

Ecologic Niche Modeling and Spatial Patterns of Disease Transmission

A. Townsend Peterson*

Ecologic niche modeling (ENM) is a growing field with many potential applications to questions regarding the geography and ecology of disease transmission. Specifically, ENM has the potential to inform investigations concerned with the geography, or potential geography, of vectors, hosts, pathogens, or human cases, and it can achieve fine spatial resolution without the loss of information inherent in many other techniques. Potential applications and current frontiers and challenges are reviewed.

The emerging and evolving field of landscape epidemiology has explored techniques for summarizing spatial patterns in disease transmission data. These techniques seek spatial patterns at some level of generalization or averaging and then summarize overall patterns and trends in the form of a smoothed surface. Techniques typically applied to these challenges include splining and kriging, as well as smoothing based on average values within coarser-grained windows across landscapes (1–3). These approaches always involve some loss of resolution to smooth the surfaces, and some degree of averaging is involved (Figure).

Although these approaches provide simple summaries of spatial patterns, they do not often succeed in illustrating true levels of complexity and heterogeneity that characterize biologic landscapes. Disease transmission cycles are composite phenomena that represent interactions between sets of species: hosts, vectors, and pathogens. The complexities of spatial occurrence of disease will represent the combination of complexities of occurrence of the component species, as well as effects of chance events. Thus, broad-trend generalizations such as those produced using the smoothing techniques mentioned above are unlikely to lead to novel insights and new understanding of complex

systems. The approach advocated in this report improves the pattern summary by estimating species-specific ecologic niches. In this way, the complex influences of environmental variation on species' distributions and their translation into disease transmission patterns can be appreciated in greater detail (Figure).

Ecologic Niche Modeling (ENM)

Joseph Grinnell originated the concept of ecologic niches and was the first to explore the connections between ecologic niches and geographic distributions of species (4). His idea, translated into more modern terminology, was that the ecologic niche of a species is the set of conditions under which the species can maintain populations without immigration of individuals from other areas. A more complete discussion of the concept of ecologic niches and their mapping onto the geographic distributions of species has been provided elsewhere (5).

Use of the ENM approach has grown considerably in the biodiversity community in recent years (6–10). The idea is that known occurrences of species across landscapes can be related to raster geographic information system coverages summarizing environmental variation across those landscapes to develop a quantitative picture of the ecologic distribution of the species. ENM characterizes the distribution of the species in a space defined by environmental parameters, which are precisely those that govern the species' geographic distribution under Grinnell's definition.

A particular strength of ENM is its independence from any particular landscape. ENM can be used to identify potential distributional areas on any landscape: unsampled or unstudied portions of the native landscape, areas of actual or potential invasion by a species with an expanding range, or changing potential distributional areas as a consequence of change (e.g., land use change or climate

*University of Kansas, Lawrence, Kansas, USA

change). Thus, ENM represents a powerful tool for characterizing ecologic and geographic distributions of species across real-world landscapes.

Applications to Disease Systems

In recent years, the ENM approach has seen several prototype applications to disease transmission systems by public health and epidemiology specialists who have been willing to explore novel ideas and approaches. I outline what the technique has to offer to the field and provide citations of example publications for each benefit and use.

