

**Spatial distribution modeling of *Dermacentor variabilis* ticks
under current and future climate change scenarios**

by

Gunavanthi Devi Yadav Boorgula

B.S., Osmania University, 2004
M.Sc., Kakatiya University, 2006

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Diagnostic Medicine and Pathobiology
College of Veterinary Medicine

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Approved by:

Major Professor
Dr. Ram K Raghavan

Copyright

© Gunavanthi Devi Yadav Boorgula 2018.

Abstract

Dermacentor variabilis (Say) (Acari: Ixodidae), commonly known as the American dog tick is a medically important tick species in N. America, which has been implicated as a competent vector for several diseases, including tularemia, bovine anaplasmosis and canine tick paralysis. This tick is also the primary suspect for the transmission of *Rickettsia rickettsii*, the causative agent of Rocky Mountain spotted fever (RMSF). The spatial distribution and geographic extent of *D. variabilis* territory in N. America is suspected to have changed in the recent times due to natural and anthropogenic, non-stationary forces. A clear understanding of the spatial distribution and environmental factors contributing to the distribution has public health significance, allowing us to make informed management decisions and for setting robust future research goals aimed at understanding vector-biology and disease management. Additionally, ongoing climate-change is expected to alter species spatial distribution and abundance within distribution range. In this research, I studied the current and likely future spatial distribution of *D. variabilis* ticks in N. America based on two representative concentration pathways, RCP 4.5 and RCP 8.5, representing lower and higher emission scenarios, respectively, under several global circulation models (GCM). The spatial distribution models were constructed using MaxEnt program and BioClim data was used as environmental data for modeling. Best models were selected based on Partial ROC curves, AIC, and omission rates. Median prediction of these models indicate a wider spread of *D. variabilis* from its currently known extent, and much further spread as a result of climate change. Different environmental variables that significantly influenced current and future *D. variabilis* distribution included annual mean temperature, mean diurnal range, maximum temperature of the warmest month, annual precipitation, precipitation seasonality, and precipitation of the wettest quarter.

Table of Contents

List of Figures	v
List of Tables	ix
Acknowledgements	x
Chapter 1 - LITERATURE REVIEW AND REASEARCH OBJECTIVE.....	1
Chapter 2 - MATERIALS AND METHODS.....	7
Chapter 3 - RESULTS	14
Chapter 4 - DISCUSSION	18
BIBLIOGRAPHY	53

List of Figures

Figure 1 : Occurrence locations, and accessible area, <i>M</i> , considered for spatial distribution modeling of <i>Dermacentor variabilis</i> (Say) in N. America. Red dots indicate locations used for model calibration, and blue dots represent locations used for model evaluation.....	22
Figure 2 : Jackknife plot representing regularized training gain for <i>Dermacentor variabilis</i> spatial distribution with all BioClim variables included in a global model.	23
Figure 3 : Jackknife plot representing regularized training gain for <i>Dermacentor variabilis</i> spatial distribution with all BioClim variables but without variables 6, 7, 11 and 13.	24
Figure 4 : Jackknife plot representing regularized training gain for <i>Dermacentor variabilis</i> spatial distribution with all BioClim variables but without variables 3, 4, 6, 7, 10, 11 and 13.	25
Figure 5 : Jackknife plot representing regularized training gain for <i>Dermacentor variabilis</i> spatial distribution with all BioClim variables but without variables 3, 4, 6, 7, 10, 11, 13, 14 and 17.....	26
Figure 6 : Current distribution of <i>Dermacentor variabilis</i> based on present-day climatic conditions. This picture represents the median of point-wise medians of the top five models. Each of the top five models were derived based on the median of the 10 replicate models applied to the environmental layers. The cell- wise probability of finding <i>D. variabilis</i> is depicted in different colors.	27
Figure 7 : Uncertainty of current distribution of <i>Dermacentor variabilis</i> distribution in N. America. This picture represents the point-wise range (maximum – minimum) of median distribution model based on the top five models. Areas in red indicate current distribution based on present – day model, and different grey colors are uncertainties.	28
Figure 8 : Current distribution of <i>Dermacentor variabilis</i> based on present-day climatic conditions with Linear feature, environmental set 3, and regularization multiplier value set to 2.0. This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding <i>D. variabilis</i> is depicted in different colors.	29
Figure 9 : Current distribution of <i>Dermacentor variabilis</i> based on present-day climatic conditions with Linear feature, environmental set 3, and regularization multiplier value set to 3.0. This picture represents the median of the 10 replicate models applied to the	

environmental layers. The cell-wise probability of finding <i>D. variabilis</i> is depicted in different colors.	30
Figure 10 : Current distribution of <i>Dermacentor variabilis</i> based on present-day climatic conditions with Linear feature, environmental set 3, and regularization multiplier value set to 4.0. This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding <i>D. variabilis</i> is depicted in different colors.	31
Figure 11 : Current distribution of <i>Dermacentor variabilis</i> based on present-day climatic conditions with Linear feature, environmental set 3, and regularization multiplier value set to 0.1. This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding <i>D. variabilis</i> is depicted in different colors.	32
Figure 12 : Current distribution of <i>Dermacentor variabilis</i> based on present-day climatic conditions with Linear feature, environmental set 3, and regularization multiplier value set to 0.2. This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding <i>D. variabilis</i> is depicted in different colors.	33
Figure 13 : The variable percentage contribution to the top five individual maximum entropy models for the spatial distribution of <i>Dermacentor variabilis</i>	34
Figure 14 : Potential distribution of <i>Dermacentor variabilis</i> under representative concentration pathway RCP 4.5 emission scenario and the Canadian Centre for Climate Modeling and Analysis (CGCM3) General Circulation Model (GCM). This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding <i>D. variabilis</i> is depicted in different colors.	35
Figure 15 : Potential distribution of <i>Dermacentor variabilis</i> under representative concentration pathway RCP 4.5 emission scenario and the CSIRO-Mk2 General Circulation Model (GCM). This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding <i>D. variabilis</i> is depicted in different colors.	36
Figure 16 : Potential distribution of <i>Dermacentor variabilis</i> under representative concentration pathway RCP 4.5 emission scenario and the MIROC_ESM General Circulation Model	

(GCM). This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding <i>D. variabilis</i> is depicted in different colors.	37
Figure 17 : Potential distribution of <i>Dermacentor variabilis</i> under representative concentration pathway RCP 4.5 emission scenario and the NCAR_CCSM3_0 General Circulation Model (GCM). This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding <i>D. variabilis</i> is depicted in different colors.	38
Figure 18 : Agreement in future spatial distribution prediction between different GCM models under 4.5 representative concentration pathway scenario.	39
Figure 19 : Potential distribution of <i>Dermacentor variabilis</i> under representative concentration pathway RCP 4.5 emission scenario and the Canadian Centre for Climate Modeling and Analysis (CGCM3) General Circulation Model (GCM). This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding <i>D. variabilis</i> is depicted in different colors.	41
Figure 20 : Potential distribution of <i>Dermacentor variabilis</i> under representative concentration pathway RCP 4.5 emission scenario and the CSIRO-Mk2 General Circulation Model (GCM). This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding <i>D. variabilis</i> is depicted in different colors.	42
Figure 21 : Potential distribution of <i>Dermacentor variabilis</i> under representative concentration pathway RCP 4.5 emission scenario and the MIROC_ESM General Circulation Model (GCM). This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding <i>D. variabilis</i> is depicted in different colors.	43
Figure 22 : Potential distribution of <i>Dermacentor variabilis</i> under representative concentration pathway RCP 8.5 emission scenario and the NCAR_CCSM3_0 General Circulation Model (GCM). This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding <i>D. variabilis</i> is depicted in different colors.	44

Figure 23 : Agreement in future spatial distribution prediction between different GCM models under 8.5 representative concentration pathway scenario.	45
---	----

List of Tables

Table 1: List of confirmed and suspected pathogens transmitted by <i>Dermacentor variabilis</i> and the diseases caused.....	47
Table 2 : List of bioclimatic variables.	48
Table 3 : Environmental variable combinations evaluated by jackknife procedure for use in development of <i>Dermacentor variabilis</i> distribution model.....	49
Table 4 : Compilation of three different environmental sets.	50
Table 5 : Estimates of relative contributions of the different Bioclim variables to individual models at the three jackknife steps.	51
Table 6 : Different model parameters and the values of model selection criteria.	52

Acknowledgements

I thank my advisor, Dr. Ram Raghavan who encouraged and guided me through the entire research process, and my committee members Dr. Jianfa Bai, Dr. Yoonseong Park, and Dr. John Harrington for their trust in me and their guidance. I also thank Dr. Town Peterson, Department of Ecology and Evolutionary Biology at University of Kansas, for his guidance and allowing us to run models from his lab. This project was funded partially by Kansas State Veterinary Diagnostic Laboratory (KSVDL) and the Center for Excellence in Vector Borne Diseases (CEVBD).

I am extremely grateful to my family for their emotional support; especially to my dear mother Anuradha and father Vijay, and my in-laws Krishna Murthy, and Padma who were always there to listen and provide encouragement whenever I needed. I thank my husband Mahesh, who helped me every step of the way despite his own busy and rather demanding work. My graduate research has helped reveal that he is a great cook!

Finally, my sincere thanks to our son Shivansh, who was extremely patient with me these past two years, who continues to be amused about the fact that I wanted to attend school at this “late” years.

Chapter 1 - LITERATURE REVIEW AND REASEARCH

OBJECTIVE

Ticks are ectoparasitic arthropods and obligate blood feeders that are known to transmit pathogens to humans, wildlife, and livestock throughout the world (Hopla & Durden 1994). There are over 900 documented species of ticks in the world, and 90 are found in North America (Barker & Murrell, 2004; Jongejan & Uilenberg, 2004). Ticks must have a blood meal after at every post-emergence stage of their life (larva, nymph and adult) for their survival and reproduction. More often, ticks feed on multiple hosts during the course of their life, and during that process ticks act as transmitting vectors for various pathogens that ultimately cause disease in humans and animals (Brites-Neto et al., 2015; Jongejan & Uilenberg, 2004). Ticks have evolved for several millions of years, and as per fossil records, ticks originated in the Crustaceous period, around 65-146 million years ago (Fuente, 2003). The earliest reports documenting ticks and tick-borne fevers occurred in the ancient Egyptian papyrus scrolls (1550 BC) (Hoogstraal, 2011). The discovery of the first fossil tick with a spirochete-like cells (*Borrelia* sp.) in hemocoel and in the lumen of alimentary tract was observed in a larval tick (*Amblyomma* sp.) embedded in Dominican amber (Poinar, 2015). (Smith & Kilborne, 1893) for the first time confirmed that ticks were the vector for pathogens, and they also determined that the tick species, *Rhipicephalus annulatus* (Boophilus) transmits *Babesia bigemina*.

The tick species, *Dermacentor variabilis* (Say), commonly known as the American dog tick or wood tick is one of the many medically important ticks in N. America that has been documented to transmit several pathogens to humans and animals (Brites-Neto et al., 2015). Some of the pathogens and the diseases transmitted by *D. variabilis* ticks are listed in (Table 1).

Dermacentor variabilis is known to carry various bacterial diseases that are responsible for several diseases in humans, including Rocky Mountain spotted fever (RMSF), tularemia, human granulocytic anaplasmosis. Rocky Mountain spotted fever is found east of the Rocky Mountains, and cases in these areas occur year around and are characterized by high incidence and case fatality rates, especially in children (CDC, 2017a). RMSF is caused by *Rickettsia rickettsia*, which infects the endothelial cells that line the blood vessels which leads to clotting in the brain and other organs (CDC, 2017a). This leads to loss of function of the organs and sometimes the organs need to be amputated. RMSF is a serious tick-borne disease and can be deadly if not treated early. The common sign of the disease is rash which are usually red splotches on the skin. Early treatment with antibiotic doxycycline can prevent death (CDC, 2017a).

Tularemia, anaplasmosis in cattle and canine tick paralysis are three other diseases transmitted by *D. variabilis* in N. America. Tularemia is often a lethal disease to humans and feline hosts caused by *Francisella tularensis* (CDC, 2017b). Rodents, hares, rabbits are maintenance hosts for this pathogen and along with *D. variabilis* the pathogen maintains its transmission cycle in the environment. Tularemia is particularly of a concern to the military, since soldiers disproportionately get exposed to ticks in the field and are often treated for this disease (CDC,2017b). The losses to cattle industry due to bovine anaplasmosis, caused by *Anaplasma marginale* is expected to be in several millions of dollars in the US and Canada alone (Kocan et al., 2010). This disease is transmitted by *D. variabilis* in N. America, and other *Dermacentor* ticks in other parts of the world. Once cattle are infected, they remain persistent

carriers of the disease, leading to perpetuation of the disease among herds *via* other ticks and fomites (Kocan et al., 2010).

Canine Tick paralysis is an uncommon, noninfectious, neurologic syndrome that occurs due to the feeding of *D. variabilis* ticks on dogs (Chan and Kaufman, 2008). The release of undefined salivary neurotoxin during the bite of the tick leads to paralytic symptoms and loss of reflex action in a canine host. Treatment options for this disease is limited and if the disease not resolving in earlier stages will lead to death (Chan and Kaufman, 2008).

Despite the name, *D. variabilis* ticks have been found to be attached to a wide variety of hosts other than dogs, and they are encountered in areas other than woodlands as well, particularly grassy fields and scrublands (Sonenshine, 2018). They feed on a variety of mammalian hosts of different sizes ranging from mice to deer. The adult female and males of *D. variabilis* are brown to reddish brown in appearance and have silver-grey marking on their back (scutum), and they are active from March through August, and are mostly found questing for a host on grass and other vegetation (Chan and Kaufman, 2008). The adult females are about 5 mm long and are slightly longer than the males, which are about 3-4 mm long. After blood feeding, the females can enlarge up to 15 mm long and 10 mm wide. The immature stage nymphs are about 1 mm long, pale yellow to brown in color and gray to black after blood-feeding. The larvae are about 0.6 mm long, pale yellow in color black after blood feeding. The larvae only have 3 pairs of legs unlike the nymphs and adults and are active late in the tick season (July, August) (Chan and Kaufman, 2008).

Humans are not the preferred hosts for *D. variabilis* but the adult ticks commonly attack humans, and the nymphs and larvae can also be occasionally found on humans (Chan and Kaufman, 2008). The *D. variabilis* ticks are 3-host ticks, meaning that the three life stages (larva,

nymph and adults) feed on separate hosts. Male ticks blood-feed for a brief time and mate with a female on the same host while she is blood-feeding. The female ticks detach from the host after feeding in about a week and fall to the ground and start egg laying. A female tick can lay up to 4000 – 6500 eggs on leaf litter before dying. Under field conditions, the eggs usually hatch in 36-57 days. The larva that hatch from eggs feed on small mammals (e.g., squirrel, mice) for 3-4 days and detach, falling to the ground where they molt into nymphs. The nymphs attach to medium size mammal (raccoons, skunks), where they feed for 4-5 days and detach. Not all ticks manage to find a host and unfed ticks in each stage can survive up to 2 years by overwintering during the winter months (Chan and Kaufman, 2008).

Dermacentor variabilis ticks are widely distributed throughout the central and eastern US, and in southern Canada, east of Saskatchewan. The current distribution of this species also includes areas in California west of the Cascade and Sierra Nevada Mountain ranges (CDC,2018c). The abundance of these ticks within its distribution range varies from region to region, with the eastern US more abundantly followed by southern and central US. The distribution map for *D. variabilis* commonly used by entomologists and public health professionals is produced by the CDC (CDC, 2018c), which was constructed based on several acarological surveys conducted by independent researchers over the years. However, the methodologies used behind construction of these maps, the current relevance of the acarological surveys based on which these maps were produced, and the time-period for which they are relevant are not readily available. The CDC warns its users that these maps are only to be used for reference purposes.

The geographic distribution of all living species change as a function of time, and such changes are largely regulated by a species' innate mechanisms and environmental and

anthropogenic influences. Ticks and other arthropods have expanded or shifted their spatial distribution patterns constantly, and to know their current distribution has scientific merit. In the case of medically important tick species such as *D. variabilis*, knowledge of its current distribution and the difference in its abundance levels across a geographic area will help in making informed public health messaging, appropriate resource allocation for managing disease(s) transmitted by this species, and for making informed future research plans.

There are several approaches developed for modeling the spatial distribution of species and for estimating their ecological niches (Soberón, 2007; Soberón & Nakamura, 2009). Some of the modeling approaches include climate envelopes, logistic regression, multivariate regression splines, and boosted regression trees. Much of these methods involve the utility of occurrence information in the form of geographic latitude and longitude coordinates. And most if not all the afore mentioned methods require absence data (where a species was searched for and not found) for modeling. Absence data are difficult to obtain and may also be unreliable (Elith et al., 2006; Phillips et al., 2006). Other approaches such as the genetic algorithm for rule-set production (GARP) and maximum entropy (MaxEnt) do not require species absence data and have been used to model ecological niches and estimate potential distributions of a wide variety of species. The inclusion of absence data, however, yields better information about prevalence than presence-only methods (Elith & Leathwick, 2009) such information is incorporated in GARP and MaxEnt methods as background or pseudo-absence data (Phillips & Dudík, 2008) (Stockwell, 1999). The quality of predictions based on different modeling approaches and their interpretation have been discussed previously (Phillips et al., 2004; Peterson et al., 2007) Some studies have shown MaxEnt to produce consistently robust species distribution estimates among

presence-only methods (Elith & Leathwick, 2009; Fera-Arroyo et al., 2014; Tsoar Asaf et al., 2007).

