Modelling the ecology and distribution of highly pathogenic avian influenza (H5N1) in the Indian subcontinent

D. Adhikari, A. Chettri and S. K. Barik*

Centre for Advanced Studies in Botany, School of Life Sciences, North-Eastern Hill University, Shillong 793 022, India

The highly pathogenic avian influenza (HPAI) type H5N1 has become a menace in the recent years, causing huge economic losses to the poultry industry. However, a proper understanding of the factors governing its transmission cycle and possible outbreaks in space and time is lacking. The ecological niche modelling tool, Genetic Algorithm for Rule-Set Prediction (GARP) has been used to model the 'black box' of H5N1 transmission in the Indian subcontinent. Ninety outbreak records for veterinary cases of H5N1 in West Bengal and Bangladesh, and 18 environmental parameters with a pixel resolution of ≈4 km were used in the modelling. The final optimized rules generated by GARP were projected on a broad geographical scale covering India, Bangladesh, Nepal and many parts of Pakistan to anticipate the potential areas for the presence of H5N1. The niche model predicted vast areas in the Indo-Bangladesh region for the presence of H5N1. The areas having high level of model agreements are in the Gangetic Plains, Assam, Tripura, West Bengal and some parts of Andhra Pradesh, Maharashtra and Gujarat. The model also predicted large areas in northern Pakistan and some parts of Nepal to be suitable for H5N1 transmission. The modelled ecological niche of H5N1 in the present study is characterized by a mean annual temperature range of 21–26°C, lower slope angle, areas with a tendency to pool water, a significant monthly variation in the degree of greenness and human population density of more than 100 persons/km^2. The resultant niche model validated 49 cases of HPAI outbreak out of the total 52 past as well as the latest outbreaks in India, Bangladesh, Pakistan and Nepal.

Keywords: Avian influenza virus, ecological niche, H5N1 virus, Indian subcontinent.

TRANSMISSION of disease across a spatial and temporal scale can be attributed to the overlapping ecological niches of the vector, host and pathogens, and their associated ecological landscape[1]. It has been established that the vector-borne diseases like chagas, malaria, dengue and West Nile virus are transmitted due to such niche overlapping[2–6]. However, in the absence of a potential vector, the widespread transmission of a disease in a geographic domain becomes difficult to understand. The highly pathogenic avian influenza (HPAI–H5N1) is one such disease for which apparently there is no vector, except a report describing two blow fly species, namely *Calliphora nigribarbis* and *Aldrichina grahami*, which are suspected[7] to be mechanical transmitters for H5N1. More recently, a mosquito species, namely *Culex tritaeniorhynchus* tested positive for H5N1, although its role as a potential vector has not been evaluated[8]. The disease has become a menace in recent times, causing death of millions of poultry birds and leading to huge economic losses[9]. Reports on sporadic human cases of H5N1 have aroused concerns on triggering of a probable global influenza pandemic[10,11].

The highly pathogenic strain H5N1 is a subtype of the influenza A virus belonging to the family Orthomyxoviridae. The strain is endemic to some parts of Asia and has established a niche in poultry[12,13]. It is reported to have multiple sublineages in Vietnam[14]. The wild migratory birds, considered to be the natural reservoir of all influenza A viruses, are also dying due to H5N1 infection[13,15]. Although aggressive control measures are being taken, including culling of more than 140 million poultry, outbreaks have recurred in different parts of the world[13,16].

The wild migratory aquatic birds are suspected to be the prime agents of geographic-scale spread of the pathogen at global level[15,17], while the infected poultry and ducks are important agents at the landscape level[17,18]. However, explanations on actual transmission of the virus are generally suggestive. For instance, water is argued to be an effective medium for the transmission of the avian influenza virus because of its high prevalence among the aquatic wild birds that congregate in the wetlands[15]. The transmission of the virus takes place through direct contact with the infected birds or by using the bird products like faeces for fertilizer and aquaculture[17]. Other studies have reported strong influences of environmental factors on the transmission of avian influenza. For example, the transmission of the HPAI in Nigeria and West Africa has been linked to differences in plant phenology and land-surface reflectance[19]. Studies in Thailand and Vietnam revealed that the transmission cycle of the pathogen is strongly associated with duck farming, paddy cultivation and human population density[20,21]. Thus, our understanding of the geographical spread and factors responsible for the transmission of HPAI virus remains poor.