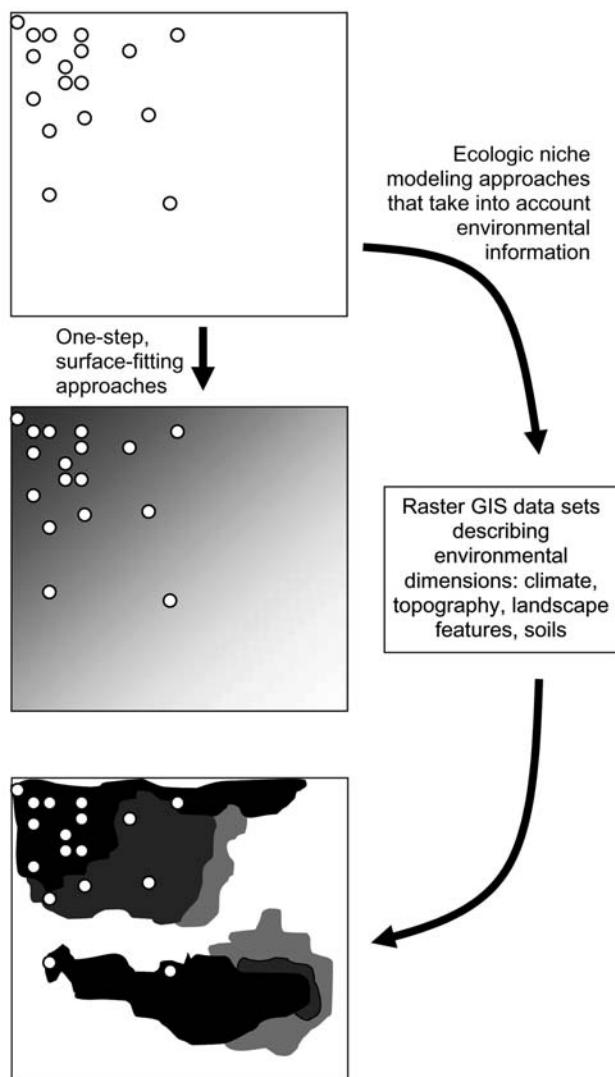


Figure. Hypothetical example of a species' known occurrences (circles) and inferences from that information. The middle panel shows the pattern that would result from a surface-fitting or smoothing algorithm, and the bottom panel shows the ability of ecologic niche modeling approaches to detect unknown patterns in biologic phenomena based on the relationship between known occurrences and spatial patterns in environmental parameters. GIS, geographic information system.

Understanding Ecology of Diseases

In many cases, the details of ecologic parameters associated with occurrences of diseases or of species participating in disease transmission (e.g., vectors, hosts, pathogens) may be unclear because of small sample sizes, biased reporting, or simply lack of detailed geographic or ecologic analysis. ENM encompasses a suite of tools that relate known occurrences of these species or phenomena to raster geographic information system layers that summarize variation in several environmental dimensions. The result is an objective, quantitative picture of how what is known about a species or phenomenon relates to environmental variation across a landscape. Studies using these approaches include an examination of ecologic differences among different Chagas disease vectors in Brazil (11) and a characterization of ecologic features of outbreaks of hemorrhagic fever caused by Ebola and Marburg viruses (12,13).

Characterizing Distributional Areas

A next step in applying ENM approaches to understanding disease systems is characterizing geographic distributions. Here, ENM (or something akin to it) is used to investigate landscapes for areas that meet the ecologic requirements of the species. The result is an interpolation between known sampling locations informed by observed associations between the species and environmental characteristics. Previous attempts to characterize geographic distributions of species in the disease realm have demonstrated the potential of the approach but have not always used the most powerful inferential techniques available (14,15). In at least 1 case (14), the methods used failed to generalize and predict into areas of sparse sampling. ENM produces statistically robust predictions of geographic distributions of species or phenomena (even in unsampled areas), greatly exceeding expectations under random (null) models. Numerous examples of applications of this functionality to disease systems have been published (11–13,16–22).

Identifying Areas of Potential Invasion in Other Regions

ENMs characterize general environmental regimes under which species or phenomena may occur. To the extent that the model is appropriately and correctly calibrated, it may be used to seek areas of potential distribution. Thus, ENMs can be used to identify areas that fit the ecologic bill for a species, even if the species is not present there. This approach has seen extensive experimentation and testing in the biodiversity realm (8,23), but applications to disease transmission have as yet been few. One study attempted to identify the particular species in the *Anopheles gambiae* complex that was responsible for the large-scale South American malaria outbreaks in the

early 20th century (19), and another evaluated the geographic potential of a possible monkeypox host (*Cricetomys* spp.) in North America (24).

Anticipating Risk Areas with Changing Climates

A logical extension of using ENMs to identify potential distributional areas is to address the question of likely geographic shifts in distributional areas of species or phenomena under scenarios of climate change or changing land use (25). This approach has seen considerable attention in the biodiversity realm, with both tests and validations (26–28), and with broad applications across faunas and floras (29–32). In the disease world, applications have been few, although 1 study used likely climate change-mediated range shifts to hypothesize the identity of *Lutzomyia* vectors of recent leishmaniasis outbreaks in southern Brazil (21).

