Journal of Medical Entomology, 2017, 1–3
doi: 10.1093/jme/tjx096
Letter to the Editor

Letter to the Editor

Response: The Geographic Distribution of *Ixodes scapularis* (Acari: Ixodidae) Revisited: The Importance of Assumptions About Error Balance

Micah B. Hahn,[1] Catherine S. Jarnevich,[2] Andrew J. Monaghan,[3] and Rebecca J. Eisen[1,4]

[1]Division of Vector-Borne Diseases, Centers for Disease Control and Prevention, 3156 Rampart Road, Fort Collins, CO 80521 (micah.hahn@gmail.com; dyn2@cdc.gov), [2]U.S. Geological Survey, Fort Collins Science Center, 2150 Centre Ave., Bldg. C, Fort Collins, CO 80526 (jarnevichc@usgs.gov), [3]National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307 (monaghan@ucar.edu), and [4]Corresponding author, e-mail: dyn2@cdc.gov

Subject Editor: William Reisen

Received 14 February 2017; Editorial decision 3 March 2017

Dear Editor,

The black-legged tick, *Ixodes scapularis* Say, is the primary vector in the eastern United States of the Lyme disease spirochetes *Borrelia burgdorferi* sensu stricto and *B. mayonii*, as well as *Anaplasma phagocytophilum* (anaplasmosis), *Ehrlichia muris*-like agent (ehrlichiosis), *Babesia microti* (babesiosis), and Powassan encephalitis virus (Goodman et al. 2005; Pritt et al. 2011, 2016). The documented distribution of this medically important tick has expanded substantially over the past two decades, paralleling the increase in the number and geographic distribution of reported human Lyme disease cases in the United States (Kugeler et al. 2015, Mead 2015, Eisen et al. 2016). Although these findings highlight the need for continued vector surveillance, resources for such work are often very limited.

Habitat suitability models can be useful for identifying areas of potential future expansion and thus aid in targeting limited public health resources. To better define the leading edge of the tick's ongoing geographic expansion, recently we published a distribution map that identified counties that were classified as suitable by at least two statistical models, but where *I. scapularis* has not yet been documented to be established (Hahn et al. 2016). Peterson and Raghavan (2017) critiqued the distribution based on the cut-point used to dichotomize a continuous probability into a binary suitability surface. The authors used a simplified version of the Hahn et al. (2016) model to demonstrate a well-known outcome that arises when continuous probabilities are dichotomized into a binary response (Fielding and Bell 1997). That is, when the probability threshold is reduced, a broader geographic area is classified as suitable and model sensitivity increases. Indeed, when Peterson and Raghavan (2017) increased sensitivity from 78% (Hahn et al. 2016) to 99% (1% omission threshold), the predicted distribution of suitable *I. scapularis* habitat increased significantly; most areas east of 98.4° W longitude were considered suitable for *I. scapularis* (Peterson and Raghavan 2017). *Ixodes scapularis*, a primarily woodland-associated tick, is remarkably well-suited to survive a broad range of climatic conditions (Eisen et al. 2015, 2016), and it is likely that if introduced, could survive at low abundance in most of the eastern United States. However, the distribution map presented by Peterson and Raghavan (2017) provides little guidance on which counties are most likely to be suitable for *I. scapularis* to establish.

Peterson and Raghavan (2017) simplified our modeling approach by ignoring forest cover and elevation, two key variables that were highly significant predictors in the models presented by Hahn et al. (2016) and that have consistently been shown to emerge as important predictors of the tick's distribution and abundance (Lane et al. 1991, Killilea et al. 2008, Diuk-Wasser et al. 2010). Furthermore, Peterson and Raghavan (2017) did not show the response curves for their selected variables, making it difficult to interpret their relevance to the biology of *I. scapularis*. In addition, they did not address multicollinearity among predictor variables. We reexamined some of the variables they included and note, for example, that two predictors—minimum temperature of the coldest month (Bio 6) and mean temperature of the coldest quarter (bio 11)—have a correlation coefficient of 0.99 at the county level. Such a high degree of correlation may result in model overfitting. The possibility of biologically inconsistent variables and inclusion of correlated variables in the model presented by Peterson and Raghavan (2017) may explain the nuanced differences in the predicted distributions when compared to the Maxent predictions from Hahn et al. (2016) based on the same dichotomization threshold. Notably, the differences between their model and ours are more apparent in representations with lower sensitivity. In addition, their focus on a single model rather than using an ensemble modeling approach as utilized in Hahn et al. (2016) underrepresents the uncertainty in predictions that arises due to the choice of modeling framework. In agreement with Peterson and Raghavan (2017), we recognize that habitat suitability models are used for multiple purposes and those purposes may be optimized based on different dichotomization thresholds. Therefore, we have reproduced the models from Hahn et al. (2016) and present the *I. scapularis* distribution maps that arise from varying dichotomization thresholds based on model sensitivity. We then show the corresponding changes in specificity that arise when changing sensitivity and display the additional information on uncertainty

