Ecological niche modelling and understanding the geography of disease transmission

A. Townsend Peterson

Summary
Methods currently used to characterise geographic patterns of disease transmission usually involve loss of resolution and do not take into account the fine-scale ecological variation that underlies transmission patterns. A new suite of tools (ecological niche modelling) that permits fine-scale characterisation of geographic patterns without loss of resolution, and forecasting of invasive potential and effects of changing climate and land use on species' distributions is presented.

Keywords
Disease transmission, Climate change, Ecological niche, Geographic distribution, Land use change, Modelling.

Creazione di modelli di nicchia ecologica e conoscenza della geografia della trasmissione delle malattie

Riassunto
I metodi correntemente utilizzati per caratterizzare il quadro geografico della trasmissione delle malattie di solito comportano una perdita di risoluzione e non tengono conto delle sottili variazioni a livello ecologico che caratterizzano le modalità di trasmissione. Viene presentata una nuova gamma di strumenti (creazione di modelli di nicchia ecologica) che permette una caratterizza-zione su piccola scala di modelli geografici, senza perdita di risoluzione, e consente inoltre una previsione del potenziale invasivo, degli effetti dei mutamenti climatici e dello sfruttamento del territorio sulla distribuzione delle specie.

Parole chiave
Creazioni di modelli, Distribuzione geografica, Mutazioni climatiche, Nicchia ecologica, Sfruttamento del territorio, Trasmissione di malattie.

Introduction

Disease transmission systems consist of interacting species. Each species, at some scale of space and time, is distributed according to its ecological potential. It is constrained by a series of evolutionary adaptations that are generally conceptualised as the ecological niche (11, 12). As such, considerable insight can be gained from a detailed understanding of the ecological niches of each species (57) in a disease transmission system (pathogens, vectors, hosts). The conjunction of these individual ecologies with features of ecological landscapes determines the geography and ecology of the risk of disease transmission.

Methodologies for evaluating the geographic risk of disease transmission (16, 21, 65) have generally focused on the overall distribution of cases as an epiphenomenon (that is, treating the transmission system as a 'black box'), and in geographic space only (34). Both of these assumptions are not without peril. Many situations exist in which the ecological circumstances of transmission are understandable only when the individual ecologies of vector and host organisms are characterised independently (6) (see

Natural History Museum and Biodiversity Research Center, The University of Kansas, Lawrence, Kansas 66045, United States of America
town@ku.edu

hypothetical illustration in Fig. 1). Similarly, analysis in geographic space will only identify broad trends. Such analyses will not generally succeed in detecting the fine-scaled details of disease transmission that depend on local conditions (see small, solid arrow in Fig. 1). As such, a revised methodology that takes into account the details of ecological patterns and that can respond to the particular characteristics of individual species' ecology is desirable. Such a methodology (ecological niche modelling or 'ENM') is presented. ENM has been developed and used in biodiversity studies (56). It focuses on reconstructing the details of species' ecological niches in ecological dimensions; these models in ecological space can then be projected onto geography to hypothesise a potential

geographic distribution. Separation of the phenomenon of geographic distributions of species into ecological and geographic phases offers many advantages for forecasting complex phenomena.

