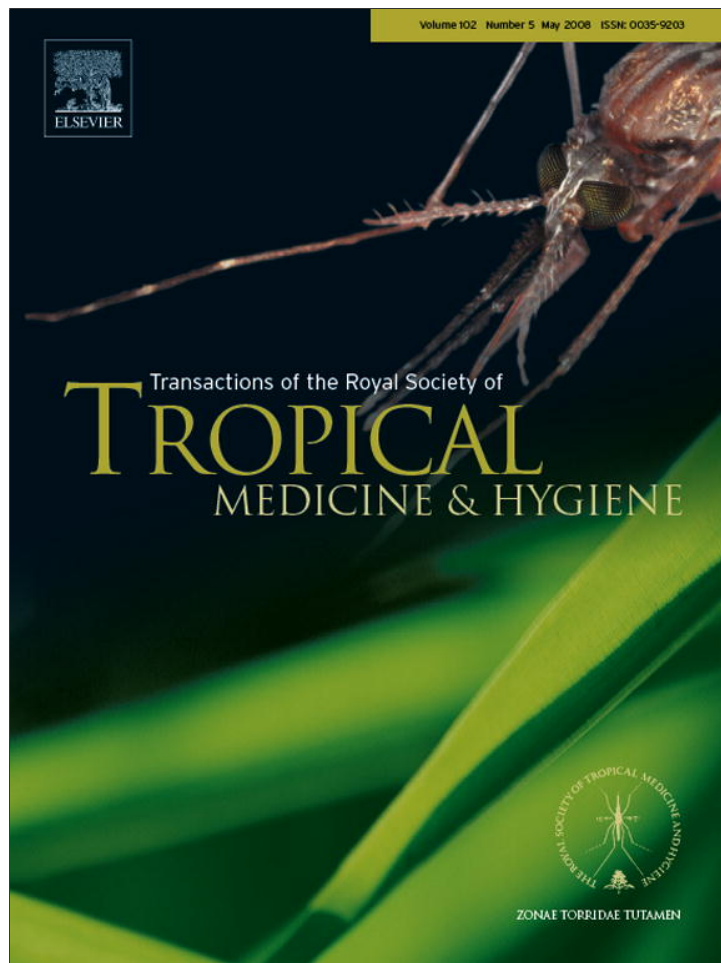


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.

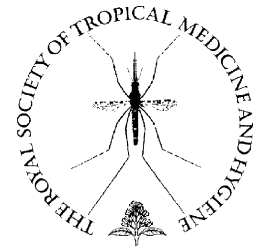


This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>

available at www.sciencedirect.comjournal homepage: www.elsevierhealth.com/journals/trst

Predictable ecology and geography of avian influenza (H5N1) transmission in Nigeria and West Africa

Richard A.J. Williams^{a,*}, Folorunso O. Fasina^b, A. Townsend Peterson^a

^a *Natural History Museum and Biodiversity Research Center, The University of Kansas, Lawrence, KS 66045, USA*

^b *Faculty of Veterinary Science, University of Pretoria, Onderstepoort 0110, South Africa*

Received 24 July 2007; received in revised form 24 January 2008; accepted 24 January 2008

Available online 17 March 2008

KEYWORDS

Avian influenza;
H5N1;
Ecological niche
model;
GARP;
Nigeria;
West Africa

Summary The emerging virus strain termed highly pathogenic H5N1 avian influenza (HP-H5N1) has spread widely in the past decade and is now the focus of considerable concern in several sectors. We tested the hypothesis that spatial distributions of veterinary and human HP-H5N1 cases are related to coarse-scale environmental features in West Africa. We used ecological niche models to associate Nigerian HP-H5N1 occurrences with 1 km resolution digital data layers summarizing parameters of surface reflectance and landform. Predictive challenges included anticipating the spatial distribution of (i) random subsamples and (ii) spatially and temporally stratified subsamples of Nigerian occurrence data, and (iii) more limited occurrence data from across West Africa. In almost all tests, we found that HP-H5N1 cases were occurring under predictable environmental conditions, suggesting that elements of the transmission cycle have some form of ecological determination, here measured as differences in land-surface reflectance and plant phenology through the year. Considerable additional work is needed to establish how these differences affect HP-H5N1 transmission.

© 2008 Royal Society of Tropical Medicine and Hygiene. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Influenza A viruses are responsible for considerable human morbidity and mortality, with annual estimates of up to 500 000 deaths globally (WHO, 2003). Influenza mortality was even higher during the three major pandemics of the twentieth century: the Spanish influenza epidemic (1917–1919) killed an estimated 20 million to 100 million

people (Beveridge, 1991; Johnson and Mueller, 2002; Taubenberger, 2006). Much speculation is focusing on a future influenza pandemic, and highly pathogenic H5N1 (HP-H5N1) is receiving attention as the prime candidate (Guan et al., 2004; Webby and Webster, 2003). At the time of writing, 337 human HP-H5N1 cases had been documented from 59 countries (WHO, 2007a), of which 207 (~60%) were fatal (WHO, 2007b).

