NDVI indices provided 12 predictor variables.

Deer population index, defined as antlered buck kill per square mile of estimated deer habitat, is used as a proxy for relative deer abundance. This (obtained from Virginia's Department of Game & Inland Fisheries) is based on reports of bucks killed during the deer hunting seasons for major deer management units, collected at the county level from 2001-2016. Many cities in Virginia as well as Arlington County do not have deer management units; therefore, there is not enough data to calculate this deer population index in these areas. Consequently, we eliminate these counties/cities from our analysis.

3. Model

3.1. Model selection

Many counties had zero counts, which poses a modeling challenge in choosing an appropriate and realistic probability distribution for the data. We considered the popular probability model choices for count data namely, the Poisson, the negative binomial models, and their zero-modified versions. In particular, we compared the fit of the following probability models to the data: negative binomial, Poisson, hurdle Poisson, hurdle negative binomial, zero-inflated Poisson, and zero-inflated negative binomial. Below, we briefly review the negative binomial and hurdle negative binomial models that are commonly used for modeling over-dispersed data with excess zeros.

3.1.1. Hurdle models

A hurdle model (Hilbe, 2014) consists of two components: the first component is a binary component that models the presence/absence using a logistic regression model for the probability of a zero count. Note that we follow a parameterization that considers a non-zero value as a failure (0) whereas a zero count is treated as a success (1). The second component is a zero-truncated component, which generates non-zero values. In this component, using a generalized linear model approach, the observed counts that are greater than zero are modeled. All the zero values are removed in the truncated component of the model.

Therefore, a hurdle count model may be described as a mixture of a count generating distribution and a zero-generating process. All the zeros are generated via just one process that assumes the zeros are "structural", in our case, it means that Lyme disease is absent and, therefore, the case count is zero. For example, a Poisson hurdle model for the Y_i (with i=1, …, n) observations can be described as a mixture of a point mass at zero with probability p_i, and a zero-truncated Poisson distribution with probability $(1 - p_i)$. In our case, we consider the observed number of cases in a county at each time period is assumed to follow a Poisson hurdle distribution with parameters μ_i and p_i : $Y_i \sim PoissonHurdle(\mu_i, p_i)$ where

$$P(Y_i = 0) = p_i, \qquad 0 \le p_i \le 1$$

$$P(Y_i = y_i) = (1 - p_i)\frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!(1 - e^{-\mu_i})} \quad y_i = 1,\ 2, \cdots;\ i = 1, \cdots, n;\ \mu_i > 0.$$

The choice of the data model may be evaluated using a model selection exercise to arrive at the "best" model among Poisson, negative binomial, and their zero-modified versions.

3.1.2. Negative binomial hurdle models

The negative binomial model addresses the issue of overdispersion if present in the data and is considered as an alternative to the Poisson model when the assumption of equi-dispersion (equal mean and variance) does not hold true. Suppose variable Y_i is distributed as negative binomial, $Y_i \sim NegBinHurdle(\mu_i, \alpha, p_i)$,

$$P(Y_i = 0) = p_i, \quad 0 \le p_i \le 1$$

$$P(Y_i = y_i) = (1 - p_i)\frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{\Gamma(y_i + 1).\Gamma\left(\frac{1}{\alpha}\right)}\left(\frac{1}{1 + \alpha\mu_i}\right)^{\frac{1}{\alpha}}\left(\frac{\alpha\mu_i}{1 + \alpha\mu_i}\right)^{y_i}$$

where α the dispersion level ($\alpha > 0$), mean E (Y_i) = μ and variance Var (Y_i) = μ + α μ^2. See Hilbe (2014) and Arab (2015) for more details on zero-modified models and their properties.

3.2. Model Implementation

Following Arab (2015), we used a hierarchical Bayesian framework, and computations were implemented in R, using the Integrated Nested Laplace Approximation (INLA, http://www.r-inla.org) package (Rue et al., 2017). The INLA method is a variational Bayes approach and is often used for conducting Bayesian inference as an alternative technique to Markov Chain Monte Carlo methods (MCMC). INLA approximates the posterior distribution as opposed to the MCMC methods where they attempt to draw from the unknown posterior distribution, and thus INLA is computationally efficient (Arab, 2015; Blangiardo and Cameletti, 2015; Brown, 2015; Khana et al., 2018). In addition, INLA is efficient in handling skewed continuous positive data with excess zeros, and this also allows us for model selection criteria using Deviance Information Criterion (Spiegelhalter et al., 2002; Quiroz et al., 2015; Rue et al., 2017). Therefore, we fit our models using INLA and the Matérn covariance function as discrete indexed Gaussian Markov Random Field (GMRF), based on Stochastic Partial Differential Equations, SPDE (Rue et al., 2017). GMRF matrix is based on a finite combination of piecewise linear functions in a triangular mesh in the model domain. GMRF defines the basis weights depending upon the Matérn parameters. The INLA implementation also allows accounting for spatial effect/spatial autocorrelations, which allows spatial predictions for all missing locations within the spatial domain (Lindgren and Rue, 2015). We used the binomial link function (i.e., the logit function) and weakly informative Gaussian priors with 0 mean and 0.001 precision. We considered a Matérn covariance function to model the spatial random effects based on the distances between the centroid of counties and an autoregressive AR (1) model for the temporal effects (i.e., Year). Fig. 3 demonstrates the INLA/SPDE mesh that approximates the spatial fields. To avoid edge effect and in order to increase accuracy in calculation, the mesh is extended well outside of the Virginia state boundary. We modeled spatio-temporal patterns using different models (zero-inflated Poisson, Poisson hurdle, Poisson hurdle with probability, negative binomial, zero-inflated negative binomial, negative binomial hurdle, and negative binomial hurdle with probability). Full models incorporated precipitation, temperature, meridional winds, zonal winds, relative humidity, NDVI, and deer population index as predictor variables. We identified the best data model using Deviance Information Criterion (DIC, Spiegelhalter et al., 2002) and Watanabe Akaike information criterion (WAIC, Quiroz et al., 2015; Vehtari et al., 2017).

