



RESEARCH  
PAPER

# Field and climate-based model for predicting the density of host-seeking nymphal *Ixodes scapularis*, an important vector of tick-borne disease agents in the eastern United States

Maria A. Diuk-Wasser<sup>1\*</sup>, Gwenaël Vourc'h<sup>2</sup>, Paul Cislo<sup>1</sup>,  
Anne Gatewood Hoen<sup>1†</sup>, Forrest Melton<sup>3</sup>, Sarah A. Hamer<sup>4</sup>,  
Michelle Rowland<sup>5</sup>, Roberto Cortinas<sup>6</sup>, Graham J. Hickling<sup>7</sup>, Jean I. Tsao<sup>4,8</sup>,  
Alan G. Barbour<sup>9</sup>, Uriel Kitron<sup>10</sup>, Joseph Piesman<sup>11</sup> and Durland Fish<sup>1</sup>

<sup>1</sup>Yale School of Public Health, 60 College Street, New Haven, CT 06520, USA, <sup>2</sup>National Institute for Agricultural Research, UR346 Animal Epidemiology, F-63122 Saint Genès Champanelle, France, <sup>3</sup>Division of Science and Environmental Policy, California State University Monterey Bay, 100 Campus Center, Seaside, CA 93955, USA, <sup>4</sup>Department of Fisheries and Wildlife, College of Agricultural and Natural Resources, Michigan State University, 13 Natural Resources Building, East Lansing, MI 48824, USA, <sup>5</sup>College of Veterinary Medicine, University of Illinois, 2001 S. Lincoln Ave., Urbana, IL 61802, USA, <sup>6</sup>Department of Entomology, University of Nebraska, Lincoln, 12BA Entomology Hall, East Campus, Lincoln, NE 68583, USA, <sup>7</sup>Department of Forestry, Wildlife and Fisheries, College of Agricultural Sciences and Natural Resources, University of Tennessee, Knoxville, TN 37996, USA, <sup>8</sup>Department of Large Animal Clinical Sciences, College of Veterinary Medicine, Michigan State University, D202 Veterinary Medical Center, East Lansing, MI 48824, USA, <sup>9</sup>Department of Microbiology and Molecular Genetics, University of California Irvine, Irvine, CA 92697, USA, <sup>10</sup>Department of Environmental Studies, Emory University, 400 Dowman Drive, Atlanta, GA 30322, USA, <sup>11</sup>Centers for Disease Control and Prevention, Division of Vector-Borne Infectious Diseases, Bacterial Zoonoses Branch, 3150 Rampart Rd., Fort Collins, CO 80521 USA

\*Correspondence: Maria A. Diuk-Wasser, 60 College Street, PO Box 208034, New Haven, CT 06520-8034, USA.

E-mail: maria.diuk@yale.edu

†Present address: Children's Hospital Informatics Program, Harvard-MIT Division of Health Sciences and Technology, 1 Autumn Street, Boston, MA 02215, USA.

## ABSTRACT

**Aim** *Ixodes scapularis* is the most important vector of human tick-borne pathogens in the United States, which include the agents of Lyme disease, human babesiosis and human anaplasmosis, among others. The density of host-seeking *I. scapularis* nymphs is an important component of human risk for acquiring *Borrelia burgdorferi*, the aetiological agent of Lyme disease. In this study we used climate and field sampling data to generate a predictive map of the density of host-seeking *I. scapularis* nymphs that can be used by the public, physicians and public health agencies to assist with the diagnosis and reporting of disease, and to better target disease prevention and control efforts.

**Location** Eastern United States of America.

**Methods** We sampled host-seeking *I. scapularis* nymphs in 304 locations uniformly distributed east of the 100th meridian between 2004 and 2006. Between May and September, 1000 m<sup>2</sup> were drag sampled three to six times per site. We developed a zero-inflated negative binomial model to predict the density of host-seeking *I. scapularis* nymphs based on altitude, interpolated weather station and remotely sensed data.

**Results** Variables that had the strongest relationship with nymphal density were altitude, monthly mean vapour pressure deficit and spatial autocorrelation. Forest fragmentation and soil texture were not predictive. The best-fit model identified two main foci – the north-east and upper Midwest – and predicted the presence and absence of *I. scapularis* nymphs with 82% accuracy, with 89% sensitivity and 82% specificity. Areas of concordance and discordance with previous studies were discussed. Areas with high predicted but low observed densities of host-seeking nymphs were identified as potential expansion fronts.

**Main conclusions** This model is unique in its extensive and unbiased field sampling effort, allowing for an accurate delineation of the density of host-seeking *I. scapularis* nymphs, an important component of human risk of infection for *B. burgdorferi* and other *I. scapularis*-borne pathogens.

## Keywords

*Borrelia burgdorferi*, climate, *Ixodes scapularis*, Lyme disease, remote sensing, tick-borne, USA.

## INTRODUCTION

*Ixodes scapularis* Say, commonly known as the deer tick or black legged tick, is the vector of *Borrelia burgdorferi*, *Babesia microti* and *Anaplasma phagocytophilum*, aetiological agents of Lyme disease, human babesiosis and human anaplasmosis, respectively, as well as other pathogenic agents. Lyme disease is the most frequently reported vector-borne disease of humans in the United States, with approximately 20,000 reported cases per year and increasing incidence. The steady increase in reported incidence is due to both increased recognition and geographic expansion of endemic areas (Bacon *et al.*, 2008). Human babesiosis and anaplasmosis, both of which have life-threatening potential, are also increasing in incidence, and co-infection of ticks with multiple pathogens can cause difficulties in diagnosis and treatment, and may result in more severe disease (Krause *et al.*, 2002).

Lyme disease surveillance based on human case reports is complicated by both underreporting and overdiagnosis (Bacon *et al.*, 2008), often resulting in inaccuracy of the true geographic distribution of disease risk. Accurate information on spatial patterns of risk for exposure to vector ticks is essential for the public to make informed decisions regarding how to avoid high-risk areas and for the medical community to consider a diagnosis of tick-borne disease (Piesman & Eisen, 2008).

The construction of an accurate map showing the spatial distribution and density of *I. scapularis* in the United States has been limited by passive and non-standardized collection methods. *Ixodes scapularis* habitat suitability models for the United States have been developed by Brownstein *et al.* (2003) and Estrada-Peña (2002). These models were based on *I. scapularis* distribution data from Dennis *et al.* (1998), in which US counties were classified into those with 'established' or 'reported' *I. scapularis* populations, based on a review of published sources and questionnaires. *Ixodes scapularis* was assumed to be absent in counties where no reports were found, which is likely to have resulted in significant biases. Another limitation of these models is that they include data from all life stages, whereas only the nymphal stage has been found to have a significant role as a vector for *B. burgdorferi* in North America (Mather *et al.*, 1996; Stafford *et al.*, 1998; Falco *et al.*, 1999). This distinction is important, because although *I. scapularis* populations are present in southern states, host-seeking nymphs are rarely collected (Cilek & Olson, 2000; MacKay & Foil, 2005; Goddard & Piesman, 2006) and very rarely bite people (Felz *et al.*, 1996; Merten & Durden, 2000). This is consistent with the very low numbers of human cases reported in southern states and the absence of culture-confirmed cases (Bacon *et al.*, 2008).

