2018

Investigating the Spatiotemporal Distribution of a Tick-Borne Pathogen, Ehrlichia Chaffeensis

Dylan Simpson

College of William and Mary - Arts & Sciences, dylantux@gmail.com

Follow this and additional works at: https://scholarworks.wm.edu/etd

Part of the Ecology and Evolutionary Biology Commons

Investigating the spatiotemporal distribution of a tick-borne pathogen, *Ehrlichia chaffeensis*

Dylan Tux Simpson

Kirkland, Washington

Bachelor of Science, Western Washington University, 2016

A thesis presented to the Graduate Faculty of The College of William & Mary in candidacy for the Degree of Master of Science

Department of Biology

College of William & Mary
January, 2019

APPROVAL PAGE

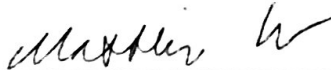This Thesis is submitted in partial fulfillment of
the requirements for the degree of

Master of Science

_____
Dylan Tux Simpson

Approved by the Committee January, 2019
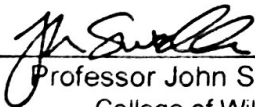
_____
Committee Chair
Associate Professor Matthias Leu, Biology
College of William & Mary

_____
Associate Professor Harmony Dalgleish, Biology
College of William & Mary

_____
Professor John Swaddle, Biology
College of William & Mary

COMPLIANCE PAGE

Research approved by

<u>Institutional Biosafety Committee</u>

Protocol number(s): IBC-2018-03-09-12629-mleu

IBC-2015-02-09-10058-mleu

IBC-2012-06-26-8010-mleu

IBC-2010-05-17-6684-mleu

Date(s) of approval: 5/29/2018

3/11/2015

6/26/2012

6/25/2010

ABSTRACT

The incidence of tick-borne diseases is on the rise in the US and around the world, due in part to emerging pathogens. However, the environmental drivers affecting these pathogens remain unclear. Most research on the topic in the US has focused on *Borrelia burgdorferi,* which causes Lyme, but it is unknown if the same conditions that affect *B. burgdorferi* also affect other pathogens, which may be carried by other ticks or reservoirs. The answer will help determine generalizable principles in tick-borne pathogen ecology, if they exist, as well as better manage for tick-borne pathogen risk in areas at risk from new and often unfamiliar pathogens. One such pathogen in the eastern US is *Ehrlichia chaffeensis*, which is transmitted by the lone star tick (*Amblyomma americanum*) and is the causative agent of monocytic ehrlichiosis, a potentially fatal illness. Here, I examine spatial and temporal variation in *E. chaffeensis* prevalence in southeastern Virginia and how this is influenced by the environment.

In Chapter 1, I used four years of data to ask how *E. chaffeensis* prevalence changed between years and how this was affected by seasonal weather patterns. Using mixed-effect models, I related *E. chaffeensis* occurrence to temperature, humidity, vapor-pressure deficit, and precipitation up to 21 months prior to sampling. Annual prevalence varied significantly from 0.9% - 3.7%, and was positively affected by temperatures during the previous winter (i.e. before the current cohort of nymphs hatched). I hypothesize this is because winter temperature affects reservoir host mortality or natality, which would in turn affect the availability of naïve reservoir hosts in the spring. Regardless of mechanism, my findings have implications for the future because winters in this region are predicted to grow warmer, which could increase *E. chaffeensis* prevalence.

In Chapter 2, I used five years of field data to ask how landscape context affects spatial variation in the prevalence of *E. chaffeensis* and interannual occupancy dynamics of its vector, *A. americanum*. Under a Bayesian framework, I created a metric- and scale-optimized model to relate *E. chaffeensis* prevalence and *A. americanum* turnover to the availability, quality, and fragmentation of habitat. Prevalence was highest and turnover was lowest in areas of low forest cover and low edge density, dominated by deciduous trees. Thus, highest disease risk is predicted in areas of forested areas that are either isolated or abutted against impermeable boundaries, both characteristic of many parks.

Many of my results highlight the complexity of tick-borne disease dynamics and the challenges inherent to the subject; some results ran counter to my predictions and *E. chaffeensis* prevalence remains rare, which makes it challenging to model. That said, my work also represents important progress in an often-neglected area of tick-borne disease ecology. To my knowledge, this is the first study to address temporal variation in *E. chaffeensis* prevalence, and is one of few studies to relate *E. chaffeensis* prevalence to landscape context at a scale relevant to the pathogen's hosts and to disease-risk management.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Chapter 1
Broad, multi-year sampling effort highlights complex dynamics of the tick-borne pathogen *Ehrlichia chaffeensis*

1.1 Introduction

Predicting the distribution of tick-borne pathogens is a vital but difficult challenge in disease ecology. Tick-borne disease risk is increasing in the United States and in much of the world (Childs and Paddock 2003, Süss 2008, Ebel 2010, Stanek et al. 2012), but complex population dynamics of ticks and their hosts (Ostfeld et al. 1995, 1996) make it difficult to accurately model pathogen dynamics. Much of the work on this subject regards the black-legged tick (*Ixodes scapularis*, Say) and *Borrelia burgdorferi*, the causative agent of Lyme disease, but other tick-borne pathogens, like *Ehrlichia chaffeensis*, have received less attention.

Ehrlichia chaffeensis is a pathogenic bacterium of increasing importance in the United States and particularly the Southeast. In humans, *E. chaffeensis* infects mononuclear phagocytes and causes a potentially fatal illness (Paddock and Childs 2003). Cases are rare relative to Lyme disease, but its incidence has doubled between 2000 and 2012 (Heitman et al. 2016). *Ehrlichia chaffeensis* is maintained in a transmission cycle between the lone star tick (*Amblyomma americanum*) as its only known vector and the white-tailed deer (*Odocoileus virginianus*) as its primary reservoir (Yabsley 2010). There is some evidence that other mammals, notably in the families Scuridae and Leporidae, also act as reservoirs for *E. chaffeensis*, but most *A. americanum* seem to be infected by *O. virginianus* (Allan et al. 2010). The apparent simplicity of this system, relative to Lyme disease, may make it easier to model and predict the prevalence of *E. chaffeensis* (i.e. proportion of ticks infected) but, to do so effectively, we must begin to consider all the dimensions over which

prevalence varies.

To date, most studies regarding *A. americanum* and *E. chaffeensis* prevalence have neglected temporal variation, conducting surveys over only one or two seasons (e.g. Harmon et al. 2015, Trout Fryxell et al. 2015, Whitlock et al. 2000). This approach overlooks the possibility of pathogen prevalence fluctuating through time, which could prove to be important to effectively predicting disease risk. For instance, Steiner et al. (1999) observed differences in tick infection rate by *E. chaffeensis* between the two seasons of their study, but were unable to conclude whether it was due to "natural variations" in yearly infection rate or to sampling error. This distinction is not trivial, and the question remains whether infection prevalence is fluctuating, and which environmental factors are involved. In this study, we begin to address these questions by examining four years of variation in *E. chaffeensis* occurrence in eastern Virginia.

The objective of this paper is, first, to describe temporal heterogeneity in *E. chaffeensis* prevalence and, second, to evaluate potential environmental drivers of that variation. We expect *E. chaffeensis* occurrence to be driven by factors influencing population density or habitat use of its primary hosts, *A. americanum* and *O virginianus*. Admittedly, there are myriad factors and interactions that can determine interannual animal population dynamics (for instance, oak masting patterns have been posited to drive dynamics within the Lyme system; Ostfeld et al. 1996), but we aimed to use environmental variables for which there are broadly continuous datasets to facilitate predictive modeling. To this end, we focused on evaluating the effects of weather, data for which are widely available and vary on relevant spatial and temporal scales.

We evaluated the effects of weather by modeling *E. chaffeensis* occurrence as a

function of seasonal weather variables across four years. We define occurrence as the detection of *E. chaffeensis* in analyzed ticks, and use the probability of occurrence as a proxy for *E. chaffeensis* prevalence in the *A. americanum* population. Specifically, we collected ticks and tested for *E. chaffeensis* occurrence at 130 plots across four years. We describe changes in *E. chaffeensis* prevalence between years and used generalized linear mixed-effect models to relate *E. chaffeensis* occurrence to temperature, humidity, vapor-pressure deficit, and precipitation while controlling for host density, and compare between candidate models using corrected Akaike's Information Criterion (AICc; Burnham et al. 2011).

1.2 Methods

1.2.1. Study Species and Area

Amblyomma americanum are hard-bodied, three-host ticks with asynchronous life stages. Peaks of activity vary between years and regions, but the general pattern is the same (Sonenshine and Levy 1971, Kollars et al. 2000): adults emerge first, peaking in the late spring, usually a month before nymphs peak in early summer. Larvae emerge in the summer and peak in late summer or early fall. This sequence likely helps to perpetuate the transmission cycle because adults, by having already fed twice in their lifetime, are the most likely to be infected (via transstadial transmission; Mixson et al. 2004, Varela-Stokes 2007) and, by emerging first, are positioned to pass the infection to naïve hosts. This is important because, through an acquired immune response, *O. virginianus* are able to purge *E. chaffeensis* from their bloodstream over time (Davidson et al. 2001, Yabsley et al. 2003, Nair et al. 2014).

We used these activity patterns to designate seasons over which to aggregate weather data in our analyses (Figure 1.1). We designate 'spring' as February through May, representing the end of quiescence and a period of increasing activity for both adults and nymphs, and 'summer' as June through October, in line with the larval activity period. Our two winters were defined differently; 'winter (*t*)' was designated as November through January, a period when all life stages of *A. americanum* are dormant, and 'winter (*t-1*)' was expanded to include October because of the schedule of *O. virginianus*' mating season and gestation period (Yarrow 2009).

This study was conducted on the Virginia and Middle Peninsulas of eastern Virginia. Our 130 plots were selected via stratified random sampling of forests on public lands and were spatially clustered within 17 distinct sites defined by a combination of management and plots' proximity (Figure A1). Because our intention was to evaluate *E. chaffeensis* prevalence across a range of environmental conditions, we erred for spatial rather than temporal replication within seasons. Plots ranged in elevation from approximately 2 – 40 m above sea level and were in mixed hardwood-coniferous forest.
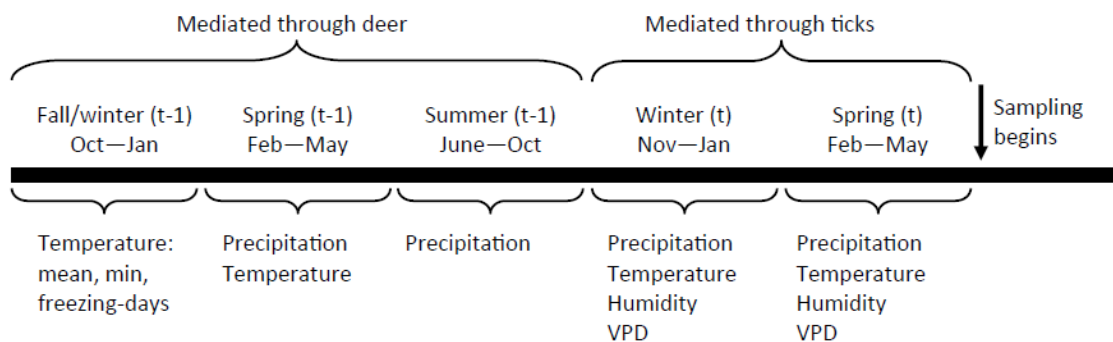
Figure 1.1: This timeline, moving from the beginning of our sampling period through the previous fall and winter, shows the seasons and variables considered in our models as they concern the nymphs used in our study. Precipitation is a cumulative measure, while temperature, humidity, and vapor pressure deficit (VPD) are daily means, except fall/winter *t-1*, for which we also include a mean daily minimum temperature and cumulative degree-days below 0 °C (i.e. freezing degree-days).

4

The most common tree species were *Pinus taeda*, *Liquidambar styraciflua*, *Liriodendron tulipifera*, *Ilex ocpaca*, *Fagus grandifola*, *Acer* spp., and *Quercus* spp.

The state of Virginia has among the highest incidence rates of ehrlichiosis in the country (Centers for Disease Control and Prevention 2018a), and the Virginia and Middle Peninsulas are an ideal region to conduct this study because there is a high density of *O. virginianus* (Virginia Department of Game and Inland Fisheries 2015) and an anthropogenically fragmented landscape indicative of high contact between ticks and humans (De Keukeleire et al. 2015, Jirinec et al. 2017).

1.2.2. Data Collection

We collected ticks during the mornings and afternoons of June through July in 2012, 2013, 2015, and 2016, visiting sites in random order. Due to a gap in funding, we did not collect ticks in 2014. We avoided sampling during or immediately following rain because, during our pilot season, we observed rainfall to reduce tick activity. We visited each plot once per year, where we collected ticks by flagging along two perpendicular 30-m transects, crossing at the 15-m midpoint (Ginsberg and Ewing 1989). We dragged a 1 m^2 square white-canvas flag along the ground and checked for ticks every 3 m. Ticks were identified to species and life stage, placed in 70% ethanol, and frozen on the same day at -80°C until extraction to prevent DNA degradation.

Our study focused on *E. chaffeensis* prevalence in nymphal ticks. This is in part because they are considered the most responsible for pathogen transmission to humans (Centers for Disease Control and Prevention 2018b), and also because our study period most closely aligns with the nymph activity period. Moreover, larvae would not have been

infected with *E. chaffeensis* because they would not have fed previously and *E. chaffeensis* is not transovarially transmitted (Yabsley 2010).

Due to inclement weather interfering with sampling and PCR runs for which tick DNA did not amplify, which sometimes occurred when testing only one or two nymphs, we did not have four years of data for every plot. Of the 130 plots, we had data for 104 in 2012, 116 in 2013, 127 in 2015, and 106 in 2016. There were 4 years of data for 81 plots, 3 years for 34 plots, 2 years for 11 plots, and only 1 year for 4 plots.

As part of an on-going study of biodiversity, we surveyed for *O. virginianus* scat (pellets) at these same sites every year from 2010 – 2016, allowing us to associate *E. chaffeensis* occurrence in a given year with deer use the previous year (when the collected ticks would have fed). We walked two 60-m transects (centered on the same point as the tick transects) and recorded number of pellets and perpendicular distance from transects. We estimated deer pellet-group density (groups ha^{-1}; group ≥ 1 pellet) in Program DISTANCE (Thomas et al. 2010), fitting detection functions on the basis of Chi-square goodness-of-fit tests and assessing fit of covariates on the basis of AIC (Burnham et al. 2011). Included covariates were: days since rain, days since above average rain (regional average within field season), Julian date, observer, and year. Our final deer-pellet DISTANCE model was a half-normal cosine ($\chi^2 = 20.5$, df = 17, P = 0.25) with Julian date and year as covariates. In our statistical analyses of *E. chaffeensis* occurrence, we consider three deer metrics: plot-level density, plot-level presence or absence, and mean site-level density (i.e. the mean of density estimates from all plots within a site).

We downloaded daily weather data from PRISM (Parameter-elevation Relationships on Independent Slopes Model; Daly et al., 2001, Oregon State University

2018). These data are produced at 4-km resolution, but we downloaded them at the point level on the basis of distance-weighted interpolation within an 8-km radius. We downloaded precipitation, temperature, and dew point values and back calculated mean vapor-pressure deficit and relative humidity from dew point and temperature values (Alduchov and Eskridge 1996). In our analyses, we included cumulative rainfall and average temperature, vapor-pressure deficit, and relative humidity because these variables describe water availability, which influences tick survival and distribution (Semtner et al. 1971, Koch 1984, Springer et al. 2015) and because temperature and precipitation have the potential to influence mammal populations via resource availability (Carroll and Brown 1977, McGinnes and Downing 1977) or fecundity (Coulson et al. 2000, Patterson and Power 2002). For the deer-breeding season (winter, $t-1$), we also included average daily-minimum temperatures and cumulative freezing-degree-days as alternative measures of metabolic demand (Fig. 1). Because of ticks' ability to avoid extreme weather events in microclimatic refugia, tick population dynamics are driven more by averages than by extremes (Ogden and Lindsay 2016) and, similarly, deer population dynamics also seem to be attune to cumulative effects (Post and Stenseth 1998). Therefore, we used sums and averages rather than individual minima or maxima.

1.2.3. Molecular Analyses

Nymphal tick and *E. chaffeensis* bacterial DNA was extracted using a DNeasy Blood and Tissue Kit (QIAGEN, Valencia, CA). Ethanol-fixed nymphal ticks from each plot were placed in screw-cap tubes with 1 mm glass beads for bead-beating in an Omni Bead Ruptor Homogenizer (Omni International, Kennesaw, GA), and DNA was extracted

following the manufacturer's protocol. The DNA was eluted in 100 μL buffer and stored at 4 °C short-term, and -80 °C long-term. Ticks were aggregated at the plot level (following previous studies, e.g. Whitlock et al. 2000, Mixson et al. 2004, Wright et al. 2014), and analyzed in groups of up to 20 ticks. When greater than 20 ticks were collected during a plot-visit (8% of observations, N = 452), a random subsample of 20 ticks were analyzed. Across all years, we used endpoint polymerase chain reaction and gel electrophoresis to determine presence of *E. chaffeensis,* and validated results of this analysis with real-time polymerase chain reaction for 2013 samples.

Presence of tick and *E. chaffeensis* DNA in extracts was confirmed using a PCR targeting species-specific regions of the 16s rRNA gene. For ticks, we used primers 16S+1: 5'CTGCTCAATGATTTTTTAAATTGCTGT-3' and 16S-1: 5'-GTCTGAACTCAGATCAAGT-3' targeting a 454-bp amplicon (Macaluso et al. 2003, Nadolny et al. 2011) and for *E. chaffeensis* we used primers HE1: 5'- CAATTGCTTA TAACCTTTTGGTTATAAAT-3' and HE3: 5'-TATAGGTACCGTCATTATCTTCCCTAT-3', targeting a 389-bp amplicon (Anderson et al. 1992, Stromdahl et al. 2000). The total PCR reaction volume was 20 μL consisting of 5 μL extracted DNA and 15 μL reaction mix, including 10 μL EconoTaq PLUS GREEN 2x Master Mix (#30033, Lucigen, Middleton, WI), 0.8 μL each forward and reverse primers (10μM) and 3.4 μL H_2O. PCRs were performed in a 2720 Thermal Cycler (Life Technologies, Grand Island, NY). The cycle parameters for the tick-specific PCR consisted of an initial step at 95°C for 4 min, 35 cycles of 95°C, 50°C, 68°C for 1 min each, followed by a final 10 min step at 68°C. The presence of *E. chaffeensis* was tested using cycle conditions of 3 cycles of 94°C for 60 s, 55°C for 120 s, 72°C for 90 s, 28 cycles of 94°C

for 30 s, 55°C for 35 s, 72°C for 40 s followed by 72°C for 7 min.

All PCR products were separated on a 1% agarose gel using either GelRed Nucleic Acid Stain (10 μL/100 mL) (RGB-4103, Phenix Research Products, Candler, NC) or ethydium bromide (5 μL/100 mL) (Thermo Fisher Scientific, Waltham, MA). Each gel included a negative (no DNA) control to test for contamination.

PCR products were spot checked for accuracy by sequencing to confirm species. Samples were purified and prepared for sequencing using a PCR/Gel Extract Mini Prep Kit (#IB47020, IBI Scientific, Peosta, IA). Sequencing was conducted by Lidia Epp at the College of William & Mary Molecular Core Facility.

1.2.4. Data Analysis

We conducted mixed-effect regression analyses in R (R Core Team 2017), using the lme4 package (Bates et al. 2015) to model *E. chaffeensis* occurrence at the plot level via logistic regression. Models were conditional on the presence of ticks (i.e. removing plot-visits where no ticks were found) and controlled for the abundance of ticks (i.e. including $\log_e$-transformed tick abundance as a covariate in all models). Candidate covariates were: mean and minimum daily temperature, mean vapor-pressure deficit and humidity, cumulative precipitation and freezing degree-days, aggregated for five seasons (Fig. 1), three measures of deer use during the previous summer (presence/absence and estimated plot- and site-level pellet-group density), and survey date, which we included in case there is within season variation in *E. chaffeensis* prevalence. We began by including nested site and plot random-intercepts, but these were removed if they did not explain any variation in the null model.

To select between models, we held random-effect structure constant and screened variables individually on the basis of AIC_c (Zuur et al. 2009, Burnham et al. 2011). We then considered multiple-regression models combining variables performing better than the null and within cumulative AIC_c weight of 0.95. We opted for this bottom-up approach to model selection because, due to the exploratory nature of our analysis, we wanted to reduce the total number of models compared. Especially because it was unknown which, if any, variables were important, we felt that a top-down approach using a global model would produce a large candidate set, which would increase the risk of spurious, overfit results (Anderson 2008). We also screened covariates for collinearity using Pearson's correlation and did not include pairs for which $|r| \geq 0.70$ (Mela and Kopalle 2002).