The number of HPAI outbreaks in the Indian subcontinent has increased over the years, causing huge economic losses[16]. Its rapid transmission in the subcontinent has been attributed to a number of causes, which are based on wild speculations without having any credible scientific basis. Thus, the transmission cycle of the virus largely remains as an incomprehensible 'black box'. In the present study, an attempt has been made to understand and model the ecological niche of the HPAI (H5N1) considering the transmission cycle as a 'black box'. The study would help in predicting the areas having potential for its

---

*For correspondence. (e-mail: sarojkbarik@gmail.com) presence in the Indian subcontinent, which would help in taking necessary precautionary measures against any future outbreak.

Considering the uncertainties associated with the transmission of the HPAI (H5N1), ecological niche modelling (ENM) was used that follows an empirical approach for niche reconstruction[1]. The technique was employed to model the potential geographic distribution of many infectious diseases earlier[22–24]. The ecological niche of the 'black box' of HPAI (H5N1) transmission may be defined as 'a set of environmental conditions in which the HPAI (H5N1) virus is capable of maintaining its number by infecting its hosts without immigration' (after ref. 25). The modelling approach in the present study used genetic algorithm as a tool to model the ecological niche of the 'black box' of HPAI (H5N1) transmission. Genetic algorithms are designed to solve problems following natural selection to get an optimized solution for a particular problem from a population of solutions. The algorithm first starts with a set of solutions for a particular problem. The best solutions then reproduce a new generation of solutions. Crossovers and mutations occur in the process of reproduction and the bad solutions are eliminated. Genetic Algorithm for Rule-Set Prediction (GARP) is a variant of the family of genetic algorithms, which has gained considerable importance because of its ability to solve problems related to species distribution using different solutions termed as rule types, i.e. atomic rule, range rule, negated range rule and logistic regression rule in an iterative and evolutionary fashion[26]. These rules are subjected to the process of natural selection of crossovers and mutations and the optimized rules are used in determining the species distribution. The ecological niche modelling in the present study was carried out using a desktop version of GARP, which was freely downloaded[27] from www.nhm.ku.edu/desktopgarp.

Ninety records for the occurrence of H5N1 in West Bengal and Bangladesh were collected through comprehensive search of news reports, field visits and from the websites of the Ministry of Agriculture, Department of Animal Husbandry, Dairying and Fisheries, Government of India (http://dahd.nic.in/birdflue.htm); the Ministry of Fisheries and Livestock, Government of the People's Republic of Bangladesh (http://www.mofl.gov.bd/), and the World Organization for Animal Health (OIE) (http://www.oie.int/downld/AVIAN%20INFLUENZA/A_AI-Asia.htm; Figure 1). The geographic coordinates of the locations of outbreaks were collected from www.fallingrain.com, as well as from field visits. These 90 records were used for the development of the model. For independent validation of the model, 52 additional independent outbreak records in different parts of the subcontinent, i.e. India, Pakistan, Bangladesh and Nepal were downloaded from the website of the World Organization for Animal Health. These records were not used as model training or testing records.

Eighteen environmental parameters pertaining to the Indian subcontinent were used for modelling. These are raster data on altitude and mean annual temperature with a spatial resolution of 2.5 arc min (http://www.worldclim.org)[28]; Normalized Difference Vegetation Index (NDVI) layers for January to December derived from AVHRR; slope, aspect, compound topographic index downloaded from the USGS Hydro-1K dataset[29], and raster data on average global human population per square kilometre with a resolution of 2.5 arc min (http://sedac.ciesin.columbia.edu/gpw). All the environmental layers were resampled to a spatial resolution of 0.04° using the spatial analyst module of ArcView 3.3. Coarsening of the resolution of environmental layers was done to match with the approximate spatial resolution of the occurrence data.

