Habitat Suitability Model for the Distribution of *Ixodes scapularis* (Acari: Ixodidae) in Minnesota

T. L. Johnson[1,2], **J. K. H. Bjork**[3], **D. F. Neitzel**[3], **F. M. Dorr**[3], **E. K. Schiffman**[3], and **R. J. Eisen**[1]

J. K. H. Bjork: jenna.bjork@state.mn.us; D. F. Neitzel: david.neitzel@state.mn.us; F. M. Dorr: franny.dorr@state.mn.us; E. K. Schiffman: elizabeth.schiffman@state.mn.us; R. J. Eisen: dyn2@cdc.gov

[1]Bacterial Diseases Branch, Division of Vector-borne Diseases, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, 3156 Rampart Rd, Fort Collins, CO 80521

[3]Minnesota Department of Health, 625 Robert St N, St. Paul, MN 55164

Abstract

Ixodes scapularis Say, the black-legged tick, is the primary vector in the eastern United States of several pathogens causing human diseases including Lyme disease, anaplasmosis, and babesiosis. Over the past two decades, *I. scapularis*-borne diseases have increased in incidence as well as geographic distribution. Lyme disease exists in two major foci in the United States, one encompassing northeastern states and the other in the Upper Midwest. Minnesota represents a state with an appreciable increase in counties reporting *I. scapularis*-borne illnesses, suggesting geographic expansion of vector populations in recent years. Recent tick distribution records support this assumption. Here, we used those records to create a fine resolution, subcounty-level distribution model for *I. scapularis* using variable response curves in addition to tests of variable importance. The model identified 19% of Minnesota as potentially suitable for establishment of the tick and indicated with high accuracy (AUC = 0.863) that the distribution is driven by land cover type, summer precipitation, maximum summer temperatures, and annual temperature variation. We provide updated records of established populations near the northwestern species range limit and present a model that increases our understanding of the potential distribution of *I. scapularis* in Minnesota.

Keywords

Ixodes scapularis; Lyme disease; Maxent; species distribution model

The black-legged tick, *Ixodes scapularis* Say, is the primary vector to humans in the eastern United States of several human pathogens including *Borrelia burgdorferi* (Lyme disease), *Anaplasma phagocytophilum* (anaplasmosis), *Babesia microti* (babesiosis), and the deer tick lineage of Powassan encephalitis virus (Dennis et al. 1998, Homer et al. 2000, Piesman and Gern 2004, Brown et al. 2005, Goodman 2005, Ebel 2010). Over the past two decades in the United States, the incidence of *I. scapularis*-borne diseases has increased, and the geographic distribution of cases has expanded (Centers for Disease Control and Prevention [CDC]

Corresponding author, uzj6@cdc.gov.

2014). Annual case counts of Lyme disease, the most commonly reported vector-borne disease in the United States, increased 101% from 1992 to 2006 (Bacon et al. 2008). While the mean number of counties reporting at least one Lyme disease case remained relatively stable in several states in the Northeast and mid-Atlantic regions (Connecticut, Delaware, Massachusetts, Maryland, New Jersey, and Rhode Island) from 1992 to 2006, during the same time period the percentage of counties reporting at least one case increased in other regions, particularly in the Upper Midwest. The greatest increase was in Minnesota, where the percentage of counties reporting at least one human case increased from 33% in 1992 to 74% in 2006 (Bacon et al. 2008). From 1996 to 2011, the number of reported *I. scapularis*-borne disease cases (including Lyme disease, anaplasmosis, and babesiosis) increased by 742% in Minnesota, which was coupled with an expanded geographic distribution of reported cases throughout the state (Robinson et al. 2015).

Although some of these national and regional changes could have occurred as a result of changing surveillance, increased awareness, and reporting inconsistencies, a true increase in the number of infections is likely. Increased rates of transmission to humans is possible as a result of *I. scapularis* range expansion, increased tick densities in some localities, increased encounter rates between ticks and humans, and possibly an increase in pathogen-infected ticks (Mather et al. 1996, Stafford et al. 1998, Bacon et al. 2008, Diuk-Wasser et al. 2012, Pepin et al. 2012, Robinson et al. 2015). At the national scale, the distribution of Lyme disease cases is closely correlated with the distribution of host-seeking nymphal *I. scapularis* (Diuk-Wasser et al. 2010). *Ixodes-scapularis* is distributed throughout much of the eastern, central, and Upper Midwestern United States, but the density of host-seeking nymphs, as assessed by drag sampling, is greatest in coastal states in the Northeast and Mid-Atlantic and in the Upper Midwest, particularly in Wisconsin and eastern-central Minnesota (Dennis et al. 1998, Diuk-Wasser et al. 2010).

Several previous studies have estimated the distribution of *I. scapularis* in Minnesota as part of efforts to map the species' distribution at a national spatial scale (Dennis et al. 1998; Estrada-Peña 2002; Brownstein et al. 2003, 2005; Diuk-Wasser et al. 2006, 2010, 2012). The earlier distribution models (Estrada-Peña 2002, Brownstein et al. 2003, 2005) were based on species records at the county scale that were collected through literature review and a survey of acarologists and public heath entomologists (Dennis et al. 1998). Later studies (Diuk-Wasser et al. 2006, 2010, 2012) conducted systematic sampling of host-seeking nymphs at 304 sites east of the 100th meridian with 26 sites in Minnesota and used these data to identify ecological correlates of acarological risk for Lyme disease (density of host-seeking *I. scapularis* nymphs and infection rates with *B. burgdorferi* in these nymphs). Recognizing that the data used to inform the existing models were collected nearly a decade ago, that Lyme disease incidence has continued to increase and *I. scapularis* distribution has expanded in Minnesota since that time (Robinson 2015), we developed a species distribution model utilizing 25 unique georeferenced established populations (as per Dennis et al. 1998) of *I. scapularis* collected from 18 counties in Minnesota from 2005 through 2014. In an attempt to identify areas at a subcounty scale that are suitable for *I. scapularis* to establish, high resolution (<1 km) predictive variables were used to develop the species distribution model. The model presented here increases our understanding of the potential distribution of *I.*

scapularis in Minnesota and presence data used to inform the model provide updated records of established populations near the northwestern species range limit.

