2023

Newer Surveillance Data Extends Our Understanding of the Niche of *Rickettsia montanensis* (Rickettsiales: Rickettsiaceae) Infection of the American Dog Tick (Acari: Ixodidae) in the United States

Catherine A. Lippi
University of Florida

Holly D. Gaff
Old Dominion University, hgaff@odu.edu

Robyn M. Nadolny
Defense Centers for Public Health

Sadie J. Ryan
University of Florida

Original Publication Citation

1   **Newer Surveillance Data Extends our Understanding of the Niche of *Rickettsia***

2   ***montanensis* (Rickettsiales: Rickettsiaceae) Infection of the American Dog Tick (Acari:**

3   **Ixodidae) in the United States.**

4   Catherine A. Lippi[1], Holly D. Gaff[2], Robyn M. Nadolny[3], Sadie J. Ryan[1]*

5   [1]Department of Geography and Emerging Pathogens Institute, University Florida, Gainesville,

6   FL 32611

7   [2]Department of Biology, Old Dominion University, Norfolk, VA 23529

8   [3]Defense Centers for Public Health-Aberdeen, Aberdeen Proving Ground, MD 21010

9

10   *correspondence to sjryan@ufl.edu

11

12   **Abstract**

13   **Background:** Understanding the geographic distribution of *Rickettsia montanensis* infections in

14   *Dermacentor variabilis* is important for tick-borne disease management in the United States, as

15   both a tick-borne agent of interest and a potential confounder in surveillance of other rickettsial

16   diseases. Two previous studies modeled niche suitability for *D. variabilis* with and without *R.*

17   *montanensis*, from 2002-2012, indicating that the *D. variabilis* niche overestimates the infected

18   niche. This study updates these, adding data since 2012.

19   **Methods:** Newer surveillance and testing data were used to update Species Distribution Models

20   (SDMs) of *D. variabilis*, and *R. montanensis* infected *D. variabilis*, in the United States. Using

21   random forest (RF) models, found to perform best in previous work, we updated the SDMs and

22   compared them with prior results. Warren's I niche overlap metric was used to compare

23   between predicted suitability for all ticks and 'pathogen positive niche' models across datasets.

24   **Results:** Warren's I indicated <2% change in predicted niche, and there was no change in order

25   of importance of environmental predictors, for *D. variabilis* or *R. montanensis* positive niche.

26   The updated *D. variabilis* niche model overpredicted suitability compared to the updated *R.*

27   *montanensis* positive niche in key peripheral parts of the range, but slightly underpredicted

28   through the northern and midwestern parts of the range. This reinforces previous findings of a

29   more constrained pathogen-positive niche than predicted by *D. variabilis* records alone.

30   **Conclusions:** The consistency of predicted niche suitability for *D. variabilis* in the United

31   States, with the addition of nearly a decade of new data, corroborates this is a species with

32   generalist habitat requirements. Yet a slight shift in updated niche distribution, even of low

33   suitability, included more southern areas, pointing to a need for continued and extended

34   monitoring and surveillance. This further underscores the importance of revisiting vector and

35   vector-borne disease distribution maps.

36   **Keywords:** Dermacentor variabilis, Rickettsia montanensis, Species distribution modeling, Tick-

37   borne disease

38 **Introduction**

39 Species distribution models (SDMs) are increasingly utilized to estimate the geographic

40 distribution of infectious diseases, particularly those caused by agents transmitted by arthropod

41 vectors. The basic methodology for constructing SDMs (or ecological niche models) consists of

42 combining species occurrence data with continuous layers of environmental predictor variables,

43 which are fed into a modeling algorithm (Elith and Franklin, 2013; Franklin, 2010; Peterson and

44 Soberón, 2012). The resulting model is projected onto a defined study area, yielding spatially

45 continuous habitat suitability estimates for areas of the landscape that were not originally

46 sampled. Species distribution modeling is an intuitive approach to delineating vector-borne

47 disease ranges that is logistically feasible, particularly when surveillance programs or capacity

48 for pathogen testing are limited. When faced with multiple unknowns (e.g., unknown

49 transmission cycles, emerging novel pathogens, etc.), the distribution of vectors on the

50 landscape are sometimes used in a public health context to approximate risk of exposure to

51 pathogens (Lippi et al., 2021b, 2021c). Yet, it is important to differentiate between the

52 distribution of the vectors and that of the pathogens they transmit. Vector presence is not in

53 itself sufficient for pathogen transmission to occur. Precise delineation of geographic risk

54 facilitates the development of targeted health policies, educational campaigns, and interventions

55 with the potential to avert the misallocation of limited resources.