Prior to predicting the areas having potential for the presence of H5N1, the fitness of the ecological niche model was evaluated following a series of calibration tests, wherein different subsets of the entire area of occurrence were subjected to reciprocal predictivity tests[19]. To enable this, the occurrence records were divided into quadrants above and below the median latitude and median longitude. The occurrence records were then divided into subsets of north and south of the median latitude, west and east of the median longitude and on- and off-diagonal.

For building different subset models of the entire area of occurrence, an algorithm threshold of 0.01 was selected with 1000 iterations as an upper limit for each run. Then 50% of the total occurrence points were used to generate the rule and the remaining 50% was used for intrinsic testing for model optimization and refinement.

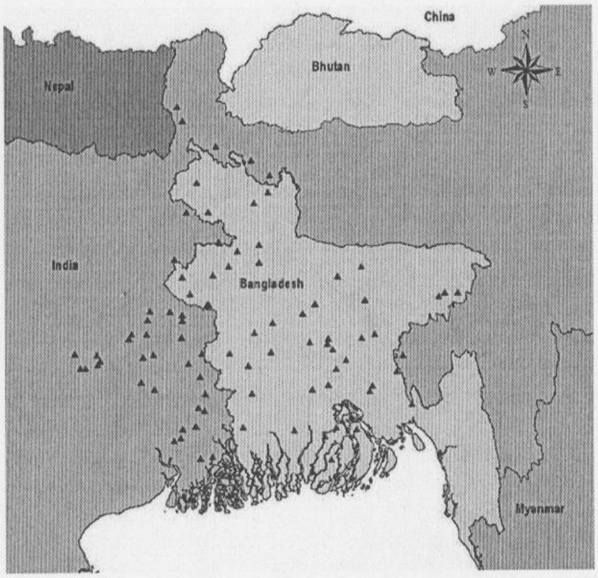

Figure 1. Locations of the 90 HPAI outbreaks in India and Bangladesh used for model calibration and generation of final rules.

Table 1. Results of model calibration test performed at all threshold levels (1–10)

Subset test	Sample size Train	 Test	No. of test points correctly predicted	Per cent of test area predicted present at all threshold levels (1–10)	Cumulative binomial probability
North-south					
North predicts south	33	37	35	52.36	<0.001
South predicts north	37	33	30	53.90	<0.001
West-east					
West predicts east	46	42	38	44.88	<0.001
East predicts west	42	46	42	57.76	<0.001
Diagonals					
On predicts off	25	23	10	25.91	0.66
Off predicts on	23	25	25	80.06	<0.001

Due to the stochastic nature of GARP in giving different outputs at different runs, the model was set to perform 100 tasks. For selecting the 10 best subset models, the in-built best subset selection radio button in Desktop GARP was checked with threshold level of 0% extrinsic hard omission and 50% commission. All the 10 best subset models were summed up using the calcgridsinview.ave script in ArcView 3.3, to create a composite distributional representation of the ecological niche of the 'black box' of HPAI (H5N1) transmission. Independent validation of the model was done by overlaying the occurrence points with the model projections. To test the predictive accuracy of the subset models, binomial probability test was performed[30].

After model calibration, the final predictions were made by projecting the rules to the Indian subcontinent covering India, Bangladesh, Nepal and some parts of Pakistan. Overall, 90 occurrence records were used in the final analysis, of which 45 were used for model building and the remaining 45 were used for model validation. As the final model projections were made in a much broader unsampled area, GARP was set to perform 500 tasks with a threshold level of 0% extrinsic hard omission and 50% commission.

To assess the predictive ability of the ecological niche model in a broader geographical range, the final country-scale projection was subjected to independent validation test. This was done by overlaying 52 independent outbreak locations on the final model projection.

To visualize the ecological niche of HPAI (H5N1) in a two-dimensional environmental space, the grid of the composite ecological niche model was combined with the original environmental data layers using the combine grids option in ArcVew 3.3. The combined grids were then exported in ASCII format. A subset of the combined grids was then imported into MS EXCEL for visualization. The entire modelling exercise was performed on a DELL Precision 690 WORKSTATION (X64 Edition).