Materials and Methods

Tick Occurrence Data

Tick occurrence data were compiled from field collections of immature and adult *I. scapularis* made by the Minnesota Department of Health (MDH) Vector-borne Diseases Unit between 2005 and 2014. The majority of tick location records were obtained by dragging a modified 1-m^2 white canvas cloth with weighted fingers over the ground in order to collect host-seeking *I. scapularis*. Locations were chosen based on the known endemic areas of *I. scapularis* as well as locations near the periphery of endemic areas. Specific sampling sites were selected based on appropriate *I. scapularis* habitat (i.e. wooded and brushy mesic areas with at least 50% canopy coverage), land manager recommendations, and ease of access. A small number of tick location records used in the development of this model were obtained via an additional effort with the Minnesota Department of Natural Resources (DNR). For this particular effort, DNR forestry staff submitted any ticks they found on themselves while performing fieldwork in wooded areas to MDH along with specific location data. All ticks collected in the field, either on the drag cloth or the person, were labeled with the date and location of collection and identified according to genus and species, life stage, and sex (if adult) at MDH. Georeferenced sampling locations and associated records of the number and life stages of *I. scapularis* were provided by the MDH and were used to inform the model of existing *I. scapularis* habitat. Only records of established tick populations, i.e. at least two life stages or at least six ticks of any one life stage (Dennis et al. 1998), were included as presence locations. All data were projected to Albers Equal Area North American Datum 1983 (NAD83). There were a total of 122 georeferenced locations from which ticks were collected across the state; however, these data were highly spatially clustered. To avoid pseudoreplication, we used nearest neighbor statistics in ArcMap 10.3 (ESRI, Redlands, CA) to reduce clustering and produce a random distribution of presence locations across the state. Among the full dataset, there were 25 main clusters of established *I. scapularis* populations, from which a single randomly chosen presence point was selected. Nearest neighbor analysis was performed and the z-score and associated p-value were calculated to ensure a random distribution was achieved.

Environmental Data

We sought to identify areas that are ecologically conducive to the establishment of *I. scapularis* populations in Minnesota using both landscape and climatic variables (covariates). We began with 69 variables including elevation (National Elevation Dataset, USGS; http://ned.usgs.gov/; accessed February 2014), land cover (US Geological Service Gap Analysis Program (GAP); http://gapanalysis.usgs.gov/; accessed February 2014), and 67 bioclimatic variables obtained from the WorldClim database (Hijmans et al. 2005; http://worldclim.org/current; accessed April 2014). To limit the number of land cover classes, the GAP National Vegetation Classification Standard, v. 2, formation classification (NVC_FORM) was used. Formation is a third-level classification that describes macroclimate conditions that are modified by altitude, seasonality of precipitation,

substrates, and hydrologic conditions (Whittaker 1975, Lincoln et al. 1998, Federal Geographic Data 2008). There are 11 total NVC_FORM classes in Minnesota, including "open water," which was masked out of the model.

We chose to eliminate correlated variables to decrease model complexity and increase the interpretability of model output (Merow et al. 2013). We identified correlated variables (Pearson's r ≥ |0. 80|), using the Band Collection Statistics Tool in ArcMap 10.3 (ESRI, Redlands, CA), which calculates the Pearson's correlation coefficient (r) between all pairs of variables. High levels of correlation were noted throughout the WorldClim dataset, which included the following five data categories: minimum, mean, and maximum monthly temperatures, average monthly precipitation, and BIOCLIM 1–19 which captures annual extremes in temperature and precipitation. A single variable was chosen from each of the five aforementioned categories, resulting in a set of seven uncorrelated climatic variables (representing all variables in the WorldClim dataset). These were combined with elevation and land cover to model the distribution of *I. scapularis* in Minnesota. All environmental layers were projected to Albers NAD 1983, resampled to 30 m to align with the spatial resolution of the land cover layer, and all layers were clipped to the extent of the Minnesota state boundary.

Species Distribution Models

We modeled the potential distribution of suitable habitat for *I. scapularis* in Minnesota using Maxent version 3.3.3k (http://www.cs.princeton.edu/schapire/maxent/, accessed April, 2014). Maxent is a machine learning algorithm based on the principle of maximum entropy and uses environmental data at occurrence and background locations to predict the distribution of a species across a landscape (Phillips et al. 2006) and routinely out-performs other presence-only models (Elith et al. 2006, Merow et al. 2013). Maxent identifies a probability distribution across the landscape that is constrained by parameter values at occurrence locations. This constraint ensures that the mean of each variable used in the model is close to the mean of the variable over occurrence sites, and a regularization parameter prevents overfitting to occurrence locations (Phillips et al. 2006).

We developed a "full model" including all nine ecological variables (Table 1) and all default Maxent settings. Based on output from the full model, we reduced the number of variables by eliminating variables not contributing to model fit and changed the features setting to only include hinge features. Hinge features are capable of modeling piecewise linear responses to variables and allow for more simple and succinct approximations of the response to environmental variables and have been shown to substantially improve model performance and smooth the fit to the data, thus simplifying the fitted features (Phillips and Dudík 2008; Elith et al. 2010, 2011). Hinge features fit the data similar to a generalized additive model with nonlinear fitted functions (Elith et al. 2011). Tenfold cross validation was used ensuring that all data points were utilized to train and test the model fit.

Final model fit was assessed using the AUC_{test} statistic. An AUC was calculated for each of the 10 models and represents the probability that a random presence point will be ranked above a randomly chosen background site (Phillips et al. 2006). Spatially autocorrelated data can artificially inflate AUC scores, especially when performing cross-validation (Hijmans

2012); therefore, it was important to remove correlation among presence points prior to modeling, as described above. Variable contribution was assessed using permutation importance and jackknife tests. Permutation importance was determined by randomly permuting the values of each variable among the presence and background training points and measuring the resulting decrease in training AUC. Values were normalized to percentages; a large decrease indicated that the model relied heavily on that variable. Jackknife tests evaluate and compare AUC values of the model utilizing all variables, with models created using only a single variable in turn and models leaving out one variable in turn. To further investigate the influence of ecological variables, we produced response curves to characterize the effect of each variable on the logistic probability of suitability.

The probability of suitability was estimated across the state of Minnesota. We used the 10th percentile logistic training threshold as the ROC cutoff value to create a binary suitability map. The 10th percentile logistic training threshold is a conservative threshold value determined by the Maxent software that assumes that 10% of the presence data may be prone to errors, thus, using only 90% of the training data to determine the final distribution of suitable habitat (Raes et al. 2009). We chose this conservative measure because the presence data were collected over several years and by multiple collectors (Raes et al. 2009). To visualize the geographic distribution of predicted suitability, we produced a binary raster representing areas with values greater (suitable) or less than (unsuitable) the 10th percentile logistic training threshold.