We estimated and report regression and prevalence estimates as follows. For regression estimates, we calculated 95% confidence intervals using likelihood profiles. For prevalence, we used a Microsoft Excel add-in (Biggerstaff 2009) to calculate bias-corrected maximum likelihood estimates with skew-corrected 95% confidence intervals. We used this program for both annual point estimates and to test for the differences between years. We considered differences significant if the confidence interval for the difference in prevalence between two years did not include 0. When reporting summary statistics of field observations, we simply report the mean ($\pm$ SD) unless otherwise noted.

Because the focus of this research was to evaluate variation in *E. chaffeensis* prevalence at broad spatial extents, requiring spatial over temporal replication, we did not include analyses of tick abundance in this paper. Our preliminary analyses revealed small effect sizes that were inconsistent between life stages. While this could represent real lack of effect or consistency, we find it just as likely that our data are not robust enough for

modeling because we did not replicate plot-level samples within years (Dobson 2013).

1.3 Results

Ehrlichia chaffeensis prevalence was variable, with significant differences between years 2012 and 2013, and between 2013 and 2015 (Figure 1.2). On the basis of 2472 nymphs, estimated prevalence was 0.89% in 2012 (n = 591), 3.7% in 2013 (n = 700), 0.93% in 2015 (n = 550), and 2.2% in 2016 (n = 631). Across years, the number of plots testing positive for E. chaffeensis were 5 of 104 (4.8%), 23 of 116 (19.8%), 5 of 127 (3.9%), and 12 of 106 (11.3%), in 2012, 2013, 2015, and 2016, respectively. Prevalence also varied spatially; no plot was positive every year it was surveyed and, of the 39 plots that did test positive, only 5 were positive more than once (Figure S1). Aggregating at the site level also shows high turnover. While 88% of sites (n = 17) tested positive at least once, 35% tested positive only once, 24% twice, 24% three times, and only one site tested positive across all four years. It is also worth noting that the only two sites to never test positive were the two northern most sites, on the Middle Peninsula. Although these plots also had among the lowest tick abundances, they had collectively more ticks than five other sites that all tested positive.

Interannual variation in tick abundance and deer pellet-group density was much less dramatic than E. chaffeensis prevalence, although within-year variation of deer pellet-group density was very high. Mean nymph abundance per plot was 8.5 (± 20.9) in 2012, 10.7 (± 25.3) in 2013, 6.5 (± 18.5) in 2015, and 8.0 (± 13.4) in 2016, with nymphs being found on 75% of plot-visits. Mean adult abundance per plot was 0.50 (±1.2) in 2012, 0.70
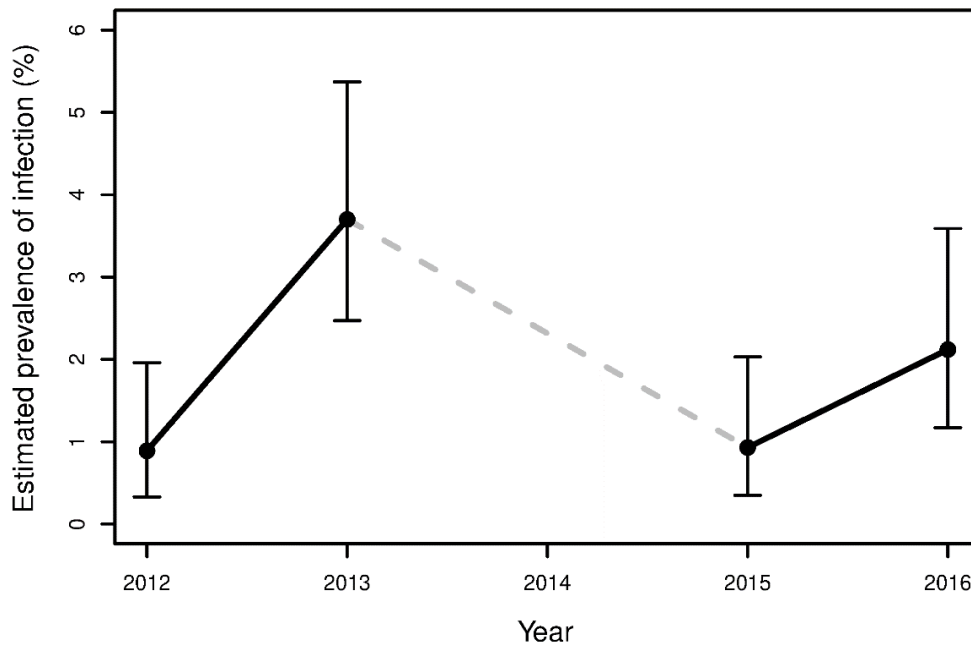
Figure 1.2: Estimated prevalence of *Ehrlichia chaffeensis* in *Amblyomma americanum* nymphs, aggregated across all plots, during the four years of our study. Error bars represent skew-corrected 95% confidence intervals (Biggerstaff 2009).

($\pm$1.8) in 2013, 0.52 ($\pm$0.91) in 2015, and 0.36 ($\pm$1.0) in 2016, with adults found on 29% of plot visits. Across years, site-level means ranged from 0.17 ($\pm$0.41) to 95.5 ($\pm$132.23) nymphs per plot and 0.0 ($\pm$0) to 3.0 ($\pm$4.64) adults per plot. Because we did not analyze all ticks found in large clusters (i.e. >20 ticks), the mean number of nymphs analyzed each year were slightly different, being 8.0 ($\pm$6.4), 7.4 ($\pm$6.3), 5.7 ($\pm$5.5), and 8.4 ($\pm$6.8) nymphs per plot in 2012, 2013, 2015, and 2016, respectively. Mean plot-level deer pellet-group density was 86 groups ha^{-1} ($\pm$150.5) in 2011, 89 groups ha^{-1} ($\pm$128.6) in 2012, 99 groups ha^{-1} ($\pm$193.5) in 2014, and 82 groups ha^{-1} ($\pm$157.1) in 2015. Across years, site-level means ranged from 0.0 – 382 groups ha^{-1}.

1.3.1. *Ehrlichia chaffeensis* occurrence models

Occurrence of *E. chaffeensis* was best explained by previous-winter (i.e. *t-1*)

temperature, with warmer winters corresponding to higher probability of occurrence. While other models performed better than the null, the variables representing winter temperature held >99% of the weight of evidence. Of these, average winter temperature and freezing degree-days held 97% cumulative AIC_c weight, and so only these were considered in a multiple-regression model, which was then added to the candidate model set (Tables 1.1, A1).

In the final candidate model set, the top model included both mean temperature and freezing degree-days during the previous winter (October – January, *t-1*), holding 56% AIC_c weight. Given the variation in temperature within our dataset, this model predicted mean temperature to cause *E. chaffeensis* occurrence probabilities of 0.033 – 0.14, and freezing-degree-days to cause a range of 0.0057 – 0.12. The simple-regression models of these two variables fell not far behind, with mean temperature $\Delta AIC_c = 1.64$ (AIC_c weight = 0.25) and freezing degree-days $\Delta AIC_c = 2.28$ (AIC_c weight = 0.18). In each case, the magnitude and precision of slope estimates increased relative to the multiple-regression model, likely due to moderate collinearity between mean temperature and freezing degree-days (Pearson's r = -0.47). Mean daily minimum temperature was the fourth model, but performance dropped considerably relative to the others ($\Delta AIC_c = 7.93$, AIC_c weight = 0.01). Each of these variables were far more important than tick abundance alone ($\Delta AIC_c = 15.27$, AIC_c weight < 0.01). Date, deer presence/absence, and plot-level deer density were unimportant explanatory variables, with ΔAIC_c values of 14.30, 14.17 and 14.56, respectively, and having performed only slightly better than the tick-only model (difference in $AIC_c \leq 1.1$, Table S1). Site-level deer density performed worse than the tick-only model ($\Delta AIC_c = 16.94$, difference from ticks-only model = -1.67).

Table 1.1: Abridged summary of the *E. chaffeensis* model set. All models are conditioned on tick presence and include tick abundance (except the null). Candidate variables were mean and minimum temperature, freezing degree-days (FDD), cumulative precipitation, mean vapor pressure deficit, mean relative humidity, estimated deer-pellet density, and deer-pellet presence/absence. Regression parameters are derived via logistic-regression of centered and scaled data. Confidence intervals represent the profiled likelihood CI's for the slope (β), except the null, for which they represent the intercept's CI. Ellipses (…) mark where models were omitted. For the complete model set, see Supp. Table S1.

Model	β_0	β_i	Lower 95% CI	Upper 95% CI	ΔAIC_c	AIC_c weight
Mean winter temp (*t-1*) + FDD (*t-1*)	-2.6	0.45 -0.50	0.026 -1.09	1.45 0.009	0.0	0.56
Mean winter temp (*t-1*)	-2.5	0.71	0.35	1.12	1.64	0.25
Winter FDD (*t-1*)	-2.51	-0.85	-1.40	-0.39	2.28	0.18
Min winter temp (*t-1*)	-2.41	0.54	0.19	0.93	7.93	0.01
…						
Deer presence/absence	-1.91	-0.60	-1.28	0.06	14.17	<0.01
Survey date	-2.26	-0.28	-0.62	0.04	14.30	<0.01
Deer-pellet density	-2.23	-0.32	-0.82	0.05	14.56	<0.01
…						
Ticks only	-2.23	0.51	0.13	0.92	15.27	<0.01
…						
Null	-1.92	-	-2.50	-1.58	20.30	<0.01

Random-effect estimates in the *E. chaffeensis* model showed negligible variation to be attributable to the site-level aggregation. Therefore, our analyses included only plot-level random-effects. In the top four models, plot-intercept standard deviation estimates ranged from 0.49 – 0.60 (95% CI: 0.0 - 1.4).

1.3.2. Weather

Aggregating across all years (2011-2016) and plots, temperature and vapor-pressure deficit displayed clear seasonal trends, while precipitation and relative humidity were relatively stable (Figure A2). Temperature rose after January (mean = 3.7 ˚C) and through July (mean

= 27.0 °C) before declining again through January. Vapor-pressure deficit followed the same trend, correlating strongly with temperature (Pearson's r = 0.94). Precipitation tended to be higher in the summer (June – October) than in other months, but there was much variation between plots and years. Relative humidity tended to peak in September and was lowest January – April, but here, too, there was much variation between plots and years. Freezing degree-days measured during October – January, accumulated most heavily in January, and none were recorded in October (Figure A3a). The number of freezing degree days also varied between years (coefficient of variation = 82%), with the most being recorded in 2014, and the fewest in 2012 (Figure A3b). Further summary of weather during the seasons defined in our models can be found in Table A2.

1.4 Discussion

We found that *E. chaffeensis* prevalence can vary significantly between years, suggesting that robust investigation of prevalence and distribution of this pathogen needs to span both space and time, and that effective modeling of disease risk must begin to account for temporal dynamics. Just as importantly, we also found the spatial distribution of *E. chaffeensis* to be inconsistent between years, with only 5 of our 130 plots testing positive multiple years. Even when plots are aggregated at the site level, *E. chaffeensis* occurrence was inconsistent between years, an observation corroborated by the fact that negligible variation was attributed to the site-level random-intercept. That said, it is noteworthy that *E. chaffeensis* was never detected in either site on the Middle Peninsula. Together, these results suggest high turnover in the regional spatiotemporal distribution of *E. chaffeensis*, which has implications in our ability to make inferences based on spatially

or temporally limited datasets. Moving forward, it will be critical to build spatially replicated, longitudinal studies to effectively account for these dynamics when trying to understand *E. chaffeensis* prevalence or predict disease risk.

In the present study, we attempted to explain variation in *E. chaffeensis* occurrence by evaluating the effects of seasonal weather. Of those included, our analysis identified winter temperature as the only plausible weather variable corresponding with *E. chaffeensis* occurrence. This effect is seen at a one-year time lag, prior to the hatching of the sampled cohort of ticks. While this could be driven by increased over-winter mortality of adult ticks, which would decrease the infection rate of naïve hosts in the spring and thus suppress the transmission cycle, there remains little to no evidence of cold winters reducing tick populations. In particular, the supercooling ability of *A. americanum* (and other ticks; Burks et al. 1996) and the lack of deep freezes in our study region make us consider this unlikely. Instead, we find it more likely the effect of winter temperature is mediated through the ticks' vertebrate hosts. While strict deduction of mechanisms is beyond the scope of this study, this trend is consistent with aspects of *E. chaffeensis* ecology. Below, we discuss these to posit biological mechanisms warranting further investigation.

Our models showed cold winters during the previous year (*t-1*) to be associated with lower *E. chaffeensis* prevalence in the current year. Most simply, this could be due to increased overwinter mortality causing decreased availability of competent reservoir hosts. Overwinter survival of Leporids can be positively associated with temperature (Rödel et al. 2004), and it seems likely Sciurds are affected similarly. Winter temperatures could also act on *O. virginianus*, though perhaps not as directly. Cold winter have been shown to decrease ungulate natality (Coulson et al. 2000), delay birthing and increase fawn mortality

(Parker et al. 2009), and to interact with density dependence to decrease overwinter survival (Sæther 1997). On the Virginia peninsula, where deer are chronically overpopulated, we hypothesize that low winter temperatures during the mating and early gestational period could decrease the abundance of juvenile (fawn and yearling) deer in the spring, by reducing natality and increasing mortality. All of this is important because young deer appear to be play an important role in the disease transmission cycle. While all age-classes of *O. virginianus* are equally likely to be seropositive for *E. chaffeensis* antibodies, suggesting they are exposed at equal rates, younger deer are far more likely to carry the bacteria in their bloodstream (Yabsley et al. 2003). This suggests a tick is more likely to become infected after feeding on a fawn or a yearling than an adult *O. virginianus*. For this reason, we hypothesize a decrease in juvenile deer availability in the spring will decrease the proportion of blood meals taken by lone star larvae from rickettsemic deer during the summer and thus lead to decreased proportion of infected nymphs the following spring. Conversely, warmer winters could cause an increase in *E. chaffeensis* transmission and subsequent prevalence. It is plausible, then, that cold winters could affect *E. chaffeensis* prevalence through all of its known vertebrate hosts, and future studies should explicitly relate *E. chaffeensis* prevalence to population dynamics of reservoir hosts, including not just *O. virginianus*, but Sciurids and Leporids.

It is worth noting that two of our predictions were not met: neither tick nor deer abundance were strong predictors of *E. chaffeensis* occurrence. Tick abundance alone did improve model performance relative to the null (decrease in AIC_c = 5.03), but the strength of this model is diminutive compared to those including winter temperature (ΔAIC_c = 15.27). Models of deer abundance performed worse, having no bearing on *E. chaffeensis*

occurrence. Given the body of evidence for deer as important hosts for both *A. americanum* and *E. chaffeensis* (Paddock and Childs 2003, Yabsley 2010), we expected areas of high deer use to have higher *E. chaffeensis* occurrence. This could be because deer may not be as solely responsible for maintaining the *E. chaffeensis* transmission cycle as traditionally thought, and/or because of unequal contribution of deer from different age classes to *E. chaffeensis* transmission. The former hypothesis is simpler, and corroborates previous studies that have suggested other animals are equal if not more competent reservoirs for *E. chaffeensis* (Allan et al. 2010, Harmon et al. 2015), but we feel the latter is also plausible and consistent with deer ecology, and we have no way of directly testing either hypothesis with our data.

Conclusion

Further substantiation of our results will provide two opportunities. First is the production of dynamic *E. chaffeensis* risk models. Any landscape analyses of *E. chaffeensis* occurrence given a single year of data would be confounded by inter-annual variation in weather (unless those variations are homogenous across the landscape) and would be generally unable account for temporal fluctuations in *E. chaffeensis* prevalence. Hence, our study provides purpose and direction for explicit spatiotemporal modeling of disease risk. The second regards climate change. The effect of climate change on emergence and prevalence of zoonotic disease is a growing concern (Patz et al. 1996), and our study provides a basis for how climate change could affect *E. chaffeensis* prevalence in the Chesapeake Lowlands of Virginia. By the end of the 21st century, days below freezing are projected to decrease across the southeast (Kunkel et al. 2013), which, according to our

results, could increase occurrence of *E. chaffeensis*. With further substantiation of the association between *E. chaffeensis* prevalence and variation in seasonal weather patterns, we can begin to make meaningful predictions of climate change's future contribution to disease risk.

In all, our study reveals high interannual variation in *E. chaffeensis* prevalence, including high spatiotemporal turnover of pathogen occurrence within our study region. This variation can be explained in part by previous winter temperature, which we hypothesize to act by reducing reservoir host abundance and thus bacterial availability during the larval questing period. However, a robust understanding of the apparently rapid spatiotemporal dynamics of *E. chaffeensis* will require continued broad-scale, longitudinal investigation that considers both spatial (i.e. landscape composition and configuration) and temporal (e.g. weather) factors.

Habitat availability, quality, and fragmentation drive prevalence of the tick-borne pathogen *Ehrlichia chaffeensis* and occupancy dynamics of its vector, *Amblyomma americanum*

2.1 Introduction

Across the globe, tick-borne diseases are becoming increasingly common. In part, this is due to increasing prevalence of familiar pathogens, like *Borrelia burgdorferi* (Stanek et al. 2012), but also due to the emergence of new pathogens, often unfamiliar to the public. One such pathogen is *Ehrlichia chaffeensis*. While still rare relative to *B. burgdorferi*, the prevalence of *E. chaffeensis* has increased greatly in the 25 years since its discovery, more than doubling in incidence since 2000 (Anderson et al. 1993, Heitman et al. 2016). An obligate intracellular bacterium, *E. chaffeensis* is transmitted by the lone star tick (*Amblyomma americanum*) as its primary vector and carried by the white-tailed deer (*Odocoileus virginianus*) as its primary reservoir (Yabsley 2010). The bacterium is distributed widely across the United States, but prevalence varies within and between states (Yabsley et al. 2003), and even within counties (Wright et al. 2014). Much previous work that explicitly investigates the distribution of *E. chaffeensis* has focused on a broad, continental scale (e.g. Wimberly et al. 2008, Liu et al. 2017), whereas local studies are typically only qualitative. However, to effectively mitigate disease risk through management or education, it is critical to understand environmental factors driving local variation in the spatial distribution of *E. chaffeensis*.

As an obligately intracellular pathogen, the distribution of *E. chaffeensis* is directly conditioned on that of its vertebrate and invertebrate hosts, requiring a community of competent vectors and reservoirs. In turn, the distribution of the invertebrate vector, *A. americanum*, is itself conditioned on that of its vertebrate hosts, which includes a broader

set of species than the set of reservoirs for *E. chaffeensis* (Bishopp and Trembely 1945, Allan et al. 2010). Environmental factors can thus affect *E. chaffeensis* directly, by altering the distribution (and thus availability) of reservoirs, or indirectly, by affecting tick population dynamics and the availability of vectors.

One potentially important driver of host species' distributions is landscape context. While fine-scale factors like microclimate can constrain tick distributions to particular habitat types (Semtner et al. 1971), tick distributions within acceptable habitats depend on that of their vertebrate hosts. For instance, the abundance and distribution of the black-legged tick (*Ixodes scapularis*) has been shown to be dependent on the density and movement patterns of *O. virginianus* (Rand et al. 2003, Kilpatrick et al. 2014). In turn, the distribution of *O. virginianus* in part is driven by landscape context, particularly patterns of fragmentation and land-use (Lovely et al. 2013). In fact, through their effect on the vertebrate community (including deer, but especially rodents), forest patch size and fragmentation can indirectly drive spatial patterns of both *I. scapularis* density and *B. burgdorferi* prevalence (Allan et al. 2003, Brownstein et al. 2005). While *B. burgdorferi* and *E. chaffeensis* are carried by different vectors and reservoirs, there are enough similarities to expect similar landscape effects on *E. chaffeensis*. *Odocoileus virginianus* are not only important host for *A. americanum* (Kollars et al. 2000), they are the primary reservoir of *E. chaffeensis* (Allan et al. 2010, Yabsley 2010). In a fragmented landscape that favors *O. virginianus*, one should expect an increase in the proportion of blood-meals taken from deer by *A. americanum* and should therefore see increased transmission and infection rates. This effect could be compounded if an increase in deer also positively affects *A. americanum*, effecting either higher densities of ticks across space or greater

persistence of ticks across time. The latter would be especially important because availability of *E. chaffeensis* from *O. virginianus*' bloodstream appears to decrease over time, especially as the deer ages (Davidson et al. 2001, Yabsley et al. 2003), and so continual re-infection of reservoir hosts should be necessary to perpetuate the transmission cycle. Thus, landscape factors that support dense and persistent host populations would be associated with higher pathogen prevalence.