We have developed a spatial model of host-seeking *I. scapularis* nymphs to more accurately reflect the risk of exposure to humans. The model is based on a 4-year survey to generate a risk map for human infection with nymphal *I. scapularis*-borne *B. burgdorferi* throughout the range of the vector in the United States (Diuk-Wasser *et al.*, 2006). Human risk of infection ('entomological risk' *sensu* Mather *et al.*, 1996) is estimated by

the product of the density of host-seeking *I. scapularis* nymphs and their infection prevalence with different pathogens. We report here on the density of host-seeking *I. scapularis* nymphs (hereafter 'nymphs') component of risk. We used the dragging methodology, which provides the most sensitive measure of potential contact between ticks and humans (Falco & Fish, 1992; Mather *et al.*, 1996).

In order to generate a continuous probability surface of the risk of encountering nymphs in the United States, we predicted nymphal density using remotely sensed data as in previous models of *I. scapularis* habitat suitability (Glass *et al.*, 1995; Dister *et al.*, 1997; Kitron & Kazmierczak, 1997; Guerra *et al.*, 2002; Ogden *et al.*, 2006) as well as interpolated weather station data. This model can be used by public health agencies to better target surveillance and prevention measures to higher-risk areas. In addition, the establishment of a baseline distribution of host-seeking *I. scapularis* nymphs in the United States and an understanding of the climate drivers of this distribution are essential to any effort to monitor future expansion/retraction of human risk of diseases associated with this tick due to changes in climate or land use.

## METHODS

### Tick sampling scheme

The details of the sampling methodology are described in Diuk-Wasser *et al.* (2006); a summary is provided here. The study area included all of the continental United States east of the 100th meridian (37 states), encompassing the known distribution of *I. scapularis*. A spatially stratified random design was used by setting up a 2° sampling grid across the study area, and state parks or other publicly accessible forested areas were randomly selected within each grid cell. We sampled a total of 304 sites between 2004 and 2006, 30 of which were repeatedly sampled in 2 or 3 years, resulting in a total of 348 site-year samples. In 2004, one site was sampled per grid cell, for a total of 96 sites. Only nine nymphs were collected in 38 grid cells in the southern United States, which were identified as a statistically significant low-density cluster (Diuk-Wasser *et al.*, 2006). We excluded these grid cells from the sampling scheme in 2005 and 2006 to focus our sampling efforts on areas with high and more variable nymphal density. We sampled two sites in 2005 and two different ones in 2006 in each of the 'northern' grid cells, for a total of five sites per grid cell. Within each site, we measured relative nymphal density by drag sampling of closed-canopy deciduous forest habitats along five 200-m transects, with ticks collected from the drag cloth every 20 m to avoid fall-off. To capture the host-seeking phenology of *I. scapularis* nymphs, we visited sites a median of five times during late spring and summer, when nymphs had been reported to actively seek hosts in the north-eastern United States (Fish, 1993). Sampling was performed between 19 May and 27 August in 2004, 9 May and 4 October in 2005 and 10 May and 30 September in 2006.

## Environmental covariates

### *Weather station-derived climate data*

We extracted meteorological data from spatially continuous, daily meteorological surfaces for the United States with a spatial resolution of 8 km × 8 km produced by the NASA Terrestrial Observation and Prediction System (TOPS) (Nemani *et al.*, 2007). TOPS utilizes the surface observation and gridding system (Jolly *et al.*, 2005) to input daily observations of maximum temperature (TMAX), minimum temperature (TMIN), dew point temperature and solar radiation from approximately 3000 meteorological stations throughout the United States. These variables are then interpolated from the station observations to spatially continuous grids following Thornton *et al.* (1997). The interpolation algorithm also accounts for effects of elevation differences on temperature and precipitation using an empirical relationship calculated for each daily grid from observations using a weighted least squares regression (Thornton *et al.*, 1997). In addition, interpolated average saturation vapour pressure deficit (VPD) surfaces were generated by estimating the average daytime VPD at each meteorological station from observations of daily average, minimum, and dew point temperatures, following Thornton *et al.* (1997):

$$\text{VPD} = e_s(T_a) - e_d \quad (1)$$

where  $e_s(T_a)$  is the saturated vapour pressure at the average daytime site temperature  $T_a$  (°C), and  $e_d$  is the ambient vapour pressure. Vapour pressures were calculated following Murray (1967).

Using ERDAS IMAGINE v.9.1 (Leica Geosystems, Atlanta, GA, USA), we computed the annual mean for the monthly average of each climate variable between 1986 and 2005. This 20-year period was chosen to characterize long-term climate. We also performed temporal Fourier analysis (TFA) on the TMAX and TMIN data sets, deriving the detrended annual amplitude and phase and biannual amplitude and phase. Amplitude is the maximum variation of the cycle around the mean and phase is the timing of the cycle. In TFA, the trajectory through the year (the sequence of 12 monthly images) of every pixel in the environmental time series is described by a series of orthogonal sine curves (cycles/harmonics) with different frequencies. TFA removes noise from the seasonal data, achieves data reduction and produces a set of uncorrelated outputs while retaining a description of seasonality. It is commonly used to study the distribution of vectors (Rogers & Randolph, 2003).

### *Remotely sensed climate and vegetation index data*

Remotely sensed data were derived from the Advanced Very High Resolution Radiometer (AVHRR) instrument on board the National Oceanographic and Atmospheric Administration (NOAA) series of satellites for the period 1982–99, at 8 km × 8 km pixel resolution. TFA was applied to channel 3 middle infrared (MIR), land surface temperature (LST) and the

normalized difference vegetation index (NDVI) by the Trypanosomiasis and Land-use in Africa (TALA) Research Group of the University of Oxford, from whom the data was obtained (Hay *et al.*, 2006). The annual and biannual mean, amplitude and phase products for MIR, LST and NDVI were used in the current study.

### *Altitude*

Altitude (ALT) was derived from a seamless digital elevation model (DEM) mosaic (Earth Resource Data Analysis System, Atlanta, GA, USA) created from about 1500 individual files of USGS 1:250,000 digital terrain data. The 300-m grid spaced raster was used in this study.

### *Forest fragmentation*

Forest fragmentation maps were derived from a global assessment of forest fragmentation (Riitters *et al.*, 2000). The fragmentation index was calculated from existing global land-cover characterization 1-km land-cover maps (Loveland *et al.*, 2000). To derive the index, the proportion of forest pixels in a 9 km × 9 km pixel moving window was calculated for all pixels in the image. The result was stored at the location of the centre pixel. Thus, a pixel value in the derived map refers to ‘among-pixel’ fragmentation around the corresponding forest location.

### *Soil texture*

Soil texture data were derived from the US General Soil Map (STATSGO2) database (United States Department of Agriculture, 2006) and gridded to an 8 km × 8 km resolution using TOPS. The percentages of sand, clay and silt were calculated for the continental United States.