56 The need for geographically conservative assessments of transmission risk is perhaps most

57 evident with cosmopolitan vectors, whose broad geographic ranges may far exceed the limits of

58 known transmission to humans. The American dog tick (*Dermacentor variabilis*) is a medically

59 important arthropod vector of several zoonotic pathogens, including the causative agents of

60 Rocky Mountain spotted fever (RMSF) (*Rickettsia rickettsii*) (Brumpt; Rickettsiales:

61 Rickettsiaceae) and tularemia (*Francisella tularensis*) (Dorofe'ev; McCoy and Chapin;

62 Thiotrichales: Francisellaceae). Both of these diseases can be fatal without medical

63   intervention, perhaps justifying medical advisories that equate risk of tick exposure with

64   transmission risk, particularly when surveillance data are scarce, or in cases where ticks

65   themselves act as reservoir hosts (CDC, 2022). In addition to RMSF, *D. variabilis* also transmits

66   other spotted fever group (SFG) rickettsial agents, as well as *R. montanensis* (Rickettsiales:

67   Rickettsiaceae), a rickettsial group agent that is suspected of causing nonfebrile rashes in

68   humans, and has caused clinical symptoms in an animal model (McQuiston et al. 2012;

69   Snellgrove et al. 2021). Although not included in the case definition for SFG pathogens, it is

70   likely that *R. montanensis* infections may account for some of the recent increases in SFG

71   reporting, as immunological cross-reactivity between rickettsial pathogens is frequently

72   observed with commonly used serologic tests (Abdad et al. 2018). Of note, *D. variabilis* has

73   recently been proposed to be split into two species, with a western portion of the population as a

74   distinct species, *D. similis* (Lado et al., 2021); however, we do not differentiate in this study.

75   Determining the geographic risk of *D. variabilis* infection with *R. montanensis* has profound

76   implications for the management of tick-borne diseases in the United States, as both a tick-

77   borne agent of interest and a potential confounder in the surveillance of other Rickettsial

78   diseases. A model of the distribution of *D. variabilis* and *R. montanensis* positive samples was

79   published by St John et al. in 2016, using MaxEnt modeling to describe and predict

80   environmental suitability in the United States, based on data obtained through the Department

81   of Defense (DoD) Human Tick Test Kit Program, now called the Military Tick

82   Identification/Infection Confirmation Kit Program (MilTICK). These data were available at the

83   time through the VectorMap online data platform (http://vectormap.si.edu/dataportal/) (St John

84   et al., 2016). The MilTICK data were human-biting ticks submitted from U.S. military installations

85   as part of a tick-testing program; test results were reported back to the bitten individuals, and

86   the data were also used as passive vector surveillance. In 2021, Lippi et al. re-examined the

87   distribution of *D. variabilis* and the *R. montanensis* infected niche in the USA, both to

88   understand whether predicted risk of suitability for tick encounters or infected tick encounters

89   were distinct, and to explore and compare multiple modeling approaches for assessing the

90   distribution of this tick vector (Lippi et al., 2021a). The 2021 study was able to leverage the

91   original dataset used in the 2016 study, and used a refined set of environmental predictors to

92   compare a suite of Species Distribution Model (SDM) approaches. Lippi et al. found support for

93   an "infected niche" within the broader distribution of *D. variabilis* which was largely consistent

94   across models, though the Random Forests (RF) approach (Breiman, 2001) provided the best

95   performing models, given the available data (Lippi et al., 2021a). Though somewhat limited in

96   terms of the full geographic distribution of *D. variabilis* ticks (i.e., few locations were reported

97   from the tick's southern extent), the dataset used in these studies provided a rare opportunity to

98   directly assess the distribution of pathogens within vectors, as every individual tick collected had

99   been tested for *R. montanensis* as part of an extensive passive surveillance network. Both of

100   these studies demonstrated that *D. variabilis* ticks infected with *R. montanensis* had estimated

101   geographic distributions that were considerably restricted compared to that of *D. variabilis*

102   alone, thus supporting an "infected niche" that exists as a subset of the vector's full range.

103   In the current study, we revise the *D. variabilis* distribution maps using occurrence data updated

104   with novel surveillance points collected since 2012, and further refine the environmental

105   variables according to current best practices using the RF approach (Escobar et al., 2014;

106   Valavi et al., 2021). We explore whether the additional data impact the estimated suitability

107   distribution, the relative importance of environmental input variables, and mapped prediction

108   outputs.

Methods

110   *Tick Surveillance Data* – Two previous studies on *D. variabilis* in the United States were

111   conducted using occurrence locations recorded in the continental United States from 2002 to

112   2012, where ticks were tested for *R. montanensis* as part of MilTICK, and are described in St

113   John et al. (2016) and Lippi et al. (2021) (Lippi et al., 2021a; St John et al., 2016).

114   Georeferenced data were openly available through VectorMap

115   (http://vectormap.si.edu/dataportal/), a project of the Walter Reed Bioinformatics Unit (WRBU),

116   housed at the Smithsonian Institution Washington DC (St John et al., 2016). All ticks submitted

117   through MilTICK are tested for rickettsial pathogens via PCR as previously described

118   (Milholland et. al., 2021, Stromdahl et al., 2011), providing information on infection status (i.e.,

119   true presence or absence) for the entire dataset. Exposure locations were determined by asking

120   MilTICK participants to self-report where the tick bite was most likely acquired, accounting for

121   travel history. If no separate information on tick-bite location was submitted, ticks were assumed

122   to be acquired on or near the military installation from which the tick was submitted.