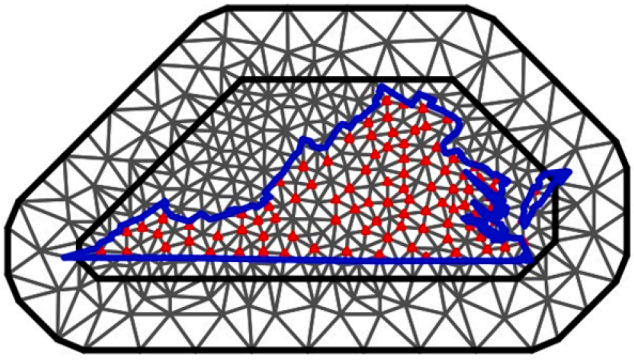

Fig. 3. INLA mesh for the study region. Blue line indicates the Virginia's state border, and the red dots indicate the county/city centroids.

Our data (cases of confirmed Lyme diseases in Virginia from 2001-2016) contains about 48% zeros (Fig. 2a). With a mean of 7.36, the expected percentage of zero counts given a standard Poisson distribution would be about 0.063%. This implies that the Poisson distribution would not provide a reasonable fit for the data given its inability to account for the excess zero counts. One reason for this discrepancy is that there are a few counties (e.g., Loudon and Fairfax Counties) with consistently large numbers of confirmed Lyme cases that have a significant effect on the mean. Furthermore, Lyme disease is relatively rare in Virginia and it is imperfectly detected, which may result in excess zero counts.

First, we extract the predictor variables on a monthly timescale, as mentioned in Section 2. The extracted variables include precipitation, temperature, meridional winds, zonal winds, relative humidity, NDVI, and deer population index.

After identifying the best data model, we also conducted variable selection (as the best subset of predictor variables) for this model. Starting with the full set of predictor variables, we eliminated the variables that were not statistically significant (i.e., 95% CI included zero) at each round until we reached the final model where all the predictor variables are significant. Finally, for the resulting best model, we conducted statistical inference and reported the results.

4. Results and discussion

Table 2 summarizes the model selection results based on DIC and WAIC. The lower the DIC and WAIC values, the better the model is (both criteria assume a combination of rewarding the best fit while penalizing the least parsimonious model). All the models were run with a full set of predictor variables, and we selected the model that has the lowest DIC/WAIC (see supplementary Table 3-9). Here, the negative binomial hurdle with regression for zero-inflation probability had the lowest DIC and WAIC values. In this case, the negative binomial model is also better than the Poisson model, which indicates the presence of overdispersion in data (See Fig. 2a). This result corroborates with Chen et al., 2006, where they found that the negative binomial performed better than the Poisson in modeling the spatio-temporal distributions of Lyme disease in the state of New York from 1990-2000.

We chose the negative binomial hurdle with the regression model for the zero-inflation probability as the best model. Two other models from the negative binomial family, the negative binomial and the zero-inflated negative binomial, had similar performances (with differences in DIC and WAIC equal to 1.87 and 1.49, less than 10%). Among these three, the simplest one, negative binomial, is more tractable for inferential purposes.

In contrast to the final model (negative binomial family) selected

Table 2
Model selection criteria results.

Spatio-Temporal Model Type	DIC	WAIC
Poisson	5586.13	5873.99
Zero-inflated Poisson	5478.36	5726.97
Poisson Hurdle	6178.77	6427.43
Poisson Hurdle with Probability	5342.98	5588.97
Negative Binomial	5040.58	5044.64
Zero-inflated Negative Binomial	5042.45	5046.13
Negative Binomial Hurdle	5847.52	5848.71
Negative Binomial Hurdle with Probability	5011.77	5010.21

here, Arab (2015) selected the Poisson family models based on DIC for Lyme diseases in Illinois. It should be noted that there are several important differences between the current study and the one conducted by Arab (2015). This study includes 17 years of data, whereas Arab (2015) used only 5 years of data and different predictor variables. Further, there are differences in the prevalence of Lyme disease between Illinois and Virginia, and the location/size of counties.

For our best model, among all the 72 predictor variables considered, only 19 variables were found to be significant. Table 3 provides a summary of all the monthly covariates that were significant in the negative binomial model. From the variable selection procedure (described in section 2), most models selected V-wind and relative humidity, and both of these predictors were significant in May. Five of the monthly V-wind variables corresponding to February, April, May, September, and December, and the relative humidity for January, February, May, June, and July were selected (see supplementary Tables 1-7 for a full list of models and variables selected). Our results corroborate results from past studies which have shown that relative humidity is positively associated with confirmed Lyme cases (Estrada-Peña, 2002; Bennet et al., 2006), and wind plays an important role in transporting disease vectors (Harrus and Baneth, 2005). Indirectly wind affects the disease's transportation via migrating birds and animals (Moen, 1976; Weisbord and Johnson, 1989; Smith et al., 1996; Jenkins et al., 2001; Brinkerhoff et al., 2011).

Three monthly variables for June (NDVI, Relative Humidity, and precipitation) were important for the disease prediction. The deer population index was not a significant predictor in any of the models we considered here. This is consistent with some past studies (Ostfeld et al., 2006; Jordan et al., 2007), however, in contrast, several past studies have found positive association between deer population density and tick populations (Wilson et al., 1990; Kitron et al., 1992; Duffy et al., 1994; Glass et al., 1994; Kilpatrick et al., 2014). This may be due to biases in our deer population index as it is based on buck kill data during hunting seasons. Another possibility is that other mammals play a larger role in disease transportation in Virginia. Studies have shown that this disease may be transported by small mammals such as mice (Strother et al., 2007; Caimano et al., 2016) and Lyme disease may be more closely associated with smaller mammals, such as red fox and coyotes, than bigger mammals like deer in some areas (Levi et al., 2012). We were unable to test these alternative ecological predictors due to the unavailability of data and given that our main focus is on environmental factors.

Based on DIC and WAIC values, we selected the negative binomial hurdle with regression for the zero-inflated probability as the best model. This model has two components: logistic regression for the zero-inflation probability to predict the probability of the binary response variable (zero vs. non-zero response; regression coefficients as displayed in Table 4a), and a zero-truncated negative binomial model for the non-zero cases (as displayed in Table 4b). Of 72 covariates, 65 variables were dropped from the logistic regression model, whereas 55 variables were dropped from the truncated model. For the logistic regression model on the zero-inflation probability, the variable selection procedure selected V-wind, specifically fall months quite often.