## Model development

### *Environmental data processing*

We projected all data sets used in the analysis to a Lambert azimuthal equal area projection, which accurately represents area in all regions of the globe, with the projection centred at –100° longitude, 45° latitude to minimize distortion of the study area. Each study site was considered as a point obtained by averaging the latitude and longitude of the start of all transects at the site. We extracted the values of the pixels corresponding to each site from all environmental rasters (ARCMAP software, v.9.1, ESRI, Redlands, CA, USA). When the site occurred at a pixel with missing information, we assigned it the average of the value in neighbouring pixels in a 3 × 3 or 5 × 5 window. The extracted values were standardized by subtracting the mean and dividing by the standard deviation prior to inclusion in predictive models of the density of *I. scapularis* nymphs.

### *Tick density data processing*

To obtain a synthetic measure of nymphal density collected throughout the season for a given site in a given year, we

calculated a weighted mean of the density of host-seeking nymphs per 200-m transect by: (1) calculating the mean number of *I. scapularis* nymphs collected per visit, averaging the collections from the five transects; (2) calculating the area under the line formed by the mean number of *I. scapularis* nymphs collected per visit over the whole season; and (3) dividing this area by the total number of days elapsed between the first and last sampling visit for a given year, to obtain a daily mean density measure. To examine whether differences in nymphal activity patterns may affect regional comparisons of this density measure, we compared the weekly distribution of mean nymphal density between sites with at least five sampled nymphs in the north-east and Midwest. We found that the distributions were not significantly different (two-sided Wilcoxon signed-rank test,  $n = 16$  weeks,  $x \geq 9$ ,  $P = 0.8$ ) (see Fig S1 in Supporting Information).

### Statistical analyses

We developed a zero-inflated negative binomial (ZINB) regression model for the expected density of nymphs using the environmental covariates as predictors (countreg procedures, SAS software, SAS Institute Inc., Cary, NC, USA). The distribution assumption behind the ZINB regression model is a mixture of two discrete probability distributions or 'components'. The first component (i.e. the 'zero-inflated' component) estimates the probability that the observation was a result of sampling from the point mass at zero, indicating a 100% probability of observing a zero when sampling from this distribution. The second component (i.e. the 'negative binomial' component) estimates the expected value assuming the observation is a result of sampling from a negative binomial distribution. The distribution of data arising from a mixture of these two distributions resembles a negative binomial distribution with a disproportionately high number of zeros. Such mixtures of distributions are common when a sample contains two types of observations probably driven by different ecological processes. In the case of host-seeking *I. scapularis* nymphs, behavioural differences in 'southern' *I. scapularis* reduce the probability of collecting a host-seeking nymph to almost nil, so these samples would follow a probability distribution with all of its mass at zero. 'Northern' sites, where there is a positive probability of observed host-seeking nymphs, would follow a negative binomial distribution. Each component of the ZINB regression model uses a component-specific set of covariates to predict the respective values. The two estimates are combined to produce an overall prediction as seen in equation (2). The model can then be viewed as a combination of a logistic regression model predicting presence/absence and a negative binomial model predicting tick density:

$$\hat{Y} = (\hat{P})(0) + (1 - \hat{P}) \left[ \exp(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots) \right] \quad (2)$$

where  $P$  is the probability that the observation was a result of sampling from the point mass at zero:

$$\hat{P} = \exp(\hat{\gamma}_0 + \hat{\gamma}_1 z_1 + \hat{\gamma}_2 z_2 + \dots) / [1 + \exp(\hat{\gamma}_0 + \hat{\gamma}_1 z_1 + \hat{\gamma}_2 z_2 + \dots)] \quad (3)$$

and  $\hat{\beta}_0$ ,  $\hat{\beta}_1$ ,  $\hat{\beta}_2$  and  $\hat{\gamma}_0$ ,  $\hat{\gamma}_1$ ,  $\hat{\gamma}_2$  are estimated coefficients for the respective negative binomial and zero-inflated components of the model and  $x_1$ ,  $x_2$ , ... and  $z_1$ ,  $z_2$ , ... are predictors for the respective negative binomial and zero-inflated part of the model.

To assess the robustness of the variables selected in the analysis, the data set was split randomly into a training ( $n = 172$ ) and a testing ( $n = 176$ ) data set. Extensive data exploration and model testing was conducted using the training data set. To make sure we did not identify any false positive relationships during the data exploration phase of analysis, we tested the best model identified with the training data by applying it to the remaining data, the 'testing' data set.

### Data exploration steps

To select the variables to include in the ZINB model, we first ran a stepwise logistic regression to identify environmental predictors of nymphal presence or absence. During the second phase of the analysis we included the selected covariates in the zero-inflated component of the ZINB regression model. If residuals were found to be significantly autocorrelated by a Moran's  $I$ -test, an autocovariate term (Dormann *et al.*, 2007) was included in the model to remove spatial autocorrelation and improve the fit. This term was calculated as an inverse distance-weighted average of the response values within a defined neighbourhood around a given site  $i$ :

$$Ai = \sum_{j \in k_i} w_{ij} y_j \quad (4)$$

where  $k_i$  is the set of neighbours included in a specified radius,  $y_j$  is the response value at site  $j$  and  $w_{ij}$  is the inverse-distance weight given to site  $j$ 's influence over site  $i$ . The radius (227,655 m) that resulted in at least one neighbour for all sites was used. Other radii were explored, but did not result in a significant reduction in the Akaike information criterion (AIC), so were not used further in the analysis. The autocovariate term removed spatial autocorrelation of the residuals and improved the model fit (lower AIC).

Improvements of the model fit were evaluated in an automated fashion by including one or two additional variables in each component of the ZINB. Those models for which all variables were significant at  $P \leq 0.05$  in the ZINB model using the training data were sorted by AIC. The top three models based on the lowest AIC were identified and evaluated using the testing data set. These three models were further evaluated by cross-validation root mean squared error (RMSE) and examination of the distribution of chi-square residuals. The final model was validated by running it with the testing data set and verifying that all variables remained significant and the trends observed with the training data set persisted. For optimal estimates of the

regression parameters, the final values were obtained by using the entire data set.

### Mapping

TOPS-derived climate and soil rasters and the AVHRR-derived rasters had an 8 km × 8 km cell size. Altitude and forest fragmentation were resampled to the same cell size using a nearest-neighbour interpolation. The values of the autocovariate term calculated for each sample point were interpolated using inverse distance weighting to obtain a continuous surface, also at 8 km × 8 km cell size. We estimated the expected value for nymphal density from the ZINB autocovariate model for each pixel using equation (2).

### RESULTS

A total of 5332 nymphs were collected at 304 sites between 2004 and 2006. Nymphs were found in 94 of these 304 sites, with total ticks collected per year in positive sites ranging from 1 to 507 (Table S1). The weighted mean ranged from 0 to 20.26 per 200 m transect. Out of the 36 states sampled, no nymphs were collected in 13 states and only 9 nymphs were collected from the other 'southern' states (Table 1). Of the 30 sites where repeated samples were obtained over 2 or 3 years, 15 were consistently negative, 9 consistently positive and 6 had both positive and negative years, but with less than 0.45 nymphs per 200-m transect in the positive year (except Foot Hills State Forest, MN) (Table S1).