Ecological niche modelling methodology

ENM can take two forms – process-based modelling and empirical reconstructions. The former is based on detailed physiological information on the species that make up the system, whereas the latter is based on associations between known geographic occurrences of species and the ecological characteristics of the landscapes in which they occur. Process-based models have the
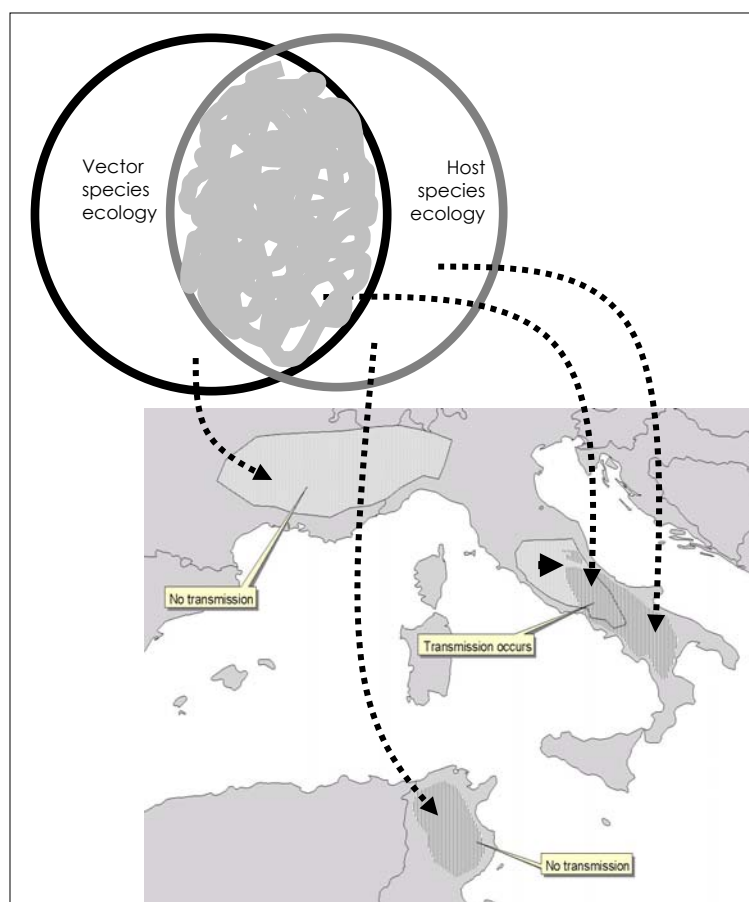

Figure 1
Hypothetical example in which vector and host ecologies must be understood separately for a synthetic understanding of the overall ecology of the transmission system
Areas in which only one of the two interacting species is present are without disease transmission, but where both are present, transmission occurs
Note that the geographic distributions (hypothetical map) are consequences of ecological niches (intersecting circles) and that the only geographic areas in which transmission can occur are a consequence of the joint ecologies of the two species – only where all 'players' in the game are present can transmission occur
The small solid arrow indicates a narrow zone of no transmission

advantage of not being based on sampling inaccuracies (50, 51), but require full knowledge (which is perhaps impossible) of the factors influencing the distributions of the species involved. Empirical reconstructions have the advantage of broad applicability, but can be biased by sampling and the influence of other species not included in the study. Attempts to minimise these biases have been the focus of numerous analyses (1, 4, 14, 22, 32, 39, 63). Empirical reconstructions are the focus of this review because they are much more broadly applicable, particularly to situations that are not well-characterised at the outset.

Empirical ENMs and their geographic predictions are developed in several steps. Known occurrences are compiled and referenced geographically, representing a sample of known occurrences of the species of interest. Although large samples are preferable, ENMs can be developed and tested based on relatively small samples of occurrences, as few as 10-20 points in some cases (31). Digital geographic information system (GIS) layers are assembled to summarise the ecological dimensions in which the species is distributed, and that may (or may not) be involved in limiting its distribution. Examples of such GIS layers include climate, substrate, topography, landform and land surface reflectance.

The relationship between occurrence data and environmental data layers is then characterised. A considerable diversity of methods has been explored, including range-based rules (5, 28), multiple regression and other linear and additive statistical models (13, 17, 27), and distance- and factor-based approaches (5, 14, 15). Perhaps most powerful have been evolutionary-computing approaches in which models can be developed that capture complex relationships between distributions and ecological dimensions with more flexibility and less bias (29, 48, 49, 58, 59).

Few comparative studies among ENM methodologies have been conducted (9, 22, 23), and these studies are in need of careful interpretation. When viewed in a rigorous theoretical framework (57), users must take great care to distinguish between modelling applications that aim to produce general models of species' ecological niches (extrapolation) versus those that aim to reconstruct species' geographic distributions (interpolation).

In disease applications of ENM, several studies have been developed using interpolation-oriented methodologies (52, 62), although not necessarily under the rubric of ENM; however, these approaches do not necessarily focus on fitting a full climatic envelope for the species. Most recent studies have used newer, evolutionary-computing approaches that place greater emphasis on extrapolation and full fitting of niche models (34).

Example applications

Predict distributions

The simplest of ENM applications is to profile species' ecological niches to identify areas matching the niche requirements of the species and to interpolate predictions of likely presence versus absence among sampling points. These applications and the relative performance of different analytical approaches are well-documented and tested (10). Several disease applications of this type have been developed (18, 19, 20, 45, 53).

Characterise ecological requirements

In many cases, for species participating in disease transmission systems, even the basics of their ecological requirements may be unknown. In some cases, even the identity of the species may remain unknown. As such, ENM approaches can be based on what is known (geographic occurrences) to obtain a first characterisation. Several detailed examples have been developed (7, 41, 60).