Results

Tick occurrence data for established *I. scapularis* populations revealed 122 locations with established populations; 25 of which were independent locations used to model the potential distribution of *I. scapularis* in Minnesota (Fig. 1). These 25 presence points were randomly distributed across the state as indicated by a z-score = −1.624 and a p-value = 0.104. The sites were distributed across 15 counties and four land cover types, with the majority (52%) classified in cool temperate forest habitat, while lowland and montane boreal forest accounted for 28% of occurrence points (Table 2). The potential distribution of *I. scapularis* in Minnesota was based on land cover, elevation, and five uncorrelated BioClim variables (Table 2). The mean AUC_{test} for the 10 replicate models was 0.86 (s.d. = 0.05), which represents the probability that a randomly chosen presence site is ranked above a random background site. In total, 19.3% of Minnesota was classified as suitable habitat for *I. scapularis* (Fig. 1). We further tested model performance by using the remaining 97 presence locations and show an overall model sensitivity of 79.5%.

Seven variables contributed to model fit (Table 3). However, land cover (GAP) overwhelmingly had the most influence on the model, having a permutation importance of almost 80 (Table 3). Other variables, including the maximum temperature during the warmest month, precipitation of the wettest quarter, and annual temperature range, had the next highest permutation importance (Table 3). In Minnesota, July is typically the warmest month while the warmest and coldest quarters correspond with June through August and December through February, respectively. Jackknife tests of variable importance showed comparable results to the permutation importance findings and indicated that land cover

alone contained the most useful information. That is, it produced the highest gain in AUC when used in the model alone. Eliminating this variable from the model caused the largest drop in AUC because it contained the most information that isn't present in other variables (Fig. 2). Both summer precipitation and annual temperature range contributed a large amount of unique information when used as single variables in the model. Eliminating the aforementioned variables from the model caused the third highest drop in AUC, indicating some unique information is contained within these variables (Fig. 2). Maximum temperature during the warmest month by itself did not provide a good fit to the data, but exclusion of this variable from the model resulted in the second largest drop in AUC. Jackknife results of the remaining variables showed minimal effects of removing each one in turn (Fig. 2).

Response curves were produced in Maxent, with each curve representing a model created using individual variables. The plots characterize the dependence of the predicted suitability on each variable as well as dependencies induced by underlying correlations between the selected variable and other variables. Although we removed correlated climate variables, we did not identify correlations between land cover and other variables; therefore, it was important to consider these effects when interpreting results. Although the cool temperate forest and temperate grassland, meadow, and shrubland categories had the highest logistic probability of suitability (Fig. 1), greater than 50% of both input and predicted points were classified as cool temperate forests while less than six percent of input and predicted points were classified as temperate grassland, meadow, and shrubland (Table 2). The predicted *I. scapularis* distribution was composed of 10 land cover types (Table 2); however, the majority (91.5%) of the distribution was predicted to occur in three types of forest containing deciduous hardwood species including maple, basswood, oak, and aspen (Table 4). The response curve for summer precipitation exhibits a threshold of 28 cm, below which it appears there is not enough moisture to support *I. scapularis* (Fig. 3B). Areas of the state with annual temperature ranges between 39°C and 42°C have the highest probability of suitability (Fig. 3C). Overall, the highest probability of suitability occurred when the maximum temperature of the warmest month is between 24.0–28.5°C (Fig. 3D). The values for the maximum temperature of the warmest month for input and predicted points range from 23.6–28.8°C and 21.8–29.9°C, respectively.

Discussion

Based on updated tick distribution records, we created a fine resolution, subcounty-level distribution model for *I. scapularis* in Minnesota that expands on previous efforts to define the spatial distribution of *I. scapularis* in the eastern United States and Upper Midwest (Dennis et al. 1998; Estrada-Peña 2002; Diuk-Wasser et al. 2006, 2010). Our study, combined with results from Diuk-Wasser et al. (2006, 2010), reports the establishment of *I. scapularis* in 21 additional counties since county records were reviewed by Dennis et al. (1998). At that time only eight Minnesota counties were classified as established (greater than six individual ticks of a single life stage) and 13 were classified as reported, although it is likely that the actual distribution of *I. scapularis* was greater than what was officially documented through field efforts. Here, we classified 18 counties as having established *I. scapularis* populations, six of these were also classified as established by Diuk-Wasser et al. (2010). Three additional counties that were not included in our study were shown to be

established by Diuk-Wasser (2010). Many of these new establishment records are further north and west than the original distribution limits (Dennis et al. 1998). Notably, the report by Dennis et al. (1998) and our current study span roughly the 16-yr surveillance period in Minnesota that was reported by Robinson et al. (2015) in a study that described the northwesterly expansion of Lyme disease and anaplasmosis in Minnesota. Our findings support the notion that geographic expansion of *I. scapularis*-borne illnesses in Minnesota are driven by range expansion of the vector. The model presented here is concordant with Minnesota counties that pose the largest threat of *I. scapularis*-borne diseases, based on the average incidence (cases/100,000 population) of Lyme disease and human anaplasmosis cases reported to the MDH between 2007 and 2013 (Fig. 1). All counties classified as posing either high or moderate risk of disease were classified as having at least some suitable habitat to support *I. scapularis* populations. However, the model predicts some areas in northern and southern Minnesota to have potentially suitable habitat, yet *I. scapularis*-borne disease risk remains low in these areas. These may represent areas 1) with potentially suitable habitat that have yet to be colonized or are unable to be colonized by the tick, 2) where suitable tick habitat exists but where there is minimal human activity, or 3) where the prevalence of infection is low in ticks, perhaps owing to host community composition in these areas. Future studies are needed to determine causes for discordance between suitable *I. scapularis* distribution and the occurrence of *I. scapularis*-borne diseases.

In agreement with each of the existing models (Estrada-Peña 2002, Brownstein et al. 2003, Diuk-Wasser et al. 2010), our model predicts that the most suitable areas for *I. scapularis* include much of the eastern border with Wisconsin. Based on our model, the distribution radiates to the northwest across one half to two-thirds of Minnesota, a finding consistent with Diuk-Wasser et al. (2010), more broad in distribution than expected by Estrada-Pena (2002) and more limited in its western limit than predicted by Brownstein et al. (2003). Much of the area predicted to be suitable is dominated by cool temperate forests (Fig. 1B). Similar to our model, each of the other models predicted the northern tier of the state to be less suitable than the central portion (Estrada-Peña 2002, Brownstein et al. 2003, Diuk-Wasser et al. 2010). Northern Minnesota is composed of a mix of boreal forest and limited temperate forest and is highly fragmented by open water or bog habitat (Fig. 1B). As a result of land cover and the lower temperatures characteristic in that part of the state, our model predicted limited suitability in the north. The model identifies some areas in north-central Minnesota, along the Canadian border, adjacent to reported locations of the acquisition of Lyme disease in Canada. However, the model does not predict suitability in north western Minnesota in an area adjacent to other risk areas in Canada (Ogden et al. 2015), suggesting that our model could be under-predicting in this area. The landscape of southern and western Minnesota is dominated by agriculture, but there are areas of aspen parkland and other upland forest that may provide suitable *I. scapularis* habitat in these areas. Our model predicted suitable habitat near the Iowa and South Dakota borders in the southwestern part of the state and near the Iowa border in southcentral Minnesota and extending up the Minnesota River Valley. These predictions are in agreement with the distribution put forth by Brownstein et al. (2003), however, in the model presented here the distribution of suitability is limited by land cover (e.g. temperate shrublands that may provide adequate habitat; Fig.