The spatial distribution of most arthropods, including *D. variabilis* ticks, are limited for the most part by climatic conditions and physical environment, such as landscape cover and landscape structure. Other influential factors that limit species' distributions include ecological forces such as predator availability and density, competition, and host abundance, which are difficult to incorporate in correlative models (Soberón & Nakamura, 2009; Thuiller et al., 2006). In this study, I modeled the spatial distribution of *D. variabilis* in N. America using the maximum entropy approach under the present-day climate (current) and future distribution based on climate change conditions.

Chapter 2 - MATERIALS AND METHODS

Occurrence data:

The Walter Reed Bio-systematics Unit (WRBU) is a unique national resource for systematics research on medically important arthropods and maintenance of the U.S. mosquito collection (WRBU, 2018). This collection also includes tick specimen records for N. America and other parts of the world. I requested occurrence data for *D. variabilis* in N. America from WRBU and the data was later provided in an Excel format. Quality analysis was conducted on this data in sequential steps. First, I checked for the presence of wrong species, misspelt species name, records that lacked specific location information, and records that did not include an error indicating accuracy of occurrence data. Any rows of information with inadequacies, viz., misspelt species names, wrong species, and with no location information and error measurement were removed.

Following the data quality analysis, the tabular data (Excel) was georeferenced in ArcGIS using the Georeferencing Tool, thereby converting it to a geographic dataset or an “ESRI shapefile”. Verbal descriptions of locations were originally converted by WRBU personnel to *X* and *Y* geographic coordinates using standard georeferencing procedures such as Biogeomancer and MANIS protocol (Mammal Networked Information System) (WRBU, 2018). The *X* and *Y* coordinates provided by WRBU were used to map point locations on N. America. It is not uncommon for museum records such as those available from WRBU to have spatial bias, i.e., some geographic areas could be more represented than others. In order to avoid spatial bias in the occurrence data, first, I removed more than one occurrence at the same location, then I followed a procedure called as rarefaction, to remove spatial autocorrelation in the dataset. For

this, I used the SDM extension for the Spatial Analyst Toolbox in ArcGIS. In the rarefaction step, and occurrence points that fall within 10 km from each other in a random order.

Estimation of accessible area (M):

For a given species of interest, the parts of the world that have been accessible to it via dispersal over relevant periods of time is an important factor that determines the region in which one can detect a species (Barve et al., 2011; Soberón & Peterson, 2005). This accessible area is denoted by M , and it provides an ideal area for model development, testing and comparison. The importance of estimating M carefully has been discussed (Barve et al., 2011). Most importantly, if M is not represented entirely in a spatial distribution model, the importance of coarse resolution factors such as climate may be underestimated, resulting in models that are largely influenced by the effect of spatial extent alone. The effects of M on model training, model validation and model comparison are discussed thoroughly in (Barve et al., 2011).

While the importance of choosing M is widely recognized, any robust methods for estimating appropriate M has not been developed thoroughly yet. Early research on this topic has called for more research in this area (Barve et al., 2011). For this study, I selected an M using a subjective evaluation of all the geographic areas that *D. variabilis* ticks would have had access to historically over time. A buffer area of 7.5 degrees surrounding rarefied occurrence points was created in ArcGIS for this purpose.

Environmental data:

I used environmental dataset that are publicly available through the WorldClim project (Fick & Hijmans, 2017) for modeling the current spatial distribution of *D. variabilis* ticks in N. America. The WorldClim data is freely available online (www.worldclim.org) for download, and

it is a commonly used data resource for spatial distribution modeling. WorldClim is a set of global climatic layers in grid format available in a spatial resolution from 30 seconds ($\sim 1\text{km}^2$) to 10 minutes ($\sim 340\text{ km}^2$). Different bioclimatic variables are derived from historic (1970 – 2000) monthly temperature and rainfall values in order to generate biologically meaningful variables. There are 19 bioclimatic variables available, which represent annual trends, seasonality, extreme environmental factors, and precipitation of the wet and dry quarters (3 months) of a year. Individual BioClim data layer represented summary conditions for different bioclimatic variables, and these are presented in (Table 2).

Similar studies investigating tick distributions and their ecological niche space have used BioClim data, and they have found this data to be highly responsive (Peterson & Raghavan, 2017a, 2017b; Raghavan et al., 2016). Once the data was downloaded, all of the BioClim raster layers were clipped to the extent of “M” using the “Extract by Mask” tool in ArcGIS. Subsequently, the clipped data layers, which were in GRID format were converted into ASCII format in ArcGIS using the “GRID to ASCII Conversion Tool”, which is the required format for MaxEnt modeling software.

Future climate:

Data representing future climate conditions in year 2050 were downloaded from the Climate Change, Agriculture and Food Security (CCAFS) - Climate data portal. This portal provides global and regional future high-resolution climate datasets that are used for assessing climate change impacts on various disciplines, including biodiversity and ecosystem services (CCAFS, 2018). The CCAFS portal has rich resources on the science behind climate change, different representative concentration pathways (RCPs), and climate models. The portal also

provide links to extensive information and documentation on methods used for spatial data downscaling, bias correction, and spatial disaggregation.

For this project, I considered two representative concentration pathways, RCP 4.5 and RCP 8.5, that corresponds to the lower and higher greenhouse gas emission scenarios, respectively. Under both scenarios, I considered four General Circulation Models (GCMs) produced by different agencies *viz.*, CSIRO, MIROC, NCAR and CCCMA. Data for these models were downloaded in ASCII Grid format for the North American region at 30 second resolution.

Spatial Distribution Modeling

To determine the spatial distribution of *D. variabilis* ticks under current and future climate change scenarios, I utilized the freely available and widely used MaxEnt software version 3.3.3 (Phillips et al., 2006). This program allows users to input occurrence data in the form of comma separated variable file for occurrence data, and takes environmental data in the form of ASCII datasets. A detailed description of Maxent modeling software and the theory of ecological niche modeling is discussed in the literature e.g. (Phillips et al., 2006). Briefly, after inputting the occurrence and environmental datasets, the program allows users to select various modeling feature options. *viz.* Linear, Quadratic, Product, Threshold, and Hinge, each representing different machine learning algorithms for model construction (Phillips et al., 2006). The feature classes represent mathematical transformation of the different covariates used in the model to allow complex relationships (Elith & Phillips, 2010). Further, the model allows users to specify options for model extrapolating and clamping, and set different regularization multipliers, which also affects model performance. The regularization multiplier is a numerical value that adds constraints to the models, in other words, is a penalty imposed to the model.

Smaller regularization multiplier values are less restrictive, potentially resulting in over prediction, and *vice versa*. It is recommended to evaluate the best potential combination of environmental variables, feature classes and regularization multiplier, to select the most appropriate model (Phillips et al., 2006).

Environmental variable selection

There are 19 bioclimatic variables available from the WorldClim project. Of these, Bioclim layers 8, 9, 18 and 19 were not considered for determining the spatial distribution of *D. variabilis* due to high autocorrelation that exist among these layers (Escobar et al., 2014). The relevance of the remaining bioclimatic layers to explaining the spatial distribution of *D. variabilis* was evaluated using the jackknife procedure in MaxEnt program. The jackknife procedure allows us to determine the importance of a given environmental data to the overall performance of the model by estimating the variable's individual contribution and contribution in the presence other variables. For this, the program was run in sequential steps with the same occurrence data, beginning with all of the bioclimatic layers, and the jackknife option selected in the MaxEnt program. The variable contribution to the model was evaluated using the jackknife plot, and the model's AUC and AIC values, and those variables contributing least to the model were removed in subsequent steps.

Model selection

Several spatial distribution models were constructed with different combinations of three factors, *viz.*, feature classes, regularization multipliers and sets of environmental (BioClim) data layers. The first environmental data set consisted only those BioClim variables selected in the last jackknife step, i.e., the model that had the least number of variables and yet indicated good model performance. Environmental set-2 included all variables retained in the last step, plus

those BioClim variables in the previous step. Likewise, a total of three sets of environmental variables were selected. Additionally, for the modeling, occurrence data was split randomly into 50% ‘calibration’ and 50% ‘evaluation’ data points using a random number generator in Excel, resulting in 90 data points for calibration and 89 points evaluation each. The models were constructed using the calibration data and the evaluation data was used for estimating model performance using omission rate estimates.

There were four feature combinations used for constructing the models. They were; Linear, (Linear + Quadratic), (Linear + Quadratic + Product), and (Linear + Quadratic + Product + Threshold + Hinge). The four environmental layer sets are shown in (Table 3). And, there were 10 regularization multipliers evaluated in the models. They were, 0.1., 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 8.0, 10, 15, and 20. In total, 285 models with unique combinations were run.

Models with all combinations of feature classes, regularization multipliers and environmental sets were constructed in MaxEnt, and there were a total of 120 models. Each model was run using the bootstrap function to create 10 replicate analyses. In order to avoid model over fitting (Burnham & Anderson, 2002) the models were thresholded using a fixed allowable omission error rate of 10%. This trimming allows 10% of records (both occurrence and environmental data) with the lowest suitability and potential errors to be omitted without affecting the results (Alkishe et al., 2017).

After all the models were constructed, three criteria were used to select best fitting models. The model statistics were first extracted into an Excel spreadsheet, where the Sort by Value function was used in sequential steps to select best fitting models. First, the models were ranked based on lowest to highest partial ROC values and only those significant models ($P \leq$

0.05) were kept, second, models with lowest omission rates was picked among the selected, significant models, and third, the top five models that had the lowest AIC value (among the significant and lowest omission rate models) were picked.

Following the selection of the top 5 performing models, the median of median predictions from the individual models was computed in ArcGIS using the “Cell Statistics” tool. In order to visualize uncertainty in the present condition model prediction, the range (difference between maximum and minimum model outputs) was calculated using the “Cell statistics” tool.

Chapter 3 - RESULTS

The occurrence records received from WRBU included a total of 6224 rows, each representing a single location from where one or more *D. variabilis* ticks of any of the three post-emergent life stages was collected. Among these records, the species name was spelt in two different ways; *variabilis* (correct) and *variabilis* (incorrect). We considered this error to indicate spelling mistakes by data-entry personnel at WRBU and that it did not indicate error in the occurrence of *D. variabilis* in those geographic locations. Following the removal of data records without location information, and those that indicated an error radius ≥ 10 km, there were 2,956 occurrence locations available for analysis. And, following the subsequent rarefaction of occurrence locations, 179 X, Y coordinates were deemed fit for ecological niche modeling. This data quality analysis helped reduce problems associated with artificial clustering of occurrence locations related to biases in sampling and reporting. The occurrence locations were equally and randomly divided for calibration and evaluation of the maximum entropy model. These occurrence locations and the accessible area for *D. variabilis*, *M*, are depicted in Fig. 1.

The jackknife procedure helped determine the degree to which different variables contributed to the spatial distribution of *D. variabilis* ticks in N. America. After fitting a global model (with all BioClim variables included) (Fig.2), the jackknife plot and model statistics revealed that variables 6, 11, 7 and 13 were the least contributors. These variables were removed and the model refit (Fig. 3), and the resulting model revealed BioClim variables 4, 10 and 3 contributed the least. A model without these BioClim variables and those removed in the previous steps (Fig. 4) further revealed that BioClim variables 14 and 17 were lowest contributors. A final model was then fit without variables 14 and 17 (Fig. 5). Based on these steps, three different environmental sets were compiled, presented in (Table 4).

Table (5) gives estimates of relative contributions of the different BioClim variables to individual model at the three Jackknife steps (1, 2 and 3). To determine the first estimate, in each iteration of the training algorithm, the increase in regularized gain is added to the contribution of the corresponding variable or subtracted from it if the change to the absolute value of lambda is negative. For the second estimate, for each environmental variable in turn, the values of that variable on training presence and background data are randomly permuted. The model is reevaluated on the permuted data, and the resulting drop in training AUC is shown in (Table 5), normalized to percentages (Phillips et al., 2006).

The top five models ranked based on partial ROC, omission rate and AIC value all included only the Linear feature class and the same set of environmental variables, Set 3, indicating a strong correlative association between the variables in this set and the occurrence data. The models only varied in the regularization multiplier, with the top three models having 2.0, 3.0 and 4.0, and last two models with 0.1 and 0.2. All five models were significant based on the partial ROC estimate (< 0.00), and all models had good omission rates ($OR = 0.11$). The AIC values for the models differed only by very small amounts. The different model parameters and the values of model selection criteria are present in (Table 6).

The spatial distribution of *D. variabilis* based on the point-wise median of the top five models, each in turn representing the point-wise medians of 10 individual replicate models is presented in Fig. 6. The uncertainty in model prediction, derived based on the range (maximum point-wise estimates – minimum point-wise estimates) based on present-day climate conditions is overlaid on the median of top five models, and present in Fig. 7. These images indicate that the current distribution of *D. variabilis* extends much towards the central plains and west of the Rocky Mountain region, and through most of southeastern Canada in the North. This distribution

is similar to the currently estimated spatial distribution provided by the CDC. However, the present model indicates that the distribution of this species is much broader in the southeastern part of the US relative to the Midwest and western US. The model uncertainty does not appear to affect the model for the most part of the contiguous US, but relatively higher uncertainty can be found in west coast and northern areas of Canada, the Florida peninsula, Southern Texas, eastern coast of Mexico, much of the Caribbean and northern parts of S. America (Fig. 7).

The median point-wise prediction of each of the top 5 models is present in Fig. (8 – 12). The percentage contribution (Fig. 13) by different variables to the top five models indicate the variables annual mean temperature (BioClim 1) and precipitation seasonality (BioClim 15) together accounted for more than 74% (median) of the models, the variable, annual precipitation (BioClim 12) contributed 16.3% (median) and the remaining variables, mean diurnal range (mean of monthly (max temp - min temp)) (BioClim 2), maximum temperature of warmest month (BioClim 5), and precipitation of the wettest quarter (BioClim16) contributed about 3.9, 3.4 and 0.1% (median).

The future distribution of *D. variabilis* based on the point-wise medians derived from maxent models using different global circulation models (GCM) under representative concentration pathway (RCP) 4.5 are presented in Figs. (14 – 17). The agreement in spatial distribution of *D. variabilis* predicted by different GCMs are presented in Fig 18. Likewise, Figs. (19 – 22) depict the individual future distribution prediction by different GCMs under RCP 8.5 scenario and Fig. 23 shows the agreement between different RCP 8.5 models. The spatial distribution based on different GCMs for both RCP scenarios indicated further expansion in the spatial distribution of *D. variabilis* and potential changes to the abundance in different areas. Potential expansions based on different GCMs under RCP 4.5 scenario varied, and the spatial

distribution of *D. variabilis* generally expanded outwards from present-day distribution, but all models predicted no loss in distribution of *D. variabilis* compared to the present-day distribution. Results were similar with the same GCMs under RCP 8.5 scenario, but the spatial distribution expanded much further into the Central Plains and Northern Canada.

Chapter 4 - DISCUSSION

Dermacentor variabilis is a medically significant tick that causes several diseases to humans and animals. The spatiotemporal pattern of county-level incidences for some of the diseases transmitted by this tick species viz., tularemia (Raghavan et al., 2013), bovine anaplasmosis (Hanzlicek et al., 2016), and Rocky Mountain spotted fever (RMSF) (Raghavan et al., 2016a) have been shown to have steadily increased over the past decade. However, the current spatial distribution of *D. variabilis* is only limitedly understood, and only preliminary information is currently available on this topic. This study used publicly available occurrence data and a well-studied spatial distribution modeling program to robustly evaluate the current and future distribution of this species under present-day climate (current) and two different future climate change scenarios. While the overarching result of this modeling effort, a depiction of present distribution of *D. variabilis* is similar to the broad guidance-map provided by the CDC (https://www.cdc.gov/ticks/geographic_distribution.html), this study provides more in-depth predictions in terms of relative probability of finding *D. variabilis* across N. America.

The current spatial distribution model (Fig. 6), indicate low or low/medium (0.25 – 0.49) probability of finding *D. variabilis* in much broader areas in the western US, compared to the current map available from CDC. This finding is significant since recently in 2003 and in the years that followed, numerous cases of RMSF were diagnosed among residents of Indian reservations in Arizona, and in general areas in eastern Arizona, and subsequently it was documented that these cases were primarily due to transmission of *R. rickettsii* to these patients by a different tick species known as the Brown Dog Tick, *Rhipicephalus sanguineus* (Drexler et al., 2014). It appears, based on the present study, that *D. variabilis* could also be present in these areas at low abundance levels, therefore may be harder to detect, but it could also be involved in

the etiology of RMSF in this region. Further field studies aimed at the collection of *D. variabilis* in eastern Arizona, in other areas endemic to RMSF, and further molecular analysis to detect *R. rickettsii* among these ticks are therefore warranted.