Published by Oxford University Press on behalf of Entomological Society of America 2017.

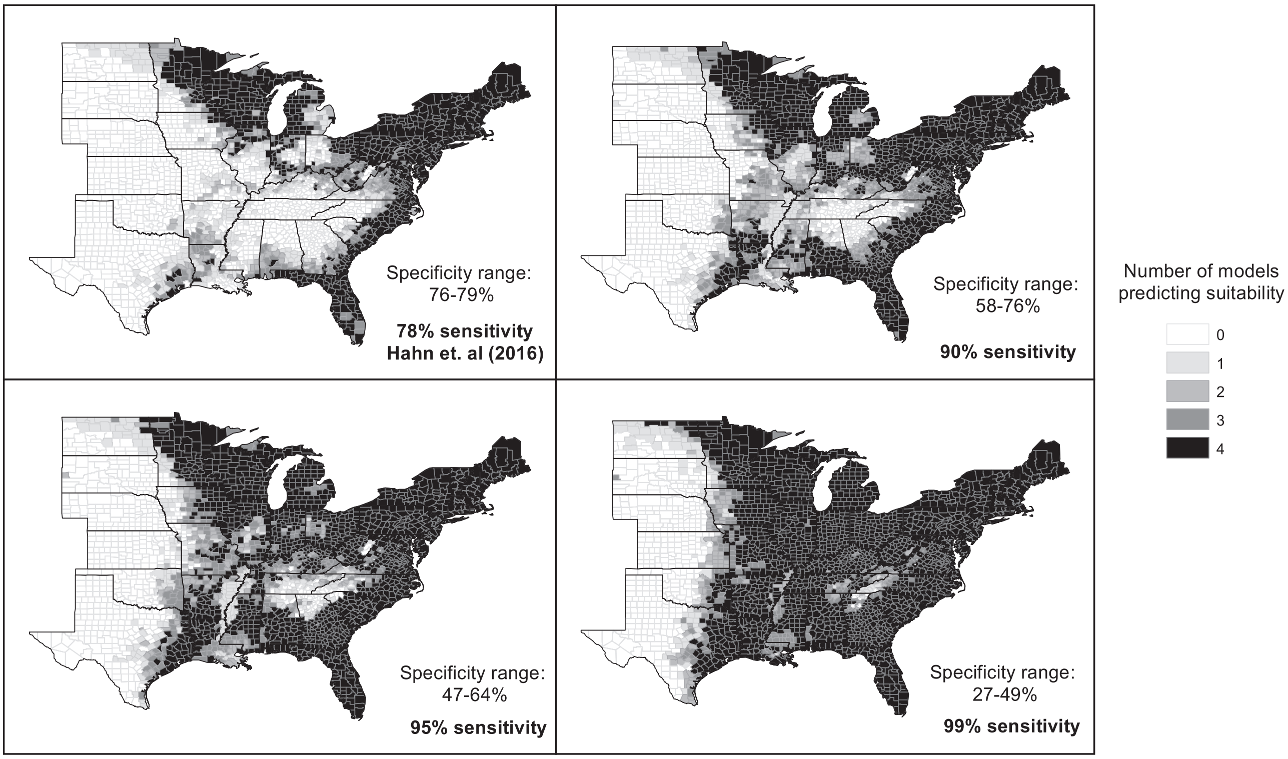

Fig. 1. Maps depicting the *Ixodes scapularis* ensemble models consensus habitat suitability scores at four sensitivity thresholds. Scores indicate the number of the four optimized models (BRT, boosted regression tree; GLM, generalized linear model; MARS, multivariate adaptive regression spline; Maxent, maximum entropy) that classified a given county as having climate, elevation, and land cover conditions suitable for the establishment of *I. scapularis*. The specificity range of the ensemble models at the given sensitivity thresholds are also shown in each map. See Supp. Fig. 1 [online only] for more details.