The wind components for months of fall were not significant in the zero-truncated model. Therefore, fall months V-wind may have different roles in the disease counts. Relative humidity (January, March, July, and September) and precipitation, especially for January, June, October, and November were important predictors. Precipitation and temperature were both found to be important for the disease prediction. For example, March temperature and four other months (January, June, October, and November) precipitation was found to be significant predictors (See Table 4b). These results are consistent with the past studies that associate Lyme disease with relative humidity (Berger et al., 2014) and precipitation (Subak, 2003; McCabe and Bunnell, 2004). The direct relationship between precipitation and Lyme disease is difficult to understand. However, there could be an indirect relationship between the two via vegetation (or greenness index, NDVI in this case). Precipitation is often positively associated with vegetation greenness (Wang et al., 2003; Tan, 2007), and vegetation supports ticks by providing habitat (Estrada-Peña, 2002; Brownstein et al., 2005; Kalluri et al., 2007; Diuk-Wasser et al., 2010; Barrios et al., 2012).

For the zero-truncated model, NDVI is an important predictor in three months (January, March, and September), and of those three, only one is common to both precipitation and NDVI (i.e., January). Past research has also shown that the herbaceous land cover is positively correlated with the disease incidence in Virginia (Seukep et al., 2015).

In addition to these environmental factors included above, there could be many socio-economic, ecological (e.g., birds and mammals), and demographic (e.g., income level, population density, as in Seukep et al., 2015) factors that may be important for Lyme disease prediction. As part of an initial exploratory data analysis, we included several demographic variables in our models (e.g., per capita income, education level, percentage of high school or college graduates, percentage of senior citizens (above the age of 65), percentage of people without health insurance, and poverty rate), but none of these variables were found to be significant for predicting the disease cases (results not shown here). Consequently, in this paper, we limited our approach to environmental factors. Also, many of the socio-economic data were available on a decadal timescale, which was not ideal in this case given the timescale of our data.

Fig. 4 displays maps for posterior predictive confirmed cases of Lyme from 2001 – 2016. Values are generally larger in northern Virginia (around Fairfax and Loudoun counties), which are the most densely populated counties in the state (USCB, 2017). In some years, mean counts are much higher in central parts (Fluvanna, Louisa, Albemarle) of the state (as in 2004, 2006). Also, there is an uptick in cases of the disease after 2007. While this could be due to the expansion of the disease or also due to improved surveillance efforts (Li et al., 2014;

Table 3
Summary of significant covariates for the log-linear regression part of the negative binomial model.

Coefficient	Mean	SD	95% CI
January NDVI	-0.20	0.06	(-0.33, -0.08)
March NDVI	0.21	0.08	(0.05, 0.37)
June NDVI	-0.40	0.16	(-0.71, -0.09)
February V-wind	-0.21	0.08	(-0.38, -0.05)
April V-wind	0.30	0.05	(0.19, 0.41)
May V-wind	-0.21	0.047	(-0.30, -0.11)
September V-wind	-0.20	0.05	(-0.30, -0.09)
December V-wind	0.47	0.09	(0.29, 0.66)
August U-wind	-0.44	0.09	(-0.61, -0.26)
October U-wind	0.35	0.06	(0.23, 0.48)
January Relative Humidity	-0.22	0.07	(-0.35, -0.09)
February Relative Humidity	0.13	0.06	(0.01, 0.25)
May Relative Humidity	-0.60	0.11	(-0.83, -0.38)
June Relative Humidity	0.73	0.13	(0.46, 0.99)
July Relative Humidity	-0.34	0.12	(-0.58, -0.10)
March Precipitation	-0.13	0.04	(-0.20, -0.06)
June Precipitation	-0.18	0.04	(-0.25, -0.11)
October Precipitation	-0.14	0.03	(-0.21, -0.07)
November Precipitation	0.12	0.04	(0.05, 0.19)

Table 4.a
Summary of significant covariates for the logistic regression part of the negative binomial hurdle model.

Coefficient	Mean	SD	95% CI
February NDVI	0.48	0.16	(0.17, 0.79)
August V-wind	0.24	0.11	(0.03, 0.45)
October V-wind	0.60	0.16	(0.28, 0.92)
November V-wind	0.57	0.21	(0.16, 0.98)
December V-wind	-1.18	0.24	(-1.66, -0.72)
August U-wind	0.67	0.22	(0.24, 1.09)
October Precipitation	0.26	0.09	(0.09, 0.43)

Table 4.b
Summary of significant covariates for the log-linear regression part of the negative binomial hurdle model.

Coefficient	Mean	SD	95% CI
January NDVI	-0.16	0.06	(-0.28, -0.03)
March NDVI	0.23	0.09	(0.06, 0.39)
September NDVI	-0.27	0.10	(-0.47, -0.07)
April V-wind	0.28	0.06	(0.16, 0.39)
May V-wind	-0.20	0.05	(-0.30, -0.11)
September V-wind	-0.19	0.05	(-0.29, -0.09)
August U-wind	-0.36	0.09	(-0.53, -0.19)
October U-wind	0.23	0.06	(0.11, 0.35)
March Temperature	-0.39	0.11	(-0.60, -0.18)
January Relative Humidity	-0.17	0.07	(-0.30, -0.04)
March Relative Humidity	-0.33	0.11	(-0.53, -0.12)
July Relative Humidity	-0.29	0.10	(-0.49, -0.09)
September Relative Humidity	0.35	0.08	(0.19, 0.51)
January Precipitation	-0.16	0.04	(-0.25, -0.08)
June Precipitation	-0.157	0.036	(-0.23, -0.08)
October Precipitation	-0.1558	0.038	(-0.09, -0.24)
November Precipitation	0.09	0.04	(0.01, 0.16)

Lantos et al., 2015; Lantos et al., 2015b). There is another region of high activity in western Virginia that includes Wythe, Pulaski, Montgomery, Carroll, and Floyd counties (as in 2008, 2009, 2010, and 2012-2016). This is consistent with the conclusions reported in the past studies (Lantos et al., 2015a; Lantos et al., 2015b, 2021). Therefore, this suggests that these regions with high activity levels should be the focus for raising awareness campaigns and providing necessary health care related to the disease. Next, we analyze the wind circulation (that is selected by both components of the hurdle model) and inference under the Bayesian framework.