The results of the model calibration tests performed at threshold levels of 1–10 are summarized in Table 1. The northern occurrence points predicted 52.4% area of the south to have potential for presence and correctly predicted 94.0% of the test points. The southern occurrences predicted 53.9% area of the northern side to have potential for presence and correctly predicted 91.0% of the test points.

Similarly, the HPAI (H5N1) occurrence points in the west predicted 44.9% area of the eastern side to be suitable for presence and correctly predicted 90.0% of the test points. The eastern occurrence points predicted 57.8% area of the west to be suitable and independently predicted 91.0% of the test points.

In the diagonal tests, the on-diagonal predicted only 25.9% area of the off-diagonal to be suitable and correctly predicted only 43.0% of the test points. However, the off-diagonal predicted 80.0% of the on-diagonal area to be suitable and correctly predicted all the test points.

The results of the cumulative binomial probability tests for model calibration revealed that five subsets out of six, correctly predicted the independent test points as evident from significant level of binomial probability ($P < 0.001$). However, the prediction through on-diagonal predicts off-diagonal subset did not yield a significant level of probability (Table 1).

The ecological niche model could independently predict 49 out of 52 outbreaks in the Indian subcontinent between the years 2006 and 2009 (Table 2). All the outbreaks in Bangladesh, Nepal and the northwestern frontier of Pakistan were successfully predicted. For India, the model could predict 19 out of 22 outbreaks (Figure 2). The three outbreaks which remained beyond the prediction of the model were located one each in Maharashtra, Sikkim and Assam.

The country-scale prediction for the potential areas of transmission of HPAI (H5N1) in India, Bangladesh, Nepal and major portion of Pakistan revealed broad dispersion of the pathogen (Figure 2). The prediction shows that the Gangetic region and West Bengal have the highest probability of HPAI transmission, as evident from the higher model agreements. Besides, substantial areas in

Table 2. Validation of the ecological niche model using independent test points across spatial and temporal scale

	Outbreak records during different years used for model validation					Outbreaks correctly predicted				
Country	2006	2007	2008	2009	Total	2006	2007	2008	2009	Total
India	2	1	16	3	22	1	1	15	2	19
Bangladesh	–	–	7	4	11	–	–	7	4	11
Pakistan	10	8	–	–	18	10	8	–	–	18
Nepal	–	–	–	1	1	–	–	–	1	1
Total	12	9	23	8	52	11	9	22	7	49

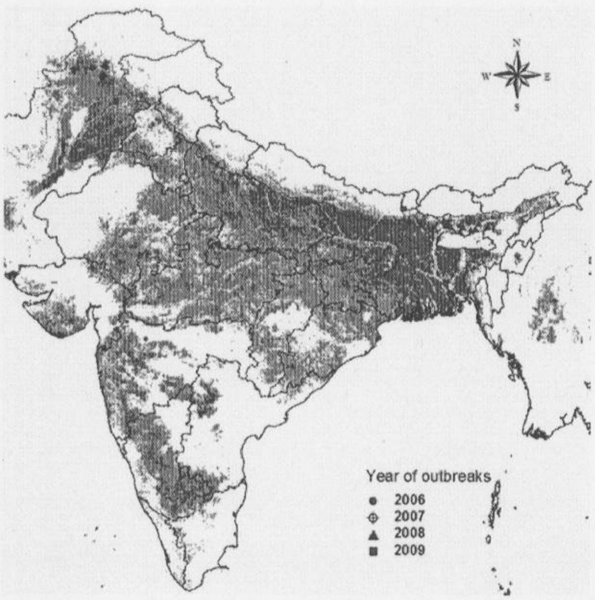

Figure 2. Geographic projection of the final rules generated by the GARP model in India, Bangladesh, Nepal and some parts of Pakistan. Grey colour depicts the areas of agreement of (1–5) models and red colour depicts agreement of (6–10) models. White colour depicts the area where all 10 best subsets predict absence. Symbols in the map denote 52 locations of HPAI outbreak during different years. These locations were used as independent test points for final validation of the model.

the northern and central India also have necessary elements for transmission of HPAI. The model also shows some areas in Andhra Pradesh, Maharashtra and Gujarat that may be suitable for HPAI transmission.