The uncertainty (Fig. 7) associated with current spatial distribution of *D. variabilis* does not indicate major problems for most of the contiguous US except for small swathes in southern Texas and Florida peninsula. The uncertainties are relatively larger in other geographic areas, which may indicate uncertainties in the prediction of presence or prediction of absence. The occurrence data used in this project did not include data points representative of the environmental conditions in these diverse climatic region, and their inclusion in future studies will help improve robust modeling specific for these regions.

Three of the environmental variables associated with the spatial distribution of *D. variabilis* are temperature related viz., annual mean temperature, mean diurnal range (mean of monthly [max temp - min temp], and maximum temperature of the warmest month. And three variables were related to precipitation, viz., precipitation seasonality, annual precipitation, and precipitation in the wettest quarter. However, the major contributors to the models were annual mean temperature and precipitation seasonality. Even though ticks have an ability to regulate body temperatures by their movement, it is conceivable that temperature extremes can limit their survival. Also, the seasonality and abundance of ticks and precipitation in spring and late fall are likely related to each other, with adequate rainfall ensuring an abundance of intermediate hosts, and increased host-finding ability of ticks. These bioclimatic variables also indirectly point to the types of vegetation and further the wildlife composition that are likely to be shaped by their unique combination.

Ongoing warming of global temperatures will likely influence the ecology and distribution of medically important ticks such as *D. variabilis*. Studies have shown that changes in climatic patterns, primarily the regional increases in temperatures and shifts in precipitation patterns, are altering the structure and function of ecosystems globally (Parmesan & Yohe, 2003). These changes are suspected to facilitate the emergence of new parasites and new disease agents transmitted by ticks (Epstein, 2001; Harvell et al., 2002), and also alter host–parasite relationships (Kutz et al., 2005). Recently, two new tick-borne viral pathogens, Heartland Virus and Bourbon Virus were identified in the Kansas, Missouri region, and it is plausible that the emergence of these viruses were mediated by cascading effects of climate change. However, definite evidence of such a role played by climate change is hard to quantify. Climatic conditions considered typical for the Central Plains region and other regions in the world have been noted to have changed in noticeable ways (Hayhoe et al., 2015; Schoof, 2012), and many such conditions are known to affect tick phenology and their spatial distribution. The diurnal temperature range, a climate-change index, an important contributor to the spatial distribution of *D. variabilis* in the present study has been decreasing steadily since the 1950s, particularly in the midwestern United States but also other regions of N. America (Karl et al., 1993; Karl et al., 1991). Increased atmospheric humidity during spring and summer months over the Northern Plains was noted for roughly the same time period (Schwartz, 1995). Other ixodid ticks (*Ixodes scapularis* and *I. ricinus*) in the northern latitudes have shown shifts in distribution patterns and abundance levels, which have been linked to climate change (Daniel et al., 2003; Descamps, 2013; Leighton et al., 2012).

Both the RCPs representing low (4.5) and high (8.5) emission scenarios under all of the different GCMs considered in this project indicated range expansion of *D. variabilis* in N. America. However, there were differences between the distributions under the two emission scenarios. The phenology and spatial distribution of ticks are highly regulated by abiotic conditions, therefore the range expansion of this species to newer territory under warmer conditions predicted by both these scenarios is expected. The disease burden due to RMSF in the Indian Reservations during 2003 alone was estimated to have cost 13.2 million (CDC,2016). The economic loss due to bovine anaplasmosis, transmitted by *D. variabilis* in the US alone is much larger, and it is in the range of several hundreds of million dollars every year. The predictive maps under these future scenarios likely indicate an exacerbation of this problem and provide us a guidance as to what extent this species may be present in the year 2050, if all other conditions stay the same.

The expansion or shifts in the spatial distribution of *D. variabilis* predicted by the models in this project must be considered in light of several factors (Alkishe et al., 2017). Importantly, ticks spend most of its life on the host and in microclimatic habitats, whose environments are difficult to accurately estimate. Particularly, the cascading effects of climate change at the resolution of microclimatic conditions are at present difficult to ascertain, and therefore models at higher resolution may or may not be effective. Second, even though climate directly regulates arthropod life cycle, other physical factors such as landscape fragmentation (Raghavan et al., 2016a) could play a role in the distribution and abundance of *D. variabilis*. Third, the spatial distribution models are only indicative that new areas predicted to be suitable, and the actual expansion will depend on dispersal mechanisms.

Figure 1 : Occurrence locations, and accessible area, M , considered for spatial distribution modeling of *Dermacentor variabilis* (Say) in N. America. Red dots indicate locations used for model calibration, and blue dots represent locations used for model evaluation.

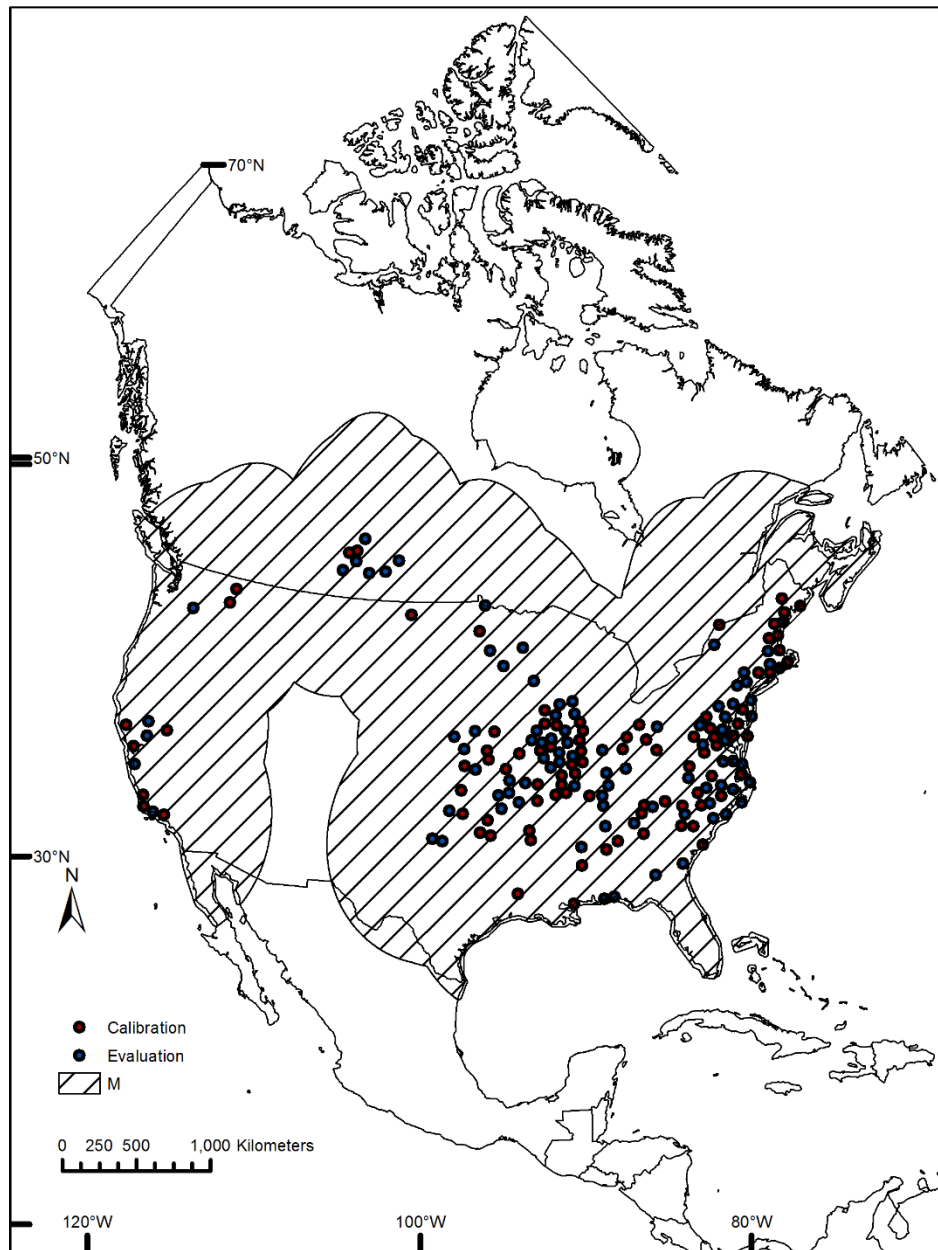


Figure 2 : Jackknife plot representing regularized training gain for *Dermacentor variabilis* spatial distribution with all BioClim variables included in a global model.

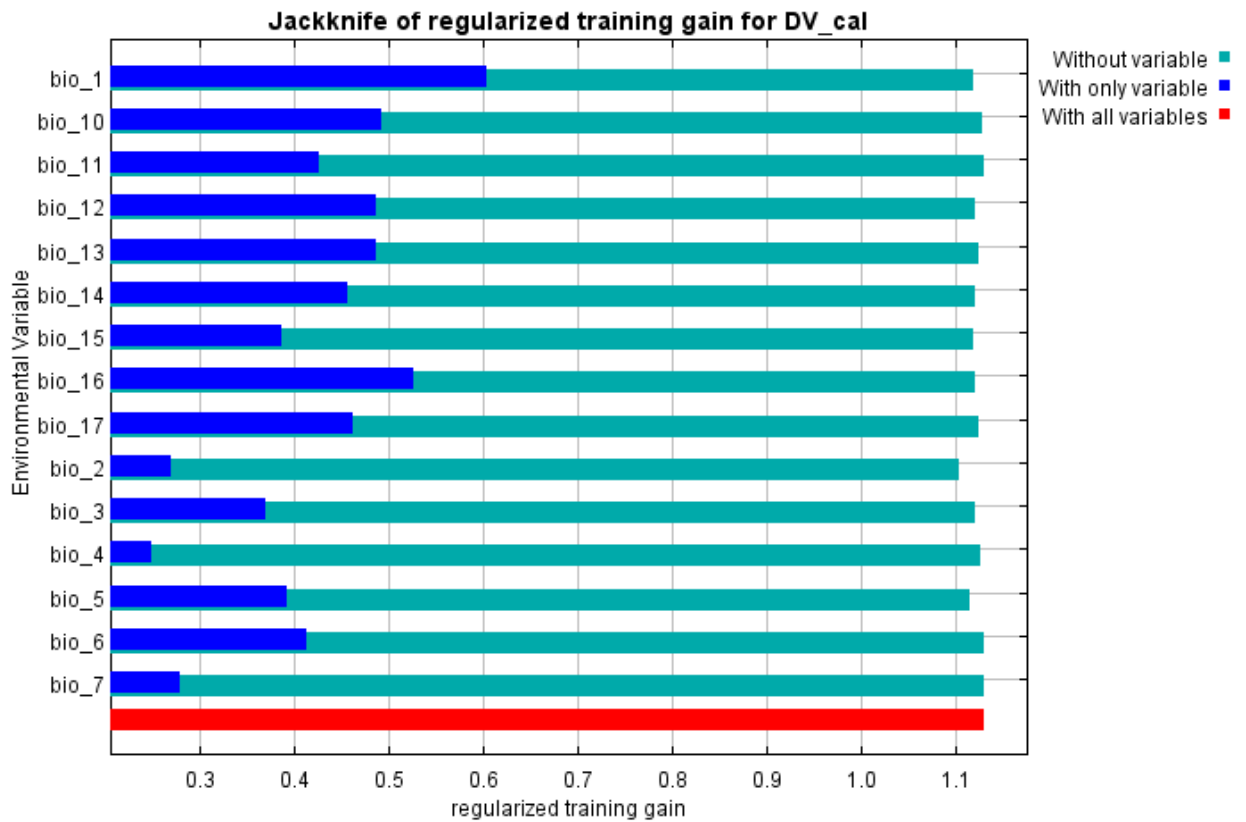


Figure 3 : Jackknife plot representing regularized training gain for *Dermacentor variabilis* spatial distribution with all BioClim variables but without variables 6, 7, 11 and 13.

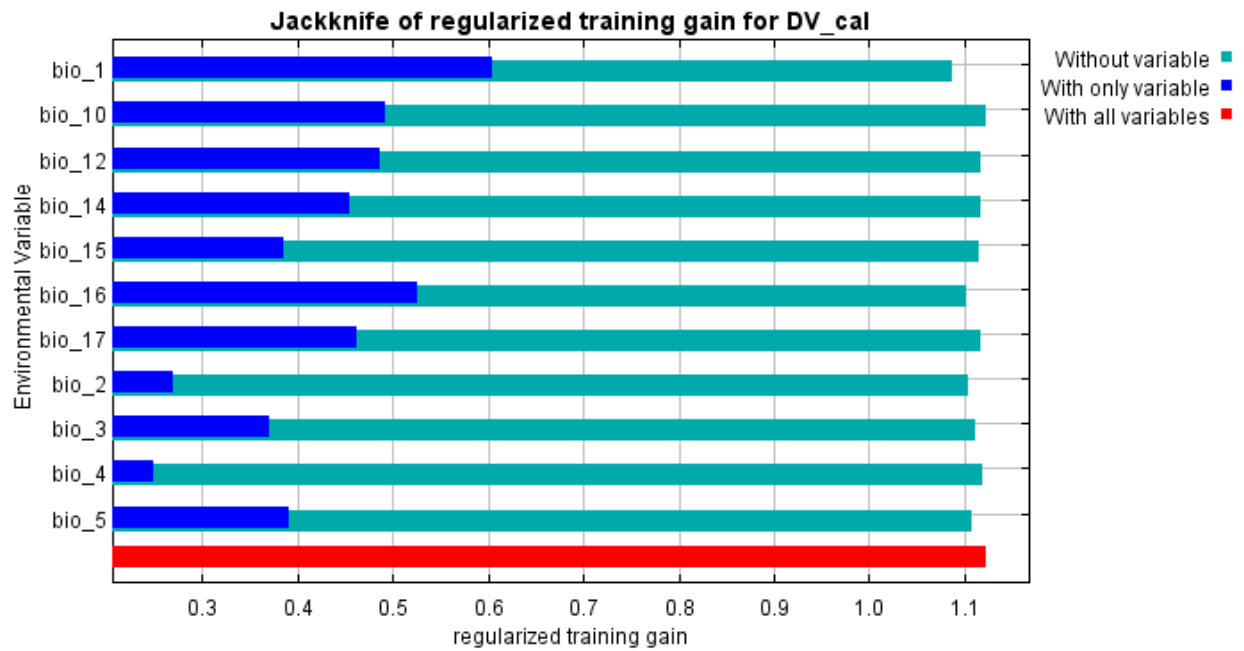


Figure 4 : Jackknife plot representing regularized training gain for *Dermacentor variabilis* spatial distribution with all BioClim variables but without variables 3, 4, 6, 7, 10, 11 and 13.

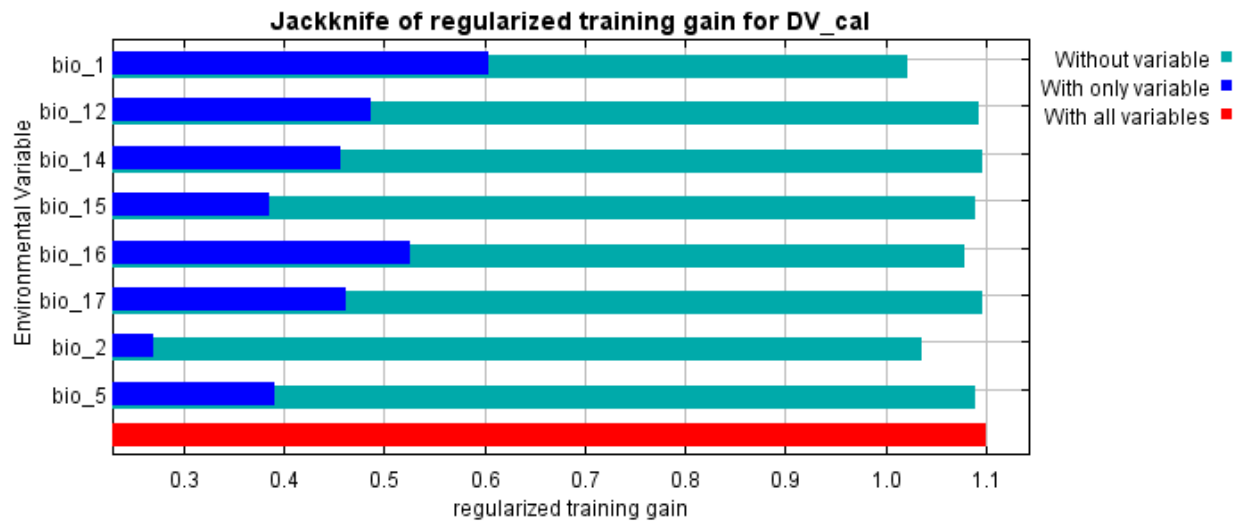


Figure 5 : Jackknife plot representing regularized training gain for *Dermacentor variabilis* spatial distribution with all BioClim variables but without variables 3, 4, 6, 7, 10, 11, 13, 14 and 17.