Our results indicate that the Lyme cases are associated with the direction of the atmospheric flows in the low levels. Years with higher cases of Lyme in the southwestern regions were also associated with stronger southwesterly winds in September (Fig. 5), for example, see years 2010-2012 (Figs. 4, 5). Likewise, larger counts in southwestern/western parts of the state in 2006 and 2013-2016 were mostly associated with a stronger westerly flows in November (not shown). Based on these results one may hypothesize the possibility that the disease vectors may be advected by the background flows from the south-southwest in the fall. This hypothesis may be investigated in future work but it is beyond the scope of this study.

In order to better understand the overall spatial structure in the data, we fitted spatial models for the data collapsed (aggregated) over time. Fig. 6 displays the posterior mean and standard deviation of the spatial fields for both linear and logistic regression models. Based on the parameterization of the hurdle model, the posterior means for the spatial random effects for the linear regression model (for non-zero counts) and the logistic regression model (for the probability of zero counts) complement each other (see Section 3.1.1). As expected, Lyme cases were larger in the northeastern region of the state in the log-linear model which corresponds to lower variability of zeros in these areas (see posterior means in Fig. 6a and c). Fig. 6b and d show posterior standard deviations for the spatial fields for the log-linear model and the logistic regression model, respectively. As discussed in Figs. 4 and 5, regions with larger cases were mostly associated with westerly flows in

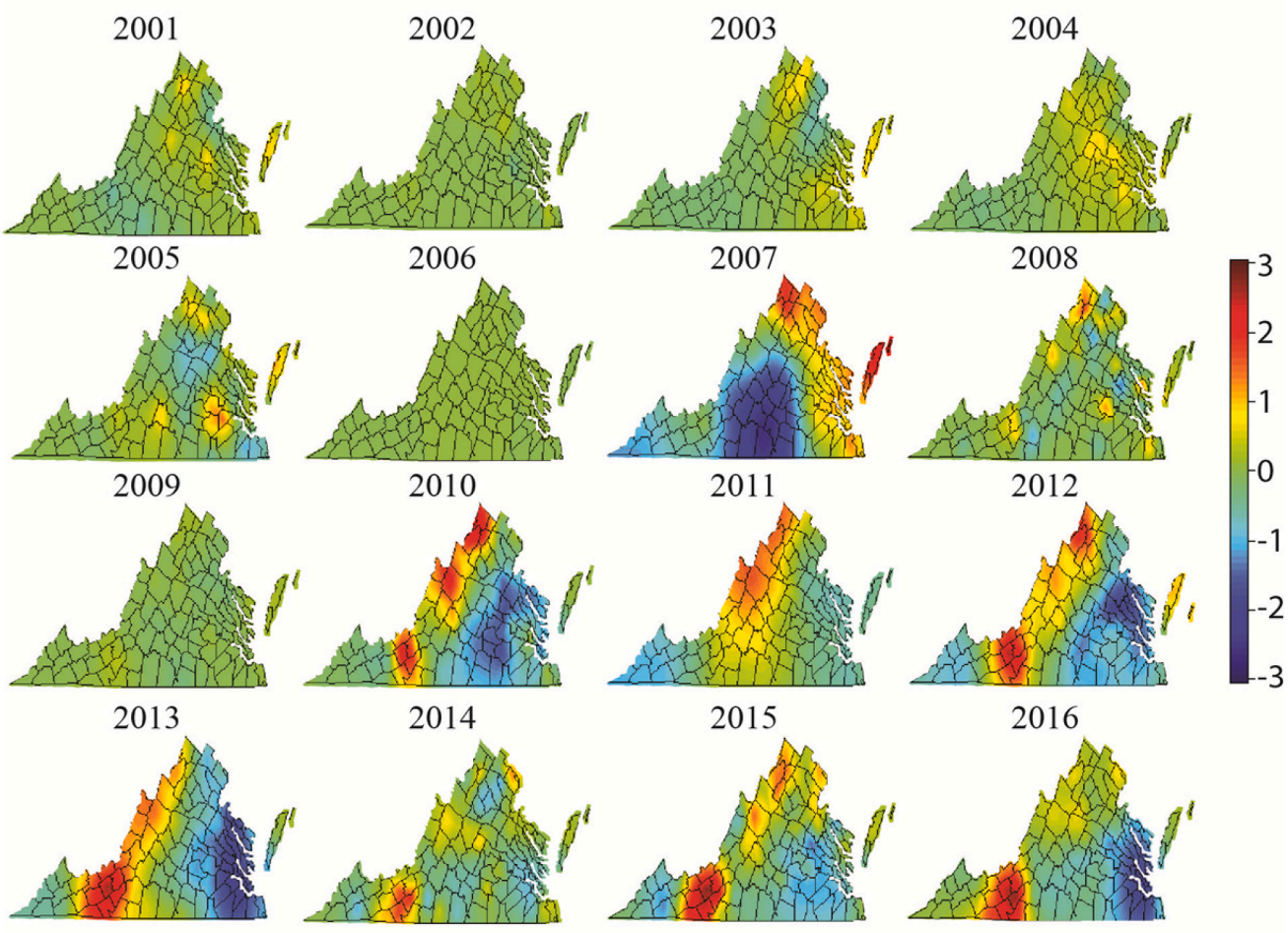

Fig. 4. Posterior means for the spatial random effects (2001-2016).