In northeastern India, some parts of Assam, Tripura and Manipur possess suitable elements for HPAI transmission. The model predicted the waterlogged areas in Assam, Tripura and Manipur to be suitable for transmission of HPAI virus. The mountainous areas of northeastern India indicate absence of these elements.

For Bangladesh, the model predicts a broad distributional range of the virus that spans across the country, indicating the extensive presence of the elements necessary for its transmission. The confirmed reports on the outbreak of avian influenza from 47 out of 64 districts in Bangladesh (http://www.mofl.gov.bd/) corroborate the model predictions.

The exploratory visualizations depict the ecological characteristics of the modelled 'black box' of the H5N1 transmission in the training landscape of Indo-Bangladesh region (Figures 3 and 4). The visualizations depict that the niche of H5N1 is restricted to areas having lower slope angle with a tendency to pool water, higher annual variation in the degree of greenness and a mean annual temperature range of 21–26°C. The combined observations on slope angle, topographic index and monthly variation in greenness index suggested specific land-use types, e.g. stagnant water bodies, water courses, and wetlands, including paddy fields, that were predicted as areas of occurrence of the pathogen. Conforming to the model prediction, the training landscapes in West Bengal and Bangladesh have extensive areas under water courses, stagnant water bodies, wetlands and agricultural lands which constitute more than 50% of the total geographical area[31,32]. The predicted areas of occurrence were also characterized by seasonal fluctuations in the greenness index due to variations in cultivation practice and cropping pattern (Figure 4). Gilbert *et al.*[20,21] observed that agricultural land use and rice crop intensity strongly drive duck husbandry, and the abundance of free-grazing ducks in intensive rice crop areas could be a potential factor for transmission of H5N1. The model training landscape also has abundant wetlands and water bodies such as shallow ponds, which are seasonal and get dried up during winter and summer months. These habitats favour aggregation of wild fowl and domestic ducks for foraging on benthic fauna, facilitating the transmission of HPAI virus.

Animal husbandry and livestock production contribute substantially to the household income in West Bengal and Bangladesh, where farm and backyard poultry are important components of the livestock production[33]. Unlike the farm poultry, backyard poultry are free-ranging and are generally reared under a scavenging system with feed available from household wastes, homestead pickings, crop residues in agricultural fields, wetlands and water bodies. In the process, they come in close contact with the wild migratory birds exposing them to the H5N1 transmission. Human population density could be another surrogate element for the active local transmission of