123   New records of *D. variabilis* reported and tested for *R. montanensis* through MilTICK since 2012

124   through 2021 were made available for this study. These data were de-identified, and though

125   general locality data were provided (e.g., military installation where reported, or towns and cities

126   where ticks were collected), positional coordinates were not provided. New surveillance data

127   were manually georeferenced for this study, following the general protocol reported in the

128   metadata of the original dataset (i.e., 2002-2012 records) georeferenced for TickMap by the

129   WRBU. Geographic coordinates (i.e., latitude and longitude) were assigned to records, taking

130   the centroid of named locations found in Google Maps. Spatial uncertainty for points was

131   established based on the spatial extent of reported locations (e.g., municipal boundaries,

132   reported area of military installations, etc.). We excluded records where the spatial uncertainty

133   exceeded 10km, ensuring that the spatial resolution of the St. John et al. (2016) and Lippi et al.

134   (2021) studies was matched for all analyses.

135   We removed duplicate records and records without pathogen testing results (n=14). Data

136   thinning on the remaining species occurrence points was performed via the 'spThin' package in

137   R (ver. 4.1.2) (R Core Team 2019), which uses a spatial thinning algorithm to randomly remove

138   excess occurrence locations within a specified distance threshold (Aiello-Lammens et al., 2015).

139   This was performed for both the original data in the Lippi et al. 2021 study and the updated

140   dataset to reduce susceptibility to geographic sampling bias, for example, when

141   overrepresented locations erroneously drive species environmental associations due to

142   repeated observations at discrete locations. Due to the passive nature of the tick surveillance

143   program, it was deemed necessary to thin occurrences and minimize the potential effect of

144   sampling bias, where locations near medical facilities and military installations may be inherently

145   overrepresented. This process resulted in one unique, randomly selected location per 10km,

146   and was performed on the full dataset of tick records, and on the subset of ticks that tested

147   positive for *R. montanensis*.

148   The original dataset used to build the distribution models reported in Lippi et al. 2021 was then

149   compared to an updated dataset, reflecting new surveillance data. Because new surveillance data

150   consisted of fewer records compared to the original study, the updated dataset was comprised of

151   both original surveillance data and new surveillance records. Following the framework of Lippi et

152   al. 2021, we estimated separate geographic distributions of *D. variabilis*, and the subset of

153   records that tested positive for *R. montanensis* infections, for both the original and updated tick

154   surveillance records. Environmental data layers used in modeling consisted of interpolated

155   bioclimatic (bioclim) layers from WorldClim (ver. 2), and gridded soil variables (0cm standard

156   depth) taken from International Soil Reference Information Centre (ISRIC) SoilGrids (Fick and

157   Hijmans, 2017; Hengl et al., 2017). Gridded environmental data inputs were used at 10km

158   resolution to match the scale of tick occurrence data. Bioclim layers with known errors (i.e.,

159   Bio8, Bio9, Bio18, and Bio19) were removed a priori, and Variance Inflation Factor (VIF) was

160   used to control for collinearity in the remaining variables (th=10) (Escobar et al., 2014). The final

161   set of variables used to build models included annual mean temperature (Bio1), mean diurnal

162   range (Bio2), temperature seasonality (Bio4), precipitation of wettest month (Bio13),

163   precipitation of driest month (Bio14), precipitation seasonality (Bio15), soil organic carbon

164   density (OCDENS), available soil water capacity until wilting point (WWP), and soil pH

165   (PHIHOX).

166   Random forests (RF) modeling, implemented in R with the package 'sdm', was used to estimate

167   tick distributions, following recommendations for settings and parameters described in Valavi et

168   al 2021 (Valavi et al., 2021). We ran 500 RF model replicates for each dataset of occurrence

169   points (i.e., original and updated records for all *D. variabilis*, and original and updated records

170   for only *D. variabilis* infected with *R. montanensis*), averaging projected model output to produce

171   four estimated distributions. Average model accuracy metrics for each experiment were

172   calculated to assess the predictive accuracy of SDMs against a random holdout of 25% data

173   from each dataset, respectively. Four measures were calculated to assess model accuracy, the

174   receiver operator characteristic (ROC) curve with area under the curve (AUC), true skill statistic

175   (TSS), model deviance, and mean omission (i.e., false negatives). We quantified the niche

176   overlap between averaged models with the Warren's I index, calculated in R with the package

177   'spatialEco' (Warren et al., 2008). The I statistic is an indicator of the similarity between two

178   distributions, with values ranging from 0 (i.e., no overlap in the niche) to 1 (i.e., the niche is

179   identical). A difference map to assess agreement in suitability predictions between the updated

180   full dataset and infected dataset models was generated in R using the packages Raster and

181   RasterVis by taking the difference of model output rasters and plotting them.

182   **Results**

183   Updated input surveillance data increased our sample sizes for the full dataset (original n=432,

184   updated n=525), and for the ticks positively identified for *R. montanensis* infection (original

185   n=44, updated n=63). We found that updating the input data increased the spatial extent of