Fragmentation has been previously shown to predict *E. chaffeensis* distribution (Manangan et al. 2007), but evidence in the literature is equivocal. Other work suggests this effect is geographically dependent, being more important in the western than in the eastern US, where *E. chaffeensis* distribution is instead constrained by climate (Wimberly et al. 2008). At the local scale, where climate is relatively homogenous, it remains unclear if fragmentation remains a helpful indicator of *E. chaffeensis* prevalence. Quantification of this relationship, if it exists, will be an important step to predicting pathogen prevalence, and thus disease risk, across landscapes.

The distribution of *E. chaffeensis* may also be driven by other landscape factors that affect its vertebrate hosts. For instance, habitat type could drive pathogen prevalence through its effect on host space-use. In our study area, coniferous stands usually lack the herbaceous undergrowth seen throughout much of the deciduous-dominated forest (one exception being *Vaccinium*, which sometimes grow densely in coniferous stands), as well as acorns, an important resource for forest mammals (Ostfeld et al. 1996). Together, these deficiencies may encourage less frequent use of conifer stands by *A. americanum*'s vertebrate hosts. Another possibility is the availability of water. If an animal's use of area within its home range is dictated by the availability of resources, it follows that there may

be more frequent use of habitat in areas nearer to water. Together, these factors could influence pathogen prevalence by affecting the availability of vertebrate hosts across the landscape.

Here, we ask how landscape context affects the prevalence of *E. chaffeensis* and the spatiotemporal dynamics of its tick host, *A. americanum,* in southeastern Virginia. Using five years of field data, we treat ticks at each of 130 plots as metapopulations to estimate local colonization and extinction rates in the context of a multi-season occupancy model, and we model the prevalence of *E. chaffeensis* within the *A. americanum* population as the probability that a single tick is infected. We made the following hypotheses:

1. *Ehrlichia chaffeensis* prevalence and *A. americanum* occupancy dynamics are both driven by host density, which is in turn driven by an interaction between habitat availability and fragmentation such that highly forested but fragmented areas will have the highest host densities and thus favor persistent tick occupancy (i.e. high colonization, low extinction) and high pathogen prevalence.

2. *Ehrlichia chaffeensis* prevalence and tick occupancy are also driven by host space-use, which is in turn driven by resource availability, or habitat quality. In our case, these are measured as the proportion of evergreen-dominant forest in the surrounding area and distance to water. Given that forests in our study area are rarely, if ever, dominated by broadleaf evergreens (*personal observation*), evergreen cover acts as a direct proxy for conifer stands, which we expect to affect *E. chaffeensis* and *A. americanum* through habitat preferences of their vertebrate hosts. We expected tick occupancy and *E. chaffeensis* prevalence to be highest in deciduous-dominated forests near water.

2.2. Methods

2.2.1 Data collection

Ticks were collected as described in Chapter 1, except for an increase in sampling effort during 2017. To get better representation of more fully urbanized areas, we added eight new plots toward the southern end of the Virginia Peninsula in the Hampton and Newport News areas. These sites were either wooded lots embedded in suburban areas or wooded areas on larger properties (namely schools). Additionally, because of the rarity of *E. chaffeensis*, we increased per-plot sampling effort by repeating each transect along a parallel ca. 2-m offset, yielding four 30-m transects instead of the two surveyed in previous years. During analysis, ticks from each transect were pooled separately, rather than together, so each plot provided up to four tick pools for molecular analyses. We analyzed all ticks collected in 2017, subdividing transects into separate pools of ≤ 20. Pathogen detection was performed as described in Chapter 1, via PCR.

We held the same nested site-plot structure as in Chapter 1, except for an aggregation of two small, adjacent sites. These sites had two and three plots and, while they were managed by different government agencies for different purposes, we felt they were similar enough to justify this change. Combining these into a site with seven plots allowed for more robust estimates of random effects. The eight new urban plots were considered a new site.

We also collected additional field data in 2017, beyond what was collected in previous years. At each plot, at the beginning and end of a survey, we measured ambient temperature and humidity to act as covariates for tick detection probability. In a

preliminary occupancy analysis of 2017 data, temperature had a positive effect on detection but humidity had no effect. To include temperature for previous years' surveys, in which we had not measured temperature, we used model estimates: we used the 2017 measurements to parameterize a quadratic linear model that predicted on-the-ground temperature as a function of mean daily temperature (taken from interpolated PRISM values; see Chapter 1 and Oregon State University 2018) and time of day ($r^2 = 0.67$), and applied used this to predict unobserved values.

2.2.2 Landcover classification

We created a landcover classification map from Landsat 8 scenes (spatial resolution = 30 m) in ENVI (Environmental for Visualizing Images; Harris Geospatial, Boulder, CO). We considered producing annual classifications but, within 3 km of our plots, interannual change in forest cover was negligible. Instead, we created a single classification in the middle of our study period and used subsequent years to improve misclassifications. The initial classification was made using ISODATA classification on a composite of a summer scene (day 227, 2014), a winter scene (day 37, 2015), and an NDVI (normalized-difference vegetation index) layer derived from the summer scene. Spectral classes were assigned to water, marsh, deciduous or evergreen forest, herbaceous (agriculture, lawn, and meadow), bare soil (typically fallow field but also recent clear cut), or impervious surface (buildings and pavement). We refined this with cluster busting and by using data from subsequent years to correct misclassifications between fallow agricultural fields and impervious surface.

Final accuracy was assessed using 420 remote ground-truth points. We used

stratified random sampling to generate ground truth points (quasi-proportional to class coverage) in ENVI and checked for the true land cover at each point using satellite imagery in Google Earth. We achieved an overall accuracy of 92%, with no individual class below 70% (Table B2).

As a final modification after accuracy assessment, we used a TIGER (Topologically Integrated Geographic Encoding and Referencing) roads layer (U.S. Census Bureau 2011) to burn in roads that were too narrow to be captured in the original classification (at 30-m resolution, only large highways were consistently captured). Quantitatively, this creates an overrepresentation of impervious surface cover by area (most roads are not 30 m wide), but improves accuracy of forest edge length and fragmentation, which is more important to our analyses.

2.2.3 Landscape metrics

To test our hypotheses, we derived landscape metrics from the land cover classification described in section 2.2.2 and from state-provided stream delineation and LIDAR-derived elevation data (Virginia GIS Clearinghouse 2011) in ArcMap 10.1 (ESRI, Badlands, CA), FRAGSTATS (Mcgarigal et al. 2012) and the Geospatial Modeling Environment (Beyer 2014). We included two potential metrics each for habitat availability and fragmentation, and also included a spatial metric that might affect tick detection probability. In all, our included metrics were 1) proportion forest cover, a measure of habitat availability; 2) proportion urban cover, a more inclusive, inverse measure of habitat availability; 3) forest edge density, a basic measure of forest fragmentation; 4) proximity, a patch level metric of isolation and fragmentation that describes the area of adjacent

patches within the focal window, penalized by distance; 5) proportion of surrounding forest that is evergreen, which is a measure of habitat type or quality that could affect vertebrate host use; 6) distance to water, which may influence animal movement and space use; and 7) average slope within the 15-m radius circle of the plot, which could affect difficulty of flagging and substrate depth and composition and in turn affect tick detection and availability. We also considered contagion, another common fragmentation metric which measures both dispersion and interspersion of land cover types, but this was almost perfectly correlated with forest cover (via an exponential relationship; $r^2 = 0.96$) and so was excluded. For a summary of relevant landscape-metric distributions, see Appendix B3.

Because the scale at which an organism is affected by and responds to its environment differs between the organism and environmental variable in question, we used an optimization method to determine the best extent at which to include each variable in each model (holding resolution constant). Landscape metrics were extracted at three extents according to a series of home-range size estimates of *O. virginianus*. The smallest extent, a 25-ha circular window, represents a compromise between mean core area (Holzenbein and Marchington 1992, Campbell et al. 2004) and the mean home range size of suburban deer, which have been shown to be much smaller than in adjacent agricultural habitats (Cornicelli et al. 1996). The mid-size extent, 100-ha, represents the mean home range size estimated in the Appalachian mountains of Virginia and West Virginia (Holzenbein and Marchington 1992, Campbell et al. 2004), and also represents an approximate average home range estimate encountered in the literature. The largest extent, 350-ha, represents the higher end of *O. virginianus* home range estimates (Hasapes and Comer 2016). These extents, particularly the larger two, also approximate home-range

estimates of raccoons (*Procyon lotor*; Glueck et al. 1988, Beasley et al. 2007, Rosatte et al. 2010), another of *A. americanum*'s wide ranging hosts (Kollars et al. 2000).

We did not expect all these metrics to fit well into simple linear models because fragmentation metrics, like edge density, can take on similar values in areas that are very different ecologically. For instance, a lightly fragmented forest will have low edge density because there are few fragment edges but, alternatively, an almost entirely urbanized area will also have low edge density because there remain few forest edges. Similarly, a large forest patch can still have low proximity values if there are no other patches near it. To account for this, we tested interactions between habitat availability and fragmentation with the prediction that fragmentation would become important when habitat availability takes on higher values.

2.2.4 Statistical analysis

Statistical models were implemented in a hierarchical Bayesian framework in JAGS 4.3.0 (Plummer 2017), using the R2jags package to interface with R 3.4.3 (Su and Yajima 2015, R Core Team 2017). Both the prevalence and occupancy analyses consisted of extent- and variable-selection followed by inference directly from a global model with the selected covariates. Each of these analyses used standardized (i.e. Z-transformed) values. Unless specified otherwise, we report parameter estimates as means with 95% highest density intervals (HDIs) in the form *mean [HDI]*, and we calculated any derived value (e.g. year-specific prevalence) from individual posterior draws before summary. Detailed model diagrams, likelihood statements, and JAGS code can be found in Appendix C.

Pathogen prevalence

We used data from all five years to estimate environmental effects on *E. chaffeensis* prevalence. While ticks were analyzed in pools, we modeled prevalence as the probability of a single tick being infected with *E. chaffeensis* according to the formula $1 - (1 - \rho)^n$, where ρ is the infection probability and n is the number of ticks in the pool. In turn, ρ is modeled as $logit(\rho) = \beta_0 + \boldsymbol{\beta}\mathbf{X}$, where β_0 is a nested site/plot random intercept in which site-level means vary around the grand mean and plot-level means vary around their respective site mean. Because we visited new plots in 2017, we only estimated a single, universal plot-level variance parameter. Additionally, to account for differences in prevalence between years, we included year as a factor. We set vague and uninformed normal priors on all beta-coefficients ($\mu = 0$, $\sigma^2 = 5$; Hooten and Hobbs 2015), and a vague but informed prior on the intercept ($\mu = logit^{-1}(0.018)$, $\sigma^2 = 5$), where the mean of the prior distribution was the minimum prevalence of infection from our entire dataset. The variance parameters associated with the random intercepts were given uniform prior distributions over the interval [0,10].

Variable- and extent- selection was performed by adapting the Bayesian Latent Indicator Scale Selection (BLISS) method described by Stuber et al. (2017). Briefly, the BLISS method uses a model in which the scale of each covariate is allowed to vary according to a categorically-distributed indicator variable. The posterior distribution of the indicator variable is used to make direct inference about the best extent at which to measure each covariate. We adapted this method to simultaneously select between extent and covariate combinations. There were two groups of covariates that we expected to interact: habitat availability (proportion forest or urban cover) and fragmentation (edge density or

proximity). Holding extent constant within each combination, there were 2 x 2 x 3 = 12 potential interactions (i.e. $\beta_1 X_{111} + \beta_2 X_{211} + \beta_3 X_{111} X_{211}$, $\beta_1 X_{112} + \beta_2 X_{212} + \beta_3 X_{112} X_{212}$, …, $\beta_1 X_{123} + \beta_2 X_{223} + \beta_3 X_{123} X_{223}$). For these variables, rather than selecting extents for each variable individually, we used an indicator variable to choose one of the 12 possible extent-specific interactions. Within the same model, we used the original BLISS method as described above to select the extent at which to measure proportion evergreen forest. To ensure these selections were conditioned on other potential effects, we also included distance to water, the only variable not being extent-optimized. The variable- and scale-selection model was run on four independent chains (initializing indicator values for each of four covariate combinations at the 100-ha scale, with random draws for proportion evergreen forest) for 120,000 iterations, thinned by 15 to reduce autocorrelation, after a 12,000-iteration burn-in period, for a total of 32,000 samples from the posterior.

Final inference was made from the global model, which included an interaction between habitat availability and fragmentation, forest type, and distance to water. With the best scale and covariate combination selected, the global model was run on three independent chains for 100,000 iterations, thinned by 10, after a 10,000-iteration burn-in, for a total of 30,000 samples from the posterior. Beta-coefficients were initialized with random values from normal distributions (μ=0, σ^2=0.5), while mean prevalence and variance parameters were given distinct, overdistributed values.

Tick occupancy dynamics

Tick occupancy dynamics between 2012 and 2017 were modeled using a multi-season occupancy model in which estimates of local colonization and extinction rate were

adjusted to account for imperfect detection (*sensu* MacKenzie et al. 2003). Each year was treated as a primary survey, in which detection and occupancy state were estimated from repeated secondary surveys, and changes in occupancy probability between years were used to infer rates of local turnover (i.e. colonization and extinction). We treated each 15-m segment of a transect as a secondary, yielding four presence/absence observations per primary. To maintain consistent sampling effort between years, we only used observations from the first two of four transects surveyed at each plot in 2017. We also only used the 101 plots that were surveyed in all five years, for a total of 101 plots x 5 years x 4 observations plot^{-1} year^{-1} = 2020 observations.

Occupancy dynamics and detection were modeled as functions of the environment via logistic regression. Colonization and extinction were estimated at the plot-level in the form $\text{logit}(\theta) = \beta_0 + \boldsymbol{\beta}\mathbf{X}$, while detection was estimated for each plot but was also allowed to vary between years using the form $\text{logit}(p) = \beta_0 + \boldsymbol{\beta}\mathbf{X} + \varsigma$, where ς is a normally distributed random variable with mean 0 that allows p to vary between plots and years beyond what is described by $\mathbf{X}$. In each, β_0 is a site-level random-intercept that is normally distributed around the grand mean. Because no surveys were conducted in 2014, the occupancy state of this year was modeled as an unobserved latent state.

Candidate variables were the same as in the prevalence model but with additional variables considered for detection probability. Colonization and extinction were modeled as functions of habitat availability and fragmentation, evergreen cover, and distance to water, while detection probability was modeled as a function of temperature and terrain slope. Variable- and extent-selection was conducted using the adapted BLISS method as described above for the prevalence model, run on three independent chains for 240,000

iterations, thinned by 20, after a 10,000-iteration burn-in, for 36,000 samples from the posterior. These chains were initialized using random values for each categorical variable. The final model was run on three chains for 100,000 iterations, thinned by 10, after a 10,000-iteration burn-in for a total of 30,000 samples from the posterior. For each of these chains, intercepts and beta-coefficients were initialized with values from standard-normal distributions, and variance parameters were initialized with random values from their priors.

2.3 Results

Considering all years, *A. americanum* were found on every plot, but tick abundance and occupancy, and *E. chaffeensis* prevalence, varied across both space and time. Adjusted for sampling effort, naïve *A. americanum* occupancy ranged from 0.69 to 0.8 and average number of ticks per plot ranged from 3.88 (±7.01) to 10.7 (±25.3). Of the original 16 sites that were sampled all years, 14 sites tested positive for *E. chaffeensis* at least once, 11 of these more than once, while the two sites on the Middle Peninsula never tested positive (for location of sites see Fig. A1). In 2017, none of the new urban plots tested positive. Within the sites that tested positive, *E. chaffeensis* occurrences were typically spread across different plots. Of 39 plots that tested positive, only six tested positive more than once and, of these, five tested positive two of five years and one tested positive three of five years.

In 2017, despite increased sampling effort and additional plots, *E. chaffeensis* was detected less frequently than any other year. A total of 1001 ticks were analyzed in 2017 (up from 550 – 631 in previous years), but only two plots tested positive. One of these plots had tested positive in a previous year and one had not. In each case, multiple pools were

analyzed, but only one pool tested positive, showing within-plot variability in infection.

2.3.1 Prevalence Model

Variable and scale selection

Between the two sets of variables, different scales were selected and with different distributions of evidence. Habitat availability and fragmentation was selected as proportion forest cover and edge length at the 350-ha extent with 31% support (of 12 possibilities; Figure B1). The same combination, but at the 100-ha extent, came in second with 26% support, and the third best was further behind with 13%. Proportion of evergreen forest was selected at the 100-ha extent with 72% support (of three possibilities; Figure B2).

Global model

In the global model, estimated mean prevalence was 0.44% [0.13%, 0.80%] (or -5.51 [-6.42, -4.72] on the log-odds scale) but varied between years and plots. Annual prevalence ranged from 0.08% - 2.0%, although with much uncertainty in each point estimate (Figure B3), and site- and plot-level standard deviations were 0.41 [0.00, 1.0] and 0.85 [0.04, 1.5] respectively (Figure B4). There was evidence for a negative effect of proportion forest cover (-0.84 [-1.5, -0.25]) and edge length (-0.81 [-1.5, -0.17]), but no evidence for an interaction between them (-0.36 [-0.86, 0.14]). There was also evidence for a negative effect of evergreen forest (-0.71 [-1.47, -0.003]), but not for distance to water (0.21 [-0.24, 0.67]) (Figure 2.1).
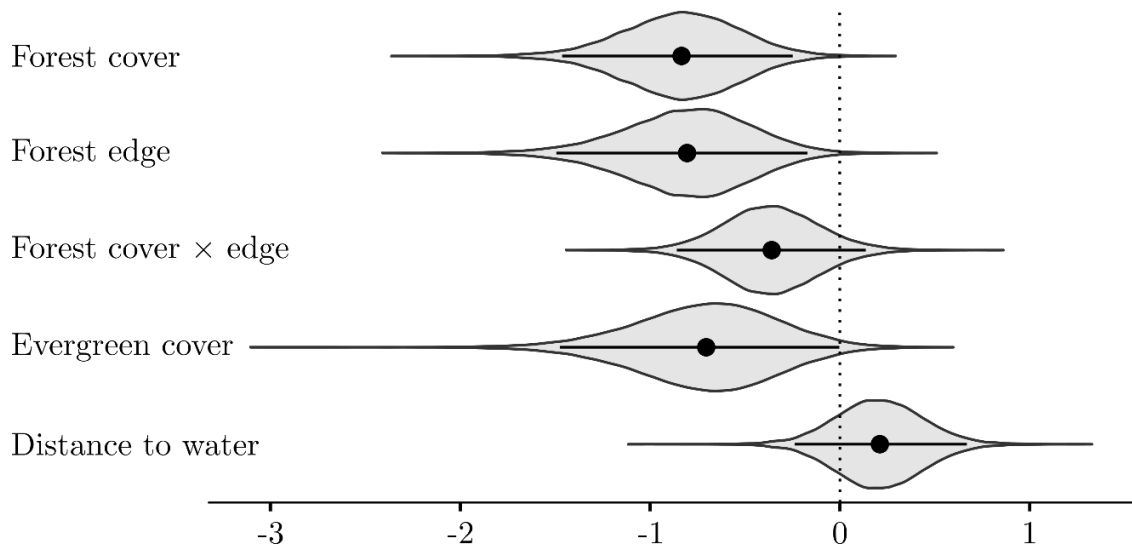
Figure 2.1: Posterior probability densities for the coefficients (on the logit scale) of environmental covariates in the prevalence model. Dots represent the posterior mean and bars represent the 95% HDI.

2.3.2 Occupancy Model

Variable and scale selection

For local extinction rate, habitat availability and fragmentation were selected as proportion forest cover and edge density at the 25-ha extent with 61% support (out of 12 possibilities; Figure B5), and proportion evergreen forest at the 350-ha extent with 62% support (of three probabilities; Figure B6). For local colonization, selections were less absolute. Habitat availability and fragmentation were selected as proportion forest cover and proximity at the 350-ha extent with 24% support (Figure B5), and proportion evergreen cover at the 350-ha extent with only 39% support (Figure B6).