Identifying Unknown Vectors or Hosts

ENM approaches can be applied to various parts of disease transmission cycles (e.g., overall case distribution, reservoir host distribution, vector distribution) to identify unknown elements in systems. The geography of overall case distributions can provide an indication of which clades are potential reservoirs and which are not. A first application was an attempt to identify mammalian hosts of the *Triatoma protracta* group of Chagas disease vectors in Mexico (22), which succeeded in anticipating the mammal hosts of 5 of 5 species for which a test was possible. Further exploration of this possible application of ENM methods has focused on the mysterious long-term reservoir of the filoviruses (Ebola and Marburg viruses) by comparing African mammal distributions with those of filovirus-caused disease outbreaks (33).

Discussion

Current Challenges in ENM

ENM, although it has old roots (4), is nonetheless a relatively new tool in distributional ecology and biogeography. Only a few recent studies have compared the performance of different methodologic approaches under the ENM rubric (34–37). As such, numerous challenges remain in terms of refining approaches toward a more powerful and synthetic methodology.

One central challenge is that of choosing modeling methods appropriate to a particular question, in the sense of discerning interpolation challenges from extrapolation challenges. In a recent comparative study focused on interpolation, which inferred details of patterns of presence and absence on a densely sampled landscape, several techniques that have internal controls on overfitting were superior (34). Extrapolative challenges, such as predicting

potential distribution of invasive species, anticipating species' responses to global climate change, and identifying unknown reservoirs or vectors, require different qualities of modeling algorithms; different methods therefore appear to emerge as superior, according to the particular challenge (5). This balance of ability to interpolate accurately versus ability to extrapolate effectively remains a challenge for the ENM methods.

A second frontier that includes yet-to-be-resolved details for ENM is that of testing and evaluating model results. Currently accepted approaches center on the ability to predict independent test occurrence data in the smallest area predicted (34,38). However, efficient predictions can be poor descriptors of a species' geographic range. Simpler techniques that place greater emphasis on minimizing the omission of known occurrences may be more appropriate. Pairing significance tests (which demonstrate that the coincidence between a prediction and test data is better than that achieved by random or null models) with setting minimum performance criteria (which ensure that that the prediction is accurate enough to meet the needs of the study) is probably the best approach (38). However, these methods have yet to be agreed upon broadly in the ENM community.

Current Challenges in Applications of ENM to Disease Systems

Beyond methodologic challenges, several issues remain to be addressed for full application of ENM methods to disease systems. The first, and perhaps most important, is understanding the role of scale in space and time. Preliminary explorations suggest that proper matching of temporal and spatial scales in analyses may offer particular opportunities for precise and accurate prediction of the behavior of disease phenomena (39). Similarly, proper choice of environmental datasets requires further exploration. Climate data provide longer temporal applicability, but remotely sensed data that summarize aspects of surface reflectance can provide finer spatial resolution, and may measure aspects of ecologic landscapes that climate parameters alone may not capture (40). Such issues will be resolved only through further exploration and testing with predictive challenges for diverse disease systems.

Finally, because disease transmission systems often represent complex interactions among multiple species (e.g., vectors, hosts, pathogens), options exist for how they should be analyzed and modeled. Simple focus on disease occurrences, such as human cases, treats the entire transmission system as a black box and as such gives an overall picture of the ecology of the transmission chain of that disease (12). An alternative, however, is modeling each component species in the transmission system and then assembling the component ENMs into a geographic

picture of the transmission system (22). Each of these approaches has its relative advantages and disadvantages, but a best-practices method has yet to be established, pending further testing and exploration.

Conclusions

The emerging field of ENM applied to questions of ecologic and geographic characteristics of disease systems has considerable potential. In particular, it can solve several problems of spatial resolution of summaries of geographic risk for disease. In sharp contrast to surface-fitting approaches to the same questions, ENM does not lose resolution to generalize and produce a result. Rather, ENM can achieve fine-scale resolution of distributions limited only by the spatial precision of the input occurrence data and the input environmental datasets. This characteristic makes possible a clear improvement in the spatial resolution that is possible in representing spatial patterns in disease risk.