1B) and thus does not encompass large contiguous portions of Minnesota as shown in Brownstein et al. (2003).

The presence of forested habitat or indirect measures of habitat quality such as the Normalized Difference Vegetation Index (NDVI) are consistently important predictors of the *I. scapularis* distribution (Estrada-Peña 2002, Killilea et al. 2008, Diuk-Wasser et al. 2010, this study). Tick survival is highly dependent on access to a stable microclimate that exists beneath a layer of leaf litter on the floor of deciduous forests with closed canopies (Ginsberg and Ewing 1989, Ginsberg and Zhioua 1996, Ginsberg et al. 2004). Although a majority of suitable habitat is characterized as cool temperate forest (Table 2), other vegetation types were also predicted to be suitable for establishment of *I. scapularis*. For example, the majority of the potentially suitable habitat in western and southwestern Minnesota is classified as temperate grassland, meadow, and shrubland (Fig. 1B). Although the tick is not yet established in most southwestern counties, wooded habitat within these areas may represent potential areas for the spread and future establishment of *I. scapularis* in western Minnesota. Further, not all areas covered with deciduous forests were classified as suitable, likely because climatic conditions were not conducive to establishment. Adequate leaf litter and canopy cover help to stabilize the tick microclimate, but tick survival, metamorphosis, and reproduction are still largely influenced by temperature and humidity (Lindsay et al. 1995, Ogden et al. 2014).

Climate variables, particularly minimum and maximum temperatures, have been recognized as limiting the distribution of *I. scapularis* (Estrada-eña 2002, Brownstein et al. 2003, Diuk-Wasser et al. 2010). Here, we identified the maximum temperature during the warmest month as the most influential climate variable. The range for maximum temperature during the warmest month with the highest predicted suitability in Minnesota was 22.9 to 29.8°C, which includes most of the variability within the state (Table 2) but also represents the range for this variable found across the eastern United States where *I. scapularis* and human incidence of Lyme disease are well documented (e.g., primarily east of the 100th meridian and at latitudes greater than 35°N). While the effects of maximum temperature may both augment developmental and hatching rates or hinder overall survival and oviposition success, low minimum temperatures can limit tick distributions by directly killing them or inhibiting host-seeking activity (Vail and Smith 1998, Perret et al. 2000, Schulze et al. 2002, Ogden et al. 2004, Rand et al. 2004). However, minimum temperature contributed minimally to defining the distribution given by the Maxent model here. The areas of Minnesota predicted to be suitable for *I. scapularis* cover nearly the entire range of this variable (Table 2), a result that might be expected when considering that winter precipitation and leaf litter both provide insulating effects during the coldest portion of the year (Lindsay et al. 1995, Lindgren and Gustafson 2001, Brunner et al. 2012). Sufficient summer precipitation is needed to offset heat-induced mortality as is indicated in the response curve showing a marked threshold below which conditions are unsuitable for *I. scapularis* (Fig. 2). Areas in southern and western Minnesota where predicted habitat suitability is low also receive the lowest amount of summer precipitation and have the highest temperatures during the summer (Fig. 3).

We present a distribution map of habitat suitability for *I. scapularis* in Minnesota, created using updated records of established tick populations and high resolution variables representing climatic conditions and vegetation type in Minnesota. We emphasize that the use of fine resolution data allowed us to recognize small-scale differences in suitability and may point to pockets within counties that pose an elevated risk for exposure to tick-borne pathogens (Eisen and Eisen 2008). In disease surveillance, reports are classified according to county of residence rather than county of exposure, which may confound results, especially in cases with potential travel-associated exposures. Further investigation is necessary to better quantify spatial risk at a subcounty spatial scale and to increase our understanding of human disease risk across Minnesota. This study increases the breadth of knowledge regarding the distribution of *I. scapularis*, a key vector for tick-borne pathogens posing significant public health risk, and adds to a growing body of knowledge on the continued range expansion of *I. scapularis* and increased distribution and incidence of tick-borne diseases, especially Lyme disease, in North America. Nonetheless, to provide an improved acarological risk assessment for this area where Lyme disease is increasing both in incidence and geographic extent, future studies are needed to quantify the density of *I. scapularis* across Minnesota, as well as the prevalence of *B. burgdorferi* infection and other tick-borne pathogens.

Acknowledgments

References Cited

Bacon RM, Kugeler KJ, Mead PS, C. Centers for Disease, and Prevention. Surveillance for Lyme disease–United States, 1992–2006. MMWR Surveillance summaries. 2008; 57:1–9.

Brown, RN.; Lane, RS.; Dennis, DT. Geographic distributions of tick-borne diseases and their vectors. In: Goodman, JL.; Dennis, DT.; Sonenshine, DT., editors. Tick-borne diseases of humans. ASM Press; Washington, DC: 2005. p. 363-391.

Brownstein JS, Holford TR, Fish D. A climate-based model predicts the spatial distribution of the Lyme disease vector Ixodes scapularis in the United States. Environ Health Persp. 2003; 111:1152–1157.

Brownstein JS, Holford TR, Fish D. Effect of climate change on Lyme disease risk in North America. EcoHealth. 2005; 2:38–46. [PubMed: 19008966]

Brunner JL, Killilea M, Ostfeld RS. Overwintering survival of nymphal *Ixodes scapularis* (Acari: Ixodidae) under natural conditions. J Med Entomol. 2012; 49:981–987. [PubMed: 23025177]

(CDC) Centers for Disease Control and Prevention. Summary of notifiable diseases-United States, 2012. MMWR. 2014; 61:1–121. [PubMed: 25233134]

Dennis DT, Nekomoto TS, Victor JC, Paul WS, Piesman J. Reported distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the United States. J Med Entomol. 1998; 35:629–638. [PubMed: 9775584]

Diuk-Wasser MA, Gatewood AG, Cortinas MR, Yaremych-Hamer S, Tsao J, Kitron U, Hickling G, Brownstein JS, Walker E, Piesman J, et al. Spatiotemporal patterns of host-seeking *Ixodes scapularis* nymphs (Acari: Ixodidae) in the United States. J Med Entomol. 2006; 43:166–176. [PubMed: 16619595]

Diuk-Wasser MA, Hoen AG, Cislo P, Brinkerhoff R, Hamer SA, Rowland M, Cortinas R, Vourc'h G, Melton F, Hickling GJ, et al. Human risk of infection with *Borrelia burgdorferi*, the Lyme disease agent, in eastern United States. Am J Trop Med. 2012; 86:320–327.