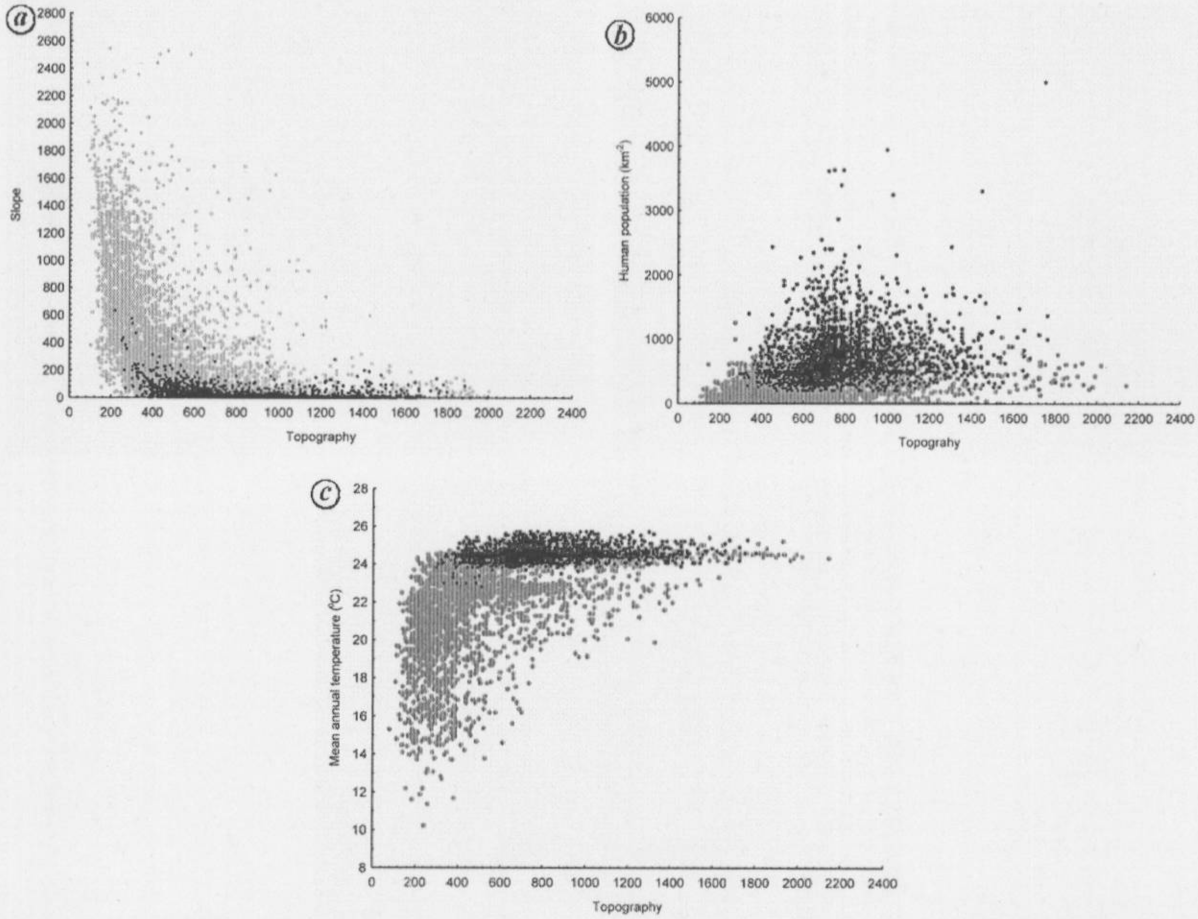

Figure 3. Exploratory visualization of H5N1 niche in two-dimensional environmental space. Grey dots indicate the available environmental conditions and black dots indicate the modelled ecological niche of the virus.

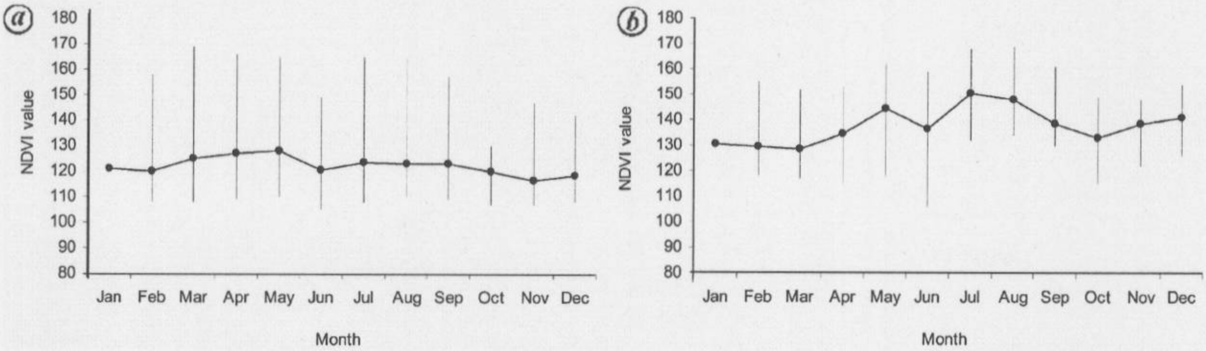

Figure 4. Comparison of mean monthly NDVI values of 100 randomly selected pixels in the Indo-Bangladesh region. Solid line depicts the variation in mean monthly NDVI values of (*a*) predicted absence (annual variance = 11.80) and (*b*) predicted presence (annual variance = 50.07). Vertical lines in both the figures represent the range of NDVI values around the mean.