The main reservoirs of influenza A viruses (such as HP-H5N1) are generally considered to be aquatic wild birds, principally Anseriformes (ducks, geese, swans) and Charadriiformes (gulls, terns, shorebirds) (Suarez, 2000; Swayne and Suarez, 2000; Webster et al., 1992), although this assumption may not be universally applicable (Kou et al., 2005; OIE,

* Corresponding author. Tel.: +1 785 864 2383;

fax: +1 785 864 5335.

E-mail address: ricw@ku.edu (R.A.J. Williams)

2005; Tanimura et al., 2006). HP-H5N1, however, is best-known in domestic birds, particularly chickens and ducks, with numerous cases documented across Asia, Europe and Africa; 140 million domestic birds have been culled owing to fears of epidemic occurrence of HP-H5N1 (Webster et al., 2006).

HP-H5N1 was first isolated in 1996 from a farm goose in Guangdong Province, People's Republic of China (Xu et al., 1999). By August, 2004, HP-H5N1 had been detected in 10 East Asian countries, and by December 2005 had been detected widely in northern Asia, the Middle East and much of Europe. It was then confirmed in Africa (Nigeria), and by April 2006 had been detected in five African countries (Burkina Faso, Cameroon, Egypt, Nigeria, Niger); the list presently also includes Sudan, Ivory Coast, Ghana, Togo and Benin.

HP-H5N1 was confirmed in Nigeria in February 2006, at a commercial poultry farm (chickens, geese, ostriches) in Kaduna Province – by the end of the initial outbreak 42 000 poultry cases had been reported (ProMED, 2006). Since that time, HP-H5N1 has been detected in >150 instances (F.O. Fasina, unpublished data) in 20 Nigerian states (particularly in the northern and southwestern parts of the country), including one confirmed human death (ProMED, 2007). Three distinct HP-H5N1 genotypes have been detected in the country (Ducatez et al., 2006), likely the result of introduction of two strains and one subsequent reassortment (Salzberg, 2007). HP-H5N1 was probably not present in Nigeria much before 2006, as serosurveys of samples collected in 1999–2004 were all negative (Owoade et al., 2006).

Factors associated with risk of HP-H5N1 transmission in local landscapes are poorly known – the only previous study (Gilbert et al., 2007) showed a clear association with domestic duck populations in rice-paddy agro-ecosystems, but such factors are unknown elsewhere in the distribution of the virus; indeed, the question of whether landscape-scale risk factors exist must remain open to testing. Here, we use novel tools [ecological niche models, (ENMs)] to provide a landscape-scale perspective on the question of HP-H5N1 risk assessment: we associate Nigerian HP-H5N1 cases with annual variation in 'greenness' [Normalized Difference Vegetation Index (NDVI) values derived from imagery from the Advanced Very High Resolution Radiometer (AVHRR) satellite], and develop and test predictive spatial models of HP-H5N1 occurrence. As such, this paper represents a first exploration of the ecological 'niche' of HP-H5N1 cases in Nigeria and more broadly across West Africa. Although we are well aware of the unlikely nature of predictable environmental correlates, given the multiple, diverse factors associated with HP-H5N1 transmission (migratory birds, poultry trade, etc.), our results indicate that HP-H5N1 occurs under consistent, predictable environmental circumstances in West Africa.

Our development of HP-H5N1 ENMs is based on the concept of ecological niches defined as the set of conditions under which a species is able to maintain populations without immigration (Grinnell, 1917, 1924). ENMs have seen considerable exploration recently as a means of estimating the dimensions of ecological niches of species based on incomplete sampling across distributions (see review in (Peterson, 2007). Widespread evolutionary conservatism in niche characteristics has been demonstrated, allowing accurate predictions of transmission of infectious diseases

(Peterson et al., 2002a, 2004), invasive species' potential distributions (Peterson and Shaw, 2003) and projections of species' responses to climate change (Peterson et al., 2002b), etc. Overall, ENMs offer a powerful approach to understanding coarse-scale environmental correlates of presence and absence of species or biological phenomena across complex landscapes.