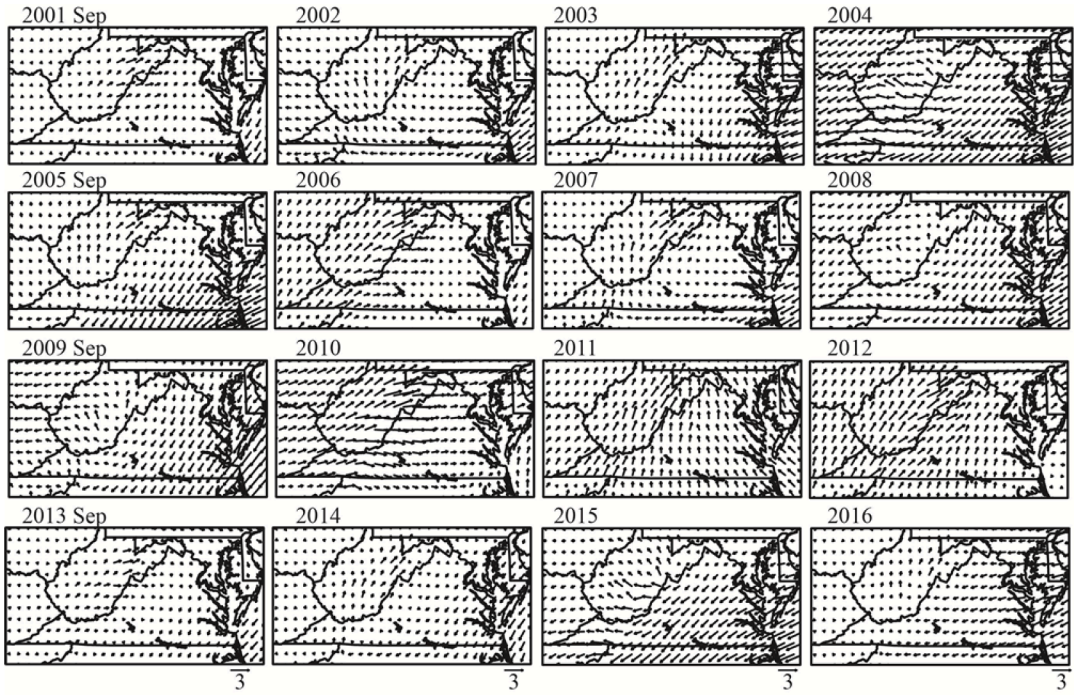

Fig. 5. Near surface (10-m above) winds (m/s) in September from the North American Regional Reanalysis for 2001-2016.

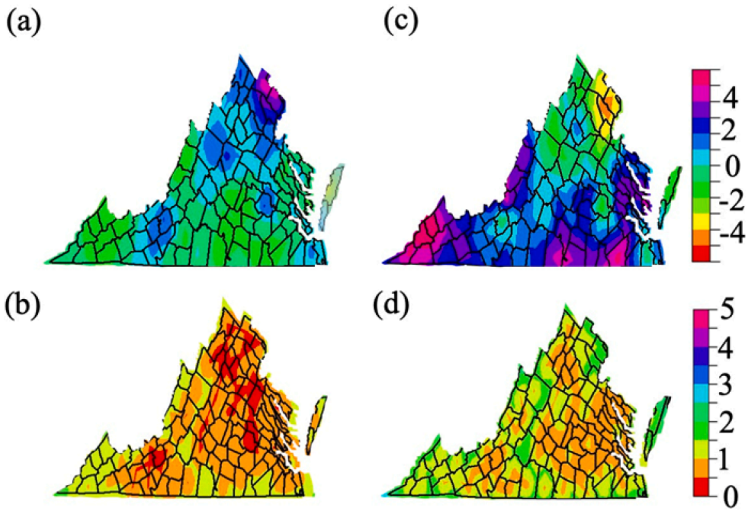

Fig. 6. Bayesian estimates for the spatial random effects (a) posterior mean (b) posterior standard deviation for the log-linear model in the negative binomial hurdle model. Similarly, (c) and (d) represent the posterior mean and posterior standard deviation for the spatial random effects for the logistic regression model in the negative binomial hurdle model, respectively.

November, as displayed in Fig. 7.

Fig. 8 (a, b) shows the posterior mean and 95% credible intervals (2.5% and 97.5% represent the lower and upper limits, respectively) of the temporal random effects. This shows an overall increasing trend of Lyme disease in Virginia after accounting for the predictors and spatial variability in the disease counts for the linear model, while the trend is decreasing for the logistic regression (probability of zero counts) model as expected by the nature of the hurdle model (i.e., over time, as the disease is spreading and increasing, the probability of zeros shows a declining trend).

5. Conclusions

We developed a spatio-temporal hierarchical Bayesian model to

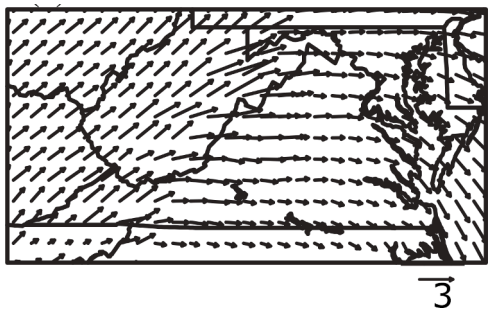

Fig. 7. Near surface (10-m above) winds (m/s) from the North American Regional Reanalysis for the 2001-2016 mean November climatology.