Global model

The global model showed very little turnover in the tick population. Across sites,
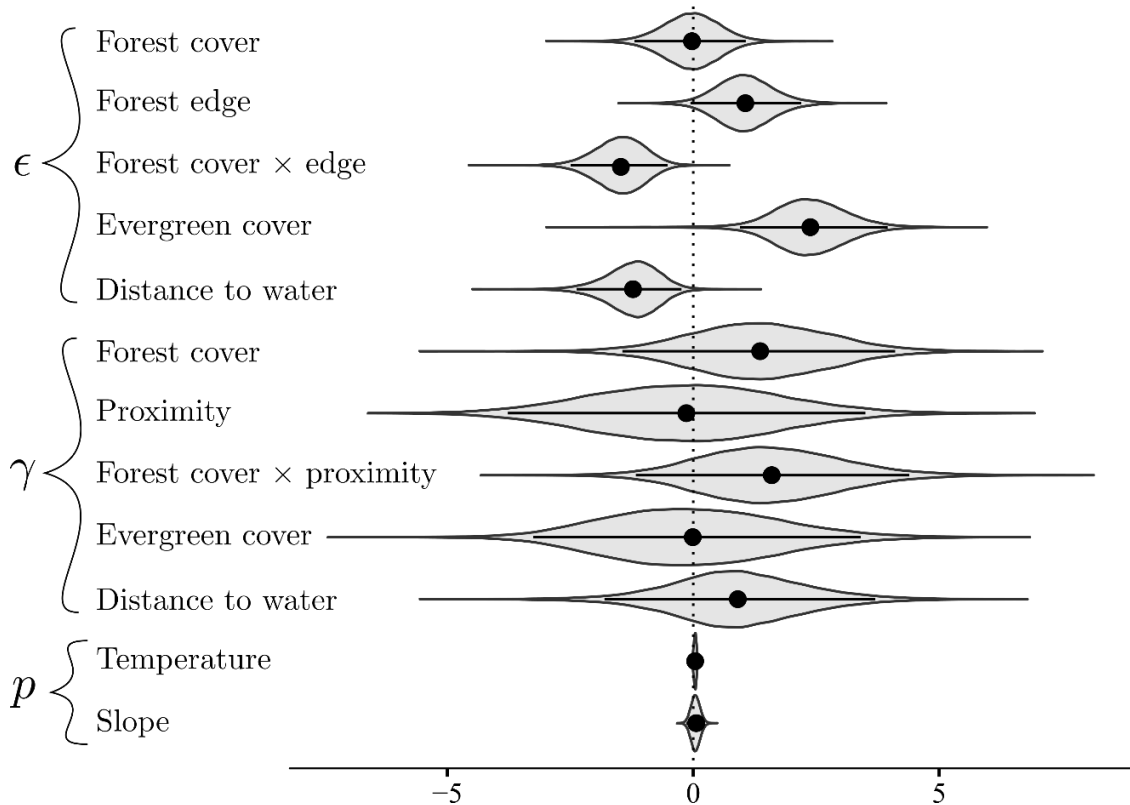
Figure 2.2: Posterior probability densities for the coefficients (on the logit-scale) of environmental covariates on extinction (ε), colonization (γ), and detection (p) in the *A. americanum* occupancy model. Dots represent posterior means and bars represent 95% HDIs.

colonization was very high, at 0.86 [0.47, 1.0] (or 2.6 [-0.60, 5.7] on the log-odds scale), and extinction was very low, at 0.006 [0.0, 0.018] (or -5.7 [-8.0, -3.6] on the log-odds scale; Figure B7). Neither of these parameter estimates varied strongly between sites; the mode of σ_γ was 0.39 [0.0, 6.2] and the mode of σ_ε was 0.36 [0.0, 2.5] (Figure B8).

There was evidence for landscape effects on extinction, but not colonization (Figure 2.2). For extinction, there was no evidence for a mean effect of forest cover (-0.04 [-1.2, 1.1]) and there was only weak evidence for a mean effect of edge density (1.0 [-0.06, 2.2]), but there was evidence for an interaction between them (-1.5, [-2.5, -0.52]). There was also evidence for an effect of evergreen forest on extinction (2.4 [0.95, 4.0]) and distance to water (-1.2 [-2.4, -0.24]). There was no evidence for any effect of these variables on

35

colonization; their posteriors were all very broad and included 0 well within their HDIs. Detection probability did vary, but was not affected by either temperature nor slope (Figures 2.2 and B8). Mean detection probability was 0.40 [0.13, 0.70] (or -0.44 [-1.8, 0.90] on the log-odds scale), with a site-level standard deviation (σ_p) of 0.84 [0.48, 1.3] and observation-level standard deviation (ς_p) of 1.0 [0.79, 1.3]. The mode of P* (the probability that a tick was observed at least once on a plot, given ticks were present) was 0.97, but with a very left-skewed, leptokurtic distribution, reflective of detection's broad posterior probability. Both beta-coefficient posteriors were narrowly distributed around 0.

2.4 Discussion

We tested two hypotheses regarding landscape drivers of *E. chaffeensis* prevalence and lone star tick occupancy, expecting parallel processes between the two systems. This was not the case. There was one parallel scenario: areas with low forest cover and low fragmentation were predicted to have the lowest turnover and the highest prevalence. However, while it makes sense that prevalence should be negatively related with turnover because a persistent vector population would enable perpetuation of the transmission cycle, this does not hold true across all combinations of forest cover and edge length. Our models predicted areas of high forest cover and fragmentation to have low turnover, in line with our expectations, but also to have the lowest prevalence, the opposite of our expectation! The interaction effect in the occupancy model predicted inconsistencies between occupancy and prevalence dynamics (Figure 2.3). While highest prevalence and lowest turnover were both associated with small, intact forests, lowest prevalence was predicted in large, fragmented forests, which were also areas of low turnover. Similarly, highest
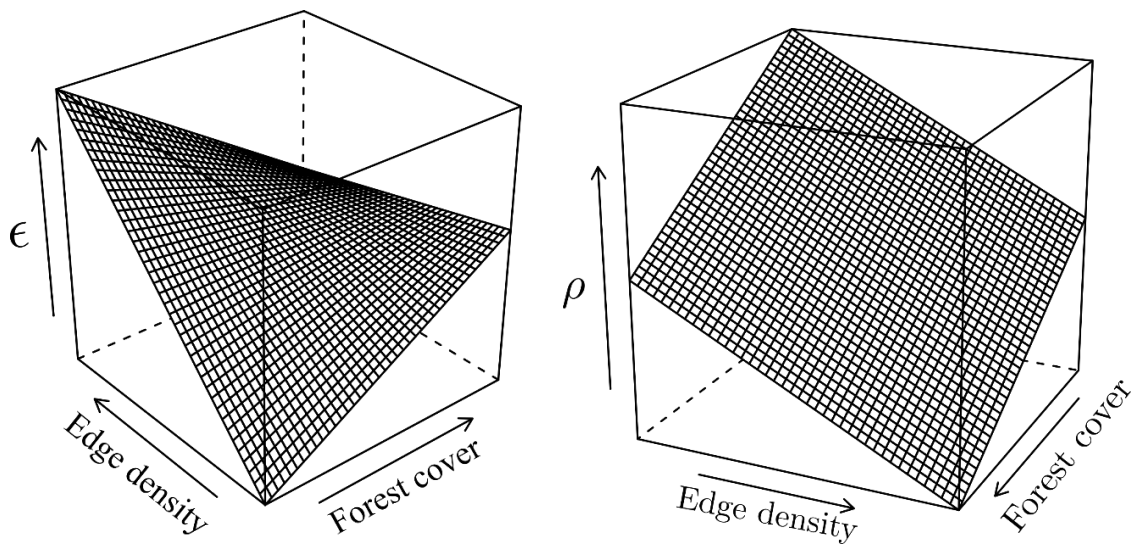
Figure 2.3: Predicted *E. chaffeensis* prevalence (ρ) and *A. americanum* (ε) extinction as a function of forest cover and edge density. Values were predicted with the mean coefficient values from their respective global models. The interaction term was included for extinction but not prevalence. Note that the graphs are viewed from different angles such that the axes run in different directions.

turnover was predicted in areas of small, fragmented forests, which was an area moderate prevalence. These inconsistencies suggest a decoupling of prevalence and occupancy dynamics such that some components should be considered separately.

Habitat availability and fragmentation were not associated with prevalence as we expected. We found a negative effect of fragmentation (as edge density) – in direct opposition of our hypothesis. However, while edge density did not interact with forest cover as we expected, previous studies *have* also shown forest cover in the surrounding landscape to be negatively associated with tick-borne pathogen prevalence (Ostfeld et al. 2018). Our expectations were based on the positive relationship between deer density and availability of edge habitat, which has been documented along suburban to rural gradients (Gaughan and DeStefano 2005), within undeveloped forests (Saïd and Servanty 2005), and even within agricultural areas of Virginia (Lovely et al. 2013), but it seems as though this relationship does not drive upstream patterns of pathogen prevalence in our study area. One

potential explanation for this defiance of expectation is that chronic overpopulation of deer on the Virginia Peninsula (Virginia Department of Game and Inland Fisheries 2015) could wash out the effect we might expect from the edge-density relationship. That is, the populations in this area may be too large for differences in density to have the expected effect. While we are unable to definitively address the mechanisms at play, below we offer discussion of potential explanations.

Rather than density, landscape context may affect *E. chaffeensis* prevalence in SE Virginia through its effects on deer movement and/or group size. Regarding movement, it may be that small, isolated forests (i.e. low forest cover and edge) act as reliable core habitat in which successive generations of ticks feed on the same animals, whereas a heavily forested and fragmented area sees more mixing. This first notion was also posited by Estrada-Peña et al. (2010), who suggested that increased tick-host contact rates in small forest fragments leads to an amplified transmission cycle, and is supported by Gaff and Schaefer (2010), in whose simulation study isolated patches maintained *E. chaffeensis* endemicity better than well-connected patches because of a diluting effect of migration. This also supports the second notion, that increased mixing decreases prevalence. That landscape context affects mixing has been shown empirically by Skuldt et al. (2008), who showed an association between forest edge and dispersal rates of juvenile white-tailed deer, and Robinson et al. (2012), who found deer in more highly forested and fragmented areas to be more genetically homogenous. Alternatively (or additionally), prevalence may be affected by deer group size, which increases with resource availability (Habib et al. 2011) and could drive something resembling an encounter-dilution effect (*sensu* Mooring and Hart 1992). While counter-intuitive, there is evidence that increasing group size can

decrease parasite load just as it can decrease an individual's predation risk (Rubenstein and Hohmann 1989, Côté and Gross 1993). Given the low transmission efficiency of *E. chaffeensis*, most ticks in any group are unlikely to be infected even if all fed from an infected host (Nair et al. 2014). Thus, in a larger group of deer, there may be more opportunities for encounters between uninfected reservoir-vector pairs.

Thus far, we have discussed areas of low forest and low edge as small, isolated forest patches, imagining a forest patch surrounded by matrix. But, this is not necessarily how low-forest/low-edge plots are situated; a plot in a large forest can have low cover because it is at the edge (e.g. Figure 2.4). However, this does not need to invalidate our above suppositions. An impermeable boundary, like a river (Figure 2.4), could have similar effects to isolation by reducing dispersal of resident animals. The resulting effect on migration can have the same effect on pathogen prevalence as isolation (Gaff and Schaefer 2010). This does introduce additional complications, though, because other landscape features, like interstate highways, could have the same effect but would not be detected using our landscape metrics.

Truly elucidating mechanisms and validating our above hypotheses will require empirical evidence. Further landscape analyses of *E. chaffeensis* prevalence are important, but studies connecting *E. chaffeensis* prevalence to behavior of its hosts may be more critical. In particular, future analyses should include movement models that account for boundaries and matrix permeability.
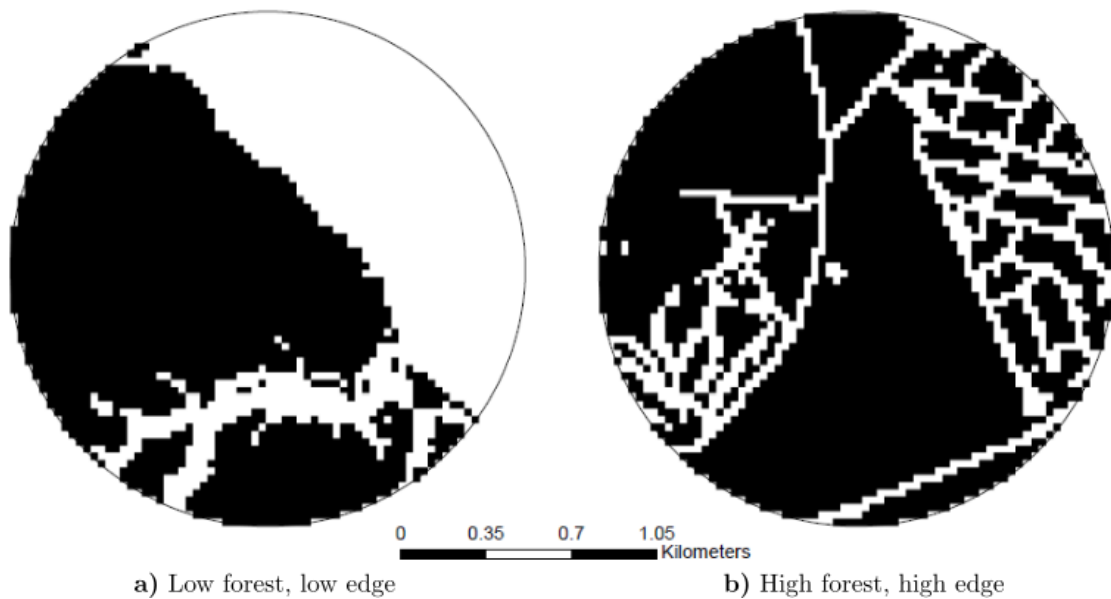
Figure 2.4: Example landscape configurations representing highest (a) and lowest (b) predicted change in *E. chaffeensis* prevalence due to forest cover and edge density. The black area represents forest. The left panel shows a plot bordered by a river and one of its tributaries. While not truly isolated, the water would create an impermeable boundary affecting animal dispersal and movement. On the right is an area with high forest cover, but with high edge density due to fragmentation.

Occupancy dynamics were, in some ways, more in line with our expectations, suggesting they are likely driven more by host density and habitat use than behavior and social dynamics. In parallel to prevalence, tick turnover was lowest in isolated, contiguous forest patches, which could be a result of the importance of these forest patches as core habitat for vertebrate hosts. However, in contrast to prevalence, tick turnover is also low in highly forested and fragmented areas. This could be because it does not matter to ticks, as it might for pathogen transmission, that there is a larger and potentially transient pool of hosts. Highest turnover is predicted in lightly forested but highly fragmented areas, which likely see low density and/or inconsistent use by *A. americanum*'s vertebrate hosts (e.g. deer, raccoons, turkeys; Kollars et al. 2000), particularly as resting areas, which would explain less consistent *A. americanum* occupancy.

Prevalence and occupancy were both negatively affected by proportion of evergreen forest cover. This is in line with our hypothesis that these parameters are driven by host habitat preference, or space-use, and our prediction that resource poor areas like coniferous dominated forests will have higher tick turnover and lower pathogen prevalence. However, our other prediction regarding space-use, concerning water availability, was not met.

Proximity to water was associated with tick turnover but not prevalence and, like some effects of forest cover and fragmentation, the effect was the opposite of our expectation. We anticipated animals would spend more time near water, thereby decreasing tick turnover and increasing pathogen prevalence. Instead, turnover decreased with distance from water. It might be that distance to water was an inadvertent proxy for the presence of wetlands or the frequency of inundation, which may negatively affect tick survival or hatching success rather than host use. While we are unaware of studies explicitly studying tick survival in wetlands, there is some work to suggest tick reproductive success decreases after extended submersion in water (Sá-Hungaro et al. 2014) and that *A. americanum* abundance is lower along wetland edge (Stein et al. 2008). Moreover, Manangan et al. (2007) did find that *E. chaffeensis* occurrence decreased with increasing wetland cover, which would be consistent with increased temporal turnover in the tick population.

Detection probability did vary across space and time, but not as a function of either included covariate. We had expected detection to increase with temperature because of increased activity levels, but there was no evidence for this effect in the model. Other work has shown *A. americanum* questing activity to be explained by saturation deficit within the

litter layer, such that the ticks are most active when the litter is dry and saturation deficit is high, and by the difference between ambient and litter temperature, such that ticks are more active when this difference is the greatest (Schulze and Jordan 2003). While both these measures are related to temperature, these authors found temperature alone to be an ineffective predictor of tick activity. We also found no effect of terrain slope, which we thought would affect the difficulty and thus efficacy of flagging, and litter depth, both of which we thought would decrease detection or availability. However, these effects may only come into play in cases of extreme slope, which only occurs in a few plots, and thus these effects, if they exist, may not be relevant to our dataset. There was, though, variation in detection, evidenced by the random-effects σ_p and ς_p. The first of these, σ_p, represents the site-level variation and suggests broad-scale patterns in detection. This is notable because there was not comparable site-level variation in tick turnover or pathogen prevalence. The second random-effect, ς_p, allows detection to vary between plots within sites, and between years by changes not described by included covariates. The partitioning of variance between both of these parameters suggests there is fine-scale spatial variation and/or temporal variation in addition to the broader spatial patterns tick detection probability. Given the scale of our sites, which are typically large enough to include unmeasured environmental variation, this is likely because of differences in tick abundance and, thus, availability.

Despite the evidence for landscape effects provided in this study, it must be considered that both prevalence and rates of turnover are very low and so the absolute effect sizes of the environmental covariates are necessarily very small. Mean prevalence and extinction (through which effects on turnover were expressed) were both less than 1%,

and so even a large relative change (e.g. from 1% to 2%) would require a large sample size to detect. So, given the relatively low number of observations of both *E. chaffeensis* occurrence and *A. americanum* extinction, our results are likely very sensitive to each occurrence or extinction event. This means our models' predictive power is likely low and, instead of providing robust spatial disease risk estimates, our study should be taken as a foundation for further analysis.

Despite the limitations of our study, our results are congruent with previous studies. In particular, previous work has shown densities of *I. scapularis* and prevalence of *B. burgdorferi*, the vectors and causative agent of Lyme disease, to be higher in smaller forest patches (Allan et al. 2003, Brownstein et al. 2005) or on plots with lower forest cover in the surrounding area (Ostfeld et al. 2018).

While results regarding fragmentation and *E. chaffeensis* have been inconsistent, this is one of few studies to the topic, especially at this scale or in this manner. Work has been done at broad, continental scales (e.g. Wimberly et al. 2008, Liu et al. 2017) and more local scales (e.g. Trout Fryxell et al. 2015), but fewer at this intermediate, inter-county scale (but see Manangan et al. 2007). Moreover, this is the first study of which we are aware to use occupancy modeling to describe interannual turnover in the tick population. Because of the inherent spatiotemporal nature of the pathogen transmission cycle, it is important to consider the temporal dynamics of actors in the system. We demonstrate this here through parallels between tick occupancy and pathogen prevalence.

Conclusion

Our results suggest that, while pathogen prevalence and tick turnover are partly

decoupled, areas with the highest prevalence and lowest turnover are those with relatively isolated but contiguous forest patches (i.e. low forest cover and low edge density). Lowest prevalence, on the other hand, is predicted in highly forested but fragmented areas. This means that forests managed for timber may not be impacted, but areas like suburban parks may be at higher risk for disease. This is similar to Lyme disease, in which fragmentation due to suburban encroachment has been implicated in increasing both pathogen prevalence (Brownstein et al. 2005) and disease incidence (Tran and Waller 2013). While prevalence of *E. chaffeensis* remains relatively low, and so remains difficult to robustly model, it is important to consider that this pathogen is becoming more common (Heitman et al. 2016) and so it is important to understand all we can about where it is found and why. Moreover, it is still unknown whether the progress made in Lyme disease ecology is generalizable to pathogens carried by different hosts. Here, we provide evidence that, like Lyme, landscape context is important to the spatial distribution of *E. chaffeensis*, reinforcing the need to consider landscape management in public-health outreach campaigns and to consider human disease risk when designing regimes of landscape management.