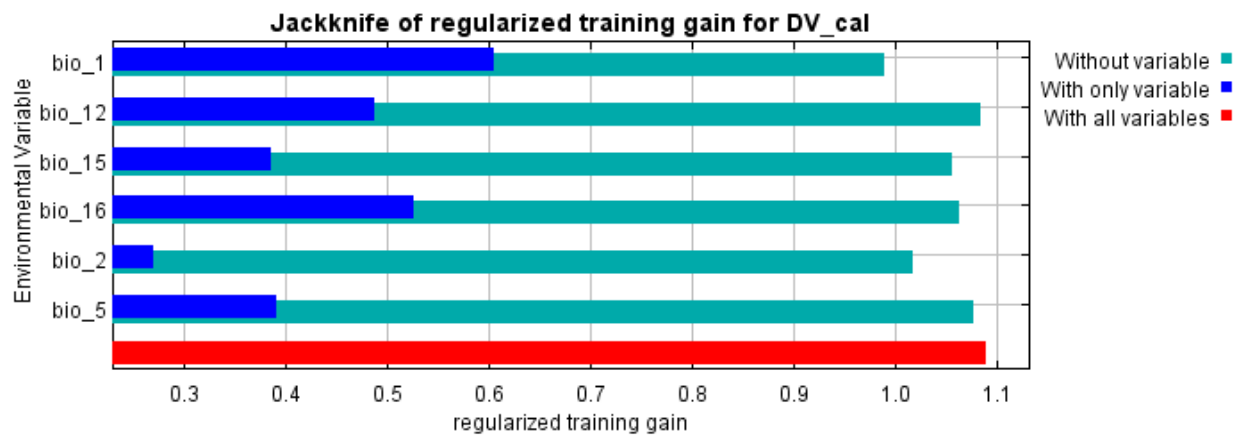


Figure 6 : Current distribution of *Dermacentor variabilis* based on present-day climatic conditions. This picture represents the median of point-wise medians of the top five models. Each of the top five models were derived based on the median of the 10 replicate models applied to the environmental layers. The cell- wise probability of finding *D. variabilis* is depicted in different colors.

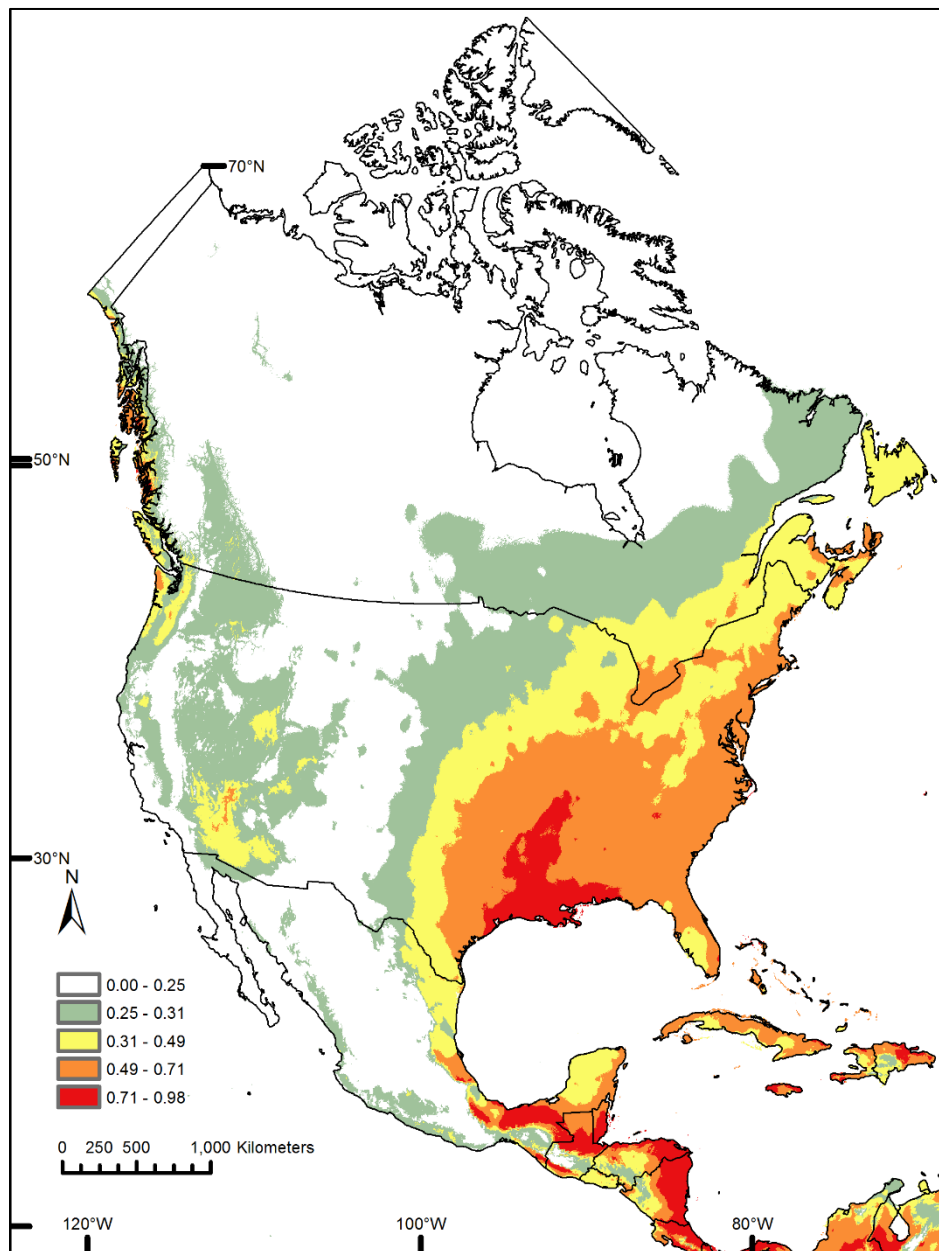


Figure 7 : Uncertainty of current distribution of *Dermacentor variabilis* distribution in N. America. This picture represents the point-wise range (maximum – minimum) of median distribution model based on the top five models. Areas in red indicate current distribution based on present – day model, and different grey colors are uncertainties.

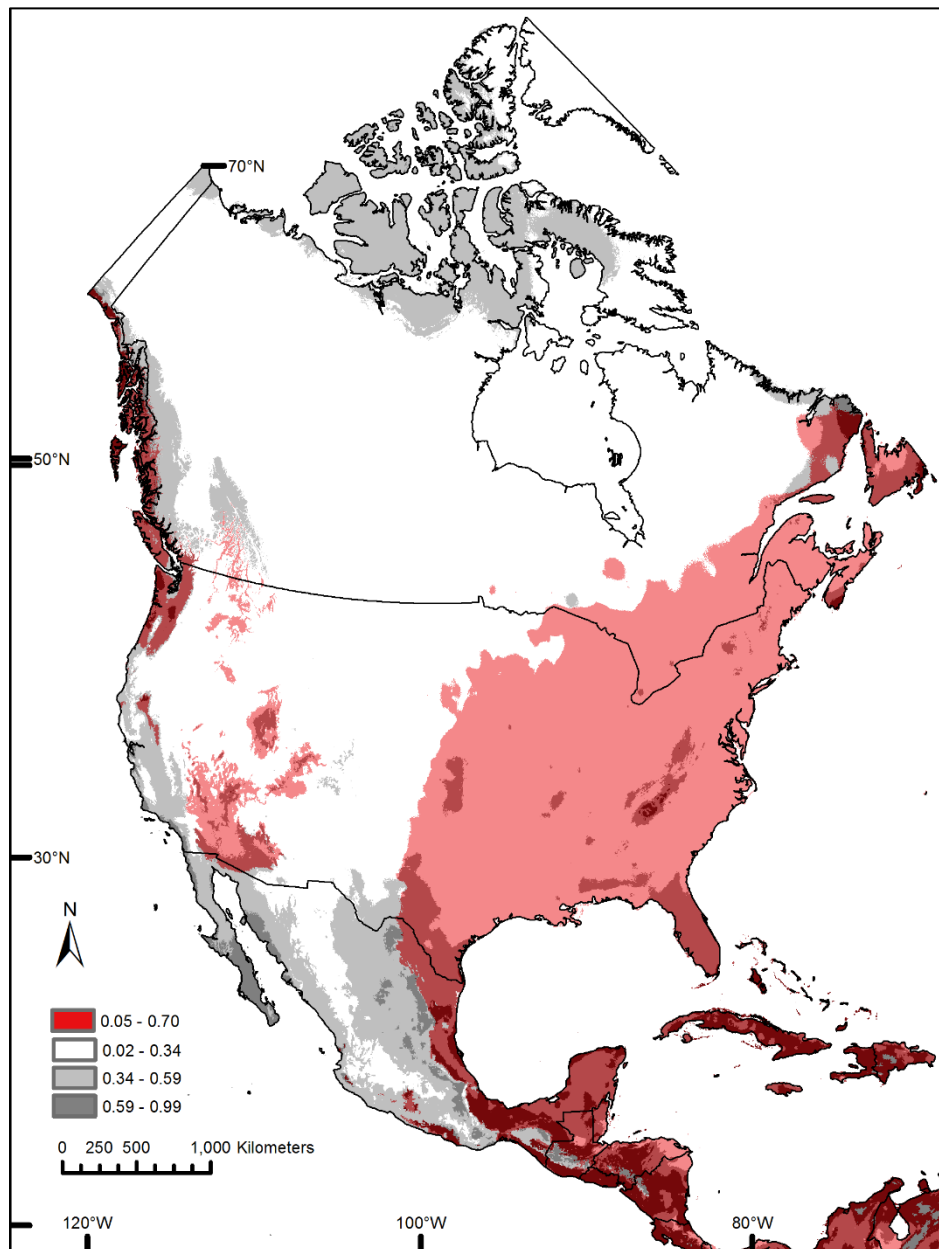


Figure 8 : Current distribution of *Dermacentor variabilis* based on present-day climatic conditions with Linear feature, environmental set 3, and regularization multiplier value set to 2.0. This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding *D. variabilis* is depicted in different colors.

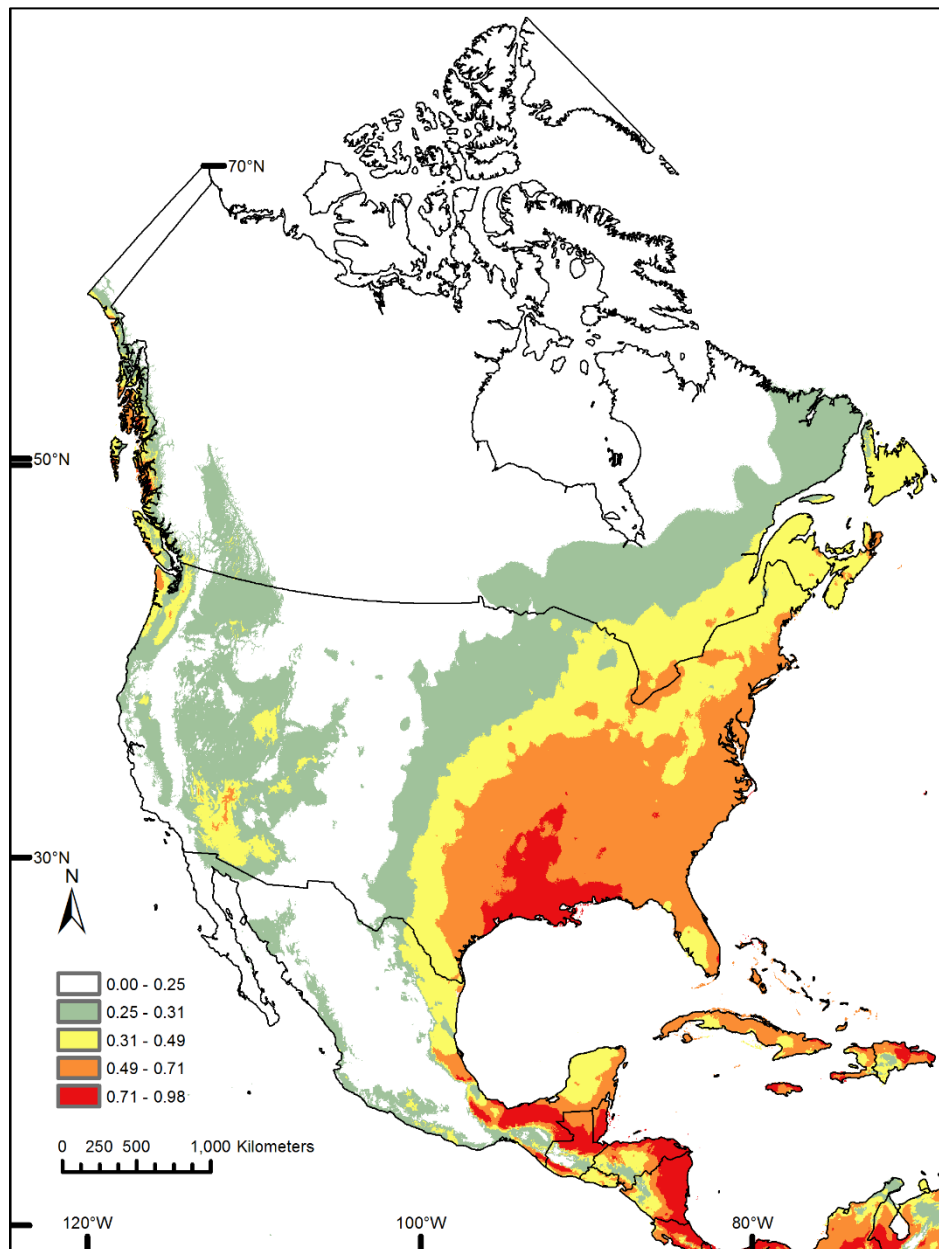


Figure 9 : Current distribution of *Dermacentor variabilis* based on present-day climatic conditions with Linear feature, environmental set 3, and regularization multiplier value set to 3.0. This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding *D. variabilis* is depicted in different colors.

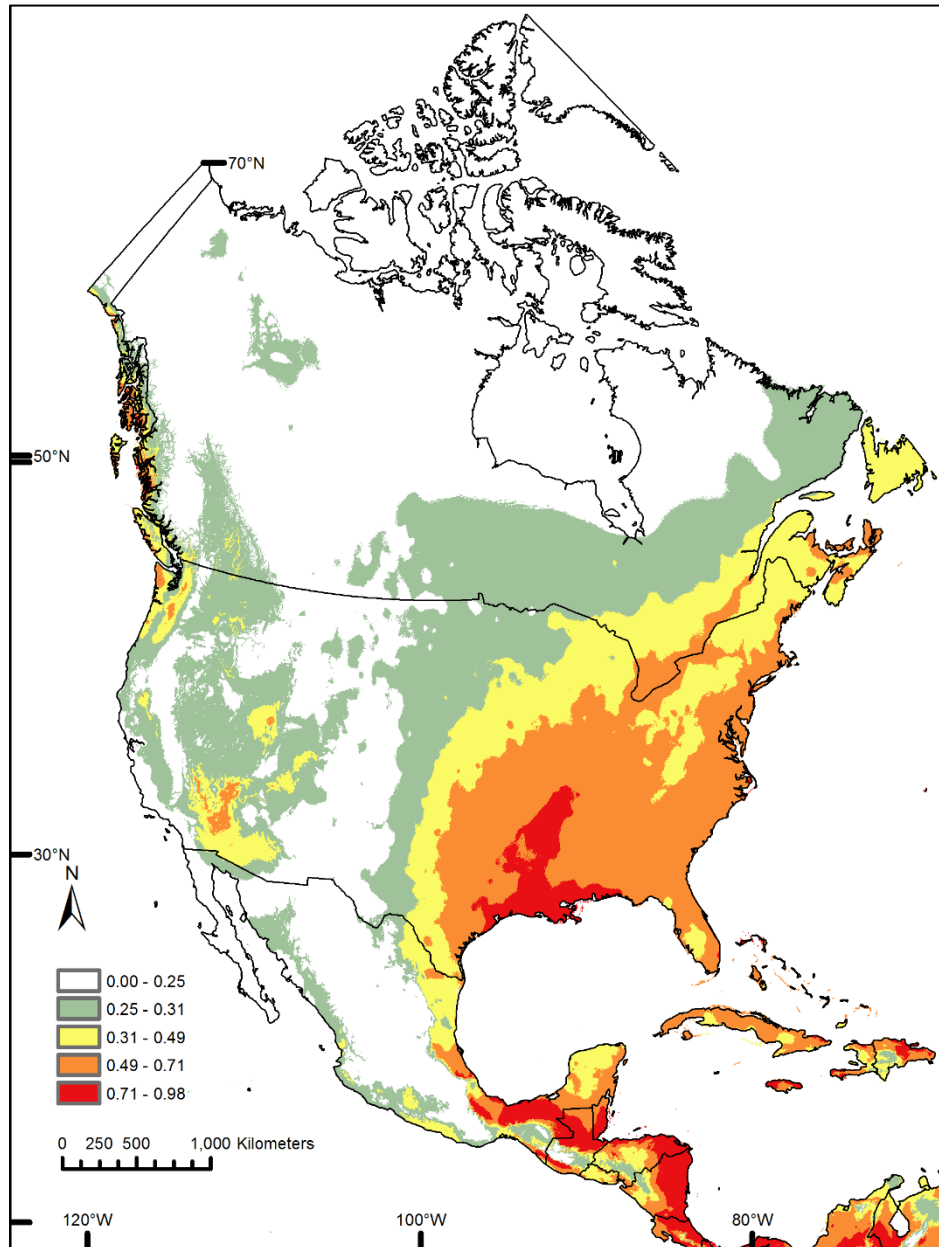


Figure 10 : Current distribution of *Dermacentor variabilis* based on present-day climatic conditions with Linear feature, environmental set 3, and regularization multiplier value set to 4.0. This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding *D. variabilis* is depicted in different colors.