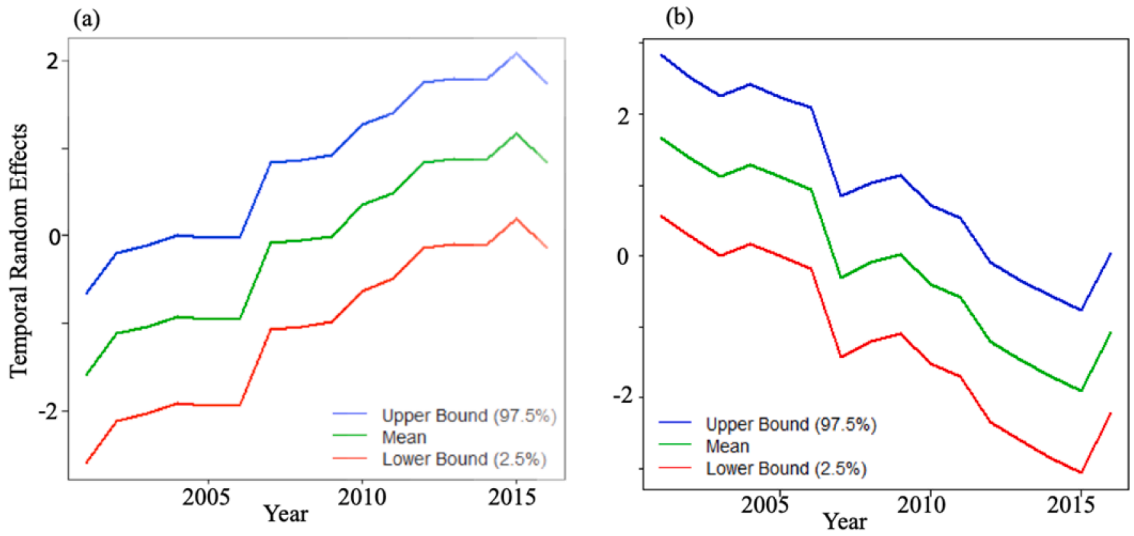

Fig. 8. Posterior mean and the 95% CI for the temporal random effects for (a) the log-linear model, (b) for the logistic regression model.

explain the spatio-temporal patterns of confirmed cases of Lyme disease in Virginia. Our model included several environmental variables such as precipitation, temperature, vegetation greenness, V-wind (winds along latitude), U-wind (winds along longitude), relative humidity, and a deer population index. We modeled the confirmed cases of Lyme using Poisson and negative binomial and their zero-modified versions (zero-inflated, hurdle, and hurdle with regression for zero-inflation probability). Among all the models we fitted, the negative binomial hurdle with regression for zero-inflation probability performed the best. Our model results identified several environmental variables (V-wind and relative humidity) as significant predictors for explaining the distribution of Lyme disease. Although the negative binomial hurdle model with logistic regression for the zero-inflation probability did perform better than the negative binomial model, it is unclear whether the marginal difference in performance outweighs the benefits of a simple model that is more tractable for inference and prediction. Our model results identified spatial "hot spots" of Lyme disease in Virginia mainly in the northern, western, and southwestern parts of the state. Also, we identified an overall increasing trend in cases of Lyme in Virginia after accounting for spatial variation in the data.

Our study has several limitations that may be addressed in future work. Mainly, we assume that the data collection and procedures remain consistent and follow standard protocols and that the cases are perfectly detected. This assumption may not be realistic and several previous studies have identified instances of incorrect diagnosis of the disease (Lantos et al., 2013, 2015b). Also, due to lack of data availability, we do not include individual-level data and the proportion of unconfirmed cases in our model which limits our ability to make inferences about the spatial and temporal patterns of expansion of Lyme disease to the county level.

Author Statement

This study is not considered for publication elsewhere. We declare that none of the co-authors (NN+ AGH + AA) * have any conflict of interest, and we did not use any animals in this study.

References

Arab, A., 2015. Spatial and spatio-temporal models for modeling epidemiological data with excess zeros. Int. J. Environ. Res. Public Health 12, 10536–10548.
Barrios, J., Verstraeten, W., Maes, P., Aerts, J.M., Farifteh, J., Coppin, P., 2012. Using the gravity model to estimate the spatial spread of vector-borne diseases. Int. J. Environ. Res. Public Health 9, 4346–4364.
Bennet, L., Halling, A., Berglund, J., 2006. Increased incidence of Lyme borreliosis in southern Sweden following mild winters and during warm, humid summers. Eur. J. Clin. Microbiol. Infect. Dis. 25, 426–432.
Berger, K.A., Ginsberg, H.S., Gonzalez, L., Mather, T.N., 2014. Relative humidity and activity patterns of *Ixodes scapularis* (Acari: *Ixodidae*). J. Med. Entomol. 51, 769–776.
Blangiardo, M., Cameletti, M., 2015. Spatial and Spatio-Temporal Bayesian Models with R-INLA. John Wiley & Sons.
Borgermans, L., Perronne, C., Balicer, R., Polasek, O., Obsomer, V., 2015. Lyme disease: time for a new approach? Br. Med. J. 351, h6520.
Brinkerhoff, R.J., Folsom-O'Keefe, C.M., Tsao, K., Diuk-Wasser, M.A., 2011. Do birds affect Lyme disease risk? Range expansion of the vector-borne pathogen *Borrelia burgdorferi*. Front. Ecol. Environ. 9, 103–110.
Brown, P.E., 2015. Model-based geostatistics the easy way. J. Stat. Softw. 63, 1–24.
Brownstein, J.S., Holford, T.R., Fish, D., 2005. Effect of climate change on Lyme disease risk in North America. Eco. Health. 2, 38–46.
Caimano, M.J., Drecktrah, D., Kung, F., Samuels, D.S., 2016. Interaction of the Lyme disease spirochete with its tick vector. Cell. Microbiol. 18, 919–927.
Center for Disease Control and Prevention: County Level Lyme Disease data from 2000–2016. Available online: https://www.cdc.gov/lyme/stats/survfaq.html (accessed 25 June 2018).
Chen, H., Stratton, H.H., Caraco, T.B., White, D.J., 2006. Spatiotemporal Bayesian analysis of Lyme disease in New York State, 1990-2000. J. Med. Entomol. 43, 777–784.
Cragg, J.G., 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica 39, 829–844.
Diuk-Wasser, M.A., Vourc'h, G., Cislo, P., Hoen, A.G., Melton, F., Hamer, S.A., Rowland, M., Cortinas, R., Hickling, G.J., Tsao, J.I., Barbour, A.G., 2010. Field and climate-based model for predicting the density of host-seeking nymphal *Ixodes scapularis*, an important vector of tick-borne disease agents in the eastern United States. Glob. Ecol. Biogeogr. 19, 504–514.
Duffy, D.C., Campbell, S.R., Clark, D, Dimotta, C, Gurney, S., 1994. *Ixodes scapularis* (Acari: *Ixodidae*) deer tick mesoscale populations in natural areas: effects of deer, area, and location. J. Med. Entomol. 31, 152–158.
Khoury, El, Y, M., Hull, R.C., Bryant, P.W., Escuyer, K.L., George, St, K, Wong, S, J., Nagaraja, A., Kramer, L., Dupuis, A.P., Purohit, T., Shah, T., 2012. Diagnosis of acute deer tick virus encephalitis. Clin. Infect. Dis. 56, 40–47.
Estrada-Peña, A., 2002. Increasing habitat suitability in the United States for the tick that transmits Lyme disease: a remote sensing approach. Environ. Health Perspect. 110, 635–640.
Fang, R., Wagner, B.D., Harris, J.K., Fillon, S.A., 2016. Zero-inflated negative binomial mixed model: an application to two microbial organisms important in oesophagitis. Epidemiol. Infect. 144, 2447–2455.
Glass, G.E., Amerasinghe, F.P., Morgan, J.M., Scott, T.W., 1994. Predicting *Ixodes scapularis* abundance on white-tailed deer using geographic information systems. Am. J. Trop. Med. Hyg. 51, 538–544.
Goldstein, E.J., Thompson, C., Spielman, A., Krause, P.J., 2001. Coinfecting deer-associated zoonoses: Lyme disease, babesiosis, and ehrlichiosis. Clin. Infect. Dis. 33, 676–685.
Harrus, S., Baneth, G., 2005. Drivers for the emergence and re-emergence of vector-borne protozoal and bacterial diseases. Int. J. Parasitol. 35, 1309–1318.
Hilbe, J.M., 2014. Modeling Count Data. Cambridge University Press.
Hoen, A.G., Margos, G., Bent, S.J., Diuk-Wasser, M.A., Barbour, A., Kurtenbach, K., Fish, D., 2009. Phylogeography of *Borrelia burgdorferi* in the eastern United States reflects multiple independent Lyme disease emergence events. Proc. Natl. Acad. Sci. 106, 15013–15018.
Huete, A., Justice, C., Van Leuwen, W., 1999. MODIS vegetation index (MOD13). Algorithm Theoretical Basis Document Version 3. NASA modis.gsfc.nasa.gov. 129.