A   Supplementary material: Chapter 1

A.1   Figures
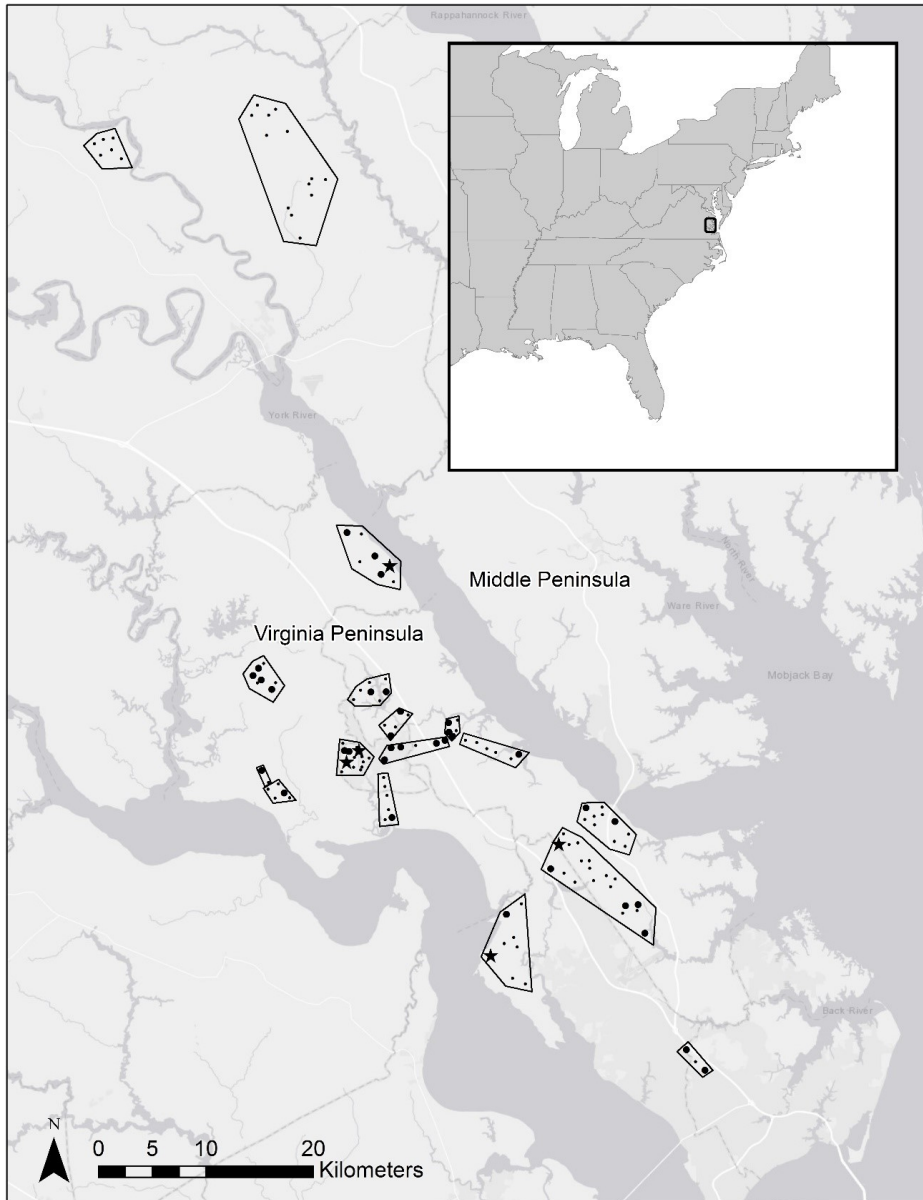
Figure A1: Location of 130 study plots (dots) and 17 study sites (polygons) on the Virginia Peninsula (n = 15) and Middle Peninsula (n = 2). Sites were defined by both proximity and management, such that adjacent sites are differentiated by management. Small dots represent plots that have never tested positive for *Erlichia chaffeensis*, larger dots represent plots that have tested positive once, and stars represent plots that have tested positive more than once. Note that polygons are meant to denote which plots belong in each site, not to delineate the actual site boundaries.

Figure A2: (a–d below) Weather variables aggregated across all plots and years (2011 2016). Asterisks designate means and whiskers designate either the minimum and maximum, or 1.5 × box length, whichever is closer to the median.
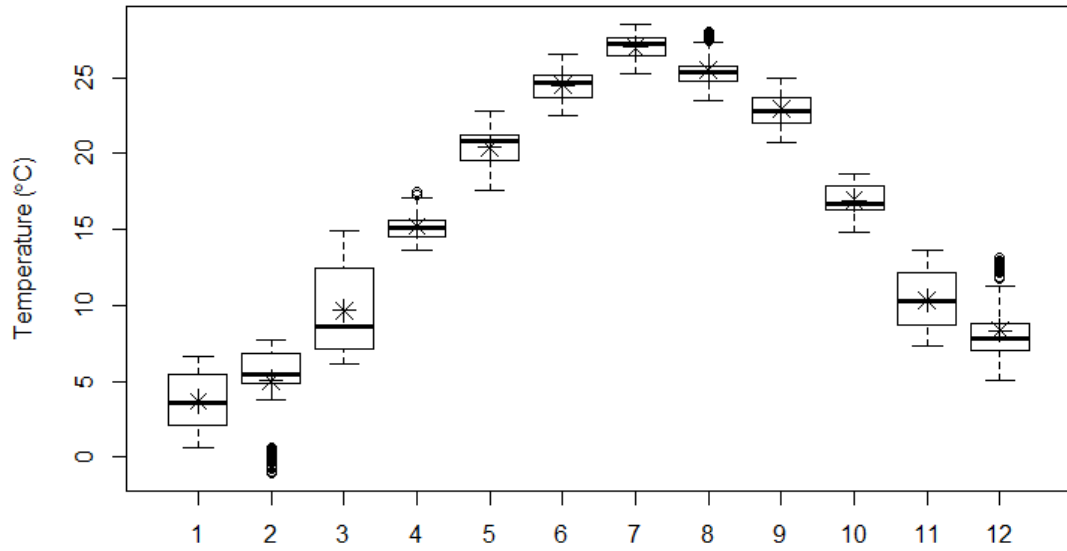
(a) Mean temperature.
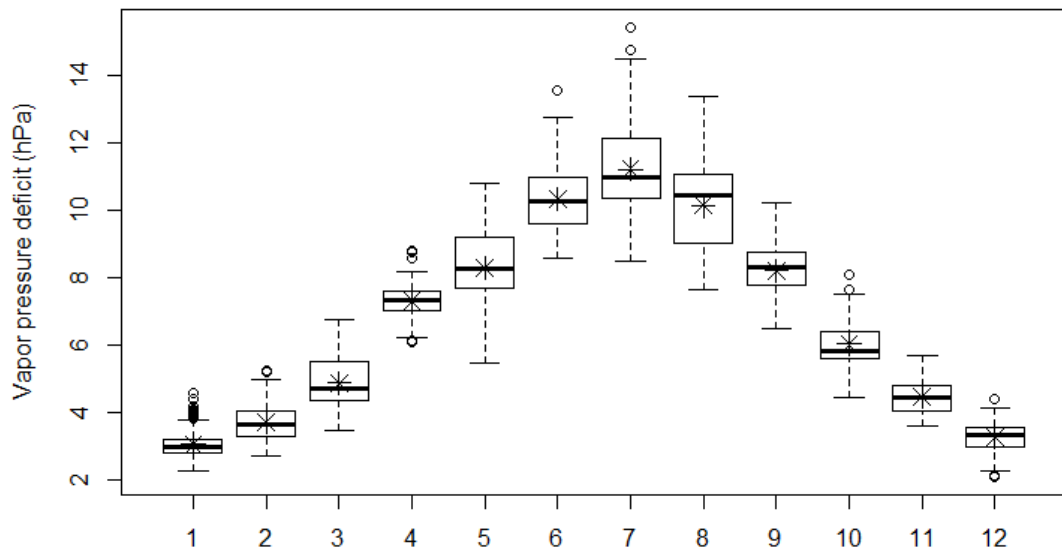
(b) Mean vapor pressure deficit.

46
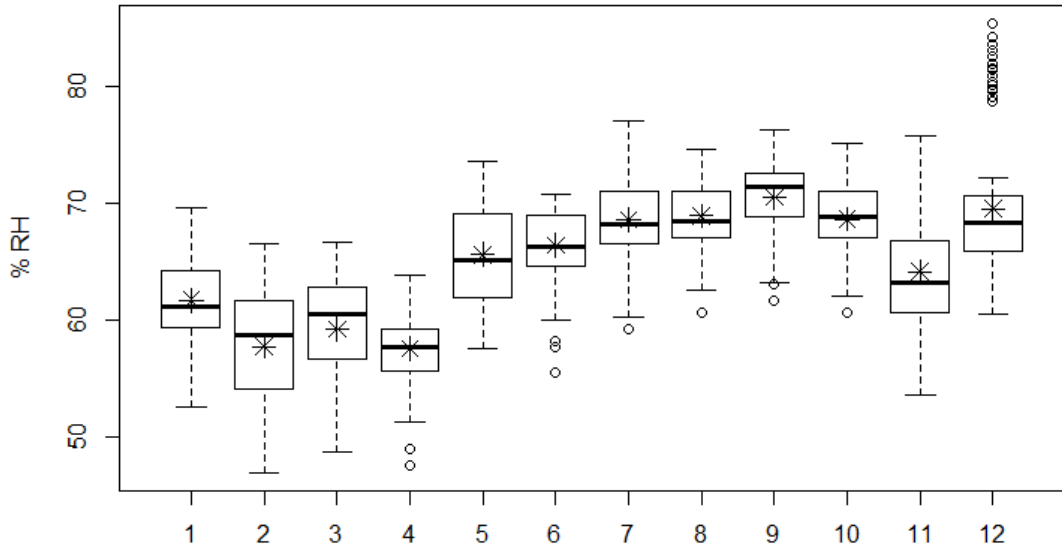
(c) Mean relative humidity.
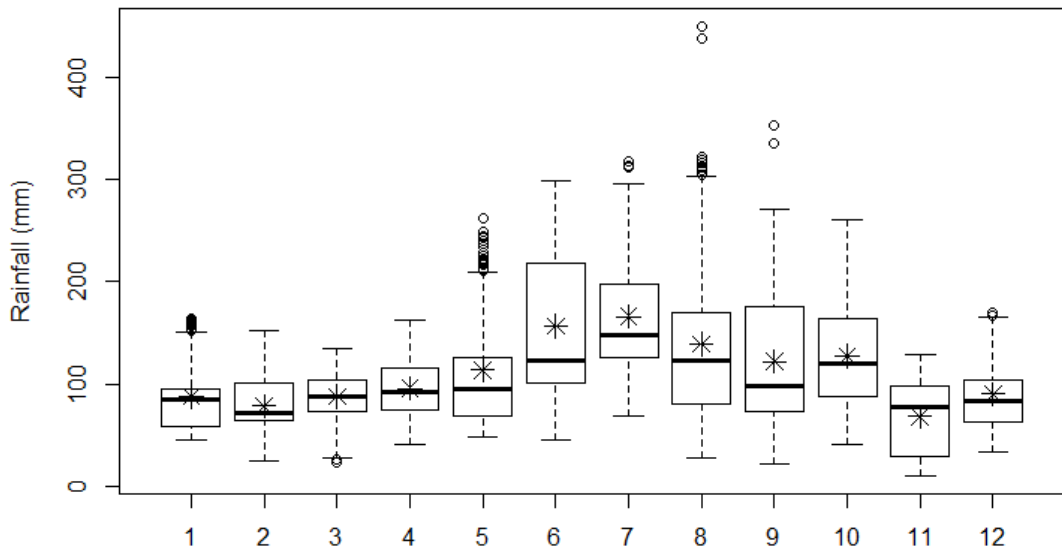
(d) Cumulative precipitation.
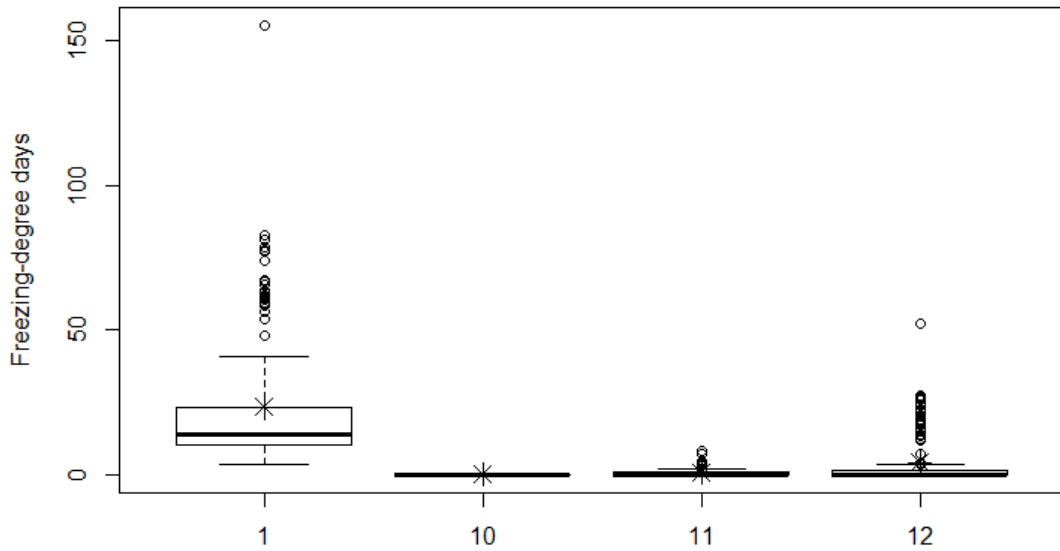
(a) Freezing-degree days by summarized by month, across all years.
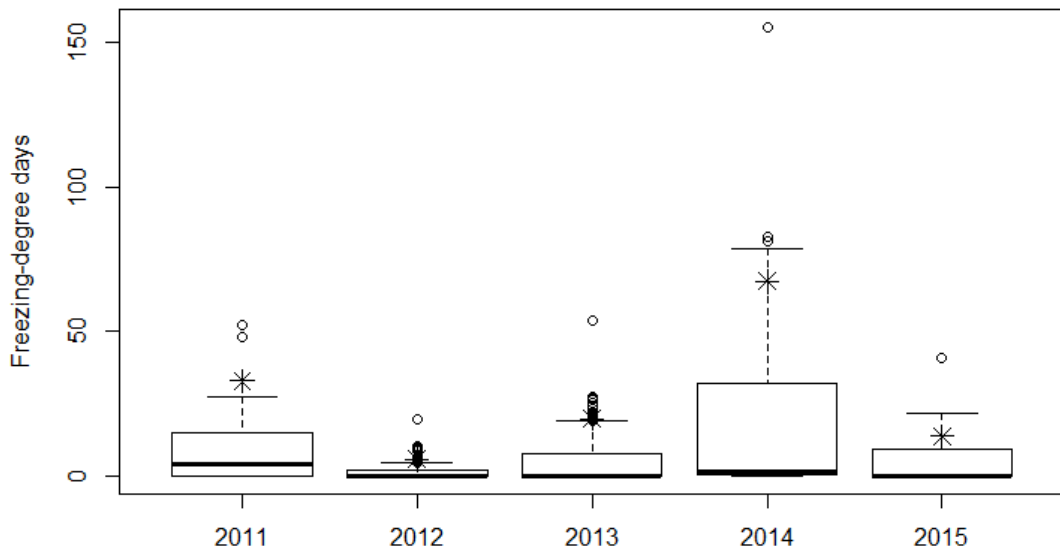
(b) Freezing-degree days summarized by year, across all months (i.e. October–January).

Figure A3: Cumulative freezing degree days summarized by month (a) and by year (b). Asterisks designate means and whiskers designate either the minimum and maximum, or $1.5 \times$ box length, whichever is closer to the median.

A.2 Tables

Table A1: *Ehrlichia chaffeensis* occurrence models as ranked by $\Delta\text{AIC}_\text{C}$.

Model	AICC
Mean T + FDD; winter (t-1)	0.00
Mean T; winter (t-1)	1.64
FDD; winter (t-1)	2.28
Minimum T; winter (t-1)	7.93
Precipitation; spring (t)	14.15
Deer presence/absence	14.17
Survey Date	14.30
Deer pellet-group density	14.56
Temperature; spring (t-1)	14.90
Ticks only	15.27
Humidity; winter (t)	15.98
Temperature; spring (t)	16.06
Temperature; winter (t)	16.64
Vapor pressure deficit; winter (t)	16.65
Precipitation; spring (t-1)	16.77
Precipitation; summer (t-1)	16.88
Vapor pressure deficit; spring (t)	17.03
Humidity; spring (t)	17.16
Precipitation; winter (t)	17.26
Null	20.30

Table A2: Seasonal weather variables for temperature (T), relative humidity (H), vapor pressure deficit (VPD), precipitation (P), and freezing-degree-days (FDD), summarized across plots and years. Coefficients of variation (CV) only given for those variables measured on ratio scales.

(a) Spring: February–May; 2011–2016

Statistic	T (°C)	H (%)	VPD (hPa)	P (mm)
Min	-1.0	47.0	2.8	22.3
Max	22.8	73.6	3.4	262.2
Range	23.8	26.5	8.0	239.9
Mean	12.6	60.1	6.0	93.6
St. Dev.	6.1	5.2	2.0	39.6
CV	-	8.7	33.4	42.3

(b) Summer: June–October; 2011–2016

Statistic	T (°C)	H (%)	VPD (hPa)	P (mm)
Min	14.8	55.5	4.5	22.2
Max	28.5	76.9	15.4	449.6
Range	13.7	21.5	10.9	427.6
Mean	23.4	68.6	9.2	142.1
St. Dev.	3.6	3.1	2.1	69.9
CV	-	4.5	22.7	49.2

(c) Winter: Novermber–January; 2012–2016

Statistic	T (°C)	H (%)	VPD (hPa)	P (mm)
Min	0.60	52.6	2.1	10.4
Max	13.6	85.3	5.7	170.3
Range	13.0	32.7	3.6	159.9
Mean	7.8	65.2	3.7	83.2
St. Dev.	3.1	6.1	0.72	38.0
CV	-	9.3	19.7	45.6

(d) Winter: October–January; 2011–2015

Statistic	Mean T (°C)	Min T (C)	FDD	H (%)	VPD (hPa)	P (mm)
Min	0.60	-5.7	3.6	52.6	2.3	10.4
Max	18.7	13.9	169	75.1	8.1	262.1
Range	18.1	19.6	165.4	22.5	5.8	251.7
Mean	9.5	4.1	27.8	64.7	4.3	89.3
St. Dev.	5.2	5.2	23.0	4.2	1.3	52.5
CV	-	-	82.7	6.6	30.7	58.9

B   Supplementary material: Chapter 2

B.1   Figures

1–3: Forest cover × edge
4–6: Urban cover × edge
7–9: Forest cover × proximity
10–12: Urban cover × proximity
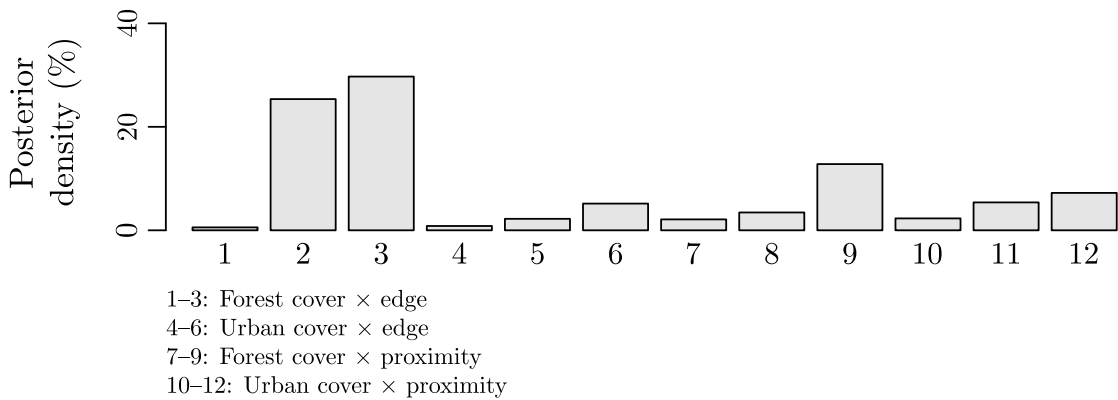
Figure B1: Posterior probability for variable and scale selection for habitat availability and fragmentation in the prevalence model. Each combination (e.g. 1–3: forest cover × edge) was derived from a 25-, 100-, and 350-ha window, respectively.

Figure B2: Posterior probability for the scale at which to include the proportion of evergreen cover in the prevalence model.