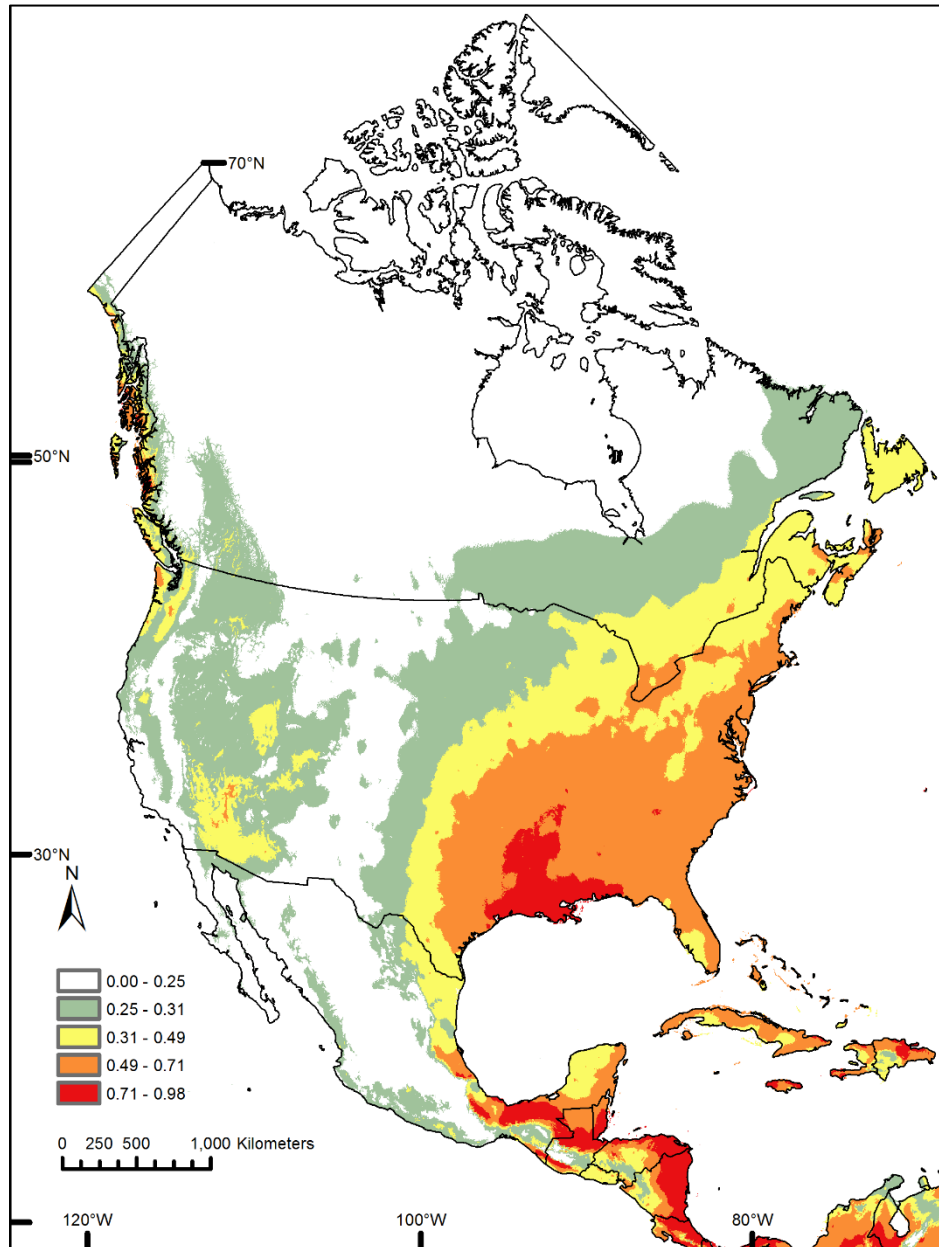


Figure 11 : Current distribution of *Dermacentor variabilis* based on present-day climatic conditions with Linear feature, environmental set 3, and regularization multiplier value set to 0.1. This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding *D. variabilis* is depicted in different colors.

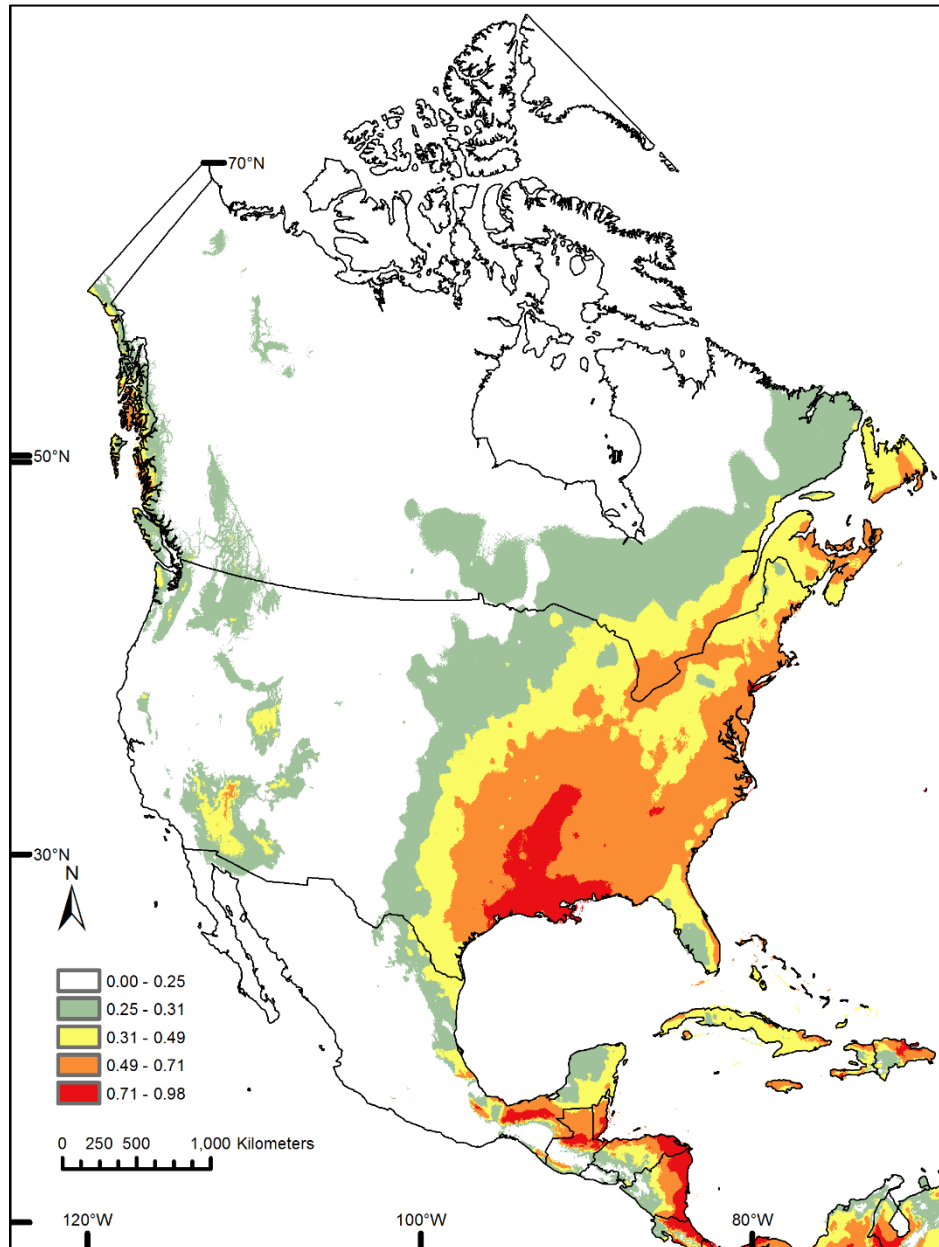


Figure 12 : Current distribution of *Dermacentor variabilis* based on present-day climatic conditions with Linear feature, environmental set 3, and regularization multiplier value set to 0.2. This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding *D. variabilis* is depicted in different colors.

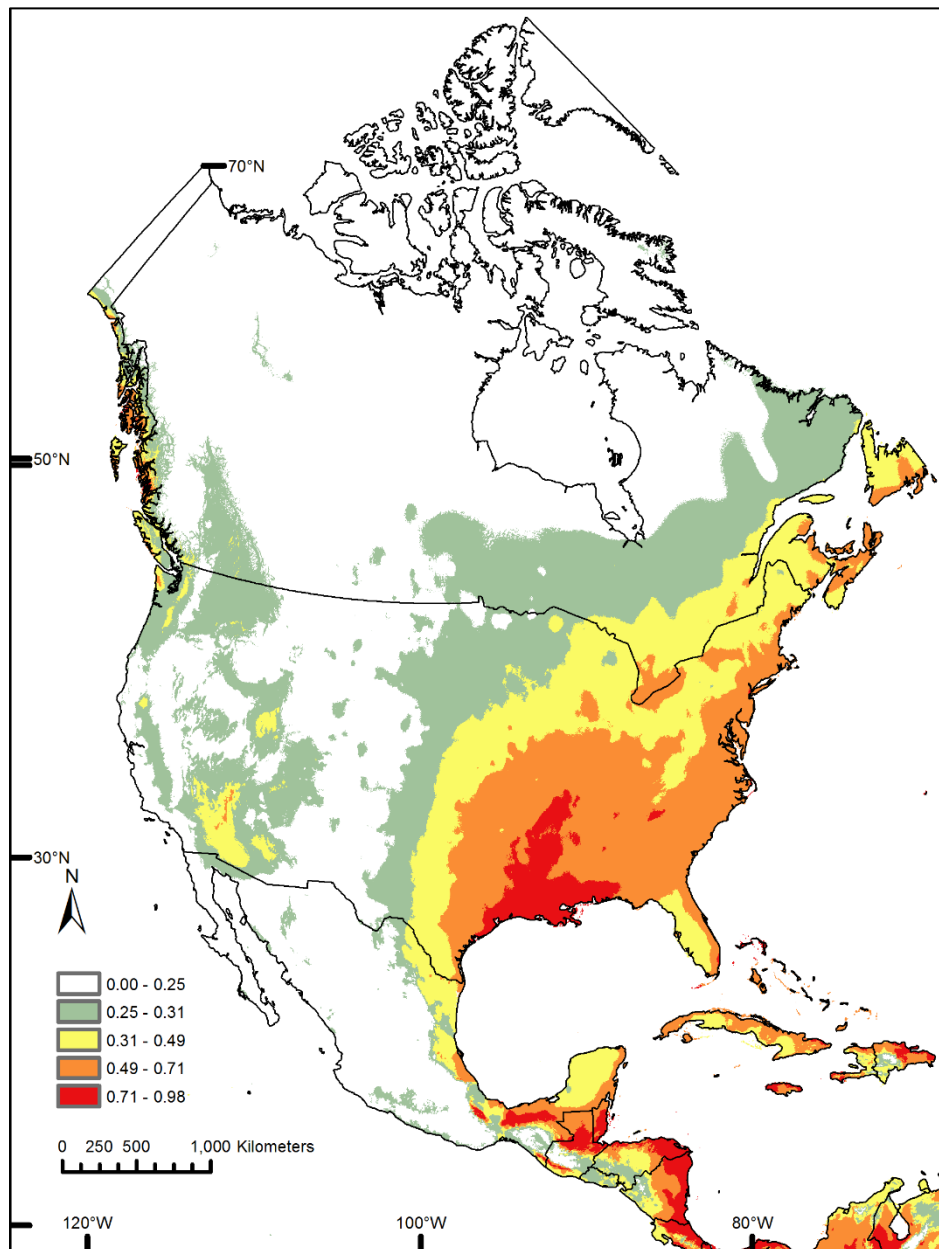


Figure 13 : The variable percentage contribution to the top five individual maximum entropy models for the spatial distribution of *Dermacentor variabilis*

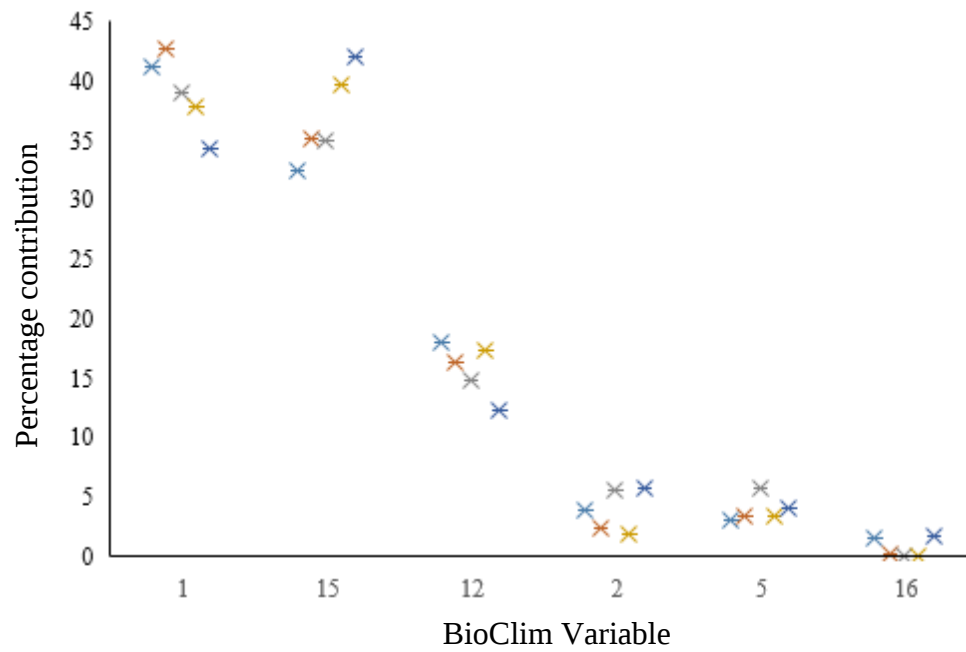


Figure 14 : Potential distribution of *Dermacentor variabilis* under representative concentration pathway RCP 4.5 emission scenario and the Canadian Centre for Climate Modeling and Analysis (CGCM3) General Circulation Model (GCM). This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding *D. variabilis* is depicted in different colors.

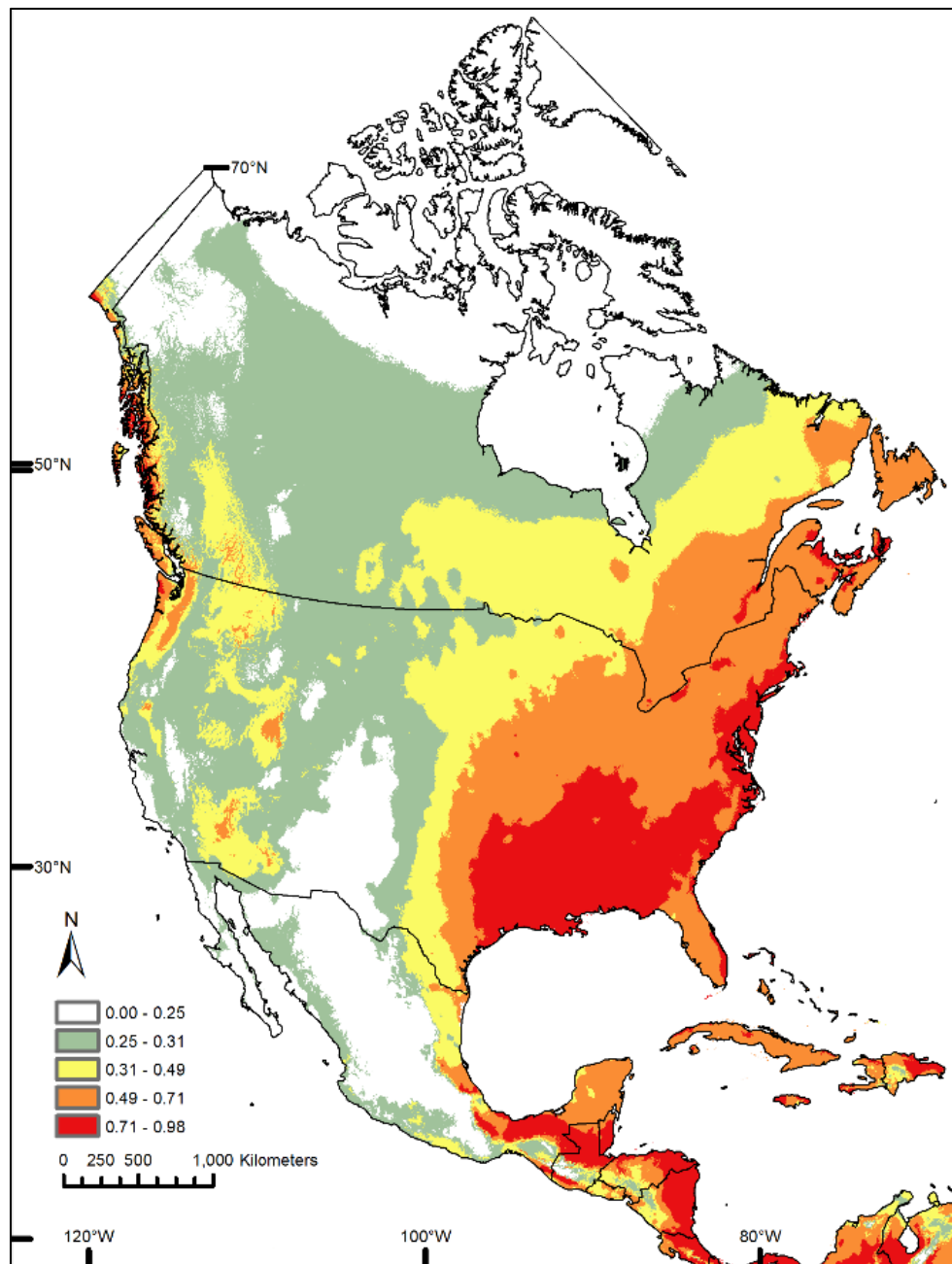


Figure 15 : Potential distribution of *Dermacentor variabilis* under representative concentration pathway RCP 4.5 emission scenario and the CSIRO-Mk2 General Circulation Model (GCM).