ENM is in the early stages of being explored for its potential for illuminating unknown phenomena in the world of disease transmission. The extensive explorations of ENM in the biodiversity field, however, serve as a benchmark of quality and acceptance for the technique. It can, once tested and prototyped extensively in the disease realm, offer a much-improved representation of spatial patterns in distributions of species or other phenomena.

Acknowledgments

I send many thanks for years of collaboration and education in the world of diseases and their geography to Ben Beard, Janine Ramsey, Jim Mills, Darin Carroll, Karl Johnson, Mark Benedict, Bex Levine, Ken Gage, Rusty Ensore, Erin Staples, Jeffrey Shaw, and Roger Nasci, as well as numerous other colleagues whose omission here is not reflective of my appreciation.

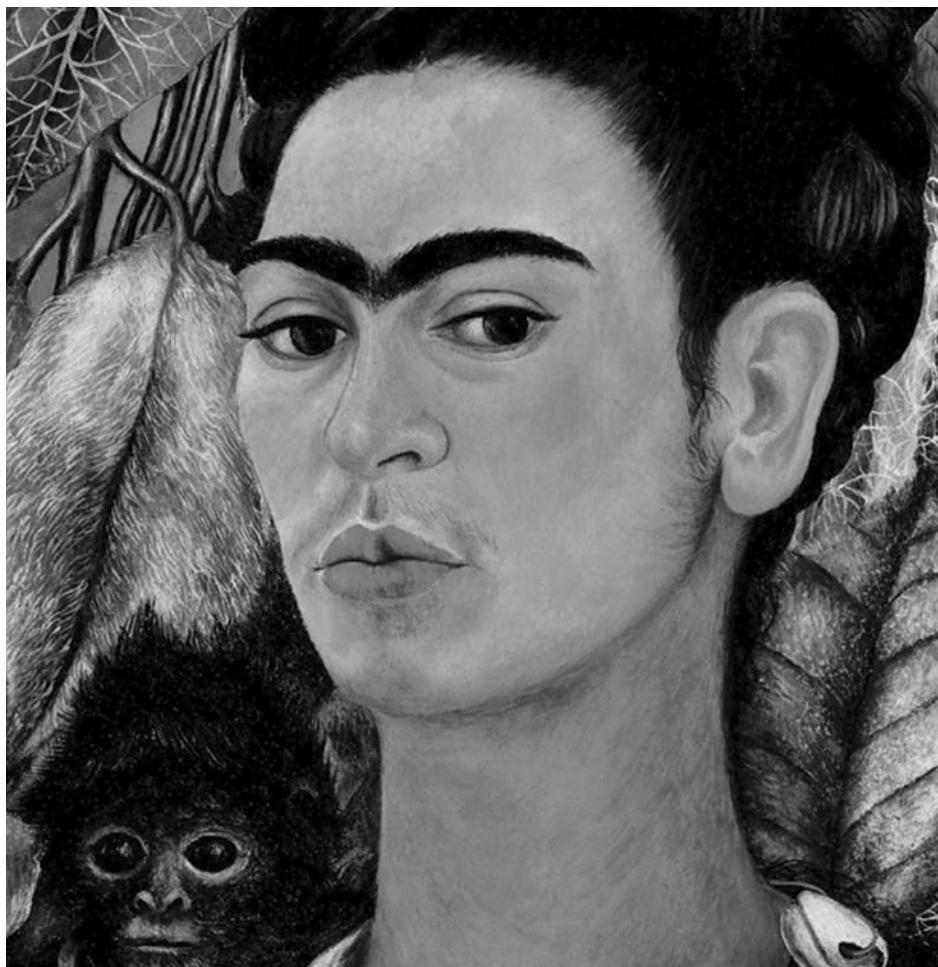
Dr Peterson is professor of ecology and evolutionary biology at the Biodiversity Institute of the University of Kansas. His research interests include many aspects of geographic distributions of species, including the geography and ecology of filoviruses and other disease systems.