Diuk-Wasser MA, Vourc'h G, Cislo P, Hoen AG, Melton F, Hamer SA, Rowland M, Cortinas R, Hickling GJ, Tsao JI. 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. Global Ecol Biogeogr. 2010; 19:504–514.

Ebel GD. Update on Powassan virus: emergence of a North American tick-borne flavivirus. Annu Rev Entomol. 2010; 55:95–110. [PubMed: 19961325]

Eisen RJ, Eisen L, Beard CB. County-scale distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the continental United States. J Med Entomol. 2016; 53:349–386. [PubMed: 26783367]

Eisen RJ, Eisen L. Spatial modeling of human risk of exposure to vector-borne pathogens based on epidemiological versus arthropod vector data. J Med Entomol. 2008; 45:181–192. [PubMed: 18402133]

Elith J, Kearney M, Phillips S. The art of modelling range-shifting species. Methods Ecol Evol. 2010; 1:330–342.

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ. A statistical explanation of MaxEnt for ecologists. Divers Distrib. 2011; 17:43–57.

Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick J, Lehmann A, et al. Novel methods improve prediction of species' distributions from occurrence data. Ecography. 2006; 29:129–151.

Estrada-Peña A. Increasing habitat suitability in the United States for the tick that transmits Lyme disease: A remote sensing approach. Environ Health Persp. 2002; 110:635.

Federal Geographic Data, C. National vegetation classification standard, version 2. Retrieved January. 2008; 8:2009.

Ginsberg HS, Ewing CP. Habitat distribution of *Ixodes dammini* (Acari: Ixodidae) and Lyme disease spirochetes on Fire Island, New York. J Med Entomol. 1989; 26:183–189. [PubMed: 2724316]

Ginsberg HS, Zhioua E. Nymphal survival and habitat distribution of *Ixodes scapularis* and *Amblyomma americanum* ticks (Acari: Ixodidae) on Fire Island, New York, USA. Exp Appl Acarol. 1996; 20:533–544.

Ginsberg HS, Zhioua E, Mitra S, Fischer J, Buckley PA, Verret F, Underwood HB, Buckley FG. Woodland type and spatial distribution of Nymphal *Ixodes scapularis* (Acari: Ixodidae). Environ Entomol. 2004; 33:1266–1273.

Goodman, JL. Human Granulocytic Anaplasmosis. In: Goodman, JL.; Dennis, DT.; Sonenshine, DE., editors. Tick-borne Diseases of Humans. ASM Press; Washington, DC: 2005. p. 218-238.

Hijmans RJ. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. Ecology. 2012; 93:679–688. [PubMed: 22624221]

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol. 2005; 25:1965–1978.

Homer MJ, Aguilar-Delfin I, Telford SR 3rd, Krause PJ, Persing DH. Babesiosis. Clin Microbiol Rev. 2000; 13:451–469. [PubMed: 10885987]

Killilea ME, Swei A, Lane RS, Briggs CJ, Ostfeld RS. Spatial dynamics of Lyme disease: a review. EcoHealth. 2008; 5:167–195. [PubMed: 18787920]

Lincoln, R.; Boxshall, G.; Clark, P. A dictionary of ecology, evolution and systematics. Cambridge University Press; New York: 1998.

Lindgren E, Gustafson R. Tick-borne encephalitis in Sweden and climate change. Lancet. 2001; 358:16–18. [PubMed: 11454371]

Lindsay LR I, Barker K, Surgeoner GA, McEwen SA, Gillespie TJ, Robinson JT. Survival and development of *Ixodes scapularis* (Acari: Ixodidae) under various climatic conditions in Ontario, Canada. J Med Entomol. 1995; 32:143–152. [PubMed: 7608920]

Mather TN, Nicholson MC, Donnelly EF, Martyas BT. Entomologic index for human risk of Lyme disease. Am J Epidemiol. 1996; 144:1066–1069. [PubMed: 8942438]

Merow C, Smith MJ, Silander JA. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography. 2013; 36:1058–1069.

Needham GR, Teel PD. Off-host physiological ecology of ixodid ticks. Annu Rev Entomol. 1991; 36:659–681. [PubMed: 2006871]

Ogden NH, Koffi JK, Lindsay LR, Flemming S, Mombourquette DC, Sandford C, Badcock J, Gad RR, Jain-Sheehan N, Moore S, et al. Surveillance for Lyme disease in Canada, 2009 to 2012. Canada Communicable Disease Report. 2015; 41:132–145.

Ogden NH, Radojevic M, Wu X, Duvvuri VR, Leighton PA, Wu J. Estimated effects of projected climate change on the basic reproductive number of the Lyme disease vector Ixodes scapularis. Environ Health Persp. 2014; 122:631–638.

Ogden NH, Lindsay LR, Beauchamp G, Charron D, Maarouf A, O'Callaghan CJ, Waltner-Toews D, Barker IK. Investigation of relationships between temperature and developmental rates of tick *Ixodes scapularis* (Acari: Ixodidae) in the laboratory and field. J Med Entomol. 2004; 41:622–633. [PubMed: 15311453]

Pepin KM, Eisen RJ, Mead PS, Piesman J, Fish D, Hoen AG, Diuk-Wasser MA. Geographic variation in the relationship between human Lyme disease incidence and density of infected host-seeking *Ixodes scapularis* nymphs in the Eastern United States. Am J Trop Med. 2012; 86:1062–1071.

Perret JL, Guigoz E, Rais O, Gern L. Influence of saturation deficit and temperature on *Ixodes ricinus* tick questing activity in a Lyme borreliosis-endemic area (Switzerland). Parasitol Res. 2000; 86:554–557. [PubMed: 10935905]

Phillips SJ, Dudík M. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography. 2008; 31:161–175.

Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol Model. 2006; 190:231–259.