186   predicted suitability for both the full dataset of all ticks (Figure 1 A (original) and B (updated))

187   and for the infected dataset (Figure 1 C (original) and D (updated)). Although we made no

188   distinction for potential records of the newly described species *D. similis*, a few occurrence

189   points were from the Western United States (original n=10, updated n=21). Model accuracy

190   metrics for averaged RF models across the four datasets are presented in Table 1. Accuracy

191   metrics across models indicated generally good performance, with AUC values exceeding 0.90,

192   and TSS values greater than 0.64. Though comparable in output, averaged models made with

193   updated data performed lower than models made with original datasets, indicated by lower AUC

194   and TSS values, and higher deviance and omission. A Warren's I index comparison of the

195   original and updated dataset suitability predictions for the full and infected niche, showed they

196   differed by less than 2% each (full dataset: full dataset =0.981, positive dataset: positive dataset

197   =0.986).
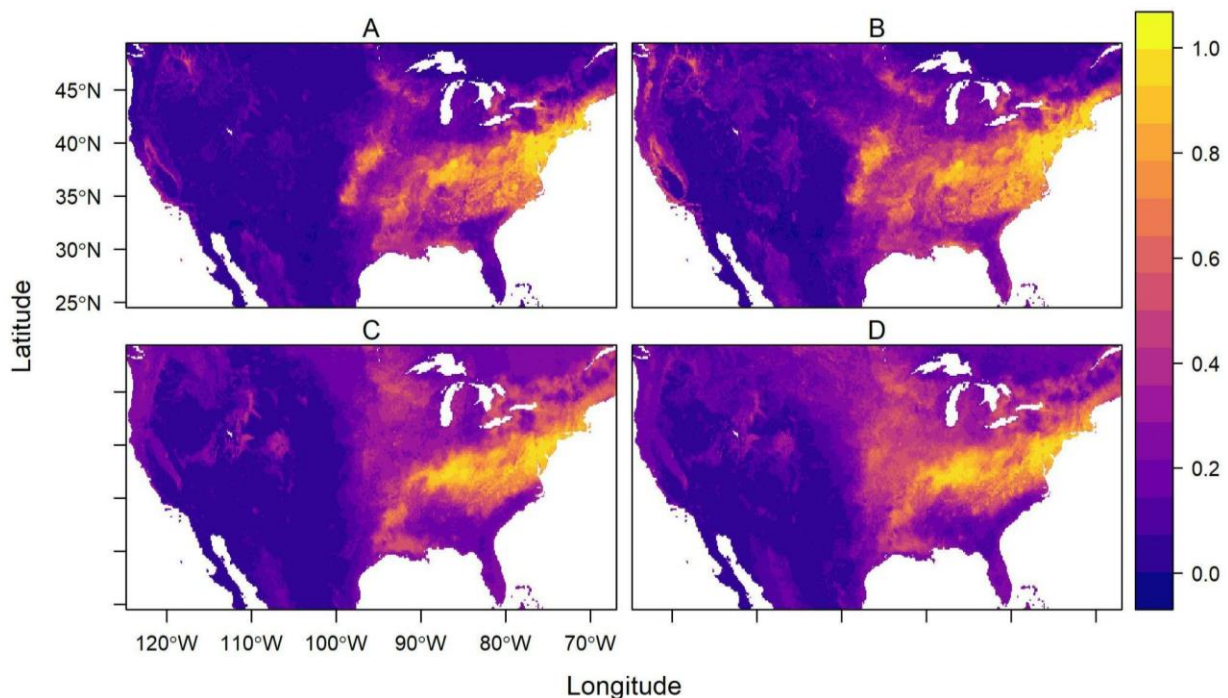
198

199   **Figure 1: Predicted habitat suitability from average output of 500 random forest models**

200   for the original (A, C) and updated (B, D) datasets for all *D. variabilis* data (A, B), and *D.*

201   *variabilis* infected with *R. montanensis* (C, D)

202   The updated *R. montanensis* positive ticks, as in the original analyses, are predicted to have a

203   niche which is a subset of the full predicted niche (Figure 1D). The Warren's I comparisons of

204   the 'infected niche' and the full datasets for original (full:infected =0.950), and updated datasets

205   (full:infected = 0.968) suggest that these are not dissimilar predicted niche distributions where

206   they overlap, yet they are not capturing identical distributions.

207   **Table 1.** Average model accuracy metrics for Random Forest models, using different datasets

208   of tick occurrences.

| Dataset	Subset	AUC	Deviance	TSS	Omission
Original*	All Ticks	0.953	0.570	0.769	0.116
Original*	Positive Ticks	0.930	0.690	0.710	0.145
Updated	All Ticks	0.918	0.742	0.692	0.154
Updated	Positive Ticks	0.905	0.812	0.643	0.179

209   *data used in Lippi et al. 2021

210   The importance of variables underlying model predictions varied across datasets, although

211   precipitation seasonality (Bio15) was the top contributing environmental predictor in all models

212   (Fig. 2). Mean diurnal range (Bio2) and precipitation of driest month (Bio14) were also relatively

213   important variables in models of both the original and updated full tick datasets, though these

214   variables did not contribute highly to the models of infected tick distributions.