Predict invasive potential

To the extent that ENM applications manage to fit general, extrapolative models of species' ecological niches, and to the extent that species' ecological niche characteristics are conservative, ENM can be used to predict the geographic potential of species as an invasive species on novel landscapes (33). These methods have been applied amply to the broader diversity of invasive species (36, 55,

64), but applications to disease systems remain relatively few (19, 46).

Predict response to change

ENM can be used to anticipate distributional responses to environmental change (44). To date, most applications have focused on responses of species to global climate change, with numerous analyses now published (2, 26, 30, 37, 61). A few explorations of the effects of land use change have also been developed (47, 54). Applications to disease systems have been scarce (40).

Predict interactions among species

Disease geography can be considered an epiphenomenon resulting from the conjunction of the respective ecologies of each of the participating species. ENM approaches can be used to infer likely species that may be participating in disease transmission systems by choosing species that present appropriate sets of ecological requirements and geographic possibilities. Example applications have been developed for Chagas disease (38) and Ebola and Marburg viruses (42), but much more exploration and experimentation is needed.

Predict dynamics in time and space

Time-specific ENMs that can predict disease transmission at fine scales both in time and in space have been developed. In these applications, remotely-sensed environmental data are essential (8), as they provide information that is specific and detailed in both time and space. For example, it is possible to anticipate where mosquito populations breed in one month based on their ecological characteristics as modelled in another month (43). These ideas, although novel and promising, are in great need of additional testing. The method requires input occurrence data that are rich in representation and fine in resolution both temporally and spatially. As such, very few usable data sets are available.

Frontiers and challenges

An optimistic view of a novel tool that can enable advances in the analysis of the ecology and geography of disease transmission has been presented. The ENM methodology can provide a species-by-species view of disease transmission elements, and offers considerable potential for forecasting complex biodiversity phenomena. Nonetheless, considerable work needs to be performed so that it can mature into a well-characterised and reliable methodology (3). A better understanding of the balance between generality and detail of prediction is essential: must a model offering detailed predictions of presence versus absence necessarily be less general? In spite of recent, intensive comparisons and surveys of ENM methods (10), these questions remain open, and beg careful study. Similarly, methods for evaluation of model predictions need to be refined to match the requirements of general, extrapolative ENMs, rather than optimisation of presence versus absence of predictions that current methods favour. Current methods are only beginning to address these questions (31).

Further testing and experimentation are required to address the limits of conservatism of ecological niches. In spite of extensive demonstration of conservatism in vertebrates and plants (24, 25, 33, 35), the degree to which these conclusions also apply to invertebrate disease vectors has not yet been established. Similarly, further experimentation is needed to establish the limits of time-space disease predictions and detection of interacting species based on ENM analysis of cases and component species. All of these forecasting possibilities need additional testing before they can be considered established methodologies that are ready for use and application to disease transmission systems.

Conclusions

Ecological niche modelling methodologies have several advantages over existing methodologies for summarising spatial patterns of disease transmission and disease risk (16, 21, 65). In particular, ENM does not result in a loss of resolution, in contrast to existing methods. Rather, resolution is limited only by the spatial precision of the occurrence data and of the environmental data sets used to characterise the ecological niche of the species. Therefore, ENM offers considerable

improvements over existing approaches to summarising geographic patterns in disease transmission research.

More generally, ENM offers the possibility of characterising the ecology and geography of a broad diversity of species. Any species for which a modicum of occurrence data is available, and for which sampling is reasonably well-distributed with regard to environmental variation, is tractable with these methods. The result is a detailed picture of the ecological and geographic distributional potential of species in a disease transmission system.

Acknowledgments

References

1. Araújo M.B., Pearson R.G., Thuiller W. & Erhard M. 2005. Validation of species-climate impact models under climate change. *Glob Change Biol*, **11**, 1504-1513.

2. Araújo M.B., Whittaker R.J., Ladle R.J. & Erhard M. 2005. Reducing uncertainty in projections of extinction risk from climate change. *Glob Ecol Biogeogr*, **14**, 529-538.

3. Araújo M.B. & Guisan A. 2006. Five (or so) challenges for species distribution modelling. *J Biogeogr*, **33**, 1677-1688.

4. Austin M.P., Nicholls A.O. & Margules C.R. 1990. Measurement of the realized qualitative niche: environmental niches of five *Eucalyptus* species. *Ecol Monogr*, **60**, 161-177.