that arises based on an ensemble modeling framework rather than a single model.

Using the same environmental predictors and methodology as presented in Hahn et al. (2016), we used VisTrails Software for Assisted Habitat Modeling (SAHM, version 2.0; Morisette et al. 2013) to develop habitat suitability models for *I. scapularis* in the eastern United States using the four modeling algorithms retained in the final ensemble model in the original publication: 1) boosted regression tree (BRT), 2) generalized linear model (GLM), 3) multivariate adaptive regression spline (MARS), and 4) maximum entropy (Maxent). After obtaining the continuous relative probability map produced by each of the modeling algorithms, we applied four different probability threshold values to dichotomize the model output onto a binary occurrence surface. The probability thresholds were based on varying the sensitivity of the models. Here we present binary habitat distribution maps for each of the modeling algorithms based on the 78% sensitivity threshold presented in Hahn et al. (2016), as well as maps using 90, 95, and 99% sensitivity thresholds. We also calculated the resulting specificity of the models at each of these sensitivity thresholds.

As expected, across modeling algorithms, as sensitivity (counties classified as suitable with records of established *I. scapularis* populations) increases, the number of counties predicted to be suitable also increases (Fig. 1; Supp. Fig. 1 [online only]). Although confidence in records showing establishment is greater than absences, by increasing sensitivity the number of counties with very low predicted relative probability of established populations are increasingly included in the predicted suitability area. As a result, with increasing sensitivity there is a corresponding increase in the number of false positives (counties classified as suitable, but where the tick has not been reported to be established). For example, both the sensitivity and specificity were ~78% across modeling algorithms in Hahn et al. (2016). Increasing the sensitivity of the model to 99% leads to substantial drops in specificity, ranging from 27 to 49% (Fig. 1). Said another way, the number of counties classified as suitable from which there are no records of established populations increases. While the distribution map showing suitable *I. scapularis* habitat based on a 99% sensitivity threshold may be useful for identifying areas where the tick could become established if introduced, it provides little insight into where to target limited resources for vector surveillance activities aimed at identifying areas with the greatest relative probability of establishment.

Although the predicted distribution of *I. scapularis* habitat using a 99% sensitivity threshold across modeling algorithms (Fig. 1) looks similar to the maps produced by Peterson and Raghavan (2017), there are two discrepancies worth noting. First, our ensemble model does not predict suitable habitat as far west as Peterson and Raghavan (2017), particularly in Nebraska and Kansas. Second, our ensemble model does not consistently predict suitable habitat along the Appalachian mountain chain from northern Georgia up along the border between Tennessee and North Carolina. A likely explanation for the model disagreement in Nebraska and Kansas is that Peterson and Raghavan (2017) did not include forest cover as a predictor in their habitat suitability model, whereas forest cover was highly significant in the Hahn et al. (2016) models. The counties predicted as suitable by Peterson and Raghavan (2017) in western Nebraska and Kansas are dominated by grasslands and contain <10% forest cover. The models in Hahn et al. (2016) show low relative probability of suitable habitat in counties with minimal forest cover. Although Peterson and Raghavan (2017) provided new *I.*

scapularis collection records in Kansas in their publication, we did not include these presence points in our reanalysis because they did not specify the number of ticks or life stages collected, so we could not determine if the surveillance records met the criteria for established populations. Likely causes for low relative probabilities of suitability in the southeastern states were discussed in detail previously (Hahn et al. 2016).