215
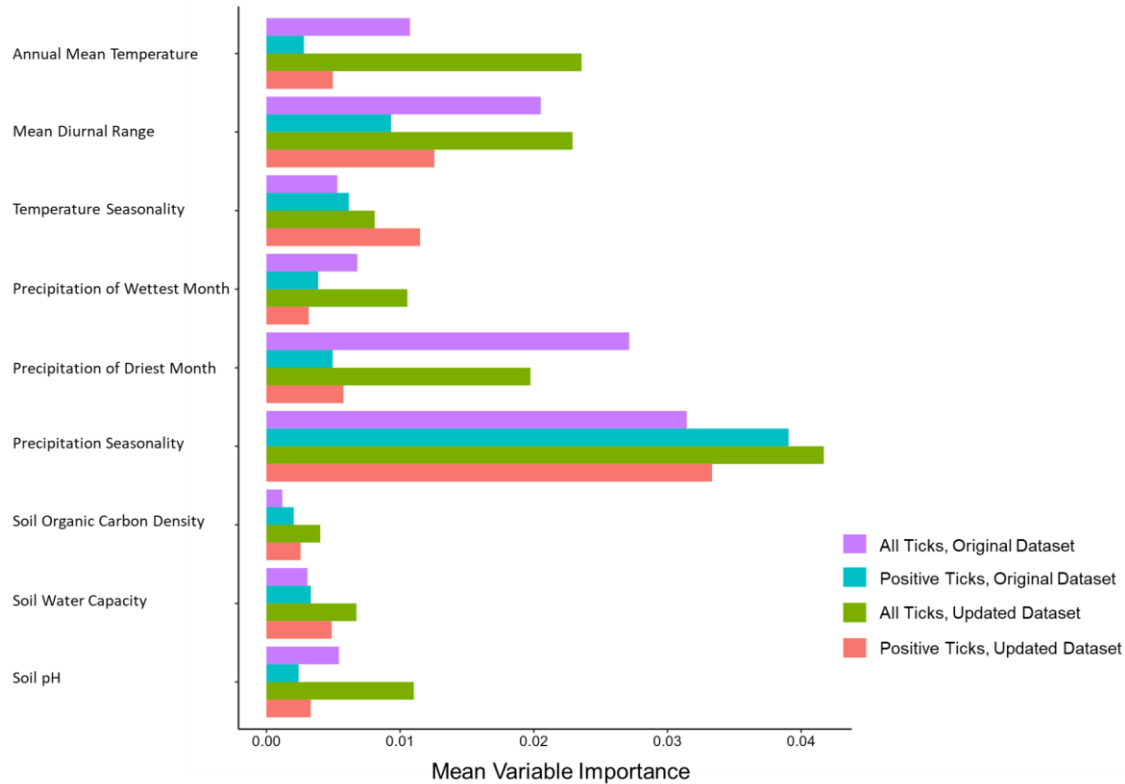
Figure 2: Relative variable importance from average output of 500 random forest models for the

original and updated datasets for all *D. variabilis* data, and *D. variabilis* infected with *R.*

montanensis.

To visualize the difference in predicted suitability for all ticks and that predicted for the

pathogen-positive ticks, we visualized the difference in mapped suitability estimates from

updated models (Fig. 3). The resulting map highlights the overprediction (redder colors) or

underprediction (darker blue colors) of a model trained on all surveilled ticks, compared to one

trained on *R. montanensis* positive ticks. Infected ticks are overpredicted by the model of all

ticks along the southeastern and western peripheries of the infected tick distribution, and

underpredicted to a lesser degree, along the northern border and through parts of the mid-