5. Carpenter G., Gillison A.N. & Winter J. 1993. DOMAIN: a flexible modeling procedure for mapping potential distributions of animals and plants. *Biodiv Conserv*, **2**, 667-680.

6. Collinge S.K. & Ray C. (eds) 2006. Disease ecology: community structure and pathogen dynamics. Oxford University Press, Oxford, 240 pp.

7. Costa J., Peterson A.T. & Beard C.B. 2002. Ecological niche modeling and differentiation of populations of *Triatoma brasiliensis* Neiva, 1911, the most important Chagas disease vector in north-eastern Brazil (*Hemiptera, Reduviidae, Triatominae*). *Am J Trop Med Hyg*, **67**, 516-520.

8. Egbert S.L., Martínez-Meyer E., Ortega-Huerta M.A. & Peterson A.T. 2002. Use of datasets derived from time-series AVHRR imagery as surrogates for land cover maps in predicting species' distributions. *In* Proceedings IEEE 2002 International Geoscience and Remote Sensing Symposium (IGARSS), **4**, 2337-2339.

9. Elith J. & Burgman M. 2002. Predictions and their validation: rare plants in the Central Highlands, Victoria. *In* Predicting Species occurrences: issues of scale and accuracy (J.M. Scott, P.J. Heglund & M.L. Morrison, eds). Island Press, Washington, DC, 303-313

10. Elith J., Graham C.H., Anderson R.P., Dudik M., Ferrier S., Guisan A., Hijmans R.J., Huettman F., Leathwick J.R., Lehmann A., Li J., Lohmann L.G., Loiselle B.A., Manion G., Moritz C., Nakamura M., Nakazawa Y., Overton J.M., Peterson A.T., Phillips S.J., Richardson K., Scachetti-Pereira R., Schapire R.E., Soberón J., Williams S.E., Wisz M.S. & Zimmermann N.E. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129-151.

11. Grinnell J. 1917. Field tests of theories concerning distributional control. *Am Nat*, **51**, 115-128.

12. Grinnell J. 1924. Geography and evolution. *Ecology*, **5**, 225-229.

13. Guisan A. & Zimmermann N.E. 2000. Predictive habitat distribution models in ecology. *Ecol Mod*, **135**, 147-186.

14. Hirzel A.H., Hausser J., Chessel D. & Perrin N. 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology*, **83**, 2027-2036.

15. Jones P.G. & Gladkov A. 1999. FloraMap, version 1: a computer tool for predicting the distribution of plants and other organisms in the wild. Centro Internacional de Agricultura Tropical, Cali, Colombia.

16. Kleinschmidt I., Bagayoko M., Clarke G.P.Y., Craig M. & Le Sueur D. 2000. A spatial statistical approach to malaria mapping. *Int J Epidemiol*, **29**, 355-361.

17. Lehmann A., Overton J.M. & Leathwick J.R. 2002. GRASP: generalized regression analysis and spatial prediction. *Ecol Mod*, **157**, 189-207.

18. Levine R.S., Benedict M.Q. & Peterson A.T. 2004. Distribution of *Anopheles quadrimaculatus* Say s.l. and implications for its role in malaria transmission in the US. *J Med Entomol*, **41**, 607-613.

19. Levine R.S., Peterson A.T. & Benedict M.Q. 2004. Geographic and ecologic distributions of the *Anopheles gambiae* complex predicted using a genetic algorithm. *Am J Trop Med Hyg*, **70**, 105-109.

20. López-Cárdenas J., González-Bravo F.E., Salazar-Schettino P.M., Gallaga-Solórzano J.C., Ramírez-Barba E., Martínez-Méndez J., Sánchez-Cordero V., Peterson A.T. & Ramsey J.M. 2005. Fine-scale predictions of distributions of Chagas disease vectors in the state of Guanajuato, Mexico. *J Med Entomol*, **42**, 1068-1081.

21. MacNab Y.C. & Dean C.B. 2001. Autoregressive spatial smoothing and temporal spline smoothing for mapping rates. *Biometrics*, **57**, 949-956.

22. Manel S., Dias J.M., Buckton S.T. & Ormerod S.J. 1999. Alternative methods for predicting species distribution: an illustration with Himalayan river birds. *J Appl Ecol*, **36**, 734-747.

23. Manel S., Dias J.M. & Ormerod S.J. 1999. Comparing discriminant analysis, neural networks, and logistic regression for predicting species distributions: a case study with a Himalayan river bird. *Ecol Mod*, **120**, 337-347.