Piesman J, Gern L. Lyme borreliosis in Europe and North America. Parasitology. 2004; 129:S191–S220. [PubMed: 15938512]

Raes N, Roos MC, Slik JW, Van Loon EE, Steege HT. Botanical richness and endemicity patterns of Borneo derived from species distribution models. Ecography. 2009; 32:180–192.

Rand PW, Holman MS, Lubelezyk C, Lacombe EH, DeGaetano AT, Smith RP. Thermal accumulation and the early development of *Ixodes scapularis*. J Vector Ecol. 2004; 29:164–176. [PubMed: 15266754]

Robinson SJ, Neitzel DF, Moen RA, Craft ME, Hamilton KE, Johnson LB, Mulla DJ, Munderloh UG, Redig PT, Smith KE, et al. Disease risk in a dynamic environment: The spread of tick-borne pathogens in Minnesota, USA. EcoHealth. 2015; 12:152–163. [PubMed: 25281302]

Schulze TL, Jordan RA, Hung RW. Effects of microscale habitat physiognomy on the focal distribution of *Ixodes scapularis* and *Amblyomma americanum* (Acari: Ixodidae) nymphs. Environ Entomol. 2002; 31:1085–1090.

Stafford KC, Cartter ML, Magnarelli LA, Ertel SH, Mshar PA. Temporal correlations between tick abundance and prevalence of ticks infected with *Borrelia burgdorferi* and increasing incidence of Lyme disease. J Clin Microbiol. 1998; 36:1240–1244. [PubMed: 9574684]

US Geologic Survey, Gap Analysis Program (GAP). National Land Cover, Version 2. 2011 Version 2.

Vail SG, Smith G. Air temperature and relative humidity effects on behavioral activity of blacklegged tick (Acari: Ixodidae) nymphs in New Jersey. J Med Entomol. 1998; 35:1025–1028. [PubMed: 9835697]

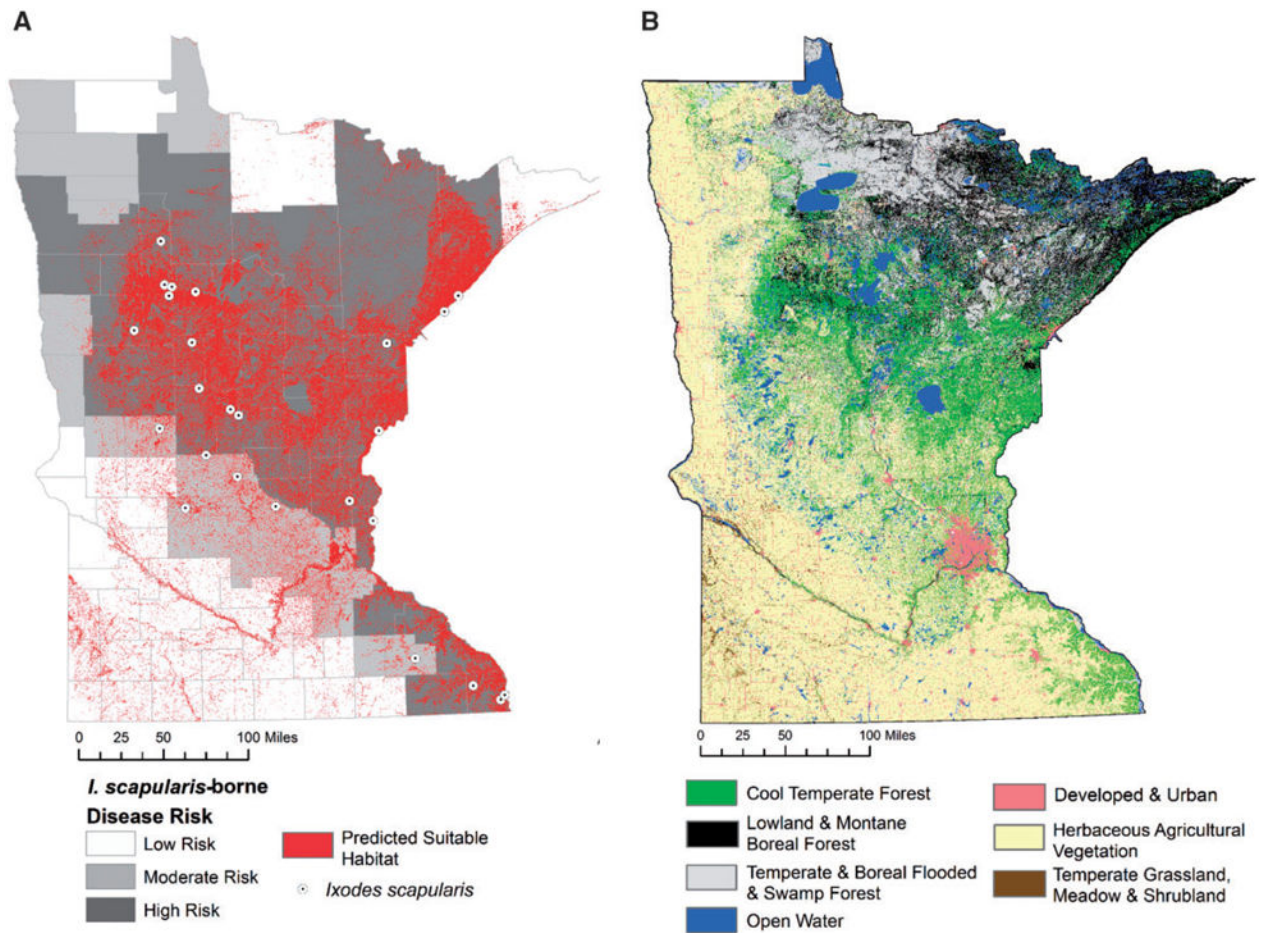Whittaker, RH. Communities and Ecosystems. 2. The Macmillan Company; New York, NY: 1975.