Atlantic to midwestern states (Figure 3).
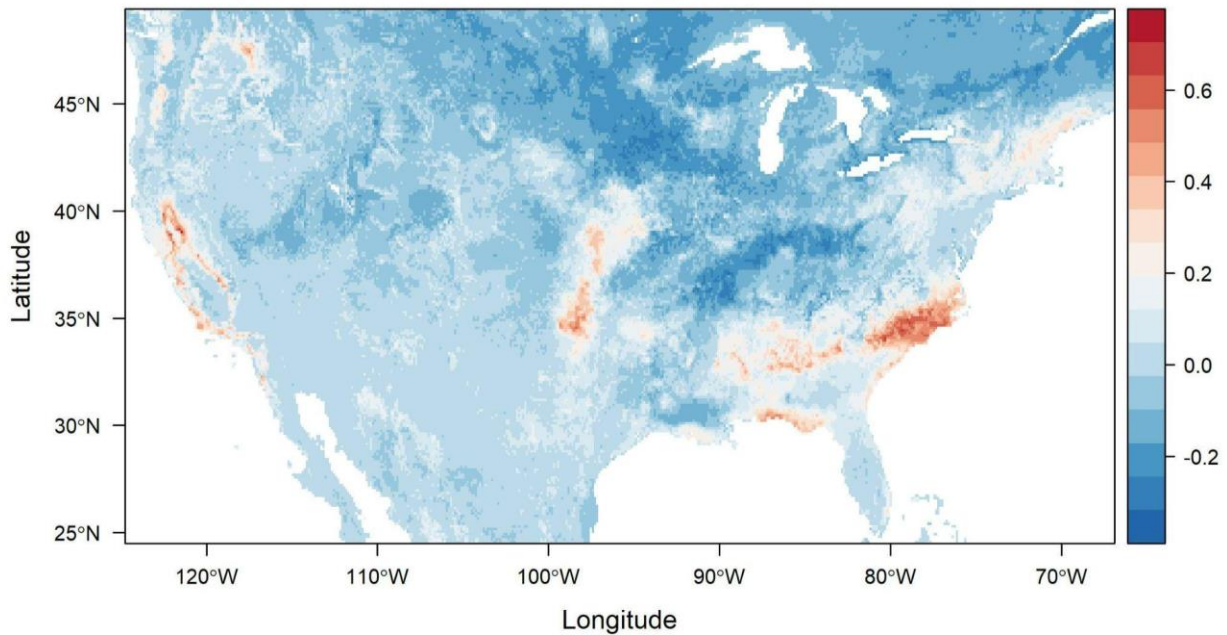
Figure 3: Assessing differences in predicted suitability for an average of 500 Random Forest models for *D. variabilis* and those infected with *R. montanesis* - redder colors depict overprediction by a tick-only model, and darker blue colors, underprediction.

Discussion

A number of factors exist that influence SDM output, including sampling bias, choice of environmental predictors, modeling algorithm, and other user-specified inputs (Araújo et al., 2019; Valavi et al., 2021). In this study, we updated previously published RF models of *D. variabilis* and *D. variabilis* infected with *R. montanensis*. This update was made possible by the addition of surveillance and testing data to the original dataset used. We thus explored what impact the additional data had on predictions found previously, via modeling both datasets and comparing predicted suitability with a niche overlap metric, Warren's I, and presenting the mapped output of modeled predictions using the original and updated datasets. We additionally

241   presented a visualization of agreement, highlighting areas of over and underprediction of the

242   infected niche by the overall niche prediction.

243   Models made with both datasets were generally high-performing, and overlap indices showed

244   that suitability predictions varied only slightly with the inclusion of novel surveillance data. The

245   estimated range of *D. variabilis* primarily extends throughout the eastern United States, with the

246   highest predicted probabilities spanning areas in the Midwest, Mid-Atlantic, and Northeast

247   regions. The southern boundary of *D. variabilis* occurrence was not well captured in Lippi et al.

248   2021, owing to limited data points from this region in the original MilTICK dataset. Although

249   records of ticks from southern locations (e.g. Texas and peninsular Florida) exist in online

250   repositories, these records were not included in efforts to directly compare distributions of ticks

251   of known infection status. Notably, the predicted geographic distribution for *D. variabilis* extends

252   further South in the updated model, indicated by higher probabilities of suitability in Texas and

253   Florida.

254   The predicted suitability distribution of *D. variabilis* infected with *R. montanensis*, or infected

255   niche, is geographically constrained, compared to the full predicted suitability distribution of *D.*

256   *variabilis*, regardless of data inputs. Areas of range disagreement, highlighted by the difference

257   map, are most prominent along the southern and western peripheries of the full *D. variabilis*

258   range in the eastern US, as well as on the west coast. A potential explanation for this kind of

259   pattern is that in the more established parts of the range - i.e. the more central parts of predicted

260   range - there may be higher *R. montanensis* exposure risk. For different tick-borne pathogens,

261   and even for different species of ticks, evidence of patterns of expansion by both the vector and

262   the pathogen, together or temporally lagged have varied (Burrows et al., 2021; Dahlgren et al.,

263   2016; Fornadel et al., 2011). This highlights the limitations inherent in using vector distribution

264   maps as proxies for transmission risk maps directly; incorporating pathogen testing results into

265   this type of distribution modeling can help constrain the area most likely to be important for

266   disease transmission exposure risk. This is particularly germane for a generalist vector such as

267   *D. variabilis*, where the presence of the pathogen in question may be patchily distributed.

268   Disagreement along the West coast may also be influenced by the inclusion of *D. variabilis*

269   records from California, Oregon, and Washington. The western population of *D.variabilis* has

270   recently been proposed as a new species (*Dermacentor similis*), and thus may have

271   fundamentally different habitat suitability requirements (Lado et al., 2021).

272   *Dermacentor* ticks are receiving increasing attention as significant vectors of zoonotic

273   pathogens, and there have been recent calls for closer monitoring of understudied species

274   (Lippi et al., 2021c; Martin et al., 2022). Species distribution modeling offers a framework for

275   rapidly estimating potential distributions of vectors when ample occurrence data are available.