References

1. Waller LA, Carlin BP, Xia H, Gelfand AE. Hierarchical spatio-temporal mapping of disease rates. *J Am Stat Assoc.* 1997;92:607–17.
2. Kleinschmidt I, Bagayoko M, Clarke GPY, Craig M, Le Sueur D. A spatial statistical approach to malaria mapping. *Int J Epidemiol.* 2000;29:355–61.
3. MacNab YC, Dean CB. Autoregressive spatial smoothing and temporal spline smoothing for mapping rates. *Biometrics.* 2001;57:949–56.
4. Grinnell J. Field tests of theories concerning distributional control. *Am Nat.* 1917;51:115–28.
5. Soberón J, Peterson AT. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics.* 2005;2:1–10.
6. Austin MP, Nicholls AO, Margules CR. Measurement of the realized qualitative niche: environmental niches of five *Eucalyptus* species. *Ecol Monogr.* 1990;60:161–77.
7. Guisan A, Zimmermann NE. Predictive habitat distribution models in ecology. *Ecological Modelling.* 2000;135:147–86.
8. Peterson AT. Predicting the geography of species' invasions via ecological niche modeling. *Q Rev Biol.* 2003;78:419–33.
9. Wiley EO, McNyset KM, Peterson AT, Robins CR, Stewart AM. Niche modeling and geographic range predictions in the marine environment using a machine-learning algorithm. *Oceanography.* 2003;16:120–7.
10. Soberón J, Peterson AT. Biodiversity informatics: managing and applying primary biodiversity data. *Philos Trans R Soc Lond B Biol Sci.* 2004;359:689–98.
11. Costa J, Peterson AT, Beard CB. Ecological niche modeling and differentiation of populations of *Triatoma brasiliensis* Neiva, 1911, the most important Chagas disease vector in northeastern Brazil (Hemiptera, Reduviidae, Triatominae). *Am J Trop Med Hyg.* 2002;67:516–20.
12. Peterson AT, Bauer JT, Mills JN. Ecologic and geographic distribution of filovirus disease. *Emerg Infect Dis.* 2004;10:40–7.
13. Peterson AT, Lash RR, Carroll DS, Johnson KM. Geographic potential for outbreaks of Marburg hemorrhagic fever. *Am J Trop Med Hyg.* 2006;75:9–15.
14. Rogers DJ, Randolph SE, Snow RW, Hay SI. Satellite imagery in the study and forecast of malaria. *Nature.* 2002;415:710–5.
15. Thomson MC, Elnaïem DA, Ashford RW, Connor SJ. Towards a kala azar risk map for Sudan: mapping the potential distribution of *Phlebotomus orientalis* using digital data of environmental variables. *Trop Med Int Health.* 1999;4:105–13.
16. Sánchez-Cordero V, Peterson AT, Martínez-Meyer E, Flores R. Distribución de roedores reservorios del virus causante del síndrome pulmonar por hantavirus y regiones de posible riesgo en México. *Acta Zoológica Mexicana.* 2005;21:79–91.
17. Peterson AT, Martínez-Campos C, Nakazawa Y, Martínez-Meyer E. Time-specific ecological niche modeling predicts spatial dynamics of vector insects and human dengue cases. *Trans R Soc Trop Med Hyg.* 2005;99:647–55.
18. Lopez-Cardenas J, González-Bravo FE, Salazar-Schettino PM, Gallaga-Solorzano JC, Ramírez-Barba E, Martínez-Mendez J, et al. Fine-scale predictions of distributions of Chagas disease vectors in the state of Guanajuato, Mexico. *J Med Entomol.* 2005;42:1068–81.
19. Levine RS, Peterson AT, Benedict MQ. Geographic and ecologic distributions of the *Anopheles gambiae* complex predicted using a genetic algorithm. *Am J Trop Med Hyg.* 2004;70:105–9.
20. Levine RS, Benedict MQ, Peterson AT. Distribution of *Anopheles quadrimaculatus* Say s.l. and implications for its role in malaria transmission in the US. *J Med Entomol.* 2004;41:607–13.
21. Peterson AT, Shaw JJ. *Lutzomyia* vectors for cutaneous leishmaniasis in southern Brazil: ecological niche models, predicted geographic distributions, and climate change effects. *Int J Parasitol.* 2003;33:919–31.
22. Peterson AT, Sánchez-Cordero V, Beard CB, Ramsey JM. Ecologic niche modeling and potential reservoirs for Chagas disease, Mexico. *Emerg Infect Dis.* 2002;8:662–7.
23. Skov F. Potential plant distribution mapping based on climatic similarity. *Taxon.* 2000;49:503–15.
24. Peterson AT, Papes M, Reynolds MG, Perry ND, Hanson B, Regnery R. Native-range ecology and invasive potential of *Cricetomys* in North America. *Journal of Mammalogy.* 2006;87:427–32.
25. Peterson AT, Tian H, Martínez-Meyer E, Soberón J, Sánchez-Cordero V, Huntley B. Modeling distributional shifts of individual species and biomes. In: Lovejoy TE, Hannah L, editors. *Climate change and biodiversity.* New Haven (CT): Yale University Press; 2005. p. 211–28.