HPAI virus in the model training landscape. Confirming this proposition, the exploratory visualization revealed that human population with density >100 persons/km^2 is susceptible to H5N1 transmission (Figure 3). Higher human population density enhances the chance of greater trade interactions, increasing the probability of HPAI transmission through trading of infected poultry[34].

The present study shows that the ecological niche model generated by GARP performed significantly better than random expectations, and was successful in capturing the coarse-scale ecological conditions associated with HPAI transmission in the Indo-Bangladesh region. The land-use pattern, occurrence of seasonal wetlands, practices of backyard poultry and animal husbandry, and density of human population were identified as risk factors for the widespread transmission of H5N1 virus. The projection of the final rules generated from the training landscape in the Indo-Bangladesh region could successfully predict the outbreaks in various parts of India, Nepal and Pakistan from a spatial perspective. However, further studies are required for predicting the time-specific outbreaks. Nevertheless, the present model could be used as a guide map for possible future outbreaks in other parts of India.

---

1. Peterson, A. T., Ecological niche modeling and understanding the geography of disease transmission. *Vet. Ital.*, 2007, **43**, 393–400.
2. Peterson, A. T., Sánchez-Cordero, V., Ben Beard, C. and Ra, J. M., Ecologic niche modeling and potential reservoirs for Chagas disease, Mexico. *Emerg. Infect. Dis.*, 2002, **8**, 662–667.
3. Peterson, A. T., Martínez-Campos, C., Nakazawa, Y. and 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**, 647–655.
4. Moffett, A., Shackelford, N. and Sarkar, S., Malaria in Africa: Vector species' niche models and relative risk maps. *PLoS ONE*, 2007, **9**, 1–14.
5. Peterson, A. T., Robbins, A., Restifo, R., Howell, J. and Nasci, R., Predictable ecology and geography of West Nile virus transmission in the central United States. *J. Vector Ecol.*, 2008, **33**, 342–352.
6. Wimberly, M. C., Hildreth, M. B., Boyte, S. P., Lindquist, E. and Kightlinger, L., Ecological niche of the 2003 West Nile virus epidemic in the northern great plains of the United States. *PLoS ONE*, 2008, **3**, 1–7.
7. Sawabe, K. *et al.*, Detection and isolation of highly pathogenic H5N1 avian influenza A viruses from blow flies collected in the vicinity of an infected poultry farm in Kyoto, Japan, 2004. *Am. J. Trop. Med. Hyg.*, 2006, **75**, 327–332.
8. Barbazan, P. *et al.*, Detection of H5N1 avian influenza virus from mosquitoes collected in an infected poultry farm in Thailand. *Vector-Borne Zoonotic Dis.*, 2008, **8**, 105–109.
9. Raynor, M. S. and Cliff, A. D., The geographical spread of avian influenza A (H5N1): panzootic transmission (December 2003–May 2006), pandemic potential, and implications. *Ann. Assoc. Am. Geogr.*, 2008, **98**, 553–582.
10. Hien, T. T. *et al.*, Avian influenza A (H5N1) in 10 patients in Vietnam. *New Engl. J. Med.*, 2004, **350**, 1179–1188.
11. Ungchusak, K. *et al.*, Probable person-to-person transmission of avian influenza A (H5N1). *New Engl. J. Med.*, 2005, **352**, 333–340.
12. Li, K. S. *et al.*, Genesis of a highly pathogenic and potentially pandemic H5N1 influenza virus in eastern Asia. *Nature*, 2004, **430**, 209–213.
13. World Health Organization, Responding to the avian influenza pandemic threat – Recommended strategic actions, WHO, Geneva, 2005.
14. Nguyen, T. D., Nguyen, T. V., Vijaykrishna, D., Webster, R. G., Guan, Y., Peiris, J. S. M. and Smith, J. D., Multiple sublineages of influenza A (H5N1), Vietnam, 2005–2007. *Emerg. Infect. Dis.*, 2008, **14**, 632–636.