276   Yet, there are considerable ramifications that may arise if models are put into public health

277   practice without thorough assessment (Erdemir et al., 2020). It is therefore necessary to

278   periodically review estimates of risk as new data or methods become available. However, in this

279   study we found that an additional nine years of passive surveillance data resulted in negligible

280   differences in distribution estimates. This points to the benefit of augmenting existing

281   surveillance to target undersampled areas, and highlights the need to expand pathogen testing

282   capabilities to other existing networks. Widespread, county-level surveillance for *D. variabilis* in

283   the United States is currently limited (Lehane et al., 2019). Pathogens with low detection rates

284   may particularly benefit from targeted, active surveillance strategies to delineate risk. In this

285   study, updated passive surveillance data yielded only 19 novel spatially unique records of

286   infected ticks after thinning. To contrast, a recent study that targeted a discrete area in Northern

287   Wisconsin, an area of low predicted suitability in our models, successfully detected *R.*

288   *montanensis* in *D. variabilis* (Vincent and Hulstrand, 2022). Focused testing efforts, particularly

289   in locations bordering areas of range disagreement, may help resolve the limits of exposure risk

290   and facilitate targeted monitoring efforts.

291 In conclusion, infected ticks are predicted to have a distribution that is a subset of the full vector

292 range, a finding which is consistent across original and updated data inputs. For a generalist

293 vector such as *D. variabilis*, ascertaining the key areas of pathogen exposure risk within such a

294 large range of predicted suitability, is an important potential tool for future surveillance and

295 monitoring. Revisiting the estimation of tick distributions is a necessary endeavor, particularly as

296 we gain more information on tick-borne transmission cycles through surveillance and laboratory

297 studies. There are few occurrence records that establish *D. variabilis* at the county level

298 throughout our predicted suitability range in the contiguous United States, pointing to a general

299 need for increased surveillance activities (Lehane et al., 2019). Yet, placing emphasis solely on

300 new data collection for the refinement of spatial risk assessments may not yield dramatic gains

301 in information. This is perhaps most evident in the passive surveillance of pathogens with low

302 detection rates. Additionally, we suggest that there is a great need to validate the data in areas

303 identified as high risk through active surveillance, particularly where passive surveillance is

304 lacking. Moving forward, efforts to further refine geographic risk estimates of tick-borne

305 pathogens will benefit from targeted surveillance to resolve distributional boundaries.

306 **References**

307 Abdad, M. Y., Abou Abdallah, R., Fournier, P. E., Stenos, J., & Vasoo, S. (2018). A concise

308 review of the epidemiology and diagnostics of rickettsioses: Rickettsia and *Orientia spp*. Journal

309 of Clinical Microbiology, 56(8), e01728-17.

310

311 Aiello-Lammens ME, Boria RA, Radosavljevic A, et al. spThin: An R Package for Spatial

312 Thinning of Species Occurrence Records for Use in Ecological Niche Models. Ecography

313 2015;38(5):541–545; doi: 10.1111/ecog.01132.

314

315 Araújo MB, Anderson RP, Márcia Barbosa A, et al. Standards for Distribution Models in

316   Biodiversity Assessments. Sci Adv 2019;5(1):eaat4858.

317

318   Breiman L. Random Forests. Mach Learn 2001;45(1):5–32.

319

320   Burrows H, Talbot B, McKay R, et al. A Multi-Year Assessment of Blacklegged Tick (*Ixodes*

321   *scapularis*) Population Establishment and Lyme Disease Risk Areas in Ottawa, Canada, 2017-

322   2019. PLoS One 2021;16(2):e0246484.

323

324   CDC. Tickborne Diseases of the United States: A Reference Manual for Health Care Providers.

325   Centers for Disease Control; 2022.

326

327   Dahlgren FS, Paddock CD, Springer YP, et al. Expanding Range of Amblyomma Americanum

328   and Simultaneous Changes in the Epidemiology of Spotted Fever Group Rickettsiosis in the

329   United States. Am J Trop Med Hyg 2016;94(1):35–42.

330

331   Elith J and Franklin J. Species Distribution Modeling. Encyclopedia of Biodiversity 2013;692–

332   705; doi: 10.1016/b978-0-12-384719-5.00318-x.

333

334   Erdemir A, Mulugeta L, Ku JP, et al. Credible Practice of Modeling and Simulation in

335   Healthcare: Ten Rules from a Multidisciplinary Perspective. J Transl Med 2020;18(1):1–18.

336

337   Escobar LE, Lira-Noriega A, Medina-Vogel G, et al. Potential for Spread of the White-Nose

338   Fungus (*Pseudogymnoascus destructans*) in the Americas: Use of Maxent and NicheA to

339   Assure Strict Model Transference. Geospat Health 2014;9(1):221–229.

340

341   Fick SE and Hijmans RJ. WorldClim 2: New 1-Km Spatial Resolution Climate Surfaces for

342   Global Land Areas. Int J Climatol 2017;(37):4302–4315.

343

344   Fornadel CM, Zhang X, Smith JD, et al. High Rates of Rickettsia Parkeri Infection in Gulf Coast

345   Ticks (*Amblyomma maculatum*) and Identification of "*Candidatus* rickettsia andeanae" from

346   Fairfax County, Virginia. Vector Borne Zoonotic Dis 2011;11(12):1535–1539.

347

348   Franklin J. Mapping Species Distributions: Spatial Inference and Prediction. Cambridge

349   University Press; 2010.

350

351   Hengl T, Mendes de Jesus J, Heuvelink GBM, et al. SoilGrids250m: Global Gridded Soil

352   Information Based on Machine Learning. PLoS One 2017;12(2):e0169748.

353

354   Lado P, Glon MG and Klompen H. Integrative Taxonomy of *Dermacentor variabilis* (Ixodida:

355   Ixodidae) with Description of a New Species, *Dermacentor similis* N. Sp. J Med Entomol

356   2021;58(6):2216–2227.