24. Martínez-Meyer E., Peterson A.T. & Hargrove W.W. 2004. Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene extinctions and climate change projections for biodiversity. *Glob Ecol Biogeogr*, **13**, 305-314.

25. Martínez-Meyer E. & Peterson A.T. 2006. Conservatism of ecological niche characteristics in North American plant species over the Pleistocene-to-Recent transition. *J Biogeogr*, **33**, 1779-1789.

26. Midgley G.F., Hannah L., Millar D., Thuiller W. & Booth A. 2003. Developing regional and species-level assessments of climate change impacts on biodiversity in the Cape Floristic Region. *Biol Conserv*, **112**, 87-97.

27. Mladenoff D.J., Sickley T.A., Haight R.G. & Wydeven A.P. 1995. A regional landscape analysis and prediction of favorable gray wolf habitat in the northern Great Lakes region. *Conserv Biol*, **9**, 279-294.

28. Nix H.A. 1986. A biogeographic analysis of Australian elapid snakes. *In* Atlas of elapid snakes of Australia (R. Longmore, ed.). Australian Government Publishing Service, Canberra. 4-15

29. Pearson R.G., Dawson T.P., Berry P.M. & Harrison P.A. 2002. SPECIES: a spatial evaluation of climate impact on the envelope of species. *Ecol Mod*, **154**, 289-300.

30. Pearson R.G. & Dawson T.P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Glob Ecol Biogeogr*, **12**, 361-371.

31. Pearson R.G., Raxworthy C., Nakamura M. & Peterson A.T. 2006. Predicting species' distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *J Biogeogr*, **34**, 102-117.

32. Pearson R.G., Thuiller W., Araújo M.B., Martinez-Meyer E., Brotons L., McClean C., Miles L., Segurado P., Dawson T.P. & Lees D.C. 2006. Model-based uncertainty in species range prediction. *J Biogeogr*, **33**, 1704-1711.

33. Peterson A.T. 2003. Predicting the geography of species' invasions via ecological niche modeling. *Q Rev Biol*, **78**, 419-433.

34. Peterson A.T. 2006. Ecological niche modeling and spatial patterns of disease transmission. *Emerg Inf Dis*, **12**, 1822-1826.

35. Peterson A.T., Soberón J. & Sánchez-Cordero V. 1999. Conservatism of ecological niches in evolutionary time. *Science*, **285**, 1265-1267.

36. Peterson A.T. & Vieglais D.A. 2001. Predicting species invasions using ecological niche modeling. *BioScience*, **51**, 363-371.

37. Peterson A.T., Ortega-Huerta M.A., Bartley J., Sánchez-Cordero V., Soberón J., Buddemeier R.H. & Stockwell D.R.B. 2002. Future projections for Mexican faunas under global climate change scenarios. *Nature*, **416**, 626-629.

38. Peterson A.T., Sánchez-Cordero V., Beard C.B. & Ramsey J.M. 2002. Ecologic niche modeling and potential reservoirs for Chagas disease, Mexico. *Emerg Inf Dis*, **8**, 662-667.

39. Peterson A.T. & Kluza D.A. 2003. New distributional modeling approaches for Gap Analysis. *Anim Cons*, **6**, 47-54.

40. Peterson A.T. & Shaw J.J. 2003. *Lutzomyia* vectors for cutaneous leishmaniasis in southern Brazil: ecological niche models, predicted geographic distributions, and climate change effects. *Int J Parasitol*, **33**, 919-931.

41. Peterson A.T., Bauer J.T. & Mills J.N. 2004. Ecological and geographic distribution of filovirus disease. *Emerg Inf Dis*, **10**, 40-47.

42. Peterson A.T., Carroll D. & Mills J.N. 2004. Potential mammalian filovirus reservoirs. *Emerg Inf Dis*, **10**, 2073-2081.

43. Peterson A.T., Martínez-Campos C., Nakazawa Y. & Martínez-Meyer E. 2005. Time-specific ecological niche modeling predicts spatial dynamics of vector insects and human dengue cases. *Trans Roy Soc Trop Med Hyg*, **99**, 647-655.

44. Peterson A.T., Tian H., Martínez-Meyer E., Soberón J., Sánchez-Cordero V. & Huntley B. 2005. Modeling distributional shifts of individual species and biomes. *In* Climate Change and Biodiversity (T.E. Lovejoy & L. Hannah, eds). Yale University Press, New Haven, Connecticut, 211-228.