15. Causey, D. and Edwards, S. V., Ecology of avian influenza in birds. *J. Infect. Dis.*, 2008, **197**, 29–33.
16. http://www.oie.int/downld/AVIAN%20INFLUENZA/A_AI-Asia.htm
17. Kilpatrick, A. M., Chmura, A. A., Gibbons, D. W., Fleischer, R. C., Marra, P. P. and Daszak, P., Predicting the global spread of H5N1 avian influenza. *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 19368–19373.
18. Gilbert, M. *et al.*, Free-grazing ducks and highly pathogenic avian influenza, Thailand. *Emerg. Infect. Dis.*, 2006, **12**, 227–234.
19. Williams, R. A. J., Fasina, F. O. and Peterson, A. T., Predictable ecology and geography of avian influenza (H5N1) transmission in Nigeria and West Africa. *Trans. R. Soc. Trop. Med. Hyg.*, 2008, **102**, 471–479.
20. Gilbert, M., Xiao, X., Chaitaweesub, P., Kalpravidh, W., Premashthira, S., Boles, S. and Slingenbergh, J., Avian influenza, domestic ducks and rice agriculture in Thailand. *Agric. Ecosyst. Environ.*, 2007, **119**, 409–415.
21. Gilbert, M. *et al.*, Mapping H5N1 highly pathogenic avian influenza risk in Southeast Asia. *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 4769–4774.
22. Peterson, A. T., Lash, R. R., Carroll, D. S. and Johnson, K. M., Geographic potential for outbreaks of Marburg hemorrhagic fever. *Am. J. Trop. Med. Hyg.*, 2006, **75**, 9–15.
23. Peterson, A. T., Bauer, J. T. and Mills, J. N., Ecological and geographic distribution of filovirus disease. *Emerg. Infect. Dis.*, 2004, **10**, 40–47.
24. Levine, R. S., Peterson, A. T., Yorita, K. L., Carroll, D., Damon, I. K. and Reynolds, M. G., Ecological niche and geographic distribution of human monkeypox in Africa. *PLoS ONE*, 2007, **1**, 1–7.
25. Grinnell, J., Field tests of theories concerning distributional control. *Am. Nat.*, 1917, **51**, 115–128.
26. Stockwell, D. R. B. and Peters, D., The GARP modelling system: problems and solutions to automated spatial prediction. *Int. J. Geogr. Inf. Syst.*, 1999, **13**, 143–158.
27. Scachetti-Pareira, R., Desktop GARP v.1.1.4. The University of Kansas Biodiversity Research Group, Kansas, USA, 2001.
28. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. and Jarvis, A., The WorldClim interpolated global terrestrial surfaces. Version 1.3, 2004.
29. United States Geological Survey, HYDRO1k elevation derivative database. United States Geological Survey, Washington DC, 2001.
30. Anderson, R. P., Lew, D. and Peterson, A. T., Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecol. Model.*, 2003, **162**, 211–232.
31. www.wbgov.com/e-gov/admin/newgovtpublications/upload/Pub-status_on_land.pdf
32. ASB Land use. In *Banglapedia: National Encyclopedia of Bangladesh*, Asiatic Society of Bangladesh, Dhaka, 2003, vol. 6, pp. 235–239.
33. FAO, Poultry sector country review. FAO Animal Production and Health Division, Food and Agriculture Organization of the United Nations, Rome, 2008.
34. Clerc, M. G., Lebarbenchon, C. and Thomas, F., Recent expansion of highly pathogenic avian influenza H5N1: a critical review. *Ibis*, 2007, **149**, 202–214.

ACKNOWLEDGEMENTS. We are grateful to the organizers of the training programme on Ecological Niche Modelling at Bangalore during 13–17 December 2006, where the first author was a participant. We thank Dr Mohammed Irfan-Ullah for valuable suggestions while preparing exploratory visualizations of the ecological niche models. We also thank the anonymous referee for useful comments. Financial assistance to the first author in the form of Research Associateship by the UPE (Biosciences) Programme under the School of Life Sciences, North-Eastern Hill University, Shillong, is acknowledged.

Received 3 March 2008; revised accepted 22 April 2009