357   Lehane A, Parise C, Evans C, et al. Reported County-Level Distribution of the American Dog

358   Tick (Acari: Ixodidae) in the Contiguous United States. J Med Entomol 2019;57(1):131–155.

359

360   Lippi CA, Gaff HD, White AL, et al. Exploring the Niche of *Rickettsia montanensis* (Rickettsiales:

361   Rickettsiaceae) Infection of the American Dog Tick (Acari: Ixodidae), Using Multiple Species

362   Distribution Model Approaches. J Med Entomol 2021a;58(3):1083–1092.

363

364   Lippi CA, Gaff HD, White AL, et al. Scoping Review of Distribution Models for Selected Ticks

365   and Rickettsial Group Pathogens. PeerJ 2021b;9:e10596.

366

367   Lippi CA, Ryan SJ, White AL, et al. Trends and Opportunities in Tick-Borne Disease

368   Geography. J Med Entomol 2021c;58(6):2021–2029.

369

370   Martin JT, Fischhoff IR, Castellanos AA, et al. Ecological Predictors of Zoonotic Vector Status

371   Among Dermacentor Ticks (Acari: Ixodidae): A Trait-Based Approach. J Med Entomol

372   2022;(tjac125); doi: 10.1093/jme/tjac125.

373

374   McQuiston, J. H., Zemtsova, G., Perniciaro, J., Hutson, M., Singleton, J., Nicholson, W. L., &

375   Levin, M. L. (2012). Afebrile spotted fever group Rickettsia infection after a bite from a

376   *Dermacentor variabilis* tick infected with *Rickettsia montanensis*. Vector-Borne and Zoonotic

377   Diseases, 12(12), 1059-1061.

378

379   Milholland, M. T., Eisen, L., Nadolny, R. M., Hojgaard, A., Machtinger, E. T., Mullinax, J. M., &

380   Li, A. Y. (2021). Surveillance of ticks and tick-borne pathogens in suburban natural habitats of

381   central Maryland. Journal of medical entomology, 58(3), 1352-1362.

382

383   Peterson AT and Soberón J. Species Distribution Modeling and Ecological Niche Modeling:

384   Getting the Concepts Right. Natureza & Conservação 2012;(10):102–107.

385

386   Snellgrove, A.N., I Krapiunaya, P Scott, and N.L. Levin. 2021. Assessment of the Pathogenicity

387   of *Rickettsia amblyommatis*, *Rickettsia bellii*, and *Rickettsia montanensis* in a Guinea Pig Model.

388   Vector Borne Zoonotic Dis; 21(4): 232-241.

389

390   St John HK, Adams ML, Masuoka PM, et al. Prevalence, Distribution, and Development of an

391   Ecological Niche Model of *Dermacentor variabilis* Ticks Positive for *Rickettsia montanensis*.

392   Vector Borne Zoonotic Dis 2016;16(4):253–263.

393

394   Stromdahl, E. Y., Jiang, J., Vince, M., & Richards, A. L. (2011). Infrequency of *Rickettsia*

395   *rickettsii* in *Dermacentor variabilis* removed from humans, with comments on the role of other

396   human-biting ticks associated with spotted fever group Rickettsiae in the United States. Vector-

397   Borne and Zoonotic Diseases, 11(7), 969-977.

398

399   Valavi R, Elith J, Lahoz-Monfort JJ, et al. Modelling Species Presence-Only Data with Random

400   Forests. 2021; doi: 10.1101/2020.11.16.384164.

401

402   Vincent and Hulstrand. Detection of *Rickettsia montanensis* in *Dermacentor variabilis* in

403   Northern Wisconsin. Vector-borne and Zoonotic Diseases 2022; doi: 10.1089/vbz.2022.0055.

404

405   Warren DL, Glor RE and Turelli M. Environmental Niche Equivalency versus Conservatism:

406   Quantitative Approaches to Niche Evolution. Evolution 2008;62(11):2868–2883.

407

408   **Author Contributions**

409   Catherine A. Lippi: Conceptualization, data duration, design of methodology, formal analysis,

410   visualization, writing – original draft, writing – reviewing and editing. Holly D. Gaff:

411   Conceptualization, writing – original draft, writing – reviewing and editing. Robyn M. Nadolny:

412   Data curation, writing – original draft, writing – reviewing and editing. Sadie J. Ryan:

413   Conceptualization, data curation, design of methodology, formal analysis, visualization, writing –

414   original draft, writing – reviewing and editing

415

431