Region	State	No. of sites	Nymphs per 200-m transect (SD)	Total nymphs per state	
Midwest	WI	22	3.16 (4.65)	1512	
	IL	19	0.65 (1.11)	225	
	MN	26	0.60 (1.45)	370	
	MI	27	0.30 (1.34)	178	
	IN	13	0.24 (0.48)	67	
	OK	4	0.04 (0.06)	5	
	MO	16	0.03 (0.04)	7	
	IA	25	0.01 (0.02)	4	
	KS	7	0.00	0	
	ND	5	0.00	0	
	NE	10	0.00	0	
	OH	14	0.00	0	
	SD	7	0.00	0	
North-east and Mid-Atlantic	CT	2	6.43 (3.40)	282	
	RI	1	5.61	77	
	MD	6	2.49 (2.84)	453	
	NY	21	2.37 (4.67)	1368	
	NJ	3	2.27 (1.57)	128	
	MA	4	1.35 (1.27)	139	
	ME	11	0.73 (1.82)	202	
	PA	17	0.69 (1.34)	217	
	VT	2	0.24 (0.34)	16	
	VA	12	0.17 (0.42)	65	
	NH	4	0.11 (0.14)	8	
	South	NC	15	0.01 (0.02)	5
		SC	7	0.01 (0.01)	2
AL		3	0.00 (0.01)	1	
GA		6	0.00 (0.01)	1	
AR		2	0.00	0	
FL		2	0.00	0	
KY		8	0.00	0	
LA		1	0.00	0	
MS		6	0.00	0	
TN		1	0.00	0	
TX		8	0.00	0	
WV	11	0.00	0		

**Table 1** Weighted mean density of host-seeking *Ixodes scapularis* nymphs per 200-m transect by state and region (2004–06).

The number of sites sampled (recorded twice when there were repeated samples) and the total number of nymphs collected per state are also reported.

Within regions, states are listed by decreasing mean number of nymphs sampled.

**Table 2** Parameter estimates for variables included in the zero-inflated and negative binomial components of the zero-inflated negative binomial (ZINB) model (model 2 in Table S3).

	Estimate	Standard error	<i>t</i> value	PR >   <i>t</i>
Zero-inflated				
Altitude	3.79	0.93	4.05	< 0.0001
VPD monthly mean	4.89	1.31	3.73	0.0002
(TMAX annual amplitude) <sup>2</sup>	1.23	0.50	2.45	0.0143
(TMIN annual phase) <sup>2</sup>	2.42	0.65	3.74	0.0002
Negative binomial				
Autocovariate term	0.44	0.10	4.28	< 0.0001
(NDVI annual amplitude) <sup>2</sup>	0.33	0.20	1.61	0.1069

Both training and testing records were included to improve model fit.

VPD, vapour pressure deficit; TMAX, maximum daily temperature; TMIN, minimum daily temperature; NDVI, normalized difference vegetation index; PR > |*t*|, 2-tailed *P*-value for testing the null hypothesis that the parameter estimate is 0.

The stepwise logistic regression performed as part of the exploratory phase resulted in a model including a negative association with ALT and VPD monthly mean and a positive association with MIR biannual amplitude and TMAX monthly mean (Table S2). No nymphs were found above a threshold altitude of 510 m (negative coefficient) but no linear relationship was observed between altitude and nymphs below that point (Fig. S2). When we entered the variables selected from the logistic regression model into the zero-inflated component of the ZINB, MIR biannual amplitude and TMAX monthly mean dropped out of the model and the residuals of the resulting model were found to be autocorrelated. We ran the ZINB model with the remaining two variables (ALT and VPD monthly mean) and an autocovariate term (AUTOCOV) included in the negative binomial component to account for spatial autocorrelation. A total of 300,459 models were generated by adding one or two additional variables to each component of the ZINB. Models containing any non-significant ( $P > 0.05$ ) covariates were rejected, leaving 1250 models. The AIC of the three top models ranged from 440.9 to 449.2 (Table S3). Both the first and third models showed higher values of the root mean square error (RMSE) and more extreme residuals than the second model. In addition, a Moran's *I*-test showed that the inclusion of the autocovariate term failed to completely remove autocorrelation of the residuals for the third model. We therefore selected the second model as the one with the 'best' fit, even if the square of NDVI annual amplitude was only significant at  $P < 0.1$  when it was run with both the training and testing data sets (Table 2).

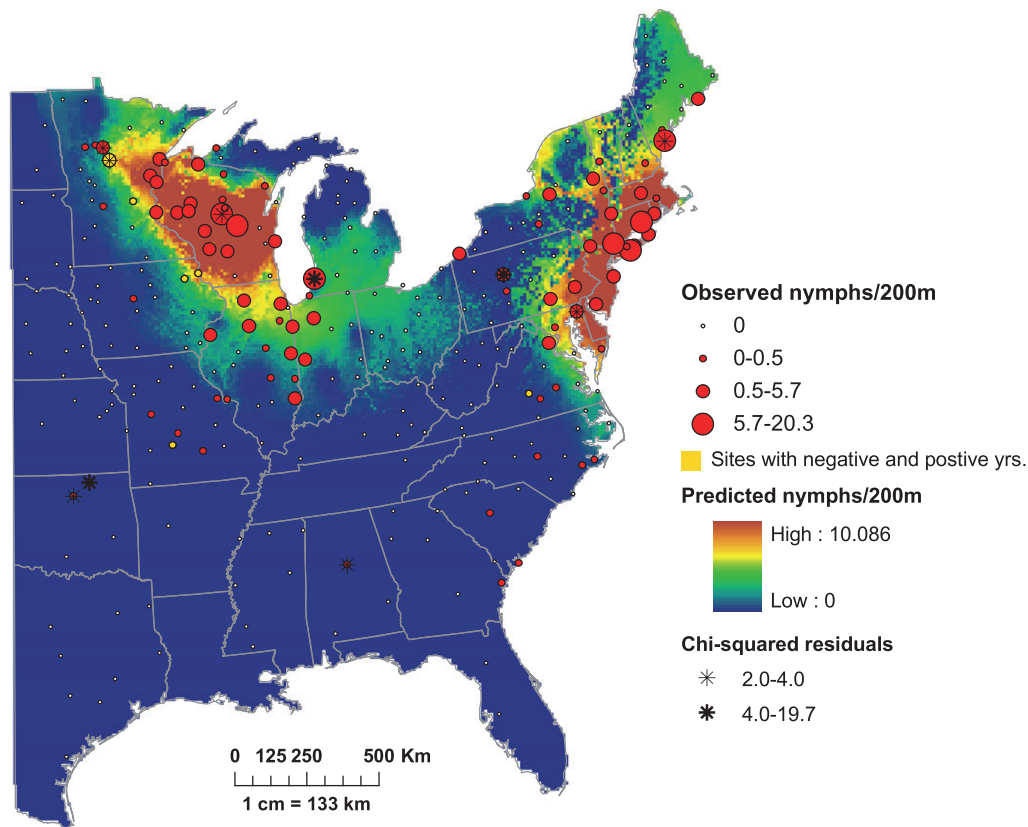
Consistent with our previous analysis of the first-year collection (Diuk-Wasser *et al.*, 2006), two main foci of host-seeking *I. scapularis* nymphs were identified, one from northern Virginia to Maine along the Atlantic seaboard and another one in the upper Midwest (Fig. 1). The final ZINB model correctly predicted 82% of sites as either positive or negative, with a sensitivity of 89% and specificity of 82%. The robustness of the model can be assessed by analysing the chi-square residuals. Of 348 site-year samples, 10 sites had chi-square residuals larger than 2 ('high'), three of which were larger than 4 ('extreme') (Fig. 1). All residuals > 2 were positive, indicating that the model tended to underestimate nymphal density. Two of the high