Figure B3: Posterior probability densities for grand mean and mean annual *E. chaffeensis* prevalence (back-transformed to percent-probability scale). Dots represent the posterior mean and bars represent the highest density interval (HDI).
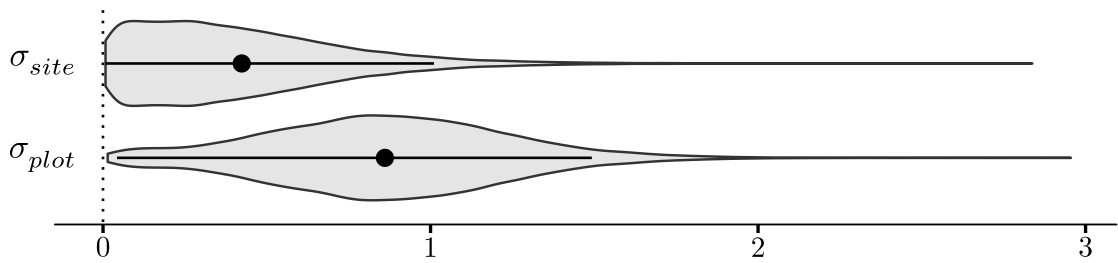
Figure B4: Posterior probability densities for the random-intercept standard deviations of the prevalence model (on the logit-scale).
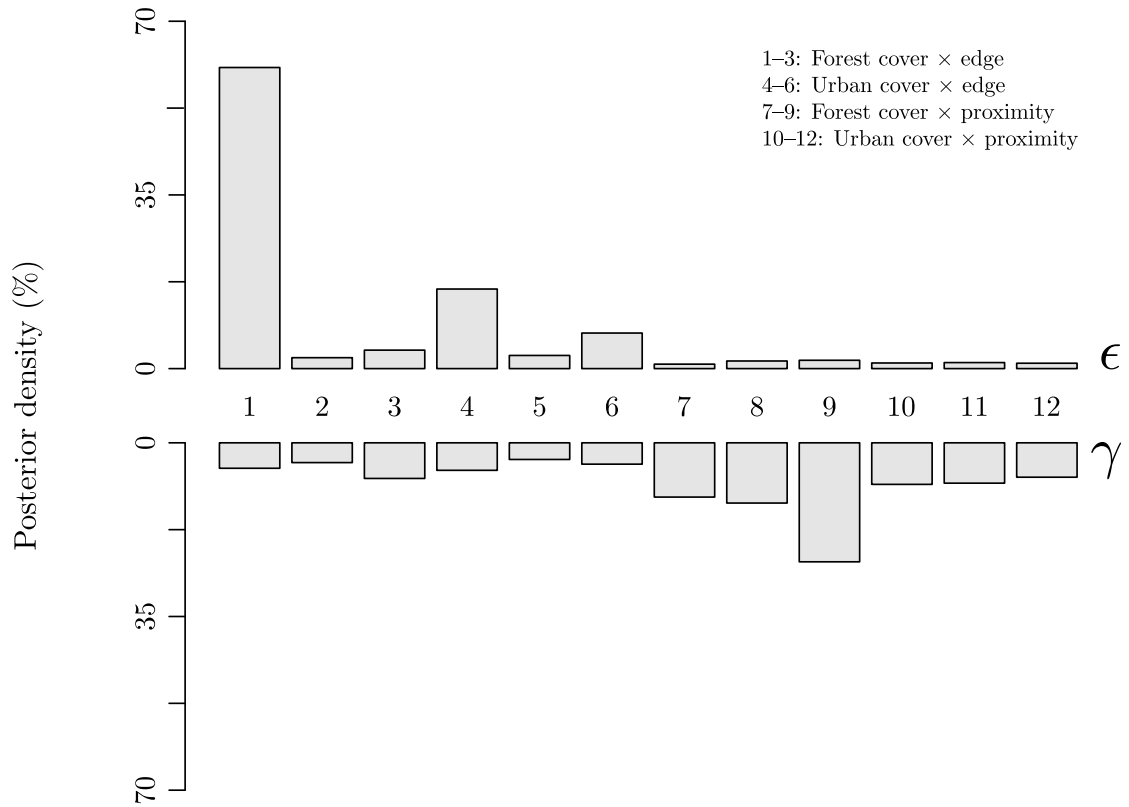
Figure B5: Posterior probability densities for variable and scale selection for habitat availability and fragmentation in the occupancy model. The upper panel regards covariates of extinction (ϵ) and the bottom panel colonization (γ). Each combination (e.g. 1–3: forest cover × edge) was derived from a 25-, 100-, and 350-ha window, respectively.

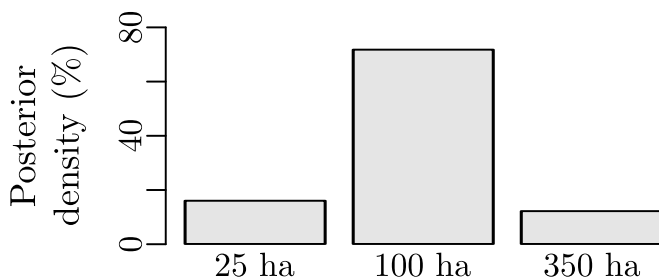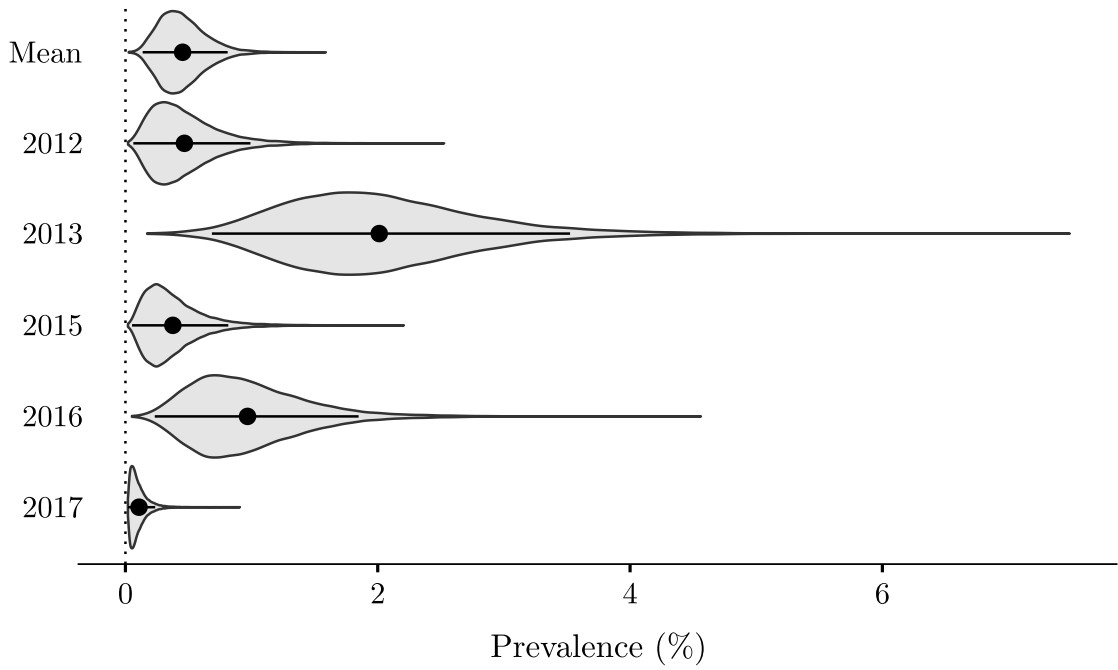**Figure B6:** Posterior probability densities for the scale at which to include the proportion of evergreen cover in the occupancy model. The left panel regards covariates of extinction (ϵ) and the right colonization (γ).
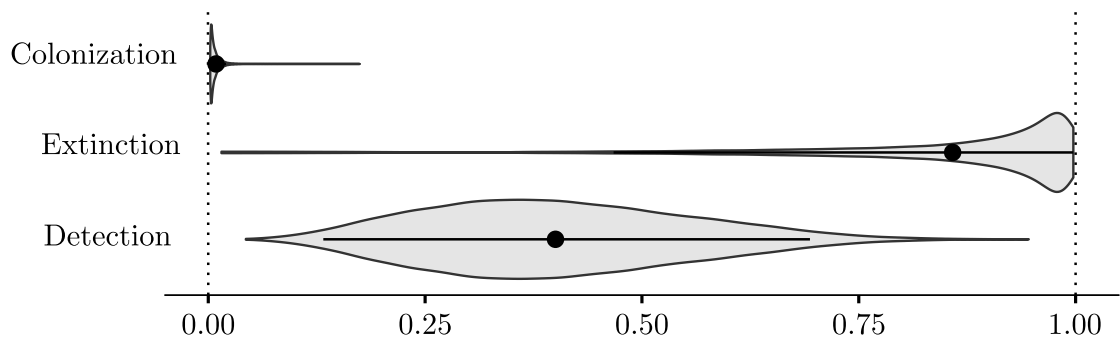
Figure B7: Posterior probability densities for colonization, extinction, and detection probability. Dots represent posterior means and bars represent HDIs.
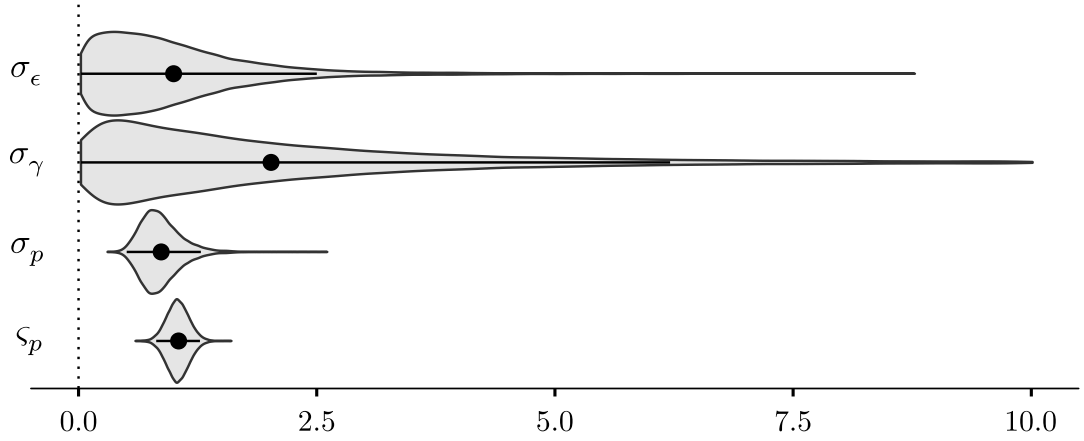
Figure B8: Posterior probability densities for the random-effects in the prevalence model. Random-intercept standard deviations are desi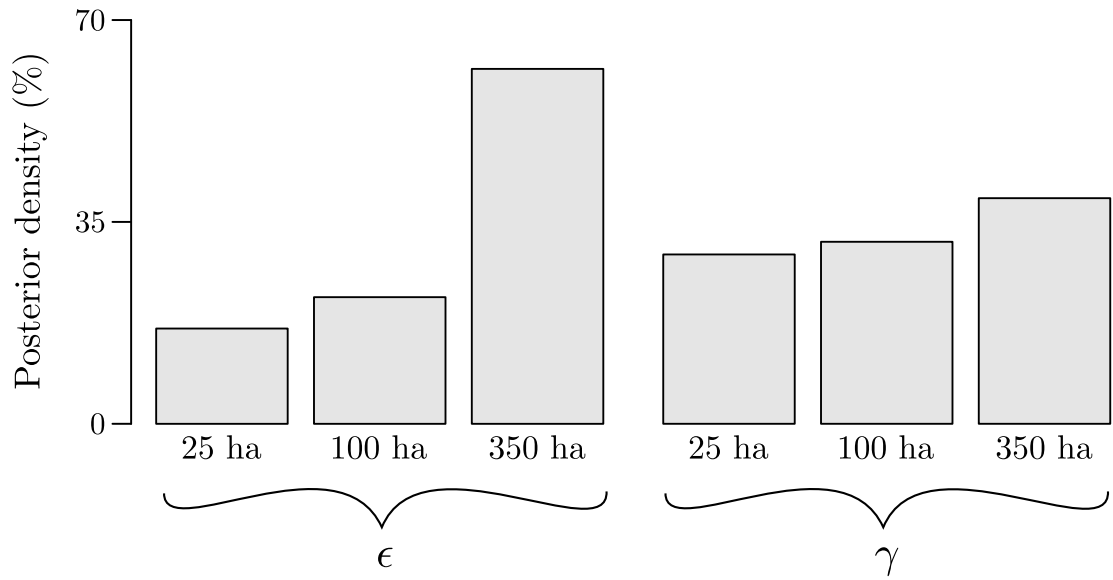gnated by σ_θ, and ς_p designates the additional variation in detection probability not explained by site or $\boldsymbol{X}_p$.

B.2 Tables

Table B1: Confusion matrix for final land-use/land-cover classification.

Class	Water	Wetland	Evergreen	Impervious surface	Deciduous	Field	Total	User's accuracy
Water	30	-	-	-	-	-	30	1.0
Wetland	5	14	-	-	1	-	20	0.7
Evergreen	-	-	86	-	6	1	93	0.92
Impervious surface	1	1	-	45	-	3	50	0.9
Deciduous	-	2	1	1	86	4	94	0.91
Field	-	-	-	2	3	128	133	0.96
Total	36	17	87	48	96	136	420	—
Producer's accuracy	0.83	0.82	0.99	0.94	0.90	0.94	—	0.93

B.3 Landscape metrics

Landscape metrics all followed skewed distributions. At the 100-ha scale, proportion forest cover was left-skewed with a range of 0.22 1.0 and a median of 0.78, and edge length is right-skewed (though less so) with a range of 102 20261 m and a median of 8361. At the 25-ha scale, proportion cover had a range of 0.26 1.0 and a median of 0.83, and edge density

had a range of 0 5984 m and a median of 1847. Proportion evergreen cover was very right skewed. At the 100-ha scale, it had a range of 0.01 0.95 and a median of 0.18, and at the 350-ha scale it had a range of 0.05 0.77 with a median of 0.20. Distance to water was also right-skewed, with a range of 0 1251 m and a median of 183 m.

B.4 Bayesian Models

Following are mathematical descriptions of the Bayesian models used in Chapter 2 together with visual representations in the form of directed acyclic graphs (DAGs). The models are shown in the form $[\boldsymbol{\theta}|y] \propto [y|\boldsymbol{\theta}][\boldsymbol{\theta}]$, in which the joint probabilities of $[y|\boldsymbol{\theta}]$ and $[\boldsymbol{\theta}]$ have been decomposed into their constituent elements. The relation between these elements is also described in the DAG: variables at the end of an arrow are conditioned on those at its start. Solid arrows describe probabilistic relationships and dotted arrows denote deterministic ones. Variables without any leading arrows are either described by observed data (and are connected to dotted arrows) or are given prior probability distributions (and are connected to solid arrows).

B.4.1 Prevalence model

$$[\rho_{ijt}, \beta_{0ij}, \boldsymbol{\beta}, \alpha_i, \sigma^2, \mu, \sigma_0^2 | y_{ijkt}] \propto$$

$$\begin{aligned}
&\Pi_{i=1}^{S} \Pi_{j=1}^{P} \Pi_{t=1}^{T} \Pi_{k=1}^{O} \\
&\text{bern}(y_{ijkt}|1 - (1 - \rho_{ijt})^{n_{ijkt}}) \times \\
&\text{norm}(\beta_{0,ij}|\alpha_i, \sigma^2) \times \\
&\text{norm}(\alpha_i|\mu, \sigma_0^2) \times \text{unif}(\sigma|0, 10) \times \\
&\text{norm}(\mu|0, 10) \times \text{unif}(\sigma_0|0, 10) \times \\
&\Pi_{h=1}^{H} \text{norm}(\beta_h|0, 10)
\end{aligned}$$

$$\text{Where}$$
$$\text{logit}(\rho_{ijt}) = \beta_{0,ij} + \boldsymbol{\beta X}$$

Here, y_{kijt} is an observation of pathogen presence or absence in a pool of n ticks from transect k of plot j in site i during year t, and ρ_{ijt} is the probability that a single tick is

56

infected at plot ij in year t (i.e. tick infection prevalence). The logistic transformation of ρ is a linear model with nested plot- and site-level random intercepts in which a plot-level mean, $\beta_{0,ij}$, varies around its site-level mean, α_i, according to variance σ^2, which in turn varies around the grand mean, μ, according to σ_0^2.
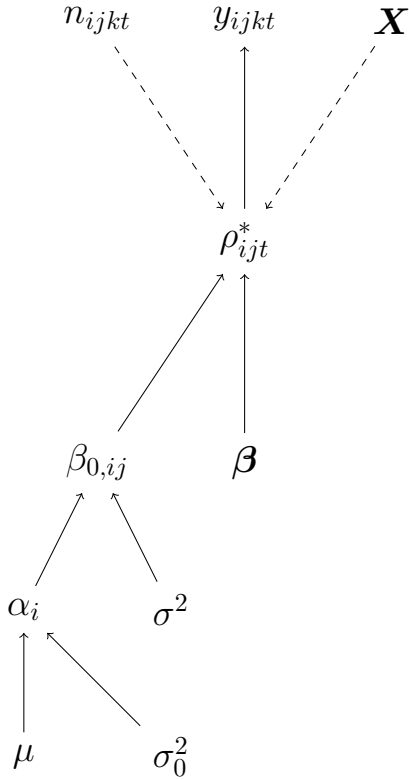

Figure B9: DAG for the pathogen prevalence model. Here, $\rho^* = 1 - (1 - \rho_{ijt})^{n_{ijkt}}$, where ρ is the probability of a single tick at plot ij being infected in year t, and n is the number of ticks collected on given transect ijk in year t.

B.4.2 Occupancy model

$[Z_{ijt},\ \gamma_{ij},\ \epsilon_{ij},\ p_{ijt},\ \boldsymbol{\beta}_\gamma,\ \boldsymbol{\beta}_\epsilon,\ \boldsymbol{\beta}_p,\boldsymbol{\alpha},\ \boldsymbol{\sigma},\ \varsigma|y_{ijkt}] \propto$

$$\Pi_{t=1}^{T}\Pi_{i=1}^{S}\Pi_{j=1}^{P}\Pi_{k=1}^{O}\text{bern}(y_{ijkt}|Z_{ijt}\times p_{ijt})\times$$
$$\text{bern}(Z_{ijt}|Z_{ij,t-1}(1-\epsilon_{ij})+(1-Z_{ij,t-1})\gamma_{ij})\times$$
$$\text{bern}(Z_{ij,t=0}|0.8)\times$$
$$\text{norm}(\text{logit}(p_{ijt})|\beta_{0,i}^{p}+\boldsymbol{\beta}_p\boldsymbol{X}_p,\varsigma)\times$$
$$\text{norm}(\beta_{0,i}^{\gamma}|\alpha_\gamma,\sigma_\gamma^2)\times\text{norm}(\beta_{0,i}^{\epsilon}|\alpha_\epsilon,\sigma_\epsilon^2)\times\text{norm}(\beta_{0,i}^{p}|\alpha_p,\sigma_p^2)\times$$
$$\text{norm}(\alpha_\gamma|0,5)\times\text{norm}(\alpha_\epsilon|0,5)\times\text{norm}(\alpha_p|0,5)\times$$
$$\text{unif}(\sigma_\gamma|0,10)\times\text{unif}(\sigma_\epsilon|0,10)\times\text{unif}(\sigma_p|0,10)\times$$
$$\Pi_{h=1}^{H_\gamma}\text{norm}(\boldsymbol{\beta}_h^{\gamma}|0,5)\times\Pi_{h=1}^{H_\epsilon}\text{norm}(\boldsymbol{\beta}_h^{\epsilon}|0,5)\times$$
$$\Pi_{h=1}^{H_p}\text{norm}(\boldsymbol{\beta}_h^{p}|0,5)\times\text{unif}(\varsigma^2|0,10)$$

Where
$$\text{logit}(\gamma_{ij})=\beta_{0,i}^{\gamma}+\boldsymbol{\beta}_\gamma\boldsymbol{X}_\gamma,$$
$$\text{logit}(\epsilon_{ij})=\beta_{0,i}^{\epsilon}+\boldsymbol{\beta}_\epsilon\boldsymbol{X}_\epsilon$$

Individual observations, k, at plot j of site i in year t are denoted by y_{ijkt}. Z_{ijt} is the true occupancy state of site i, plot j in year t, γ_{ij} is the probability of plot ij being occupied in year t given it was unoccupied in $t-1$, ϵ_{ij} is the probability that plot ij is unoccupied in year t, given it was occupied in $t-1$, and p_{ijt} is detection probability of plot ij in year t. The intercepts in the linear models, $\beta_{0,i}^{\theta}$, represent site-level means, which vary around the grand means, α_θ, according to σ_θ^2. Detection probability was also allowed to vary beyond what is described by $\boldsymbol{X}_p$, according to ς^2, in order to describe differences changes in detection not described by the included environmental variables. H_θ denotes the number of covariates associated with θ.