Jenkins, A., Kristiansen, B.E., Allum, A.G., Aakre, R.K., Strand, L., Kleveland, E.J., van de Pol, I., Schouls, L., 2001. *Borrelia burgdorferi* Sensu Lato and *Ehrlichia spp*. in *Ixodes* Ticks from Southern Norway. J. Clin. Microbiol. 39, 3666–3671.

Jordan, R.A., Schulze, T.L., Jahn, M.B., 2007. Effects of reduced deer density on the abundance of *Ixodes scapularis* (Acari: *Ixodidae*) and Lyme disease incidence in a northern New Jersey endemic area. J. Med. Entomol. 44, 752–757.

Kalluri, S., Gilruth, P., Rogers, D., Szczur, M., 2007. Surveillance of arthropod vector-borne infectious diseases using remote sensing techniques: a review. PLoS Pathog. 3, 116.

Khana, D., Rossen, L.M., Hedegaard, H., Warner, M., 2018. A Bayesian spatial and temporal modeling approach to mapping geographic variation in mortality rates for subnational areas with R-INLA. J. Data Sci. 16, 147.

Khatchikian, C.E., Prusinski, M.A., Stone, M., Backenson, P.B., Wang, I.N., Foley, E., Seifert, S.N., Levy, M.Z., Brisson, D., 2015. Recent and rapid population growth and range expansion of the Lyme disease tick vector, *Ixodes scapularis*, in North America. Evolution 69, 1678–1689.

Kilpatrick, H.J., Labonte, A.M., Stafford, III, K.C., 2014. The relationship between deer density, tick abundance, and human cases of Lyme disease in a residential community. J. Med. Entomol. 51, 777–784.

Kitron, U., Jones, C.J., Bouseman, J.K., Nelson, J.A., Baumgartner, D.L., 1992. Spatial analysis of the distribution of *Ixodes dammini* (Acari: *Ixodidae*) on white tailed deer in Ogle County, IL. J. Med. Entomol. 29, 259–266.

Lambert, D., 1992. Zero-inflated Poisson regression with an application to defects in manufacturing. Technometrics 34, 1–14.

Lantos, P.M., Nigrovic, L.E., Auwaerter, P.G., Jr, Fowler, V, G., Ruffin, F., Brinkerhoff, R. J., Reber, J., Williams, C., Broyhill, J., Pan, W.K., Gaines, D.N., 2015a. Geographic Expansion of Lyme Disease in the Southeastern United States, 2000–2014. Open Forum Infectious Diseases. Oxford University Press, p. 143, 2.

Lantos, P.M., Branda, J.A., Boggan, J.C., Chudgar, S.M., Wilson, E.A., Ruffin, F., Fowler, V., Auwaerter, P.G., Nigrovic, L.E., 2015b. Poor positive predictive value of Lyme disease serologic testing in an area of low disease incidence. Clin. Infect. Dis. 61, 1374–1380.

Lantos, P.M., Brinkerhoff, R.J., Wormser, G.P., Clemen, R., 2013. Empiric antibiotic treatment of erythema migrans like skin lesions as a function of geography: a clinical and cost effectiveness modeling study. Vector-Borne Zoonotic Dis. 13, 877–883.

Lantos, P.M., Tsao, J., Janko, M., Arab, A., von Fricken, M.E., Auwaerter, P.G., Nigrovic, L.E., Fowler, V., Ruffin, F., Gaines, D., Broyhill, J., 2021. Environmental Correlates of Lyme Disease Emergence in Southwest Virginia, 2005–2014. Journal of medical entomology.

Levi, T., Kilpatrick, A.M., Mangel, M., Wilmers, C.C., 2012. Deer, predators, and the emergence of Lyme disease. Proc. Natl. Acad. Sci. 109, 10942–10947.

Li, J., Kolivras, K.N., Hong, Y., et al., 2014. Spatial and temporal emergence pattern of Lyme disease in Virginia. Am. J. Trop. Med. Hyg. 91, 1166–1172.

Lindgren, F., Rue, H., 2015. Bayesian spatial modelling with R-INLA. J. Stat. Softw. 63, 1–25.