residuals resulted from a prediction of zero nymphs and one observed nymph in two southern locations (Pawnee Bill Ranch and Oak Mountain State Park). Another southern site (Osage Hill State Park, OK) also presented an extreme residual (predicted 0, observed 4). Extreme residuals were also found in a site in Pennsylvania (Parker Dam State Park) and in Michigan (Van Buren State Park).

## DISCUSSION

We developed a predictive model for the density of host-seeking *I. scapularis* nymphs in the United States that integrates different environmental drivers and accounts for spatial autocorrelation. This is the first zoonotic disease risk-modelling project based on standardized field data collected throughout the national distribution of a vector species. The unbiased nature of the sampling design, including a large number of both positive and negative sites, results in a robust data set to validate prediction of species presence/absence. The large sample of positive sites allowed for the modelling of tick density in addition to presence/absence, in contrast to the more typical approaches based on logistic regression and discriminant analysis techniques (Kalluri *et al.*, 2007). The use of a zero-inflated regression model allowed us to discriminate among environmental drivers of presence/absence and those driving the density of host-seeking *I. scapularis* nymphs.

The broad risk areas defined by the nymphal density model were generally consistent with previous studies, with two main foci in the north-east and the upper Midwest. In the north-east, the highest nymphal densities were observed in Connecticut, Rhode Island, Massachusetts, southern New York, New Jersey, eastern Pennsylvania and Delaware. The counties predicted as having high nymphal density in New York State matched the reported endemic counties based on data for human Lyme disease cases and *I. scapularis* (White *et al.*, 1991) or data on mammal infection with *B. burgdorferi* (Oliver *et al.*, 2006). The model performed poorly in western Pennsylvania, where it failed to predict the presence of nymphs in a high-altitude area near Elk County (extreme residual in Parker Dam State Park in Clearfield County) and in Presque Isle State Park in Erie County. *Borrelia burgdorferi* has been found in *Peromyscus leucopus* in



**Figure 1** Weighted mean of host-seeking *Ixodes scapularis* nymphs observed and predicted under the second model with the lowest Akaike information criterion (AIC) value (geographic coordinate system, GCS WGS 1984; projection, Lambert azimuthal equal area). Host-seeking *I. scapularis* nymphs were collected between May and September of 2004–06. For sites sampled in multiple years, the average of the annual weighted mean is provided. Sites where nymphs were collected in one year but not in others and those with model chi-square residuals larger than two and four ('extreme') are shown.

Elk County (Lord *et al.*, 1994) and both *I. scapularis* and *B. burgdorferi* have been previously reported in Erie County (Courtney *et al.*, 2003). More sampling is warranted to delimit the exact distribution of *I. scapularis* in these areas and their relationship with the north-eastern focus.

In the upper Midwest, the model predicts high nymphal density in all of Wisconsin, consistent with a proposed eastern expansion of *I. scapularis* (Riehle & Paskewitz, 1996) from the western Wisconsin risk areas predicted by Guerra *et al.* (2002). The model prediction of nymphal presence in Minnesota closely matched the reported areas of risk based on human case data (Minnesota Department of Health, 2008). The presence of four sites with both positive and negative years in the western limit of the upper Midwest foci indicates an unstable presence of nymphs in these border areas. A southward expansion front was observed in the Illinois–Indiana border, with records south of reported *I. scapularis* by Pinger *et al.* (1996), which was predicted by the model, although the density was underestimated. We also observed high abundance of nymphs in Van Buren State Park in south-western Michigan, reported as an invasion area by Hamer *et al.* (2007). The model, however, underestimates the density in that site ('extreme' residual) and underestimates the extent of the northern expansion as reported by the Michigan

Department of Community Health (2004). The model also predicts higher nymphal densities than observed in southern Michigan, northern Indiana and parts of Ohio, northern Maine and coastal Virginia, which could represent potential future expansion areas.

The absence and rarity of host-seeking nymphs in the southern portion of the range of *I. scapularis* is consistent with other studies (Cilek & Olson, 2000; Goddard & Piesman, 2006). While *I. scapularis* populations are present in this region, the nymphs have an altered feeding behaviour, presumably adapted to lizards and skinks, and cannot be sampled by the drag-cloth method used in this study. However, the drag-cloth collecting method is directly correlated with human contact with host-seeking ticks and is a direct measure of risk for tick bites (Falco & Fish, 1989). The absence of host-seeking nymphal *I. scapularis* in most southern states suggests that reported cases from this region are due to either misdiagnosis or previous travel to an endemic area.

The probability of finding nymphs was strongly driven by altitude, VPD and seasonal variation in temperature. In positive sites, the main driver for predicting nymphal density was the presence of host-seeking nymphs in neighbouring sites, with a weak association with environmental factors derived from

remote sensing. Unlike previous studies performed at a more detailed scale, we did not find significant effects of forest fragmentation (Brownstein *et al.*, 2005b) or soil texture (Guerra *et al.*, 2002), indicating that climate overrides the potential effect of these variables at a continental scale.

The density of host-seeking ticks of Lyme disease vectors in Europe has been found to decrease with increasing altitude (Jouda *et al.*, 2004). Decreases in tick density with altitude have sometimes been attributed to a simple effect of temperature (Jouda *et al.*, 2004; Cadenas *et al.*, 2007). However, the fact that altitude is significant in our model while controlling for temperature-derived variables suggests that other unmeasured variables may be affected along an altitudinal gradient. The strong predictive power of VPD on the density of host-seeking nymphs is consistent with other studies (Estrada-Peña, 2002; Guerra *et al.*, 2002; Brownstein *et al.*, 2003). Water stress and high temperatures are hypothesized to regulate tick populations by decreasing tick survival during off-host periods (Ogden *et al.*, 2004). Temperature and relative humidity can also regulate host-seeking activity in *Ixodes* spp., with increased activity negatively linked to saturation deficit and temperature and positively to relative humidity (Vail & Smith, 1998; Randolph & Storey, 1999; Perret *et al.*, 2000; Schulze & Jordan, 2003).