45. Peterson A.T., Lash R.R., Carroll D.S. & Johnson K.M. 2006. Geographic potential for outbreaks of Marburg hemorrhagic fever. *Am J Trop Med Hyg*, **75**, 9-15.

46. Peterson A.T., Papeş M., Reynolds M.G., Perry N.D., Hanson B., Regnery R.L., Hutson C.L., Damon I.K. & Carroll D.S. 2006. Invasive potential of *Cricetomys* rats in North America. *J Mammal*, **87**, 427-432.

47. Peterson A.T., Sánchez-Cordero V., Martínez-Meyer E. & Navarro-Sigüenza A.G. 2006. Tracking population extirpations via melding ecological niche modeling with land-cover information. *Ecol Mod*, **195**, 229-236.

48. Phillips S.J., Dudik M. & Schapire R.E. 2004. A maximum entropy approach to species distribution modeling. *In* Proceedings 21st International Conference on machine learning (ICML), 4-8 July, Banff, Alberta. ICML, Banff.

49. Phillips S.J., Anderson R.P. & Schapire R.E. 2006. Maximum entropy modeling of species geographic distributions. *Ecol Mod*, **190**, 231-259.

50. Porter W.P., Budaraju S., Stewart W.E. & Ramankutty N. 2000. Calculating climate effects on birds and mammals: impacts on biodiversity, conservation, population parameters, and global community structure. *Am Zool*, **40**, 597-630.

51. Porter W.P., Sabo J.L., Tracy C.R., Reichman O.J. & Ramankutty N. 2002. Physiology on a landscape scale: plant-animal interactions. *Int Comp Biol*, **42**, 431-453.

52. Rogers D.J., Randolph S.E., Snow R.W. & Hay S.I. 2002. Satellite imagery in the study and forecast of malaria. *Nature*, **415**, 710-715.

53. Ron S. 2005. Predicting the distribution of the amphibian pathogen *Batrachochytrium dendrobatidis* in the New World. *Biotropica*, **37**, 209-221.

54. Sánchez-Cordero V., Illoldi-Rangel P., Linaje M., Sarkar S. & Peterson A.T. 2005. Deforestation and extant distributions of Mexican endemic mammals. *Biol Cons*, **126**, 465-473.

55. Skov F. 2000. Potential plant distribution mapping based on climatic similarity. *Taxon*, **49**, 503-515.

56. Soberón J. & Peterson A.T. 2004. Biodiversity informatics: managing and applying primary biodiversity data. *Phil Trans Roy Soc London B*, **359**, 689-698.

57. Soberón J. & Peterson A.T. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiv Inform*, **2**, 1-10.

58. Stockwell D.R.B. & Noble I.R. 1992. Induction of sets of rules from animal distribution data: a robust and informative method of analysis. *Math Comp Simul*, **33**, 385-390.

59. Stockwell D.R.B. & Peters D.P. 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *Int J Geogr Inf Sci*, **13**, 143-158.

60. Sweeney A.W., Beebe N.W., Cooper R.D., Bauer J.T. & Peterson A.T. 2006. Environmental factors associated with the distribution and range limits of the malaria vector *Anopheles farauti sensu stricto* in Australia. *J Med Entomol*, **43**, 1068-1075.

61. Thomas C.D., Cameron A., Green R.E., Bakkenes M., Beaumont L.J., Collingham Y.C., Erasmus B.F.N., Ferreira de Siqueira M., Grainger A., Hannah L., Hughes L., Huntley B., Van Jaarsveld A.S., Midgely G.E., Miles L., Ortega-Huerta M.A., Peterson A.T., Phillips O.L. & Williams S.E. 2004. Extinction risk from climate change. *Nature*, **427**, 145-148.

62. Thomson M.C., Elnaiem D.A., Ashford R.W. & Connor S.J. 1999. 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*, **4**, 105-113.

63. Thuiller W., Brotons L., Araújo M. & Lavorel S. 2004. Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, **27**, 165-172.

64. Thuiller W., Richardson D.M., Pysek P., Midgley G.F., Hughes G.O. & Rouget M. 2005. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Glob Change Biol*, **11**, 2234-2250.

65. Waller L.A., Carlin B.P., Xia H. & Gelfand A.E. 1997. Hierarchical spatio-temporal mapping of disease rates. *J Am Stat Ass*, **92**, 607-617.