This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding *D. variabilis* is depicted in different colors.

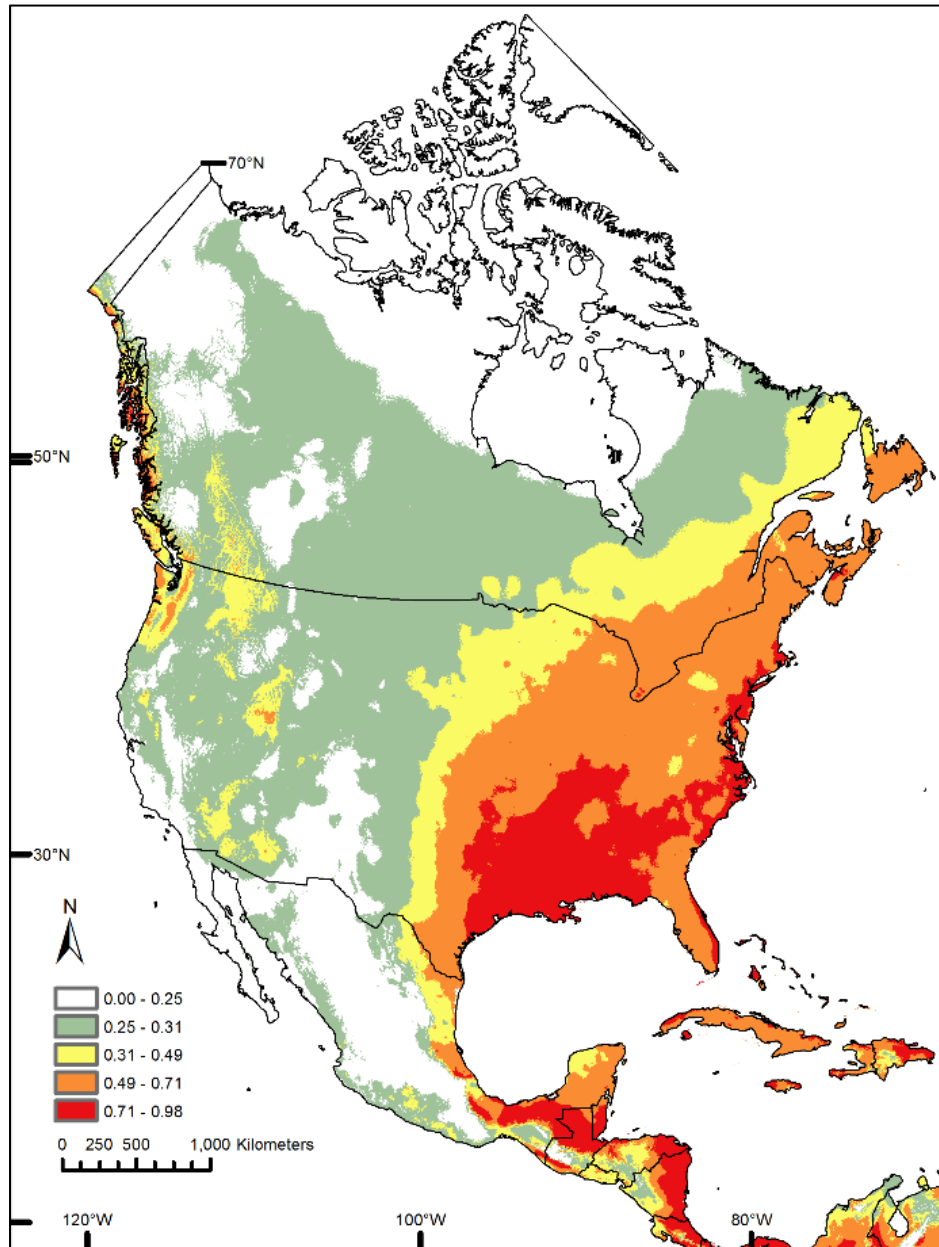


Figure 16 : Potential distribution of *Dermacentor variabilis* under representative concentration pathway RCP 4.5 emission scenario and the MIROC_ESM General Circulation Model (GCM). This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding *D. variabilis* is depicted in different colors.

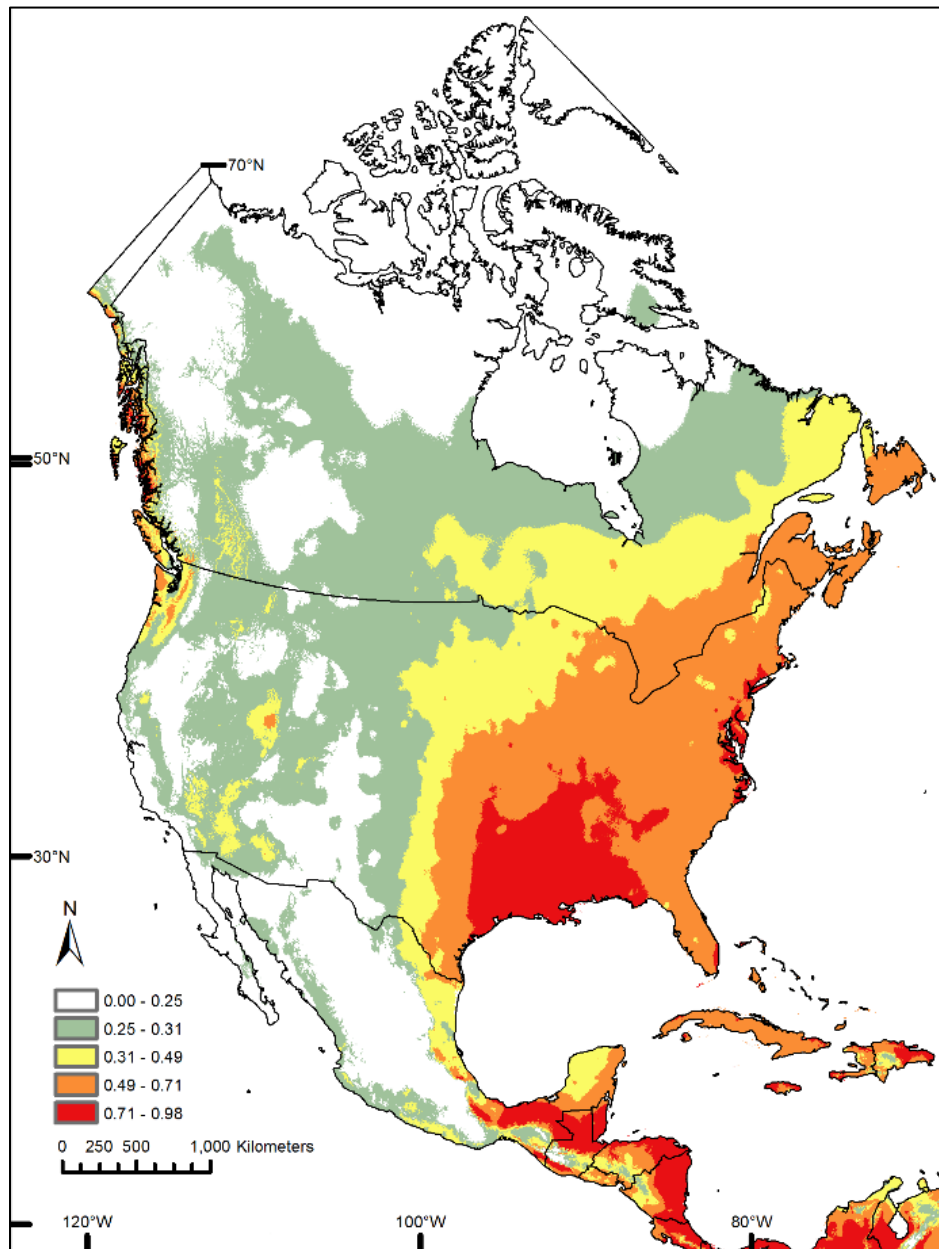


Figure 17 : Potential distribution of *Dermacentor variabilis* under representative concentration pathway RCP 4.5 emission scenario and the NCAR_CCSM3_0 General Circulation Model (GCM). This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding *D. variabilis* is depicted in different colors.

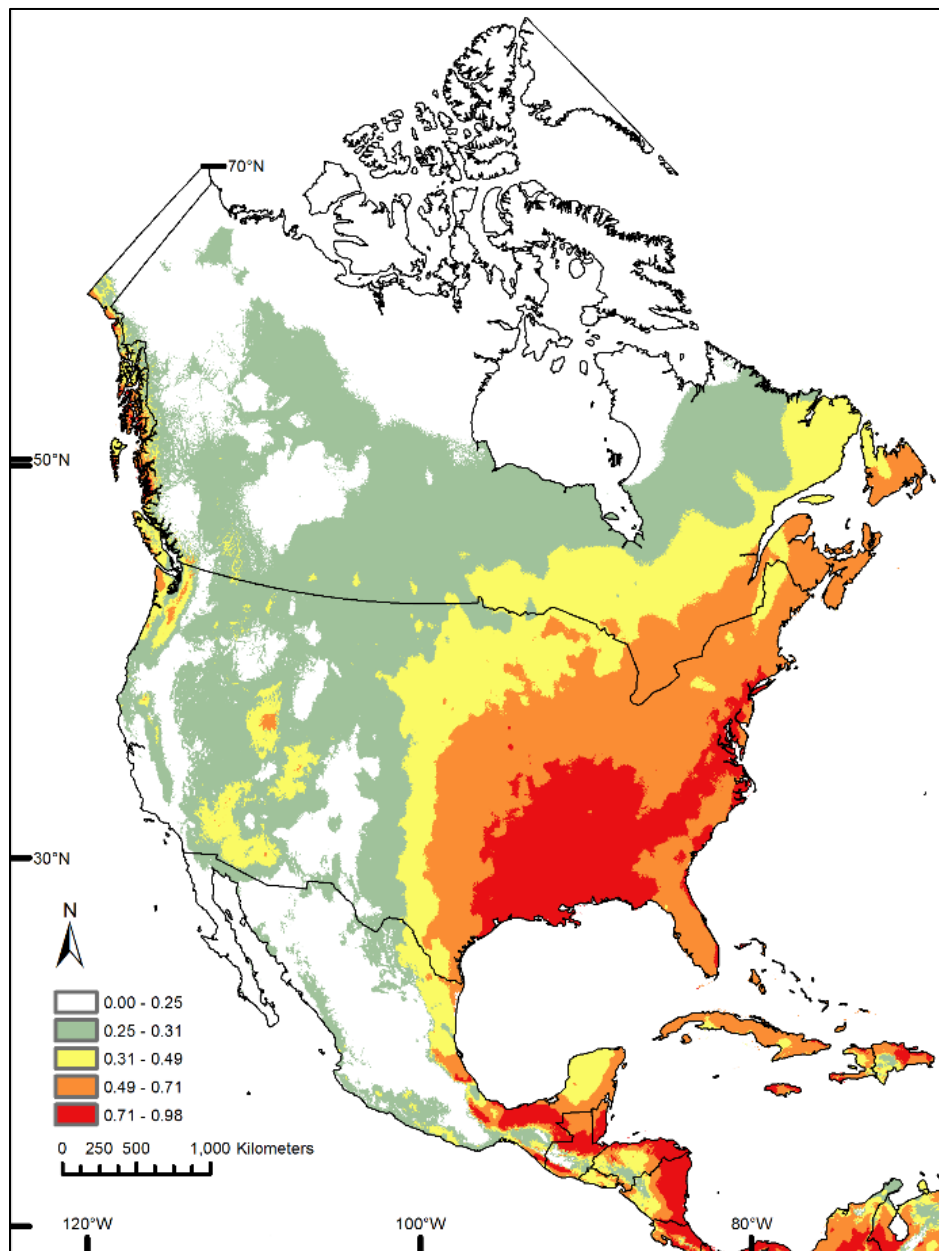


Figure 18 : Agreement in future spatial distribution prediction between different GCM models under 4.5 representative concentration pathway scenario.

(0 = species is not present in current time, and not predicted to be present in any of the future scenario models (no change); 1, 2, 3, and 4 = species not present in current time, and it is predicted to be present in at least one, two, three or of the four future scenario models, respectively (expansion); 5 = species is present in current time, and it is not predicted to be present in any of the future scenario models (entire loss of territory); 6, 7, 8 = species is present in current time, and not present in one, two, and three of the future scenario models (gradual loss of territory); 9 = species is present currently, and it is predicted to be present in all future scenario models.

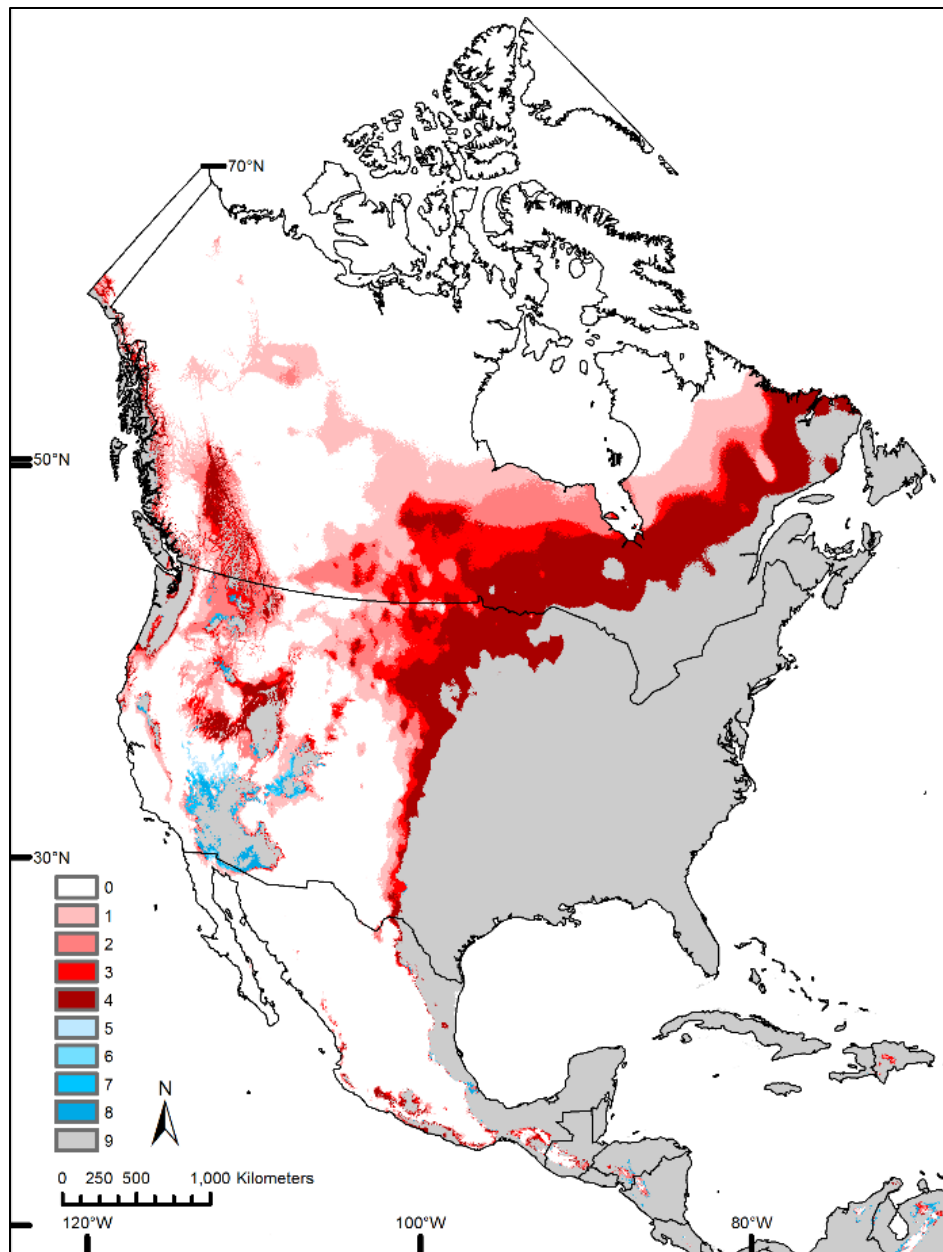


Figure 19 : Potential distribution of *Dermacentor variabilis* under representative concentration pathway RCP 4.5 emission scenario and the Canadian Centre for Climate Modeling and Analysis (CGCM3) General Circulation Model (GCM). This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding *D. variabilis* is depicted in different colors.