The presence of nymphal *I. scapularis* was also associated with the annual amplitude and phase of TMAX and TMIN, respectively. TFA of climate data have been successfully used in studies of the distribution of vectors and disease (reviewed by Scharlemann *et al.*, 2008), including tick-borne diseases (Randolph *et al.*, 2000). The effect of TFA-processed variables is complex. TFA captures both the extremes in temperature and the rates of autumn cooling and spring warming. Extreme winter temperatures can limit the northern distribution of ticks by directly killing the ticks (Ogden *et al.*, 2004; Rand *et al.*, 2004), inhibiting host-seeking activity (Vail & Smith, 1998; Perret *et al.*, 2000; Schulze & Jordan, 2003) or limiting the availability of hosts (Lindsay *et al.*, 1995). A high rate of autumn cooling can also affect tick population dynamics by limiting the time available for larvae to find a host in the autumn, thus entering diapause unfed, which potentially increases their mortality rate (Yuval & Spielman, 1990). However, high rates of spring warming may have an opposite effect, since it would result in faster accumulation of degree-days for development, potentially leading to earlier egg deposition and larval emergence (Lindsay *et al.*, 1995).

Remotely sensed environmental variables were found to be less predictive of nymphal density than ground-based climate data sets. Fourier-transformed NDVI and LST derived from AVHRR, however, were present in two of the models with the lowest AIC, in the negative binomial component of the equation. NDVI has been found in the past to be the most consistently significant variable for predicting tick distributions and has a sound biological basis in that it is related to availability of moisture to free-living ticks and is correlated with tick mortality rates (Randolph, 2000). The relatively low predictive value of LST may be due to the presence of temperature in the same model, to which LST is correlated (Green & Hay, 2002).

In sites with nymphs, the strongest predictor of nymphal density was the autocovariate term. This term is intended to capture spatial autocorrelation originating from endogenous processes such as conspecific attraction, limited dispersal, contagious population growth and movement of individuals between sampling sites (Dormann *et al.*, 2007). There is extensive evidence indicating a recent and ongoing population expansion of *I. scapularis* from past refuges throughout the north-east (White *et al.*, 1991) and the upper Midwest (Pinger *et al.*, 1996; Hamer *et al.*, 2009), suggesting that the spatial dependence in tick numbers may be due to a delay in the ticks occupying all suitable habitats due to recent environmental changes, which include reforestation, suburbanization and reintroduction of deer (Barbour & Fish, 1993).

This map provides the first step towards a large-scale risk model for Lyme disease. Strong links to climatic factors indicate the potential for changes in future distribution of risk due to climate change, as predicted by Brownstein *et al.* (2005a) and Ogden *et al.* (2008). However, the exact response of host-seeking nymphs may follow a different pattern from the distribution of all stages modelled by previous studies, warranting further research. The large geographic area covered by the map also necessarily limits its spatial resolution. More localized studies in areas identified as having high risk in this map are necessary to identify environmental factors operating at finer spatial scales and to provide guidance to local public health agencies and individual homeowners regarding their risk.

*Ixodes scapularis* host-seeking density predicted by this model has been shown to be positively correlated to Lyme disease incidence (Mather *et al.*, 1996; Stafford *et al.*, 1998; Hubalek *et al.*, 2003). Despite the limitation of spatial resolution, our model is useful for the education of the public and physicians on disease risk at the state and in some areas perhaps county level. Host-seeking nymphs were not found within most of the known range of *I. scapularis* despite current and past reports of human cases of Lyme disease. This model should be useful to public health agencies to improve the targeting of surveillance, prevention and potential future control efforts – such as human or wildlife vaccination programmes, to areas with evidence of entomological risk.

## ACKNOWLEDGEMENTS

We gratefully acknowledge the 80 field assistants who made this project possible. Special thanks to Tim Andreadis, Katherine Hansen, Laura Krueger, Jessica Payne, Elizabeth Racz, Kelly Liebman, Liza Lutzker and David Boozer for tick identification and logistic support; Carlos Diuk for database support; Brad Lobitz and Andrew Michaelis for technical assistance; and Dennis Grove, Lindsay Rollend and Russell Barbour for arranging collection permits. We also acknowledge Corrine Folsom-O'Keefe for manuscript editing, Robert Brinkerhoff and Kimberly Tsao for helpful comments, and John Brownstein and Ben Beard for their early contributions to this work. This project was funded by CDC-Division of Vector-Borne Infectious Diseases Cooperative Agreement no. C100171-01.