26. Martínez-Meyer E, Peterson AT. Conservatism of ecological niche characteristics in North American plant species over the Pleistocene-to-recent transition. *Journal of Biogeography*. 2006;33:1779–89.
27. Martínez-Meyer E, Peterson AT, Hargrove WW. Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene extinctions and climate change projections for biodiversity. *Global Ecology and Biogeography*. 2004;13:305–14.
28. Araujo MB, Pearson RG, Thuiller W, Erhard M. Validation of species-climate impact models under climate change. *Global Change Biology*. 2005;11:1504–13.
29. Thuiller W, Lavorel S, Araujo MB, Sykes MT, Prentice IC. Climate change threats to plant diversity in Europe. *Proc Natl Acad Sci U S A*. 2005;102:8245–50.
30. Bakkenes M, Alkemade JR, Ihle F, Leemansand R, Latour JB. Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. *Global Change Biology*. 2002;8:390–407.
31. Erasmus BFN, Van Jaarsveld AS, Chown SL, Kshatriya M, Wessels KJ. Vulnerability of South African animal taxa to climate change. *Global Change Biology*. 2002;8:679–93.
32. Peterson AT, Ortega-Huerta MA, Bartley J, Sánchez-Cordero V, Soberón J, Buddemeier RH, et al. Future projections for Mexican faunas under global climate change scenarios. *Nature*. 2002;416:626–9.
33. Peterson AT, Carroll D, Mills JN. Potential mammalian filovirus reservoirs. *Emerg Infect Dis*. 2004;10:2073–81.
34. Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, et al. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*. 2006;29:129–51.
35. Manel S, Dias JM, Ormerod SJ. Comparing discriminant analysis, neural networks, and logistic regression for predicting species distributions: a case study with a Himalayan river bird. *Ecological Modelling*. 1999;120:337–47.
36. Stockwell DR, Peterson AT. Effects of sample size on accuracy of species distribution models. *Ecological Modelling*. 2002;148:1–13.
37. Stockwell DR, Peterson AT. Comparison of resolution of methods used in mapping biodiversity patterns from point occurrence data. *Ecological Indicators*. 2003;3:213–21.
38. Fielding AH, Bell JF. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*. 1997;24:38–49.
39. Peterson AT, Martínez-Campos C, Nakazawa Y, Martínez-Meyer E. Time-specific ecological niche modeling predicts spatial dynamics of vector insects and human dengue cases. *Trans R Soc Trop Med Hyg*. 2005;99:647–55.
40. Roura-Pascual N, Suarez AV, McNyset K, Gómez C, Pons P, Wild TO, et al. Potential geographic distribution, ecological niche differentiation, and fine-scale regional projections for Argentine ants based on remotely-sensed data. *Ecological Applications*. 2006;16 (in press).

Address for correspondence: A. Townsend Peterson, Natural History Museum and Biodiversity Research Center, University of Kansas, Lawrence, KS 66045, USA; email: town@ku.edu



Search
past issues
EID
Online
www.cdc.gov/eid