McCabe, G.J., Bunnell, J.E., 2004. Precipitation and the occurrence of Lyme disease in the northeastern United States. Vector-Borne Zoonotic Dis. 4, 143–148.

Mesinger, F., Coauthors, 2006. North American Regional Reanalysis. Bull. Amer. Meteor. Soc. 87, 343–360.

Moen, A.N., 1976. Energy conservation by white-tailed deer in the winter. Ecology 57, 192–198.

Moore, S.M., Eisen, R.J., Monaghan, A., Mead, P., 2014. Meteorological influences on the seasonality of Lyme disease in the United States. Am. J. Trop. Med. Hyg. 90, 486–496.

Mullahy, J., 1986. Specification and testing of some modified count data models. J. Econom. 33, 341–365.

Olsen, B., Duffy, D.C., Jaenson, T.G., Gylfe, A., Bonnedahl, J., Bergström, S., 1995. Transhemispheric exchange of Lyme disease spirochetes by seabirds. J. Clin. Microbiol. 33, 3270–3274.

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, 145.

Perret, J.L., Guerin, P.M., Diehl, P.A., Vlimant, M., Gern, L., 2003. Darkness induces mobility, and saturation deficit limits questing duration, in the tick *Ixodes ricinus*. J. Exp. Biol. 206, 1809–1815.

Piesman, J., 2006. Strategies for reducing the risk of Lyme borreliosis in North America. Int. J. Med. Microbiol. 296, 17–22.

Quiroz, Z.C., Prates, M.O., Rue, H.A., 2015. Bayesian approach to estimate the biomass of anchovies off the coast of Perú. Biometrics 71, 208–217.

Rue, H., Riebler, A., Sørbye, S.H., Illian, J.B., Simpson, D.P., Lindgren, F.K., 2017. Bayesian computing with INLA: a review. Annu. Rev. Stat. Appl. 4, 395–421.

Sadilek, A., Hswen, Y., Bavadekar, S., Shekel, T., Brownstein, J.S., Gabrilovich, E., 2020. Lyme light: Forecasting Lyme disease risk using web search data. Digit. Med. 3, 1–12.

Scott, J.D., Durden, L.A., 2009. First isolation of Lyme disease spirochete, *Borrelia burgdorferi*, from ticks collected from songbirds in Ontario, Canada. North Am. Bird Bander. 34, 97–101.

Seukep, S.E., Kolivras, K.N., Hong, Y., Li, J., Prisley, S.P., Campbell, J.B., Gaines, D.N., Dymond, R.L., 2015. An examination of the demographic and environmental variables correlated with Lyme disease emergence in Virginia. Ecohealth 12, 634–644.

Shapiro, E.D., Gerber, M.A., Holabird, N.B., Berg, A.T., Jr, Feder, H, M., Bell, G.L., Rys, P. N., Persing, D.H., 1992. A controlled trial of antimicrobial prophylaxis for Lyme disease after deer-tick bites. N. Engl. J. Med. 327, 1769–1773.

Jr, Smith, P, R., Rand, P.W., Lacombe, E.H., Morris, S.R., Holmes, D.W., Caporale, D.A., 1996. Role of bird migration in the long-distance dispersal of *Ixodes dammini*, the vector of Lyme disease. J. Infect. Dis. 174, 221–224.

Spiegelhalter, D., Best, N., Carlin, B., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. J. R. Stat. Soc. 64, 583–639.

Strother, K.O., Hodzic, E., Barthold, S.W., De Silva, A.M., 2007. Infection of mice with Lyme disease spirochetes constitutively producing outer surface proteins A and B. Infect. Immun. 75, 2786–2794.

Subak, S., 2003. Effects of climate on variability in Lyme disease incidence in the northeastern United States. Am. J. Epidemiol. 157, 531–538.

Tan, S.Y., 2007. The influence of temperature and precipitation climate regimes on vegetation dynamics in the US Great Plains: a satellite bioclimatology case study. Int. J. Remote Sens. 28, 4947–4966.

Tugwell, P., Dennis, D.T., Weinstein, A., Wells, G., Shea, B., Nichol, G., Hayward, R., Lightfoot, R., Baker, P., Steere, A.C., 1997. Laboratory evaluation in the diagnosis of Lyme disease. Ann. Intern. Med. 127, 1109–1123.

USCB (United States Census Bureau), 2017. American Community Survey 5-Year Estimates. American Community Survey Office, U.S. Census Bureau. http://www.census.gov/.

Vail, S.C., Smith, G., 2002. Vertical movement and posture of blacklegged tick (Acari: *Ixodidae*) nymphs as a function of temperature and relative humidity in laboratory experiments. J. Med. Entomol. 39, 842–846.

Vehtari, A., Gelman, A., Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat. Comput. 27, 1413–1432.

Waddell, L.A., Greig, J., Mascarenhas, M., Harding, S., Lindsay, R., Ogden, N., 2016. The accuracy of diagnostic tests for Lyme disease in humans, a systematic review and meta-analysis of North American research. PloS ONE 11, e168613.

Wang, J., Rich, P.M., Price, K.P., 2003. Temporal responses of NDVI to precipitation and temperature in the central Great Plains, USA. Int. J. Remote Sens. 24, 2345–2364.

Weisbrod, A.R., Johnson, R.C., 1989. Lyme disease and migrating birds in the Saint Croix River Valley. Appl. Environ. Microbiol. 55, 1921–1924.

Welsh, A.H., Cunningham, R.B., Donnelly, C.F., Lindenmayer, D.B., 1996. Modelling the abundance of rare species: statistical models for counts with extra zeros. Ecol. Model. 88, 297–308.

Wilson, M.L., Ducey, A.M., Litwin, T.S., Gavin, T.A., Spielman, A., 1990. Microgeographic distribution of immature *Ixodes dammini* ticks correlated with that of deer. Med. Vet. Entomol. 4, 151–159.

Yabsley, M.J., Davidson, W.R., Stallknecht, D.E., Varela, A.S., Swift, P.K., Devos Jr, J.C., Dubay, S.A., 2005. Evidence of tick-borne organisms in mule deer (*Odocoileus hemionus*) from the western United States. Vector-Borne Zoonotic Dis. 5, 351–362.