## REFERENCES

- Bacon, R.M., Kugeler, K.J. & Mead, P.S. (2008) Surveillance for Lyme disease – United States, 1992–2006. *Morbidity and Mortality Weekly Report Surveillance Summaries*, **57**, no. SS10.
- Barbour, A.G. & Fish, D. (1993) The biological and social phenomenon of Lyme disease. *Science*, **260**, 1610–1616.
- Brownstein, J.S., Holford, T.R. & Fish, D. (2003) A climate-based model predicts the spatial distribution of the Lyme disease vector *Ixodes scapularis* in the United States. *Environmental Health Perspectives*, **111**, 1152–1157.
- Brownstein, J.S., Holford, T.R. & Fish, D. (2005a) Effect of climate change on Lyme disease risk in North America. *Eco-Health*, **2**, 38–46.
- Brownstein, J.S., Skelly, D.K., Holford, T.R. & Fish, D. (2005b) Forest fragmentation predicts local scale heterogeneity of Lyme disease risk. *Oecologia*, **146**, 469–475.
- Cadenas, F.M., Rais, O., Jouda, F., Douet, V., Humair, P.F., Moret, J. & Gern, L. (2007) Phenology of *Ixodes ricinus* and infection with *Borrelia burgdorferi sensu lato* along a north- and south-facing altitudinal gradient on Chaumont Mountain, Switzerland. *Journal of Medical Entomology*, **44**, 683–693.
- Cilek, J.E. & Olson, M.A. (2000) Seasonal distribution and abundance of ticks (Acari : Ixodidae) in northwestern Florida. *Journal of Medical Entomology*, **37**, 439–444.
- Courtney, J.W., Dryden, R.L., Wyleto, P., Schneider, B.S. & Massung, R.F. (2003) Characterization of *Anaplasma phagocytophila* and *Borrelia burgdorferi* genotypes in *Ixodes scapularis* ticks from Pennsylvania. *Annals of the New York Academy of Science*, **990**, 131–133.
- Dennis, D.T., Nekomoto, T.S., Victor, J.C., Paul, W.S. & Piesman, J. (1998) Reported distribution of *Ixodes scapularis* and in *Ixodes pacificus* (Acari : Ixodidae) in the United States. *Journal of Medical Entomology*, **35**, 629–638.
- Dister, S.W., Fish, D., Bros, S.M., Frank, D.H. & Wood, B.L. (1997) Landscape characterization of peridomestic risk for Lyme disease using satellite imagery. *American Journal of Tropical Medicine and Hygiene*, **57**, 687–692.
- Diuk-Wasser, M.A., Gatewood, A.G., Cortinas, M.R., Yaremych-Hamer, S., Tsao, J., Kitron, U., Hickling, G., Brownstein, J.S., Walker, E., Piesman, J. & Fish, D. (2006) Spatiotemporal patterns of host-seeking *Ixodes scapularis* nymphs (Acari : Iodidae) in the United States. *Journal of Medical Entomology*, **43**, 166–176.
- Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M. & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609–628.
- Estrada-Peña, A. (2002) Increasing habitat suitability in the United States for the tick that transmits Lyme disease: a remote sensing approach. *Environmental Health Perspectives*, **110**, 635–640.
- Falco, R.C. & Fish, D. (1989) Potential for exposure to tick bites in recreational parks in a Lyme-disease endemic area. *American Journal of Public Health*, **79**, 12–15.
- Falco, R.C. & Fish, D. (1992) A comparison of methods for sampling the deer tick, *Ixodes dammini*, in a Lyme-disease endemic area. *Experimental and Applied Acarology*, **14**, 165–173.
- Falco, R.C., McKenna, D.F., Daniels, T.J., Nadelman, R.B., Nowakowski, J., Fish, D. & Wormser, G.P. (1999) Temporal relation between *Ixodes scapularis* abundance and risk for Lyme disease associated with erythema migrans. *American Journal of Epidemiology*, **149**, 771–776.
- Felz, M.W., Durden, L.A. & Oliver, J.H. (1996) Ticks parasitizing humans in Georgia and South Carolina. *Journal of Parasitology*, **82**, 505–508.
- Fish, D. (1993) Population ecology of *Ixodes dammini*. *Ecology and environmental management of Lyme disease* (ed. by H.S. Ginsberg), pp. 25–42. Rutgers University Press, New Brunswick, NJ.
- Glass, G.E., Schwartz, B.S., Morgan, J.M., Johnson, D.T., Noy, P.M. & Israel, E. (1995) Environmental risk factors for Lyme disease identified with geographic information systems. *American Journal of Public Health*, **85**, 944–948.
- Goddard, J. & Piesman, J. (2006) New records of immature *Ixodes scapularis* from Mississippi. *Journal of Vector Ecology*, **31**, 421–422.
- Green, R.M. & Hay, S.I. (2002) The potential of Pathfinder AVHRR data for providing surrogate climatic variables across Africa and Europe for epidemiological applications. *Remote Sensing of Environment*, **79**, 166–175.
- Guerra, M., Walker, E., Jones, C., Paskewitz, S., Cortinas, M.R., Stancil, A., Beck, L., Bobo, M. & Kitron, U. (2002) Predicting the risk of Lyme disease: habitat suitability for *Ixodes scapularis* in the north central United States. *Emerging Infectious Diseases*, **8**, 289–297.
- Hamer, S.A., Roy, P.L., Hickling, G.J., Walker, E.D., Foster, E.S., Barber, C.C. & Tsao, J.I. (2007) Zoonotic pathogens in *Ixodes scapularis*, Michigan. *Emerging Infectious Diseases*, **13**, 1131–1133.
- Hamer, S.A., Tsao, J.I., Walker, E.D., Mansfield, L.S., Foster, E.S. & Hickling, G.J. (2009) Use of tick surveys and serosurveys to evaluate pet dogs as a sentinel species for emerging Lyme disease. *American Journal of Veterinary Research*, **70**, 49–56.
- Hay, S.I., Tatem, A.J., Graham, A.J., Goetz, S.J. & Rogers, D.J. (2006) Global environmental data for mapping infectious disease distribution. *Advances in Parasitology*, **62**, 37–77.
- Hubalek, Z., Halouzka, J. & Juricova, Z. (2003) Longitudinal surveillance of the tick *Ixodes ricinus* for borreliae. *Medical and Veterinary Entomology*, **17**, 46–51.
- Jolly, W.M., Graham, J.M., Michaelis, A., Nemani, R. & Running, S.W. (2005) A flexible, integrated system for generating meteorological surfaces derived from point sources across multiple geographic scales. *Environmental Modelling and Software*, **20**, 873–882.
- Jouda, F., Perret, J.L. & Gern, L. (2004) *Ixodes ricinus* density, and distribution and prevalence of *Borrelia burgdorferi sensu lato*