For 2011, which was unobserved, $Z_{ij,t=0}$ was a Bernoulli-distributed variable with a probability equal to the adjusted occupancy estimate from a preliminary analysis in Program PRESENCE. In 2014, which was also unobserved, Z was treated as in any other year but was not directly informed by any other data. Instead, the probability of occupancy in 2014

58

was only conditioned on the occupancy state of 2013 and on estimates γ and ϵ.

Note that the superscripts on some βs are not exponents but alternative indices to describe the parameter with which they are associated.
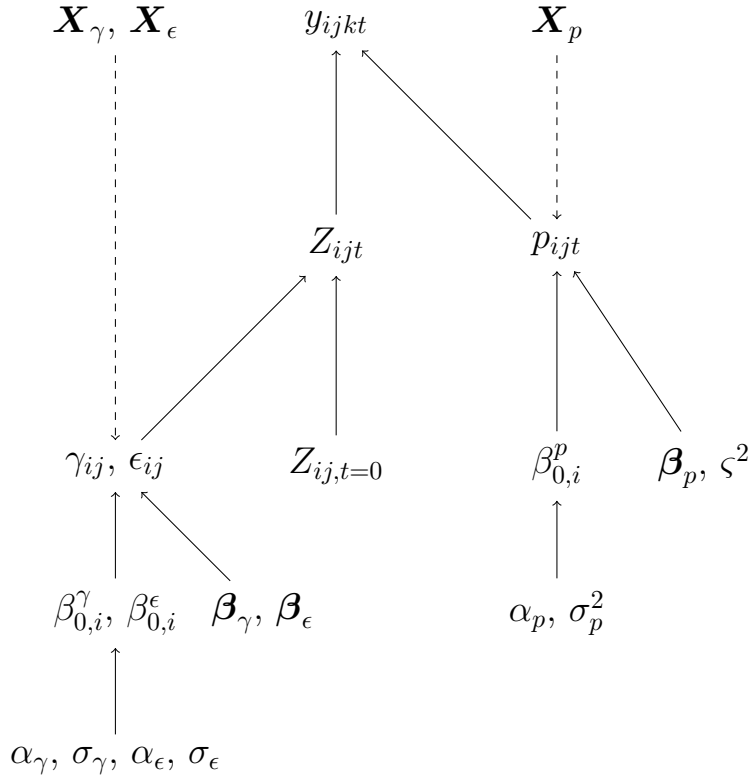
Figure B10: Directed acyclic graph of the multi-season occupancy model.

B.5 JAGS code

B.5.1 Global prevalence model

```
# Likelihood:
for(i in 1:nsite){
  for(j in 1:nplot[i]){ # The number of plots in i
    for(t in 1:nyear[i,j]){ # Number of years ij was surveyed
      for(k in 1:nobs[i,j,t]){ # Number of observations at ijt
        y[i,j,k,t] ~ dbern(1-(1-rho[i,j,t])^n[i,j,k,t])
      }
      logit(rho[i,j,t]) <- beta0[i,j]
      + inprod(beta[], X[i,j,])
      + inprod(beta.year.raw[], year[i,j,t,])
    }
    beta0[i,j] ~ dnorm(alpha[i], sigma^-2)
```

```
    }
  alpha[i] ~ dnorm(mu.raw, sigma0^-2)
}


# Priors:
for(z in 1:nbetas){
  beta[z] ~ dnorm(0, 5^-2)
}


for(t in 1:nyear.max){
  beta.year.raw[t] ~ dnorm(0, 5^-2)
}


mu.raw ~ dnorm(logit(.018), 5^-1)

sigma0 ~ dunif(0, 10)
sigma ~ dunif(0,10)

# Sum-to-zero constraint for mu and beta.year
for(t in 1:nyear.max){
  mean.year[t] <- mu.raw + beta.year.raw[t]
}


mu <- mean(mean.year[])


for(t in 1:nyear.max){
  beta.year[t] <- mean.year[t] - mu
}
```

B.5.2 Prevalence model scale- and variable-selection

```
# Likelihood:
for(i in 1:nsite){
  for(j in 1:nplot[i]){ # Number of plots in each site
    for(t in 1:nyear[i,j]){ # Number of years ij was surveyed
      for(k in 1:nobs[i,j,t]){ # Observations at ijt
        y[i,j,k,t] ~ dbern(1-(1-rho[i,j,t])^n[i,j,k,t])
      }
      logit(rho[i,j,t]) <-
        beta0[i,j] +
        inprod(beta.year.raw[], year[i,j,t,]) +

        equals(I1, 1)*(beta[1]*prop.forest[i,j,1] +
          beta[2]*edge.forest[i,j,1] +
          beta[3]*prop.forest[i,j,1]*edge.forest[i,j,1]) +
        equals(I1, 2)*(beta[1]*prop.forest[i,j,2] +
          beta[2]*edge.forest[i,j,2] +
          beta[3]*prop.forest[i,j,2]*edge.forest[i,j,2]) +
        equals(I1, 3)*(beta[1]*prop.forest[i,j,3] +
          beta[2]*edge.forest[i,j,3] +
          beta[3]*prop.forest[i,j,3]*edge.forest[i,j,3]) +

        equals(I1, 4)*(-beta[1]*prop.urban[i,j,1] +
```

```
                 beta [2]* edge . forest [i,j,1] -
                 beta [3]* prop . urban [i,j,1]* edge . forest [i,j,1]) +
             equals ( I1 , 5)*( - beta [1]* prop . urban [i,j,2] +
                 beta [2]* edge . forest [i,j,2] -
                 beta [3]* prop . urban [i,j,2]* edge . forest [i,j,2]) +
             equals ( I1 , 6)*( - beta [1]* prop . urban [i,j,3] +
                 beta [2]* edge . forest [i,j,3] -
                 beta [3]* prop . urban [i,j,3]* edge . forest [i,j,3]) +

             equals ( I1 , 7)*( beta [1]* prop . forest [i,j,1] -
                 beta [2]* prox . forest [i,j,1] -
                 beta [3]* prop . forest [i,j,1]* prox . forest [i,j,1]) +
             equals ( I1 , 8)*( beta [1]* prop . forest [i,j,2] -
                 beta [2]* prox . forest [i,j,2] -
                 beta [3]* prop . forest [i,j,2]* prox . forest [i,j,2]) +
             equals ( I1 , 9)*( beta [1]* prop . forest [i,j,3] -
                 beta [2]* prox . forest [i,j,3] -
                 beta [3]* prop . forest [i,j,3]* prox . forest [i,j,3]) +

             equals ( I1 , 10)*( - beta [1]* prop . urban [i,j,1] -
                 beta [2]* prox . forest [i,j,1] +
                 beta [3]* prop . forest [i,j,1]* prox . forest [i,j,1]) +
             equals ( I1 , 11)*( - beta [1]* prop . urban [i,j,2] -
                 beta [2]* prox . forest [i,j,2] +
                 beta [3]* prop . forest [i,j,2]* prox . forest [i,j,2]) +
             equals ( I1 , 12)*( - beta [1]* prop . urban [i,j,3] -
                 beta [2]* prox . forest [i,j,3] +
                 beta [3]* prop . forest [i,j,3]* prox . forest [i,j,3]) +

             beta [4]* prop . eg [i,j,I2]
             + beta [5]* dist . water [i,j]
      }

      beta0 [i,j] ~ dnorm ( alpha [i], sigma ^ -2)
   }
   alpha [i] ~ dnorm ( mu . raw , sigma0 ^ -2)
}

# Priors :

prec <- 5^ -1 # Precision for parameter estimates :

for(i in 1:5){
   beta [i] ~ dnorm (0 , prec )
}

for(t in 1: nyear . max ){
   beta . year . raw [t] ~ dnorm (0 , prec )
}

I1 ~ dcat ( rep (1 ,12)) # Categorical indicator variables
I2 ~ dcat ( rep (1 ,3))
```

```
mu.raw ~ dnorm(logit(.018), prec)

sigma0 ~ dunif(0,10)
sigma ~ dunif(0,10)


# Sum-to-zero constraint for mu and beta.year
for(t in 1:nyear.max){
  mean.year[t] <- mu.raw + beta.year.raw[t]
}

mu <- mean(mean.year[])

for(t in 1:nyear.max){
  beta.year[t] <- mean.year[t] - mu
}
```

B.5.3 Global occupancy model

```
# Unobserved years:

## Year 0:
for(i in 1:nsites){ # Total number of sites
  for(j in 1:nplots[i]){ # Number of plots in each site
    z[i,j,1] ~ dbern(0.8) # Prior on year 0
  }
}

## 2014
for(i in 1:nsites){
  for(j in 1:nplots[i]){
    z[i,j,4] ~ dbern(z[i,j,3]*(1-eps[i,j])+(1-z[i,j,3])*gam[i,j])
  }
}

# Observed years:
for(i in 1:nsites){ # Total number of sites
  for(j in 1:nplots[i]){ # Number of plots in each site
    for(t in yearsObs[]){ # Years observed
      for(k in 1:nobs){ # Number secondaries (all 4 in this case)
        obs[i,j,k,t] ~ dbern(z[i,j,t]*p[i,j,t])
      }
      z[i,j,t] ~ dbern(z[i,j,t-1]*(1-eps[i,j]) +
              (1-z[i,j,t-1])*gam[i,j])
      logit(p[i,j,t]) <- det.odds[i,j,t]
      det.odds[i,j,t] ~ dnorm(b0.p[i] + inprod(beta.p[], X.p[i,j,t,]),
                det.error^-2)
    }
    logit(eps[i,j]) <- b0.e[i] + inprod(beta.eps[], X.e[i,j,])
    logit(gam[i,j]) <- b0.g[i] + inprod(beta.gam[], X.g[i,j,])
  }
  # Random intercepts:
  b0.e[i] ~ dnorm(alpha.e, sigma.e^-2) # Normal distributions defined
```

```
  b0.g[i] ~ dnorm(alpha.g, sigma.g^-2) # using   precision,
  b0.p[i] ~ dnorm(alpha.p, sigma.p^-2) # the inverse of variance.
}

# Priors & derivations
prec <- 5^-1 # Precision for betas and alphas

for(i in 1:nbeta.e){ # Number of covariates for epsilon
  beta.eps[i] ~ dnorm(0, prec)
}
for(i in 1:nbeta.g){
  beta.gam[i] ~ dnorm(0, prec)
}

for(i in 1:nbeta.p){
  beta.p[i] ~ dnorm(0, prec)
}

alpha.e ~ dnorm(0, prec)
alpha.g ~ dnorm(0, prec)
alpha.p ~ dnorm(0, prec)

sigma.e ~ dunif(0,10)
sigma.g ~ dunif(0,10)
sigma.p ~ dunif(0,10)

det.error ~ dunif(0, 10)
```

B.5.4 Occupancy scale- and variable-selection

```
# Unobserved years:
for(i in 1:nsites){
  for(j in 1:nplots[i]){
    z[i,j,1] ~ dbern(0.8)
  }
}

for(i in 1:nsites){
  for(j in 1:nplots[i]){
    z[i,j,4] ~ dbern(z[i,j,3]*(1-eps[i,j])+(1-z[i,j,3])*gam[i,j])
  }
}

# Observed years:
for(i in 1:nsites){
  for(j in 1:nplots[i]){
    for(t in yearsObs[]){
      for(k in 1:nobs){
        obs[i,j,k,t] ~ dbern(z[i,j,t]*p[i,j,t])
      }
      z[i,j,t] ~ dbern(z[i,j,t-1]*(1-eps[i,j]) +
              (1-z[i,j,t-1])*gam[i,j])
      logit(p[i,j,t]) <- det.odds[i,j,t]
```

```
        det.odds[i,j,t] ~ dnorm(b0.p[i] +
                    inprod(beta.p[], X.p[i,j,t,]),
                    det.error^-2)
}
logit(eps[i,j]) <-
    b0.e[i] +
    equals(ie1,  1)*(beta.eps[1]*prop.forest[i,j,1] +
        beta.eps[2]*edge.forest[i,j,1] +
        beta.eps[3]*prop.forest[i,j,1]*edge.forest[i,j,1]) +
    equals(ie1,  2)*(beta.eps[1]*prop.forest[i,j,2] +
        beta.eps[2]*edge.forest[i,j,2] +
        beta.eps[3]*prop.forest[i,j,2]*edge.forest[i,j,2]) +
    equals(ie1,  3)*(beta.eps[1]*prop.forest[i,j,3] +
        beta.eps[2]*edge.forest[i,j,3] +
        beta.eps[3]*prop.forest[i,j,3]*edge.forest[i,j,3]) +

    equals(ie1,  4)*(-beta.eps[1]*prop.urban[i,j,1] +
        beta.eps[2]*edge.forest[i,j,1] -
        beta.eps[3]*prop.urban[i,j,1]*edge.forest[i,j,1]) +
    equals(ie1,  5)*(-beta.eps[1]*prop.urban[i,j,2] +
        beta.eps[2]*edge.forest[i,j,2] -
        beta.eps[3]*prop.urban[i,j,2]*edge.forest[i,j,2]) +
    equals(ie1,  6)*(-beta.eps[1]*prop.urban[i,j,3] +
        beta.eps[2]*edge.forest[i,j,3] -
        beta.eps[3]*prop.urban[i,j,3]*edge.forest[i,j,3]) +

    equals(ie1,  7)*(beta.eps[1]*prop.forest[i,j,1] -
        beta.eps[2]*prox.forest[i,j,1] -
        beta.eps[3]*prop.forest[i,j,1]*prox.forest[i,j,1]) +
    equals(ie1,  8)*(beta.eps[1]*prop.forest[i,j,2] -
        beta.eps[2]*prox.forest[i,j,2] -
        beta.eps[3]*prop.forest[i,j,2]*prox.forest[i,j,2]) +
    equals(ie1,  9)*(beta.eps[1]*prop.forest[i,j,3] -
        beta.eps[2]*prox.forest[i,j,3] -
        beta.eps[3]*prop.forest[i,j,3]*prox.forest[i,j,3]) +

    equals(ie1, 10)*(-beta.eps[1]*prop.urban[i,j,1] -
        beta.eps[2]*prox.forest[i,j,1] +
        beta.eps[3]*prop.forest[i,j,1]*prox.forest[i,j,1]) +
    equals(ie1, 11)*(-beta.eps[1]*prop.urban[i,j,2] -
        beta.eps[2]*prox.forest[i,j,2] +
        beta.eps[3]*prop.forest[i,j,2]*prox.forest[i,j,2]) +
    equals(ie1, 12)*(-beta.eps[1]*prop.urban[i,j,3] -
        beta.eps[2]*prox.forest[i,j,3] +
        beta.eps[3]*prop.forest[i,j,3]*prox.forest[i,j,3]) +

    beta.eps[4]*prop.eg[i,j,ie2] +
    beta.eps[5]*dist.water[i,j]

logit(gam[i,j]) <-
    b0.g[i] +
    equals(ig1,  1)*(beta.eps[1]*prop.forest[i,j,1] +
        beta.eps[2]*edge.forest[i,j,1] +
```

```
          beta.eps[3]*prop.forest[i,j,1]*edge.forest[i,j,1]) +
      equals(ig1, 2)*(beta.eps[1]*prop.forest[i,j,2] +
        beta.eps[2]*edge.forest[i,j,2] +
        beta.eps[3]*prop.forest[i,j,2]*edge.forest[i,j,2]) +
      equals(ig1, 3)*(beta.eps[1]*prop.forest[i,j,3] +
        beta.eps[2]*edge.forest[i,j,3] +
        beta.eps[3]*prop.forest[i,j,3]*edge.forest[i,j,3]) +

      equals(ig1, 4)*(-beta.eps[1]*prop.urban[i,j,1] +
        beta.eps[2]*edge.forest[i,j,1] -
        beta.eps[3]*prop.urban[i,j,1]*edge.forest[i,j,1]) +
      equals(ig1, 5)*(-beta.eps[1]*prop.urban[i,j,2] +
        beta.eps[2]*edge.forest[i,j,2] -
        beta.eps[3]*prop.urban[i,j,2]*edge.forest[i,j,2]) +
      equals(ig1, 6)*(-beta.eps[1]*prop.urban[i,j,3] +
        beta.eps[2]*edge.forest[i,j,3] -
        beta.eps[3]*prop.urban[i,j,3]*edge.forest[i,j,3]) +

      equals(ig1, 7)*(beta.eps[1]*prop.forest[i,j,1] -
        beta.eps[2]*prox.forest[i,j,1] -
        beta.eps[3]*prop.forest[i,j,1]*prox.forest[i,j,1]) +
      equals(ig1, 8)*(beta.eps[1]*prop.forest[i,j,2] -
        beta.eps[2]*prox.forest[i,j,2] -
        beta.eps[3]*prop.forest[i,j,2]*prox.forest[i,j,2]) +
      equals(ig1, 9)*(beta.eps[1]*prop.forest[i,j,3] -
        beta.eps[2]*prox.forest[i,j,3] -
        beta.eps[3]*prop.forest[i,j,3]*prox.forest[i,j,3]) +

      equals(ig1, 10)*(-beta.eps[1]*prop.urban[i,j,1] -
        beta.eps[2]*prox.forest[i,j,1] +
        beta.eps[3]*prop.forest[i,j,1]*prox.forest[i,j,1]) +
      equals(ig1, 11)*(-beta.eps[1]*prop.urban[i,j,2] -
        beta.eps[2]*prox.forest[i,j,2] +
        beta.eps[3]*prop.forest[i,j,2]*prox.forest[i,j,2]) +
      equals(ig1, 12)*(-beta.eps[1]*prop.urban[i,j,3] -
        beta.eps[2]*prox.forest[i,j,3] +
        beta.eps[3]*prop.forest[i,j,3]*prox.forest[i,j,3]) +

      beta.gam[4]*prop.eg[i,j,ig2] +
      beta.gam[5]*dist.water[i,j]
  }
  # Random intercepts:
  b0.e[i] ~ dnorm(alpha.e, sigma.e^-2)
  b0.g[i] ~ dnorm(alpha.g, sigma.g^-2)
  b0.p[i] ~ dnorm(alpha.p, sigma.p^-2)
}


# Priors & derivations

for(i in 1:5){
  beta.eps[i] ~ dnorm(0, 5^-1)
}
```

```
for(i in 1:5){
  beta.gam[i] ~ dnorm(0, 5^-1)
}

for(i in 1:2){
  beta.p[i] ~ dnorm(0, 5^-1)
}

ie1 ~ dcat(rep(1,12))
ie2 ~ dcat(rep(1,3))
ig1 ~ dcat(rep(1,12))
ig2 ~ dcat(rep(1,3))

alpha.e ~ dnorm(0, 5^-1)
alpha.g ~ dnorm(0, 5^-1)
alpha.p ~ dnorm(0, 5^-1)

sigma.e ~ dunif(0,10)
sigma.g ~ dunif(0,10)
sigma.p ~ dunif(0,10)

det.error ~ dunif(0,10)

meanEps <- ilogit(alpha.e)
meanGam <- ilogit(alpha.g)
meanP <- ilogit(alpha.p)
```

References

Alduchov, O. a., and R. E. Eskridge. **1996**. Improved magnus form approximation of saturation vapor pressure. J. Appl. Meteorol. 35: 601–609.

Allan, B. F., L. S. Goessling, G. A. Storch, and R. E. Thach. **2010**. Blood meal analysis to identify reservoir hosts for *Amblyomma americanum* ticks. Emerg. Infect. Dis. 16: 433–440.

Allan, B. F., F. Keesing, and R. Ostfeld. **2003**. Effect of forest fragmentation on Lyme disease risk. Conserv. Biol. 17: 267–272.

Anderson, B. E., K. G. Sims, J. G. Olson, J. E. Childs, J. F. Piesman, C. M. Happ, G. O. Maupin, and B. J. B. Johnson. **1993**. Amblyomma americanum: A potential vector of human ehrlichiosis. Am. J. Trop. Med. Hyg. 49: 239–244.

Anderson, B. E., J. W. Sumner, J. E. Dawson, T. Tzianabos, C. R. Greene, J. G. Olson, D. B. Fishbein, M. Olsen-rasmussen, B. P. Holloway, E. H. George, and A. F. Azad. **1992**. Detection of the etiologic agent of human ehrlichisosis by polymersase chain reaction. J. Clin. Microbiol. 30: 775–780.

Anderson, D. R. 2008. Model Based Inference in the Life Sciences: A Primer on Evidence, 1st ed. Springer, New York.

Bates, D., M. Machler, B. Bolker, and S. Walker. **2015**. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67: 1–48.