Fig. 1.
(**A**) The predicted distribution of *I. scapularis* in Minnesota is shown in red and is overlaid onto the risk of *I. scapularis*-borne disease based on the average incidence (cases/100,000 population) of Lyme disease and human anaplasmosis cases reported to the Minnesota Department of Health, 2007–2013. Low risk is defined as <10, moderate risk is defined as 10–25, and high risk is defined as 25 cases/100,000 population. (http://www.health.state.mn.us/divs/idepc/diseases/lyme/highrisk.html, accessed August, 2014). (**B**) The distribution of dominant land cover types (United States Geological Service, National Gap Analysis Program (GAP) National Vegetation Classification—formation or land use classification) across Minnesota. The majority (67%) of predicted suitable habitat for *I. scapularis* in Minnesota is located in cool temperate forests, shown in bright green.
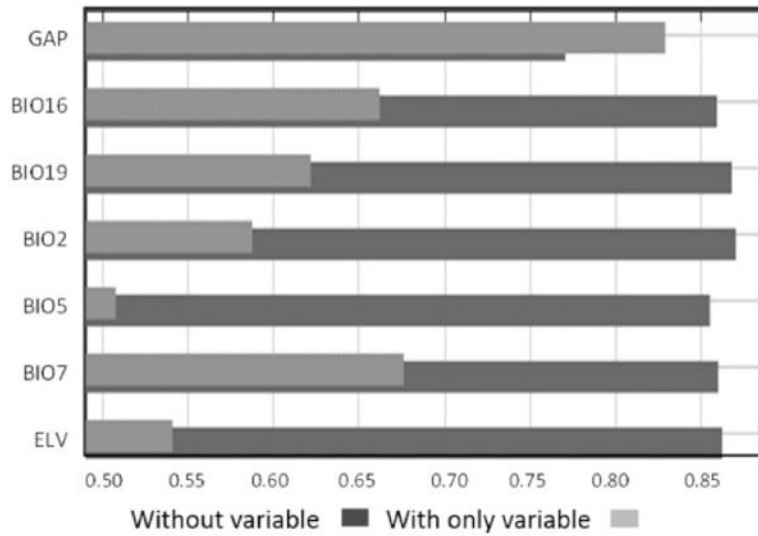
Fig. 2.
Jackknife tests were used to estimate variable importance based on the drop in AUC when excluding each variable in turn as well as models including only single variables. Decreases in AUC when variables are excluded (dark bars) indicate that unique information is contained in a variable that is not present in other variables. Models using land cover as the only variable have the largest AUC, demonstrating that land cover has the most useful information when used alone, as indicated by the large light-colored bar. GAP = United States Geological Service, National Gap Analysis Program (GAP) National Vegetation Classification—formation or land use classification, BIO16 = precipitation of the warmest quarter, BIO19 = precipitation of the coldest quarter, BIO2 = mean diurnal range (mean of monthly (maximum temperature – min temperature)), BIO5 = maximum temperature of the warmest month, BIO7 = annual temperature range, ELV = elevation.
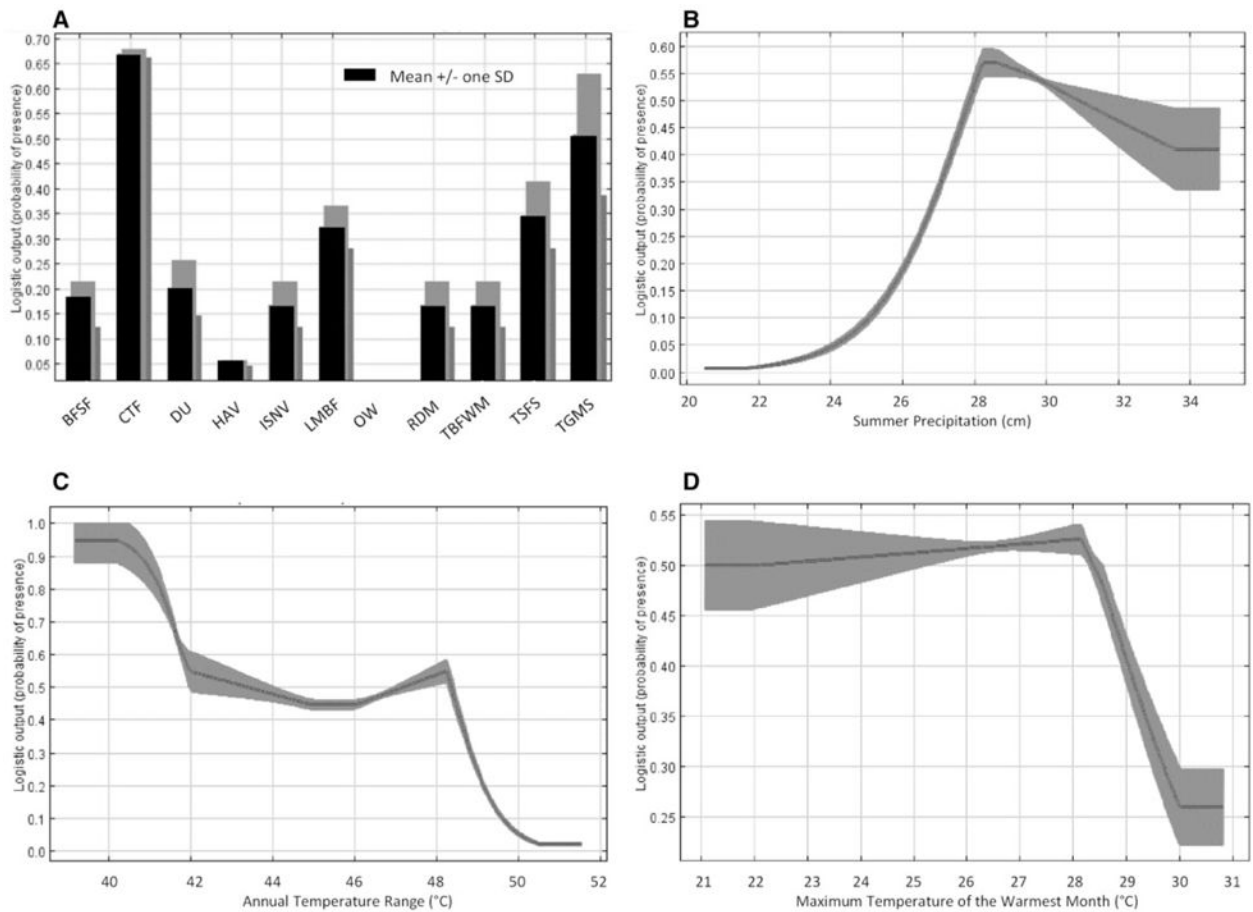
Fig. 3.
Response curves reflect the dependence of the predicted suitability on each variable and on dependencies caused by correlations between each selected variable and other variables. Shading indicates the standard deviation of the 10 replicate model runs. Four variables were most important in describing the potential distribution of *I. scapularis* in Minnesota. The y-axis shows the logistic probability of presence; higher logistic probabilities indicate higher suitability. The variables include: (**A**) United States Geological Service, National Gap Analysis Program (GAP) National Vegetation Classification—formation or land use classification: BSFS = boreal flooded and swamp forest, CTF = cool temperate forest, DU = developed and urban, HAV = herbaceous agricultural vegetation, ISNV = introduced and seminatural vegetation, LMBF = lowland and montane boreal forest, OW = open water, RDM = recently developed or modified, TBFWMM = temperate and boreal freshwater wet meadow and marsh, TFSF = temperate flooded and swamp forest, TGMS = temperate grassland, meadow and shrubland, (**B**) summer precipitation, (**C**) annual temperature range, and (**D**) the maximum temperature during the warmest month.