- infection along an altitudinal gradient. *Journal of Medical Entomology*, **41**, 162–169.
- Kalluri, S., Gilruth, P., Rogers, D. & Szczur, M. (2007) Surveillance of arthropod vector-borne infectious diseases using remote sensing techniques: a review. *PLoS Pathogens*, **3**, 1361–1371.
- Kitron, U. & Kazmierczak, J.J. (1997) Spatial analysis of the distribution of Lyme disease in Wisconsin. *American Journal of Epidemiology*, **145**, 558–566.
- Krause, P.J., McKay, K., Thompson, C.A., Sikand, V.K., Lentz, R., Lepore, T., Closter, L., Christianson, D., Telford, S.R., Persing, D., Radolf, J.D. & Spielman, A. (2002) Disease-specific diagnosis of coinfecting tickborne zoonoses: babesiosis, human granulocytic ehrlichiosis, and Lyme disease. *Clinical Infectious Diseases*, **34**, 1184–1191.
- Lindsay, L.R., Barker, I.K., Surgeoner, G.A., McEwen, S.A., Gillespie, T.J. & Robinson, J.T. (1995) Survival and development of *Ixodes-scapularis* (Acari, Ixodidae) under various climatic conditions in Ontario, Canada. *Journal of Medical Entomology*, **32**, 143–152.
- Lord, R.D., Lord, V.R., Humphreys, J.G. & McLean, R.G. (1994) Distribution of *Borrelia burgdorferi* in host mice in Pennsylvania. *Journal of Clinical Microbiology*, **32**, 2501–2504.
- Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L. & Merchant, J.W. (2000) Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, **21**, 1303–1330.
- MacKay, A. & Foil, L. (2005) Seasonal and geographical distribution of adult *Ixodes scapularis* Say (Acari : Ixodidae) in Louisiana. *Journal of Vector Ecology*, **30**, 168–170.
- Mather, T.N., Nicholson, M.C., Donnelly, E.F. & Matyas, B.T. (1996) Entomologic index for human risk of Lyme disease. *American Journal of Epidemiology*, **144**, 1066–1069.
- Merten, H.A. & Durden, L.A. (2000) A state-by-state survey of ticks recorded from humans in the United States. *Journal of Vector Ecology*, **25**, 102–113.
- Michigan Department of Community Health (2004) Black-legged tick populations in Michigan. Michigan Department of Community Health, Lansing. Available at: [http://www.michigan.gov/emergingdiseases/0,1607,7-86-25890\\_26143-95445--,00.html](http://www.michigan.gov/emergingdiseases/0,1607,7-86-25890_26143-95445--,00.html) (accessed 14 January 2010).
- Minnesota Department of Health (2008) *High risk areas for tick-borne diseases in Minnesota*. MDH Infectious Disease Epidemiology, Prevention and Control Division, St Paul.
- Murray, F.W. (1967) On the computation of saturation vapor pressure. *Journal of Applied Meteorology*, **6**, 203–204.
- Nemani, R., Votava, P., Michaelis, A., White, M., Melton, F., Milesi, C., Pierce, L., Golden, K., Hashimoto, H., Ichii, K., Johnson, L., Jolly, M., Myneni, R., Tague, C., Coughlan, J. & Running, A.S. (2007) Remote sensing methodologies for ecosystem management. *Food and water security* (ed. by U. Aswathanarayana), pp. 1–19. Taylor and Francis Group, London.
- Ogden, N.H., Lindsay, L.R., Beauchamp, G., Charron, D., Maarouf, A., O'Callaghan, C.J., Waltner-Toews, D. & Barker, I.K. (2004) Investigation of relationships between temperature and developmental rates of tick *Ixodes scapularis* (Acari : Ixodidae) in the laboratory and field. *Journal of Medical Entomology*, **41**, 622–633.
- Ogden, N.H., Barker, I.K., Beauchamp, G., Brazeau, S., Charron, D.F., Maarouf, A., Morshed, M.G., O'Callaghan, C.J., Thompson, R.A., Waltner-Toews, D., Waltner-Toews, M. & Lindsay, L.R. (2006) Investigation of ground level and remote-sensed data for habitat classification and prediction of survival of *Ixodes scapularis* in habitats of southeastern Canada. *Journal of Medical Entomology*, **43**, 403–414.
- Ogden, N.H., St-Onge, L., Barker, I.K., Brazeau, S., Bigras-Poulin, M., Charron, D.F., Francis, C.M., Heagy, A., Lindsay, L.R., Maarouf, A., Michel, P., Milord, F., O'Callaghan, C.J., Trudel, L. & Thompson, R. (2008) Risk maps for range expansion of the Lyme disease vector, *Ixodes scapularis*, in Canada now and with climate change. *International Journal of Health Geographics*, **7**, 24.
- Oliver, J., Means, R.G., Kogut, S., Prusinski, M., Howard, J.J., Layne, L.J., Chu, F.K., Reddy, A., Lee, L. & White, D. (2006) Prevalence of *Borrelia burgdorferi* in small mammals in New York state. *Journal of Medical Entomology*, **43**, 924–935.
- Perret, J.L., Guigoz, E., Rais, O. & Gern, L. (2000) Influence of saturation deficit and temperature on *Ixodes ricinus* tick questing activity in a Lyme borreliosis-endemic area (Switzerland). *Parasitology Research*, **86**, 554–557.
- Piesman, J. & Eisen, L. (2008) Prevention of tick-borne diseases. *Annual Review of Entomology*, **53**, 323–343.
- Pinger, R.R., Timmons, L. & Karris, K. (1996) Spread of *Ixodes scapularis* (Acari : Ixodidae) in Indiana: collections of adults in 1991–1994 and description of a *Borrelia burgdorferi*-infected population. *Journal of Medical Entomology*, **33**, 852–855.
- Rand, P.W., Holman, M.S., Lubelczyk, C., Lacombe, E.H., Degaetano, A.T. & Smith, R.P. (2004) Thermal accumulation and the early development of *Ixodes scapularis*. *Journal of Vector Ecology*, **29**, 164–176.
- Randolph, S.E. (2000) Ticks and tick-borne disease systems in space and from space. *Advances in Parasitology*, **47**, 217–243.
- Randolph, S.E. & Storey, K. (1999) Impact of microclimate on immature tick–rodent host interactions (Acari : Ixodidae): implications for parasite transmission. *Journal of Medical Entomology*, **36**, 741–748.
- Randolph, S.E., Green, R.M., Peacey, M.F. & Rogers, D.J. (2000) Seasonal synchrony: the key to tick-borne encephalitis foci identified by satellite data. *Parasitology*, **121**, 15–23.
- Riehle, M. & Paskewitz, S.M. (1996) *Ixodes scapularis* (Acari : Ixodidae): status and changes in prevalence and distribution in Wisconsin between 1981 and 1994 measured by deer surveillance. *Journal of Medical Entomology*, **33**, 933–938.
- Riitters, K., Wickham, J., O'Neill, R., Jones, B. & Smith, E. (2000) Global-scale patterns of forest fragmentation. *Conservation Ecology*, **4**, 3.
- Rogers, D.J. & Randolph, S.E. (2003) Studying the global distribution of infectious diseases using GIS and RS. *Nature Reviews Microbiology*, **1**, 231–237.
- Scharlemann, J.P., Benz, D., Hay, S.I., Purse, B.V., Tatem, A.J., Wint, G.R. & Rogers, D.J. (2008) Global data for ecology and

- epidemiology: a novel algorithm for temporal Fourier processing MODIS data. *PLoS ONE*, **3**, e1408.
- Schulze, T.L. & Jordan, R.A. (2003) Meteorologically mediated diurnal questing of *Ixodes scapularis* and *Amblyomma americanum* (Acari: Ixodidae) nymphs. *Journal of Medical Entomology*, **40**, 395–402.
- Stafford, K.C., Cartter, M.L., Magnarelli, L.A., Ertel, S.H. & Mshar, P.A. (1998) Temporal correlations between tick abundance and prevalence of ticks infected with *Borrelia burgdorferi* and increasing incidence of Lyme disease. *Journal of Clinical Microbiology*, **36**, 1240–1244.
- Thornton, P.E., Running, S.W. & White, M.A. (1997) Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology*, **190**, 214–251.
- United States Department of Agriculture (2006) *US general soil map (STATSGO2)*. Available at: <http://soildatamart.nrcs.usda.gov/USDGSM.aspx>.
- Vail, S.G. & Smith, G. (1998) Air temperature and relative humidity effects on behavioral activity of blacklegged tick (Acari: Ixodidae) nymphs in New Jersey. *Journal of Medical Entomology*, **35**, 1025–1028.
- White, D.J., Chang, H.G., Benach, J.L., Bosler, E.M., Meldrum, S.C., Means, R.G., Debbie, J.G., Birkhead, G.S. & Morse, D.L. (1991) The geographic spread and temporal increase of the Lyme-Disease epidemic. *Journal of the American Medical Association*, **266**, 1230–1236.
- Yuval, B. & Spielman, A. (1990) Duration and regulation of the developmental cycle of *Ixodes dammini* (Acari, Ixodidae). *Journal of Medical Entomology*, **27**, 196–201.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Figure S1** Weighted mean density of host-seeking *Ixodes scapularis* nymphs sampled in all sites east and west of  $-82.96^\circ$  W (centre of Ohio).

**Figure S2** Relationship between weighted mean density of host-seeking *Ixodes scapularis* nymphs per 200-m transect and altitude.

**Table S1** Weighted mean density of host-seeking *Ixodes scapularis* nymphs per 200-m transect at sites sampled between 2004 and 2006.

**Table S2** Parameter estimates for variables included in the logistic regression model for the presence of host-seeking *Ixodes scapularis* nymphs.

**Table S3** Three candidate zero-inflated negative binomial (ZINB) models selected by the lowest Akaike information criterion (AIC) values.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

## BIOSKETCH

**Maria Diuk-Wasser** is interested in modelling the environmental and ecological drivers of vector-borne and zoonotic diseases using intensive field and laboratory-derived data, with the ultimate goal of generating spatio-temporal predictions of human exposure risk. The other authors are interested in various aspects of tick-borne disease ecology and epidemiology. They are part of a consortium assembled to generate a spatial risk model for human risk for exposure to *Borrelia burgdorferi*, the agent of Lyme disease, in the United States.

Editor: Tim Blackburn