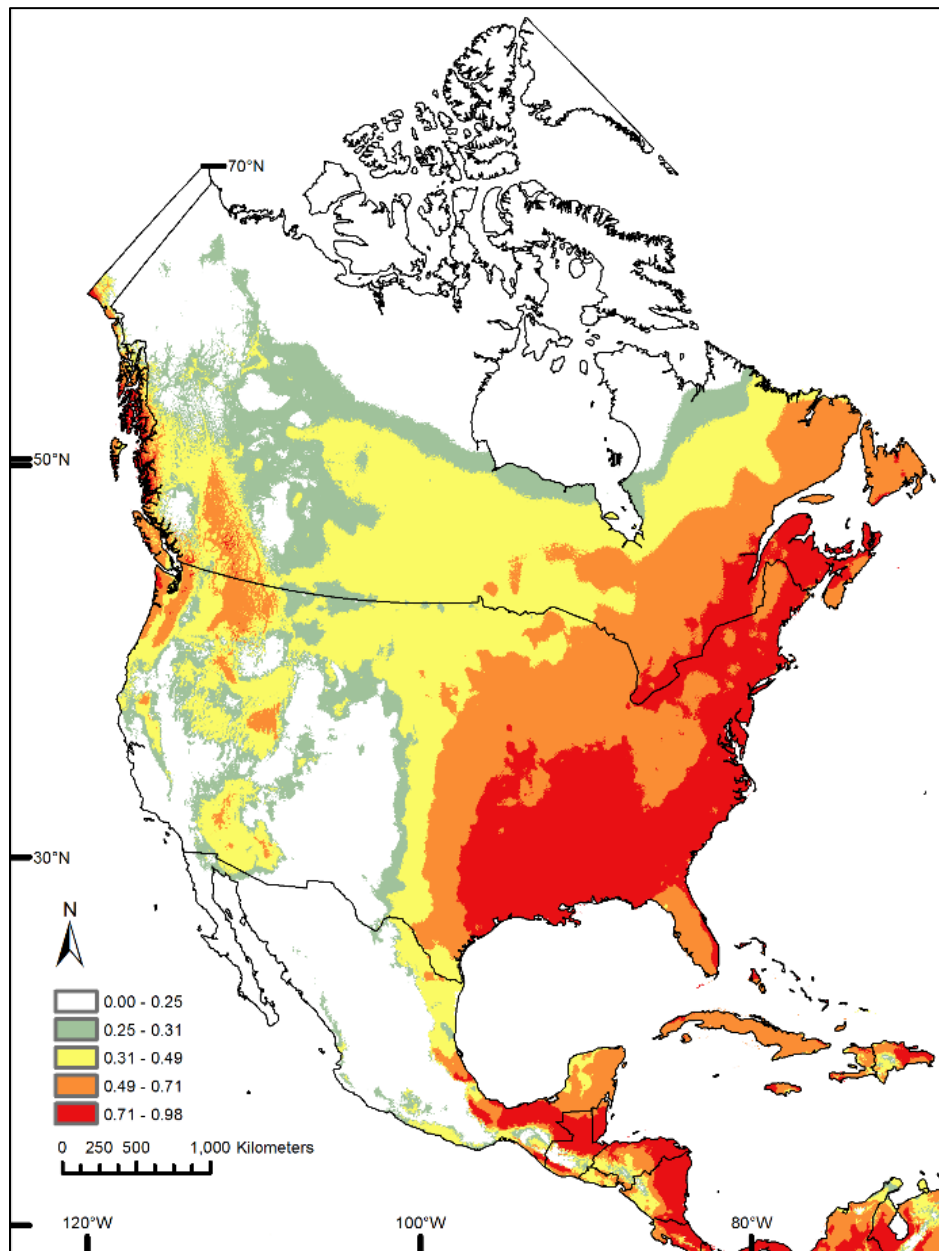


Figure 20 : Potential distribution of *Dermacentor variabilis* under representative concentration pathway RCP 4.5 emission scenario and the CSIRO-Mk2 General Circulation Model (GCM).

This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding *D. variabilis* is depicted in different colors.

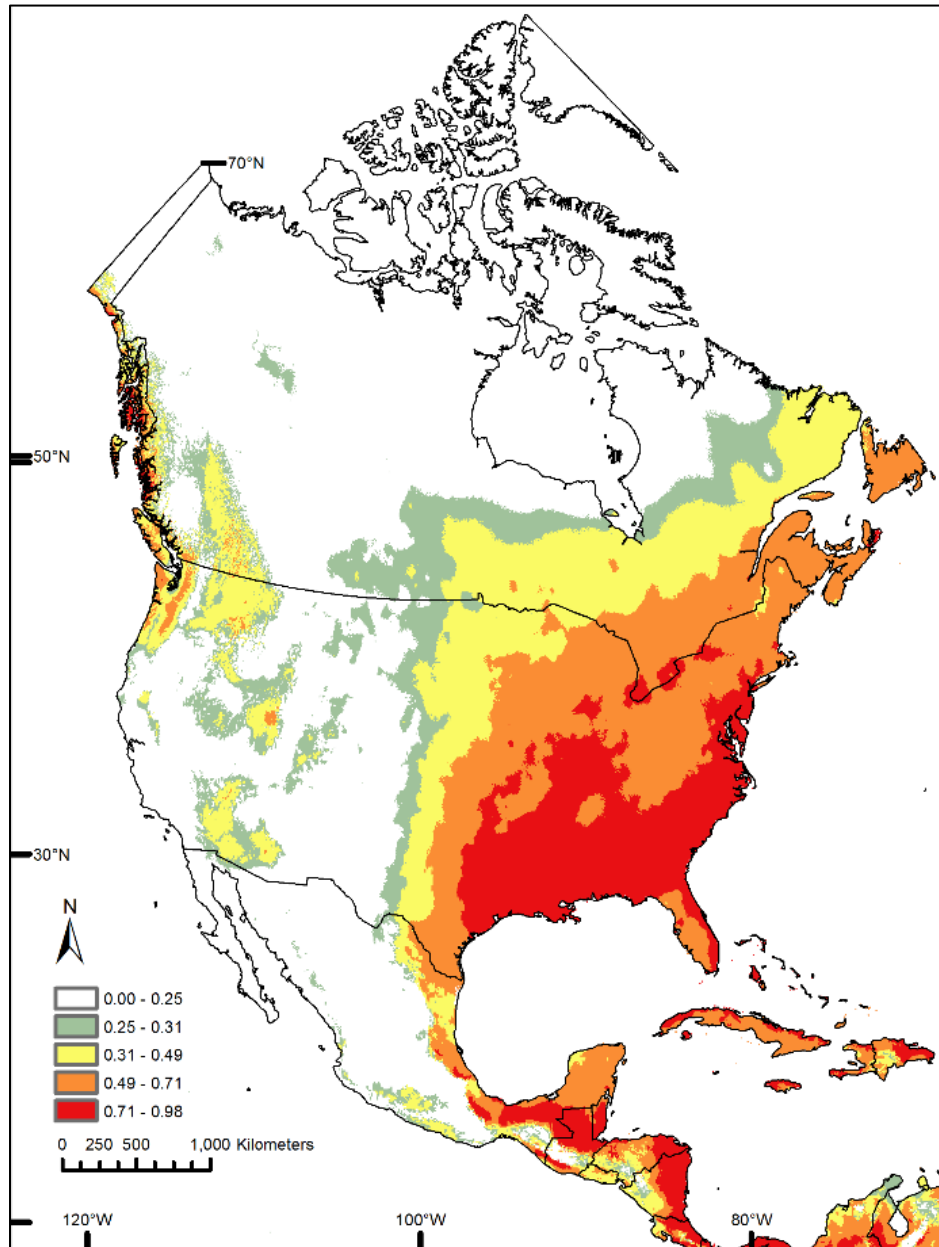


Figure 21 : Potential distribution of *Dermacentor variabilis* under representative concentration pathway RCP 4.5 emission scenario and the MIROC_ESM General Circulation Model (GCM). This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding *D. variabilis* is depicted in different colors.

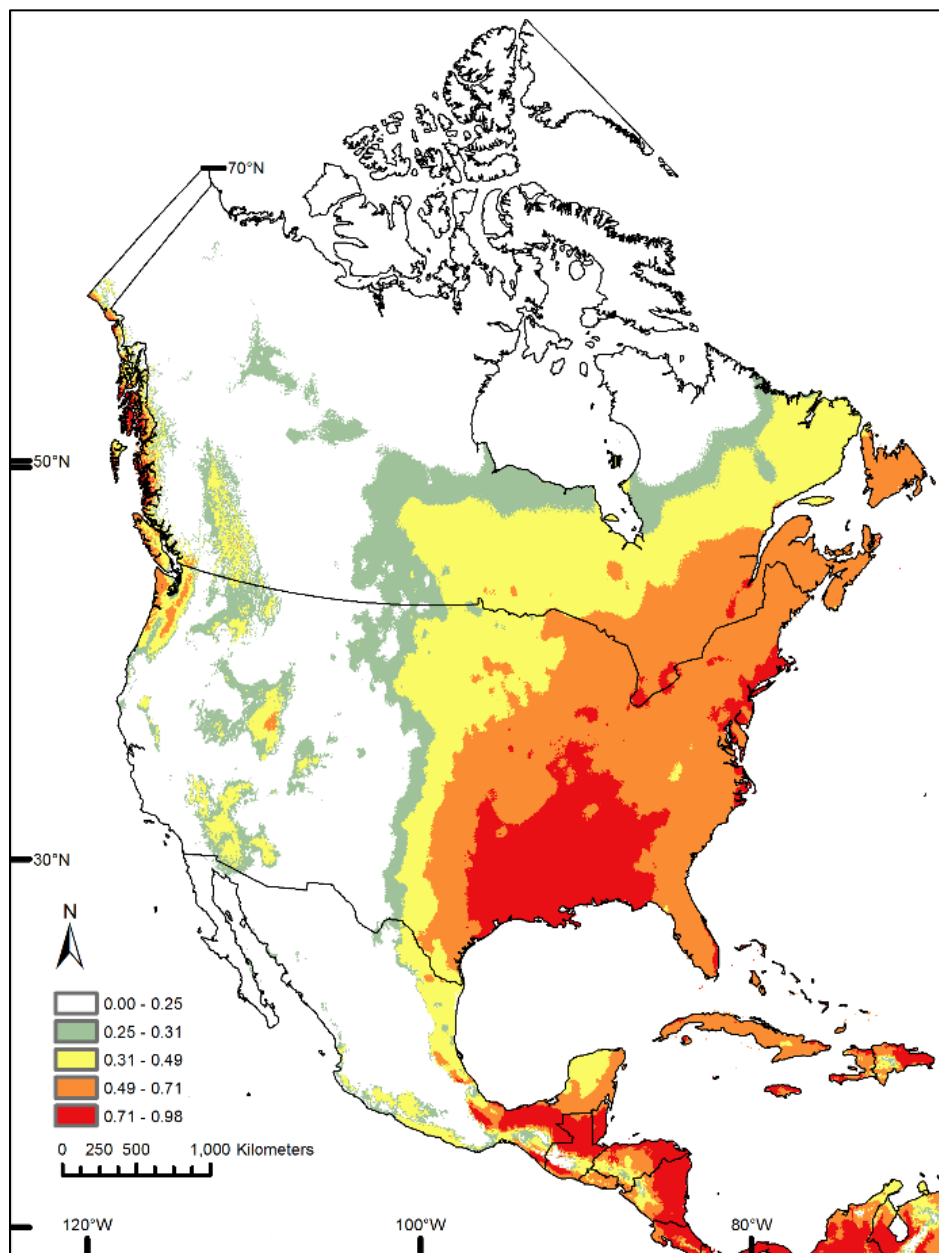


Figure 22 : Potential distribution of *Dermacentor variabilis* under representative concentration pathway RCP 8.5 emission scenario and the NCAR_CCSM3_0 General Circulation Model (GCM). This picture represents the median of the 10 replicate models applied to the environmental layers. The cell-wise probability of finding *D. variabilis* is depicted in different colors.

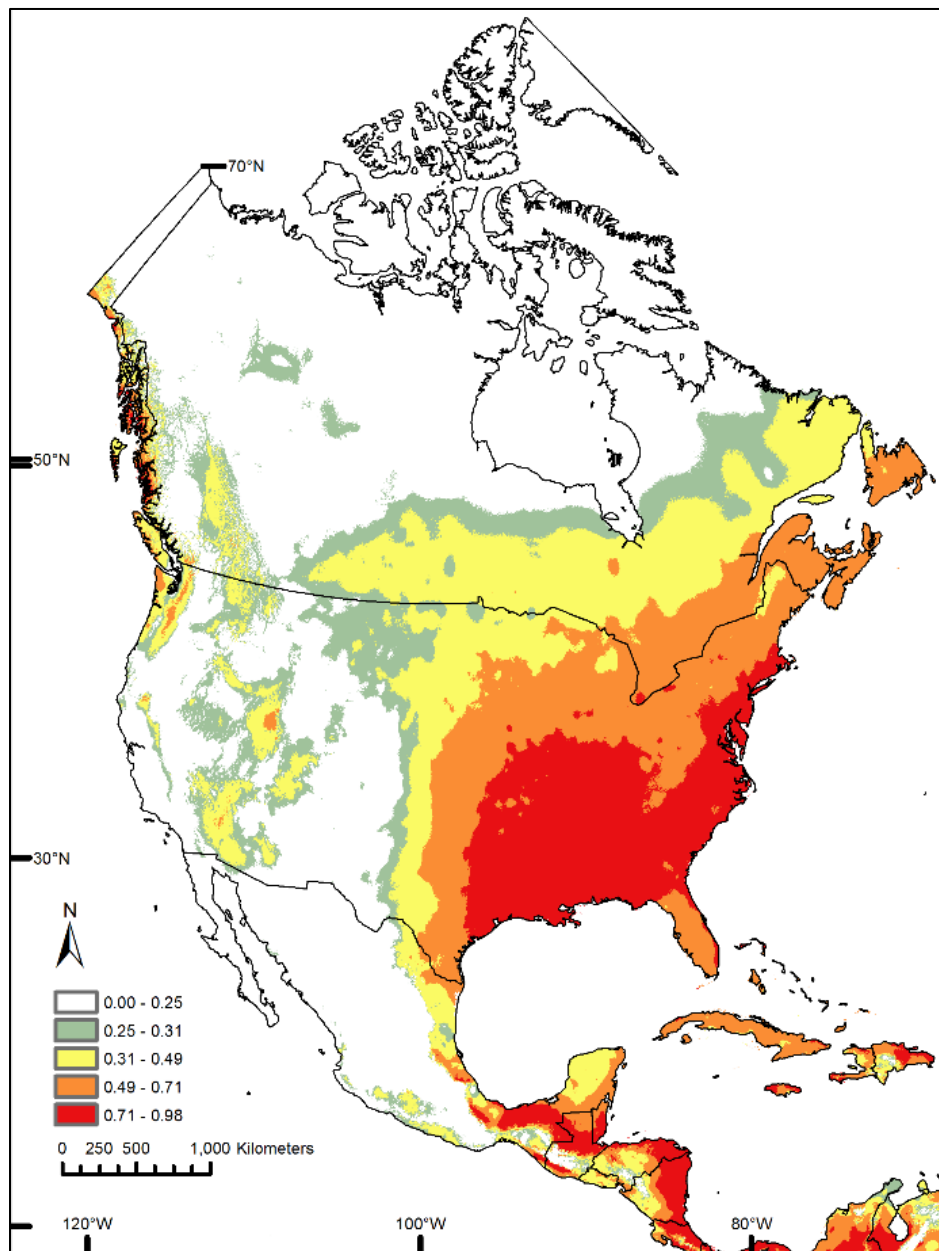


Figure 23 : Agreement in future spatial distribution prediction between different GCM models under 8.5 representative concentration pathway scenario.

(0 = species is not present in current time, and not predicted to be present in any of the future scenario models (no change); 1, 2, 3, and 4 = species not present in current time, and it is predicted to be present in at least one, two, three or of the four future scenario models, respectively (expansion); 5 = species is present in current time, and it is not predicted to be present in any of the future scenario models (entire loss of territory); 6, 7, 8 = species is present in current time, and not present in one, two, and three of the future scenario models (gradual loss of territory); 9 = species is present currently, and it is predicted to be present in all future scenario models.