In summary, habitat suitability models are useful tools in public health. Depending on the intended use of the model (e.g., defining the ecological niche or climatic envelope of a species, or allocating resources to areas where the species of interest is most likely to occur), end users may have different tolerances for false negatives or false positives. Additionally, habitat suitability models can be useful for gaining insight on the climatic or ecological factors that determine species distributions. An essential part of this work is interpreting whether the fitted explanatory variables are consistent with what is known about the species in order to verify the biological plausibility of the models. Finally, compared with the use of a single modeling framework, the ensemble modeling approach provides additional information that can be useful for end users. Specifically, areas predicted by all models to be suitable should provide greater confidence of suitability than the prediction from a single model. In an effort to make the models of greater utility to a broad audience, we present our original ensemble models dichotomized based on 78, 90, 95, and 99% sensitivity.

Disclaimer

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References Cited

Diuk-Wasser, M. A., G. Vourc'h, P. Cislo, A. G. Hoen, F. Melton, S. A. Hamer, M. Rowland, R. Cortinas, G. J. Hickling, J. I. Tsao, et al. 2010. Field and climate-based model for predicting the density of host-seeking nymphal *Ixodes scapularis*, an important vector of tick-borne disease agents in the eastern United States. Glob. Ecol. Biogeogr. 19: 504–514.

Eisen, R. J., L. Eisen, and C. Beard. 2016. County-scale distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the continental United States. J. Med. Entomol. 53: 349–386.

Eisen, R. J., L. Eisen, N. Ogden, and C. Beard. 2015. Linkages of weather and climate with *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae), enzootic transmission of *Borrelia burgdorferi*, and Lyme disease in North America. J. Med. Entomol. 53: 250–261.

Fielding, A., and J. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24: 38–49.

Goodman, J. L., D. T. Dennis, and D. E. Sonenshine. 2005. Tick-borne diseases of humans. ASM Press, Washington, DC.

Hahn, M. B., C. S. Jarnevich, A. J. Monaghan, and R. J. Eisen. 2016. Modeling the geographic distribution of *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae) in the contiguous United States. J. Med. Entomol. 53: 1176–1191.

Killilea, M. E., A. Swei, R. S. Lane, C. J. Briggs, and R. S. Ostfeld. 2008. Spatial dynamics of Lyme disease: A review. Ecohealth 5: 167–195.

Kugeler, K. J., G. M. Farley, J. D. Forrester, and P. S. Mead. 2015. Geographic distribution and expansion of human Lyme disease, United States. Emerg. Infect. Dis. 21: 1455–1457.

Lane, R., J. Piesman, and W. Burgdorfer. 1991. *Lyme borreliosis*: Relation of its causative agent to its vectors and hosts in North America and Europe. Annu. Rev. Entomol. 35: 587–609.

Mead, P. S. 2015. Epidemiology of Lyme disease. Infect. Dis. Clin. North Am. 29: 187–210.

Morisette, J., C. Jarnevich, T. Holcombe, C. Talbert, D. Ignizio, M. Talbert, C. Silva, D. Koop, A. Swanson, and N. Young. 2013. VisTrails SAHM: Visualization and workflow management for species habitat modeling. Ecography 36: 129–135.

Peterson, T., and R. Raghavan. 2017. The geographic distribution of *Ixodes scapularis* (Acari: Ixodidae) revisited: The importance of assumptions about error balance. J. Med. Entomol.

Pritt, B. S., L. B. Respicio-Kingry, L. M. Sloan, M. E. Schriefer, A. J. Replogle, J. Bjork, G. Liu, L. C. Kingry, P. S. Mead, D. F. Neitzel, et al. 2016. *Borrelia mayonii* sp. nov., a member of the *Borrelia burgdorferi* sensu lato complex, detected in patients and ticks in the upper midwestern United States. Int. J. Syst. Evol. Microbiol. 66: 4878–4880.

Pritt, B. S., L. M. Sloan, S. M. Johnson, D.K.H. Munderloh, U. G. Paskewitz, K. M. McElroy, J. D. McFadden, M. J. Binnicker, D. F. Neitzel, G. Liu, and W. L. Nicholson. 2011. Emergence of a new pathogenic *Ehrlichia* species, Wisconsin and Minnesota. N. Engl. J. Med. 365: 422–429.