Beasley, J. C., T. L. Devault, and O. E. J. Rhodes. **2007**. Home-Range Attributes of Raccoons in a Fragmented Agricultural Region of Northern Indiana. J. Wildl. Manage. 71: 844–850.

Beyer, H. L. 2014. Geospatial Modeling Environment.

Biggerstaff, B. J. 2009. PooledInfRate: a Microsoft Office Excel Add-In to compute

prevalence estimates from pooled samples.

(https://www.cdc.gov/westnile/resourcepages/mosqSurvSoft.html).

Bishopp, F. C., and H. L. Trembely. **1945**. Distribution and hosts of certain North

American ticks. J. Parasitol. 31: 1–54.

Brownstein, J. S., D. K. Skelly, T. R. Holford, and D. Fish. **2005**. Forest fragmentation

predicts local scale heterogeneity of Lyme disease risk. Oecologia. 146: 469–475.

Burks, C. S., R. L. Stewart, G. R. Needham, and R. E. Lee. **1996**. The role of direct

chilling injury and inoculative freezing in cold tolerance of *Amblyomma*

americanum, *Dermacentor variabilis* and *Ixodes scapularis*. Physiol. Entomol. 21:

44–50.

Burnham, K. P., D. R. Anderson, and K. P. Huyvaert. **2011**. AIC model selection and

multimodel inference in behavioral ecology: Some background, observations, and

comparisons. Behav. Ecol. Sociobiol. 65: 23–35.

Campbell, T. A., B. R. Laseter, W. M. Ford, and K. V. Miller. **2004**. Feasibility of

localized management to control white-tailed deer in forest regeneration areas.

Wildl. Soc. Bull. 32: 1124–1131.

Carroll, B. K., and D. L. Brown. **1977**. Factors affecting neonatal fawn survival in

southern-central Texas. J. Wildl. Manage. 41: 63–69.

Centers for Disease Control and Prevention. **2018a**. Ehrlichiosis - Statistics and

Epidemiology. (https://www.cdc.gov/ehrlichiosis/stats/index.html).

Centers for Disease Control and Prevention. **2018b**. Lyme Disease - Transmission.

(https://www.cdc.gov/lyme/transmission/index.html).

Childs, J. E., and C. D. Paddock. **2003**. The ascendancy of *Amblyomma americanum* as a vector of pathogens affecting humans in the Unites States. Annu. Rev. Entomol. 48: 307–337.

Cornicelli, L., A. Woolf, and J. L. Roseberry. **1996**. White-tailed deer use of a suburban environment in southern Illinois. Trans. illinois State Acad. Sci. 89: 93–103.

Côté, I. M., and M. R. Gross. **1993**. Reduced disesae in offpsring: a benefit of coloniality in sunfish. Behav. Ecol. Sociobiol. 33: 269–274.

Coulson, T., E. J. Milner-Gulland, and T. Clutton-Brock. **2000**. The relative roles of density and climatic variation on population dynamics and fecundity rates in three contrasting ungulate species. Proc. R. Soc. London. 267: 1771–1779.

Daly, C., M. Halbleib, and J. I. Smith. **2001**. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. Encycl. Atmos. Sci. 4: 1549–1555.

Davidson, W. R., J. M. Lockhart, D. E. Stallknecht, E. W. Howerth, J. E. Dawson, and Y. Rechav. **2001**. Persistent *Ehrlichia chaffeensis* infection in white-tailed deer. J. Wildl. Dis. 37: 538–546.

Dobson, A. D. M. 2013. Ticks in the wrong boxes: assessing error in blanket-drag studies due to occasional sampling. Parasit. Vectors. 6: 344.

Ebel, G. D. 2010. Update on Powassan virus: emergence of a North American tick-borne flavivirus. Annu. Rev. Entomol. 55: 95–110.

Estrada-Peña, A., Z. Vatansever, A. Gargili, and Ö. Ergönul. **2010**. The trend towards habitat fragmentation is the key factor driving the spread of Crimean-Congo

haemorrhagic fever. Epidemiol. Infect. 138: 1194–1203.

Gaff, H., and E. Schaefer. **2010**. Disease Transmission Modelling, pp. 51–65. *In* Michael, E., Spear, R.C. (eds.), Model. Parasite Transm. Control. Landes Bioscience and Springer Science+Business Media.

Gaughan, C. R., and S. DeStefano. **2005**. Movement patterns of rural and suburban white-tailed deer in Massachusetts. Urban Ecosyst. 8: 191–202.

Glueck, T. F., W. R. Clark, and R. D. Andrews. **1988**. Raccoon movement and habitat use during the fur harvest season. Wildl. Soc. Bull. 16: 6–11.

Habib, T. J., E. H. Merrill, M. J. Pybus, and D. W. Coltman. **2011**. Modelling landscape effects on density-contact rate relationships of deer in eastern Alberta: Implications for chronic wasting disease. Ecol. Modell. 222: 2722–2732.

Harmon, J. R., M. C. Scott, E. M. Baker, C. J. Jones, and G. J. Hickling. **2015**. Molecular identification of *Ehrlichia* species and host bloodmeal source in Amblyomma americanum L. from two locations in Tennessee, United States. Ticks Tick. Borne. Dis. 6: 246–252.

Hasapes, S. K., and C. E. Comer. **2016**. Adult white-tailed deer seasonal home range and habitat composition in Northwest Louisiana. J. Southeast. Assoc. Fish Wildl. Agencies. 3: 243–252.

Heitman, N., F. Scott Dahlgren, N. A. Drexler, R. F. Massung, and C. B. Behravesh. **2016**. Increasing incidence of ehrlichiosis in the United States: A summary of national surveillance of *Ehrlichia chaffeensis* and *Ehrlichia ewingii* infections in the United States, 2008-2012. Am. J. Trop. Med. Hyg. 94: 52–60.

Holzenbein, S., and L. R. Marchington. **1992**. Spatial integration and maturing-male

white-tailed deer into the adult population. J. Mammal. 73: 326–334.

Hooten, M. B., and N. T. Hobbs. **2015**. A guide to Bayesian model selection for

ecologists. Ecol. Monogr. 85: 3–28.

Jirinec, V., D. A. Cristol, and M. Leu. **2017**. Songbird community varies with deer use

in a fragmented landscape. Landsc. Urban Plan. 161: 1–9.

De Keukeleire, M., S. O. Vanwambeke, E. Somassè, B. Kabamba, V. Luyasu, and A.
Robert. **2015**. Scouts, forests, and ticks: Impact of landscapes on human-tick

contacts. Ticks Tick. Borne. Dis. 6: 636–644.

Kilpatrick, H. J., A. M. Labonte, and K. C. Stafford. **2014**. The relationship between

deer density, tick abundance, and human cases of lyme desease in a residential

community. J. Med. Entomol. 51: 777–784.

Koch, H. G. 1984. Survival of the lone star tick, *Amblyomma americanum* (Acari:

Ixodidae), in contrasting habitats and different years in southeastern Oklahoma,

USA. J. Med. Entomol. 21: 69–79.

Kollars, T. M. J., J. H. J. Oliver, L. A. Durden, and P. G. Kollars. **2000**. Host

associations and seasonal activity of *Amblyomma americanum* (Acari: Ixodidae) in

Missouri. J. Parasitol. 86: 1156–1159.

Kunkel, K. E., L. E. Stevens, S. E. Stevens, L. Sun, E. Janssen, D. Wuebbles, C. E. I.
Konrad, C. M. Fuhrman, B. D. Keim, M. C. Kruk, A. Billot, H. Needham, M.
Shafer, and J. G. Dobson. **2013**. Regional Climate Trends and Scenarios for the U .

S . National Climate Assessment: Climate of the Southeast. Washington, D.C.

Liu, Y., R. B. Lund, S. K. Nordone, M. J. Yabsley, and C. S. McMahan. **2017**. A

Bayesian spatio-temporal model for forecasting the prevalence of antibodies to

Ehrlichia species in domestic dogs within the contiguous United States. Parasit. Vectors. 10: 138.

Lovely, K. R., W. J. McShea, N. W. Lafon, and D. E. Carr. **2013**. Land parcelization and deer population densities in a rural county of Virginia. Wildl. Soc. Bull. 37: 360–367.

Macaluso, K. R., A. Mulenga, J. a Simser, A. F. Azad, and D. Pcr. **2003**. Differential expression of genes in uninfected and *Rickettsia*-infected *Dermacentor variabilis* ticks as assessed by differential-display PCR. 71: 6165–6170.

MacKenzie, D. D., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. **2003**. Estimating Site Occupancy, Colonization, and Local Extinction When a Species Is Detected Imperfectly. Ecology. 84: 2200–2207.

Manangan, J. S., S. H. Schweitzer, N. Nibbelink, M. J. Yabsley, S. E. J. Gibbs, and M. C. Wimberly. **2007**. Habitat Factors Influencing Distributions of *Anaplasma phagocytophilum* and *Ehrlichia chaffeensis* in the Mississippi Alluvial Valley. Vector-Borne Zoonotic Dis. 7: 563–574.

Mcgarigal, K., S. A. Cushman, and E. Ene. **2012**. Fragstats v.4: Spatial pattern analysis program for categorical and continuous maps.

McGinnes, B. S., and R. L. Downing. **1977**. Factors affecting the peak of white-tailed deer fawning in Virginia. J. Wildl. Manage. 41: 715–719.

Mela, C. F., and P. K. Kopalle. **2002**. The impact of collinearity on regression analysis: The asymmetric effect of negative and positive correlations. Appl. Econ. 34: 667–677.

Mixson, T. R., H. S. Ginsberg, S. R. Campbell, J. W. Sumner, and C. D. Paddock.

2004. Detection of *Ehrlichia chaffeensis* in adult and nymphal *Amblyomma americanum* (Acari: Ixodidae) ticks from Long Island, New York. J. Med. Entomol. 41: 1104–1110.

Mooring, M. S., and B. L. Hart. **1992**. Gropuing for protection from parasites: selfish herd and encounter-dilution effects. Behavior. 123: 173–193.

Nadolny, R. M., C. L. Wright, W. L. Hynes, D. E. Sonenshine, and H. D. Gaff. **2011**. *Ixodes affinis* (Acari: Ixodidae) in southeastern Virginia and implications for the spread of *Borrelia burgdorferi*, the agent of Lyme disease. J. Vector Ecol. 36: 464–467.

Nair, A. D. S., C. Cheng, D. C. Jaworski, L. H. Willard, M. W. Sanderson, and R. R. Ganta. **2014**. *Ehrlichia chaffeensis* infection in the reservoir host (white-tailed deer) and in an incidental host (dog) is impacted by its prior growth in macrophage and tick cell environments. PLoS One. 9.

Ogden, N. H., and L. R. Lindsay. **2016**. Effects of climate and climate change on vectors and vector-borne diseases: ticks are different. Trends Parasitol. 32: 646–656.

Oregon State University. **2018**. PRISM Climate Group. (www.prism.oregonstate.edu).

Ostfeld, R. S., O. M. Cepeda, K. R. Hazler, and M. C. Miller. **1995**. Ecology of Lyme disease: habitat associations of ticks (*Ixodes scapularis*) in a rural landscape. Ecol. Appl. 5: 353–361.

Ostfeld, R. S., C. Jones, and J. Wolff. **1996**. Of mice and mast: ecological connections in eastern deciduous forests. Bioscience. 46: 323–330.

Ostfeld, R. S., T. Levi, F. Keesing, K. Oggenfuss, and C. D. Canham. **2018**. Tick-borne disease risk in a forest food web. Ecology. 99: 0–1.

Paddock, C. D., and J. E. Childs. **2003**. *Ehrlichia chaffeensis*: a prototypical emerging

pathogen. Clin. Microbiol. Rev. 16: 37–64.

Parker, K. L., P. S. Barboza, and M. P. Gillingham. **2009**. Nutrition integrates

environmental responses of ungulates. Funct. Ecol. 23: 57–69.

Patterson, B. R., and V. A. Power. **2002**. Contributions of forage competition, harvest,

and climate fluctuation to changes in population growth of northern white-tailed

deer. Oecologia. 130: 62–71.

Patz, J. A., P. R. Epstein, T. A. Burke, and J. M. Balbus. **1996**. Global climate change

and emerging infectious diseases. J. Am. Med. Assoc. 275: 217–223.

Plummer, M. 2017. JAGS: A program for analysis of Bayesian graphcial models using

Gibbs sampling.

Post, E., and N. C. H. R. Stenseth. **1998**. Large-scale climatic fluctuation and

population dynamics of moose and white-tailed deer. J. Anim. Ecol. 67: 537–543.

R Core Team. **2017**. R: A Language and Environment for Statistical Computing.

**Rand, P. W., C. Lubelczyk, G. R. Lavigne, S. Elias, M. S. Holman, E. H. Lacombe,
and R. P. Smith**. **2003**. Deer density and the abundance of Ixodes scapularis (Acari:

Ixodidae). J. Med. Entomol. 40: 179–184.

Robinson, S. J., M. D. Samuel, D. L. Lopez, and P. Shelton. **2012**. The walk is never

random: Subtle landscape effects shape gene flow in a continuous white-tailed deer

population in the Midwestern United States. Mol. Ecol. 21: 4190–4205.

Rödel, H. G., A. Bora, P. Kaetzke, M. Khaschei, H. Hutzelmeyer, and D. Von Holst.

2004. Over-winter survival in subadult European rabbits: Weather effects, density

dependence, and the impact of individual characteristics. Oecologia. 140: 566–576.

Rosatte, R., M. Ryckman, K. Ing, S. Proceviat, M. Allan, L. Bruce, D. Donovan, and J. C. Davies. **2010**. Density, movements, and survival of raccoons in Ontario, Canada: implications for disease spread and management. J. Mammal. 91: 122–135.

Rubenstein, D. I., and M. E. Hohmann. **1989**. Parasites and Social Behavior of Island Feral Horses. Oikos. 55: 312–320.

Sá-Hungaro, I. J. B. de, V. de A. Raia, M. da C. Pinheiro, C. C. D. U. Ribeiro, and K. M. Famadas. **2014**. Amblyomma auricularium (Acari: Ixodidae): underwater survival of the non-parasitic phase of feeding females. Brazilian J. Vet. Parasitol. 23: 387–392.

Sæther, B. E. 1997. Environmental stochasticity and population dynamics of large herbivores: A search for mechanisms. Trends Ecol. Evol. 12: 143–147.

Saïd, S., and S. Servanty. **2005**. The influence of landscape structure on female roe deer home-range size. Landsc. Ecol. 20: 1003–1012.

Schulze, T. L., and R. a Jordan. **2003**. Meteorologically mediated diurnal questing of Ixodes scapularis and Amblyomma americanum (Acari: Ixodidae) nymphs. J. Med. Entomol. 40: 395–402.

Semtner, P. J., R. W. Barker, and J. A. Hair. **1971**. The ecology and behavior of the lone star tick (Acarina : Ixodidae). II. Activity and survival in different ecological habitats. J. Med. Entomol. 8: 719–725.

Skuldt, L. H., N. E. Mathews, and A. M. Oyer. **2008**. White-Tailed Deer Movements in a Chronic Wasting Disease Area in South-Central Wisconsin. J. Wildl. Manage. 72: 1156–1160.

Sonenshine, D. E., and G. F. Levy. **1971**. The ecology of the lone star tick, *Amblyomma*

americanum (L.), in two contrasting habitats in Virginia (Acarina : Ixodidae). J. Med. Entomol. 8: 623–635.

Springer, Y. P., C. S. Jarnevich, D. T. Barnett, A. J. Monaghan, and R. J. Eisen. **2015**. Modeling the present and future geographic distribution of the lone star tick, *Amblyomma americanum* (Ixodida: *Ixodidae*), in the continental United States. Am. J. Trop. Med. Hyg. 93: 875–890.

Stanek, G., G. P. Wormser, J. Gray, and F. Strle. **2012**. Lyme borreliosis. Lancet. 379: 461–473.

Stein, K. J., M. Waterman, and J. L. Waldon. **2008**. The effects of vegetation density and habitat disturbance on the spatial distribution of ixodid ticks (acari: Ixodidae). Geospat. Health. 2: 241–252.

Steiner, F. E., R. R. Pinger, and C. N. Vann. **1999**. Infection rates of *Amblyomma americanum* (Acari: Ixodidae) by *Ehrlichia chaffeensis* (Rickettsiales: Ehrlichieae) and prevalence of *E. chaffeensis*-reactive antibodies in white-tailed deer in southern Indiana, 1997. J. Med. Entomol. 36: 715–719.

Stromdahl, E. Y., M. P. Randolph, J. J. O'Brien, A. G. Gutierrez, J. J. O'Brien, and A. G. Gutierrez. **2000**. *Ehrlichia chaffeensis* (Rickettsiales: Ehrlichieae) infection in *Amblyomma americanum* (Acari: Ixodidae) at Aberdeen Proving Ground, Maryland. J. Med. Entomol. 37: 349–56.

Stuber, E. F., L. F. Gruber, and J. J. Fontaine. **2017**. A Bayesian method for assessing multi-scale species-habitat relationships. Landsc. Ecol. 32: 2365–2381.

Su, Y.-S., and M. Yajima. **2015**. R2jags: Using R to Run "JAGS."

Süss, J. 2008. Tick-borne encephalitis in Europe and beyond - the epidemiological

situation as of 2007. Eurosurveillance. 13: 2–9.

Thomas, L., S. T. Buckland, E. A. Rexstad, J. L. Laake, S. Strindberg, S. L. Hedley, J. R. B. Bishop, T. A. Marques, and K. P. Burnham. **2010**. Distance software: Design and analysis of distance sampling surveys for estimating population size. J. Appl. Ecol. 47: 5–14.

Tran, P. M., and L. Waller. **2013**. Effects of landscape fragmentation and climate on lyme disease incidence in the Northeastern United States. Ecohealth. 10: 394–404.

Trout Fryxell, R. T., J. E. Moore, M. D. Collins, Y. Kwon, S. R. Jean-Philippe, S. M. Schaeffer, A. Odoi, M. Kennedy, and A. E. Houston. **2015**. Habitat and vegetation variables are not enough when predicting tick populations in the Southeastern United States. PLoS One. 10: 1–17.

U.S. Census Bureau. **2011**. Transportation Geodatabase. TIGER Prod. (https://www.census.gov/geo/maps-data/data/tiger.html).

Varela-Stokes, A. S. 2007. Transmission of *Ehrlichia chaffeensis* from lone star ticks (*Amblyomma americanum*) to white-tailed deer (*Odocoileus virginianus*). J. Wildl. Dis. 43: 376–381.

Virginia Department of Game and Inland Fisheries. **2015**. White-tailed Deer Management Plan, 2015-2024. Richmond.

Virginia GIS Clearinghouse. **2011**. Virginia LIDAR.

Whitlock, J. E., Q. Q. Fang, L. A. Durden, and J. H. J. Oliver. **2000**. Prevalence of *Ehrlichia chaffeensis* (Rickettsiales: Rickettsiaceae) in *Amblyomma americanum* (Acari: Ixodidae) from the Georgia coast and Barrier Islands. J. Med. Entomol. 37: 276–280.

Wimberly, M. C., M. J. Yabsley, A. D. Baer, V. G. Dugan, and W. R. Davidson. **2008**. Spatial heterogeneity of climate and land-cover constraints on distributions of tick-borne pathogens. Glob. Ecol. Biogeogr. 17: 189–202.

Wright, C. L., H. D. Gaff, and W. L. Hynes. **2014**. Prevalence of *Ehrlichia chaffeensis* and *Ehrlichia ewingii* in *Amblyomma americanum* and *Dermacentor variabilis* collected from southeastern Virginia, 2010-2011. Ticks Tick. Borne. Dis. 5: 978–982.

Yabsley, M. J. 2010. Natural history of *Ehrlichia chaffeensis*: Vertebrate hosts and tick vectors from the United States and evidence for endemic transmission in other countries. Vet. Parasitol. 167: 136–148.

Yabsley, M. J., V. G. Dugan, D. E. Stallknecht, S. E. Little, J. M. Lockhart, J. E. Dawson, and W. R. Davidson. **2003**. Evaluation of a prototype *Ehrlichia chaffeensis* surveillance system using white-tailed deer (*Odocoileus virginianus*) as natural sentinels. Vector-borne zoonotic Dis. 3: 195–207.

Yarrow, G. 2009. White-tailed Deer Biology & Management. Clemson Univ. Coop. Extension's For. Nat. Resour. Fact Sheet 34.

Zuur, A. F., E. N. Ieno, N. J. Walker, A. A. Saveliev, G. M. Smith, Z. I. Walker, and S. Smith. **2009**. Mixed Effects Models and Extensions in Ecology with R, Ecology. Springer-Verlag, New York.