Table 1

Uncorrelated ecological variables used to model the distribution of potentially suitable habitat for *I. scapularis* in Minnesota

Layer	Original datum	Original resolution	Source
Land cover	NAD83	30 m^2	USGS, National Gap Analysis Project[a]
Elevation	WGS84	28 m^2	National Elevation Dataset[b]
Mean diurnal temperature range	WGS84	0.77 km^2	Worldclim Dataset[c]
Isothermality[d]	WGS84	0.77 km^2	Worldclim Dataset[c]
Maximum temperature of the warmest month	WGS84	0.77 km^2	Worldclim Dataset[c]
Annual temperature range	WGS84	0.77 km^2	Worldclim Dataset[c]
Mean temperature of the coldest quarter[d]	WGS84	0.77 km^2	Worldclim Dataset[c]
Precipitation of the wettest quarter	WGS84	0.77 km^2	Worldclim Dataset[c]
Precipitation of the coldest quarter	WGS84	0.77 km^2	Worldclim Dataset[c]

[a] http://gapanalysis.usgs.gov/gaplandcover/, accessed April, 2014

[b] http://ned.usgs.gov/, accessed April, 2014

[c] http://worldclim.org/current, accessed April, 2014

[d] Nine uncorrelated ecological variables were included in the full model. All variables except isothermality and mean temperature of the coldest quarter were included in the reduced model.

Table 2

Ecological variables included in the Maxent model to predict the potential distribution of *I. scapularis* in Minnesota

Variable	MN range[a]	Presence points range[a]	Maxent range[a]
Mean diurnal temperature range (°C)	9.8–13.5	10.0–13.0	9.9–13.5
Maximum temperature during the warmest month (°C)	21.8–30.1	23.6–28.8	21.8–29.9
Annual temperature range (°C)	40.2–50.5	40.6–48.5	40.2–50.1
Precipitation during the wettest quarter (mm)	214–336	275–324	221–336
Precipitation during the coldest quarter (mm)	39–116	50–80	39–116
Elevation (m)	175–668	190–492	177–660
Land cover[b]	% MN	% Presence points	% Maxent predicted presence
Cool temperate forest	14.3	52.0	67.1
Lowland and montane boreal forest	13.6	28.0	15.7
Temperate flooded and swamp forest	5.8	16.0	8.7
Temperate grassland, meadow and shrubland	1.4	4.0	5.6
Developed and urban	5.4	0.0	1.9
Recently disturbed or modified	1.8	0.0	0.5
Temperate and boreal freshwater wet meadow and marsh	1.5	0.0	0.4
Boreal flooded and swamp forest	10.8	0.0	0.1
Herbaceous agricultural vegetation	45.5	0.0	< 0.1
Introduced and seminatural vegetation	< 0.1	0.0	< 0.001

[a] The range for the entire state is reported in addition to the range of values of each variable at the presence points used to train the model and the range of each variable chosen to represent suitable habitat.

[b] National Vegetation Classification—formation or land use classification

Table 3

Contribution of eight ecological variables to the distribution of *I. scapularis* in Minnesota

Variable	Permutation importance[a]
Land cover[b]	79.5
Maximum temperature during the warmest month	8.2
Annual temperature range	3.6
Precipitation during the wettest quarter	3.0
Mean diurnal temperature range	2.5
Elevation	1.7
Precipitation during the coldest quarter (mm)	1.5

[a]Permutation importance is the resulting drop in AUC when each variable is permuted in turn. Isothermality and mean temperature during the coldest quarter had permutation importance values of zero in the full model; thus, both variables were excluded from the reduced model.

[b]National Vegetation Classification—formation or land use classification (See Table 2).

Table 4

Land cover classification of the predicted suitable areas for *I. scapularis* occurrence with deciduous tree species

Proportion	Formation[a]	Macrogroup	Ecosystem group	Dominant tree species
0.23	Cool temperate forest	Northern mesic hardwood and conifer forest	Laurentian-Acadian northern hardwoods forest	*Acer* spp., *Betula* spp.
0.10	Cool temperate forest	Central mesophytic hardwood forest	North-central interior maple-basswood forest	*Acer* spp., *Tilia* spp.
0.10	Lowland and montane boreal forest	Eastern and Central North American boreal conifer and hardwood forest	Boreal aspen-birch forest	*Populus* spp., *Betula* spp.
0.071	Cool temperate forest	Central oak-hardwood and pine forest	North-central interior dry-mesic oak forest and woodland	*Quercus* spp.
0.07	Lowland and montane boreal forest	Eastern and Central North American boreal conifer and hardwood forest	Boreal white spruce-fir-hard wood forest	*Populus* spp., *Betula* spp.
0.05	Cool temperate forest	Northern and eastern pine–oak forest, woodland and barrens	Laurentian pine-oak barrens	*Quercus* spp.
0.04	Cool temperate forest	Northern mesic hardwood and conifer forest	Laurentian-Acadian northern pine-(oak) forest	*Quercus* spp.
0.04	Boreal flooded and swamp forest	North American boreal swamp forest	Boreal-Laurentian conifer acidic swamp and treed poor fen	*Betula* spp.
0.03	Temperate flooded and swamp forest	Northern and central floodplain forest and scrub	Laurentian-Acadian swamp systems	*Fraxinus* spp.
0.02	Temperate flooded and swamp forest	Northern and central floodplain forest and scrub	Laurentian-Acadian floodplain systems	*Acer* spp., *Quercus* spp.
0.02	Developed and urban	Developed and urban	Developed, open space	*Quercus* spp.
0.02	Temperate flooded and swamp forest	Northern and central swamp forest	North-central interior and Appalachian rich swamp	*Acer* spp.
0.01	Temperate grassland, and shrubland	Great Plains tallgrass prairie and shrubland	Northern tallgrass prairie	*Quercus* spp., *Populus* spp.
0.01	Temperate grassland, meadow and shrubland	Great Plains tallgrass prairie and shrubland	Central tallgrass prairie	*Quercus* spp.
0.01	Temperate flooded and swamp forest	Northern and central floodplain forest and scrub	Central interior and Appalachian floodplain systems	*Acer* spp.

[a]National Vegetation Classification—formation or land use classification (See Table 2). Table shows Ecosystem groups representing areas of at least 1% of the predicted suitable habitat in Minnesota.