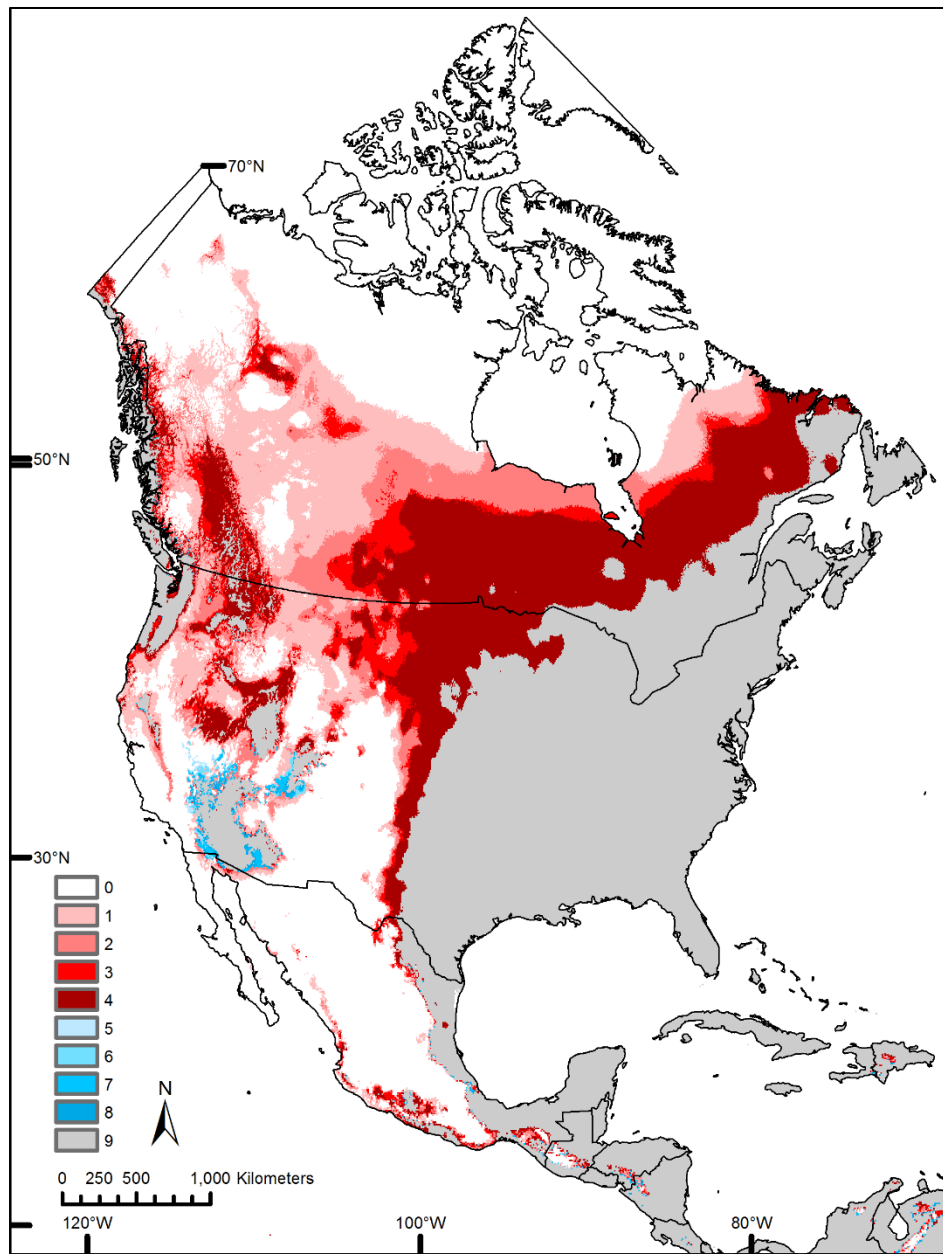


Table 1: List of confirmed and suspected pathogens transmitted by *Dermacentor variabilis* and the diseases caused.

Pathogen	Disease	Reference
<i>Anaplasma marginale</i>	Anaplasmosis	(Kocan et., 2010)
<i>Francisella tularensis</i>	Tularemia	(CDC, 2017)
<i>Rickettsia rickettsii</i>	Rocky Mountain Spotted Fever	(CDC, 2017)
<i>No pathogen (Salivary neurotoxins)</i>	Canine tick Paralysis	(Chan and Kaufman 2008)

Table 2 : List of bioclimatic variables.

Layer name	Variable represented
BIO1	Annual Mean Temperature
BIO2	Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO3	Isothermality (BIO2/BIO7) (* 100)
BIO4	Temperature Seasonality (standard deviation *100)
BIO5	Max Temperature of Warmest Month
BIO6	Min Temperature of Coldest Month
BIO7	Temperature Annual Range (BIO5-BIO6)
BIO8	Mean Temperature of Wettest Quarter
BIO9	Mean Temperature of Driest Quarter
BIO10	Mean Temperature of Warmest Quarter
BIO11	Mean Temperature of Coldest Quarter
BIO12	Annual Precipitation
BIO13	Precipitation of Wettest Month
BIO14	Precipitation of Driest Month
BIO15	Precipitation Seasonality (Coefficient of Variation)
BIO16	Precipitation of Wettest Quarter
BIO17	Precipitation of Driest Quarter
BIO18	Precipitation of Warmest Quarter
BIO19	Precipitation of Coldest Quarter

Table 3 : Environmental variable combinations evaluated by jackknife procedure for use in development of *Dermacentor variabilis* distribution model.

Environmental set	BioClim layers
Environmental variable set 1	BIO1, BIO2, BIO10, BIO15
Environmental variable set 2	SET 1 + BIO5, BIO16
Environmental variable set 3	SET 2 + BIO14, BIO17
Environmental variable set 4	SET 3 + BIO6, BIO11, BIO13

Table 4 : Compilation of three different environmental sets.

Environmental set	AUC	BioClim variables
Global model	0.902	1, 2, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 15, 16, 17
Set 1	0.901	1, 2, 3, 4, 5, 10, 12, 14, 15, 16, 17
Set 2	0.897	1, 2, 5, 12, 14, 15, 16, 17
Set 3	0.896	1, 2, 5, 12, 15, 16

Table 5 : Estimates of relative contributions of the different Bioclim variables to individual models at the three jackknife steps.

BioClim variable	Percentage contribution			Permutation contribution		
	Jackknife 1	Jackknife 2	Jackknife 3	Jackknife 1	Jackknife 2	Jackknife 3
1	41	42.5	43.4	5.8	37.2	40.6
2	7.7	8.8	10	22.3	17.7	16.2
3	2.2	-	-	9.7	-	-
4	2.4	-	-	1.5	-	-
5	6.5	6.0	5.4	29.9	9.6	7.8
10	0.1	-	-	1.8	-	-
12	23.9	24.1	23.8	2.7	5.3	4.8
14	2.8	1	-	11.2	9.3	-
15	1	4.2	8.7	1.9	2.1	6
16	10.2	9.8	8.7	12.3	18.3	24.6
17	2.3	3.6	-	0.8	0.5	-

Table 6 : Different model parameters and the values of model selection criteria.

Model	Feature class	Regularization multiplier	Environmental variable	Partial ROC	Omission Rate	AIC value
1	Linear	2.0	Set 3	0.00	0.11	1828.57
2	Linear	3.0	Set 3	0.00	0.11	1828.65
3	Linear	4.0	Set 3	0.00	0.11	1828.79
4	Linear	0.1	Set 3	0.00	0.11	1831.30
5	Linear	0.2	Set 3	0.00	0.11	1831.33

BIBLIOGRAPHY

- Alkishe, A. A., Peterson, A. T., & Samy, A. M. (2017). Climate change influences on the potential geographic distribution of the disease vector tick *Ixodes ricinus*. *PLOS ONE*, 12(12), e0189092.
- Barker, S. C., & Murrell, A. (2004). Systematics and evolution of ticks with a list of valid genus and species names. *Parasitology*, 129 Suppl, S15-36.
- Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., Soberón, J., et al. (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, 222(11), 1810–1819.
- Brites-Neto, J., Duarte, K. M. R., & Martins, T. F. (2015). Tick-borne infections in human and animal population worldwide. *Veterinary World*, 8(3), 301–315.
- Burnham, K. & Anderson, D. (2002) Model Selection and Multi-Model Inference: A Practical Information–Theoretic Approach, 2nd edn. Springer, New York, NY.
- CDC (2017). Transmission | Rocky Mountain Spotted Fever (RMSF) | CDC. Retrieved April 1, 2018, from <https://www.cdc.gov/rmsf/transmission/index.html>
- CDC Press Releases. (2016, January 1). CDC. Retrieved April 10, 2018, from <https://www.cdc.gov/media/releases/2015/p0619-RMSF.html>
- CDC, 2017 Transmission | Tularemia | CDC. Retrieved April 1, 2018, from <https://www.cdc.gov/tularemia/transmission/index.html>
- CDC, 2018(2018, January 19). Geographic distribution of ticks that bite humans | Ticks | CDC. Retrieved March 16, 2018, from https://www.cdc.gov/ticks/geographic_distribution.html
- Chan and Kaufman 2008. American dog tick - *Dermacentor variabilis* (Say). Retrieved March 11, 2018, from http://entnemdept.ufl.edu/creatures/urban/medical/american_dog_tick.htm

- Daniel, M., Danielová, V., Kríz, B., Jirsa, A., & Nozicka, J. (2003). Shift of the tick *Ixodes ricinus* and tick-borne encephalitis to higher altitudes in central Europe. *European Journal of Clinical Microbiology & Infectious Diseases: Official Publication of the European Society of Clinical Microbiology*, 22(5), 327–328.
- Descamps, S. (2013). Winter Temperature Affects the Prevalence of Ticks in an Arctic Seabird. *PLOS ONE*, 8(6), e65374.
- Drexler, N., Miller, M., Gerding, J., Todd, S., Adams, L., Dahlgren, F. S., Bryant, N., et al. (2014). Community-Based Control of the Brown Dog Tick in a Region with High Rates of Rocky Mountain Spotted Fever, 2012–2013. *PLOS ONE*, 9(12), e112368.
- Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R. J., et al. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 129–151.
- Elith, J., & Leathwick, J. R. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, 40(1), 677–697.
- Elith Jane, Kearney Michael, & Phillips Steven. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 1(4), 330–342.
- Epstein, P. R. (2001). Climate change and emerging infectious diseases. *Microbes and Infection*, 3(9), 747–754.
- Escobar, L. E., Lira-Noriega, A., Medina-Vogel, G., & Townsend Peterson, A. (2014). Potential for spread of the white-nose fungus (*Pseudogymnoascus destructans*) in the Americas: use of Maxent and NicheA to assure strict model transference. *Geospatial Health*, 9(1), 221–229.

- Feria-Arroyo, T. P., Castro-Arellano, I., Gordillo-Perez, G., Cavazos, A. L., Vargas-Sandoval, M., Grover, A., Torres, J., et al. (2014). Implications of climate change on the distribution of the tick vector *Ixodes scapularis* and risk for Lyme disease in the Texas-Mexico transboundary region. *Parasites & Vectors*, 7, 199.
- Fick, S & Hijmans, R. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*.
- Fuente, J. de la. (2003). The fossil record and the origin of ticks (Acari: Parasitiformes: Ixodida). *Experimental & Applied Acarology*, 29(3–4), 331–344.
- Hanzlicek, G. A., Raghavan, R. K., Ganta, R. R., & Anderson, G. A. (2016). Bayesian Space-Time Patterns and Climatic Determinants of Bovine Anaplasmosis. *PLOS ONE*, 11(3), e0151924.
- Harvell, C. D., Mitchell, C. E., Ward, J. R., Altizer, S., Dobson, A. P., Ostfeld, R. S., & Samuel, M. D. (2002). Climate Warming and Disease Risks for Terrestrial and Marine Biota. *Science*, 296(5576), 2158–2162.
- Hayhoe, K., VanDorn, J., Naik, V., & Wuebbles, D. (2015). Climate Change in the Midwest, 24.
- Hoogstraal, H. (2011). Physiology of Ticks. Edited by F. D. Obenchain and R. Galun (1982). (Current Themes in Tropical Science, Vol. 1), ix + 509 pages. Pergamon Press, Oxford, U.K. (US\$ 125.00). *International Journal of Tropical Insect Science*, 4, 417–420.
- Hopla C.E & Durden L.A (1994). Ectoparasites and classification - OIE - PDF Free Download. *zdoc.site*. Retrieved April 9, 2018, from <https://zdoc.site/ectoparasites-and-classification-oie.html>
- Jongejan, F., & Uilenberg, G. (2004). The global importance of ticks. *Parasitology*, 129 Suppl, S3-14.

- Karl, T. R., Jones, P. D., Knight, R. W., Kukla, G., Plummer, N., Razuvayev, V., Gallo, K. P., et al. (1993). A New Perspective on Recent Global Warming: Asymmetric Trends of Daily Maximum and Minimum Temperature. *Bulletin of the American Meteorological Society*, 74(6), 1007–1024.
- Karl Thomas R., Kukla George, Razuvayev Vyacheslav N., Changery Michael J., Quayle Robert G., Heim Richard R., Easterling David R., et al. (1991). Global warming: Evidence for asymmetric diurnal temperature change. *Geophysical Research Letters*, 18(12), 2253–2256.
- Kocan, K. M., Blouin, E. F., Coetzee, J. F., & Simpson, K. M. (2010). Current Challenges of the Management and Epidemiology of Bovine Anaplasmosis, 44(2), 10.
- Kutz, S., Hoberg, E. ., Polley, L., & Jenkins, E. . (2005). Global warming is changing the dynamics of Arctic host–parasite systems. *Proceedings of the Royal Society B: Biological Sciences*, 272(1581), 2571–2576.
- Leighton Patrick A., Koffi Jules K., Pelcat Yann, Lindsay L. Robbin, & Ogden Nicholas H. (2012). Predicting the speed of tick invasion: an empirical model of range expansion for the Lyme disease vector *Ixodes scapularis* in Canada. *Journal of Applied Ecology*, 49(2), 457–464.
- Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421(6918), 37–42.
- Peterson, A. T., & Raghavan, R. (2017a) (2017/09/01). The Leading Edge of the Geographic Distribution of *Ixodes scapularis* (Acari: Ixodidae). *Journal of Medical Entomology*, 54(5), 1103–1103.

- Peterson, A. T., & Raghavan, R. K. (2017b). The Geographic Distribution of *Ixodes scapularis* (Acari: Ixodidae) Revisited: The Importance of Assumptions About Error Balance. *Journal of Medical Entomology*, 54(4), 1080–1084.
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3), 231–259.
- Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31(2), 161–175.
- Phillips, S. J., Dudík, M., & Schapire, R. E. (2004). A maximum entropy approach to species distribution modeling (p. 83). ACM Press. Retrieved April 10, 2018, from <http://portal.acm.org/citation.cfm?doid=1015330.1015412>
- Poinar, G. P. (2015). Spirochete-like cells in a Dominican amber *Ambylomma* tick (Arachnida: Ixodidae). *Historical Biology*, 27(5), 565–570.
- Raghavan, R. K., Goodin, D. G., Neises, D., Anderson, G. A., & Ganta, R. R. (Mar 4, 2016a). Hierarchical Bayesian Spatio–Temporal Analysis of Climatic and Socio–Economic Determinants of Rocky Mountain Spotted Fever. *PLOS ONE*, 11(3), e0150180.
- Raghavan, R. K., Harrington, J., Anderson, G. A., Hutchinson, J. M. S., & Debey, B. M. (2013). Environmental, climatic, and residential neighborhood determinants of feline tularemia. *Vector Borne and Zoonotic Diseases (Larchmont, N.Y.)*, 13(7), 449–456.
- Schoof, J. (2012). Historical and projected changes in human heat stress in the Midwestern United States. *Climate Change in the Midwest: Impacts, Risks, Vulnerability, and Adaptation*, 146–157.

- Schwartz, M. D. (1995). Detecting Structural Climate Change: An Air Mass-Based Approach in the North Central United States, 1958–1992. *Annals of the Association of American Geographers*, 85(3), 553–568.
- Smith, T., & Kilborne, F. L. (1893). Investigations into the Nature, Causation, and Prevention of Texas or Southern Cattle Fever. *The Portal to Texas History*. Book, . Retrieved March 14, 2018, from <https://texashistory.unt.edu/ark:/67531/metaph143538/m1/1/>
- Soberón, J., & Nakamura, M. (2009). Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences of the United States of America*, 106(Suppl 2), 19644–19650.
- Soberón, J., & Peterson, A. T. (2005). Interpretation of Models of Fundamental Ecological Niches and Species' Distributional Areas. Retrieved April 19, 2018, from <https://kuscholarworks.ku.edu/handle/1808/20560>
- Soberón Jorge. (2007). Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, 10(12), 1115–1123.
- Sonenshine, D. E. (2018). Range Expansion of Tick Disease Vectors in North America: Implications for Spread of Tick-Borne Disease. *International Journal of Environmental Research and Public Health*, 15(3), 478.
- Stockwell, D. (1999). The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, 13(2), 143–158.
- Thuiller, W., Lavorel, S., Sykes, M. T., & Araujo, M. B. (2006). Using niche-based modelling to assess the impact of climate change on tree functional diversity in Europe. *Diversity & Distributions*, 12(1), 49–60.

- Townsend Peterson A., Papeş Monica, & Eaton Muir. (2007). Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography*, 30(4), 550–560.
- Tsoar Asaf, Allouche Omri, Steinitz Ofer, Rotem Dotan, & Kadmon Ronen. (2007). A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and Distributions*, 13(4), 397–405.
- WRBU, 2018. Walter Reed Biosystematics Unit. Retrieved April 1, 2018, from <http://www.wrbu.org/>