2. Methods

2.1. Input data

The principal suite of occurrence information for this study was HP-H5N1 case-occurrence data for January–April 2006 from the National Veterinary Research Institute, Nigeria, which consisted of 132 HP-H5N1 detections (including isolations from two wild birds, the remainder from poultry; Figure 1); an additional 12 occurrences were drawn from the same source for November 2006–January 2007 for model testing. Textual descriptions of occurrence localities were converted to geographic coordinates accurate to the nearest 0.001° using the Alexandria Digital Library Gazetteer (<http://middleware.alexandria.ucsb.edu/client/gaz/adl/index.jsp>), GEOnet Names Server (<http://gnswww.nga.mil/geonames/GNS/index.jsp>), and other sources (Rand McNally, 1998). In all, we were able to convert 72 of these locations into unique geographic coordinates; the attrition from 132 to 72 localities is largely because 38 duplicate occurrences from the same localities were discarded, and in 22 cases we were unable to confirm the coordinates of the locality. Finally, to characterize the broader distribution of HP-H5N1 in West Africa, we searched the archives of the International Society for Infectious Disease (ProMED Avian Influenza archive) for West Africa (14 occurrences; Figure 1): Burkina Faso (four points), Ivory Coast (three), Ghana (two), Niger (two) and Cameroon (one), excluding two duplicated localities and

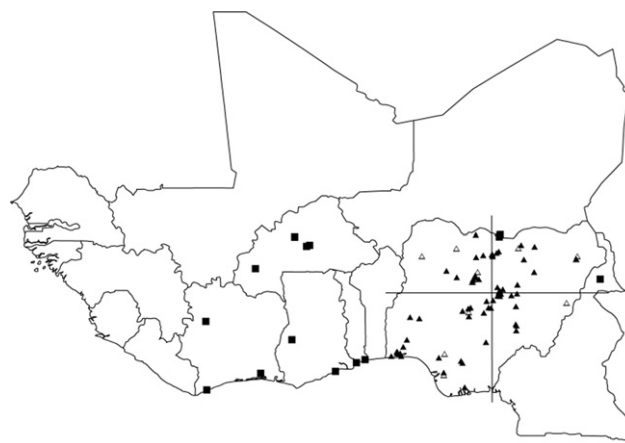


Figure 1 Occurrence data for highly pathogenic H5N1 in West Africa used in this study. Filled triangles = 2006 cases in Nigeria; open triangles = Nigerian YEAR cases (November 2006–January 2007); squares = occurrences from elsewhere in West Africa. The two dashed lines overlain on Nigeria indicate the median latitude and longitude used for spatial subsets of Nigerian occurrence data (EW, NS and DIAG tests).

four localities (three from Niger, one from Ivory Coast) for which we were unable to locate coordinates of the reported site. Throughout, although the geographic coordinates assigned may not always fix the exposure point precisely, they represent a best guess as to its position, and are likely to represent the coarse-scale ecological conditions under which HP-H5N1 transmission occurs. We believe that georeferencing imprecision is of a magnitude smaller than the resolution of our environmental grids, so the modelling process is not compromised.

Environmental data sets included twenty-four 1 km² resolution monthly composite remotely sensed data layers (April 1992–March 1993 and February 1995–January 1996), in each case presenting values of the NDVI (native spatial resolution 1 km). NDVI is derived from reflectance in the visible and near-infrared domains and as such is sensitive to photosynthetic activity and is closely correlated with photosynthetic mass (Tucker, 1979) – the time series of NDVI values used here thus profile differences in land cover and plant phenology across landscapes. We also included four data sets summarizing aspects of topography – elevation, slope, aspect and compound topographic index (which summarizes tendency to pool water) – from the U.S. Geological Survey's Hydro-1K data set (<http://eros.usgs.gov/products/elevation/gtopo30/hydro/africa.html>; native resolution 1 km). Climate data were not included in these analyses for lack of sufficiently high-resolution data sets across the region of interest. We purposefully included NDVI series from both an El Niño year (1992/1993) and a normal year (1995/1996) to take into account any effects that these global climate phenomena might have on West African landscapes.

2.2. The GARP algorithm

We used the Genetic Algorithm for Rule-set Prediction (GARP) (Stockwell and Peters, 1999) for ENM development. GARP uses an evolutionary-computing genetic algorithm to search for non-random associations between environmental variables and known occurrences of species, as contrasted with environmental characteristics across the overall study area. GARP has been applied widely to questions of disease transmission (Costa et al., 2002; Peterson et al., 2007a), and its predictive ability has been tested under diverse circumstances (Anderson et al., 2003; Elith et al., 2006; Peterson et al., 2003). Although GARP was highlighted for its relatively poor performance in recent comparative studies (Elith et al., 2006; Stockman et al., 2006), other recent studies have indicated considerably better performance (McNyset and Blackburn, 2006; Tsoar et al., 2007) and some artifactual causation of the previous negative results (McNyset and Blackburn, 2006; Peterson et al., 2007b). As such, we employed GARP in these analyses, although we highlight algorithm choice as an important issue remaining in risk analysis applications.

In general, we develop tests based on subsets of available occurrence information set aside prior to model development. Of data provided to GARP, the program divides occurrence data randomly into three subsets: training data (for rule development), intrinsic testing data (for evaluation of rules) and extrinsic testing data (for evaluation of model

quality, see below). Spatial predictions of presence versus absence can include two types of error: false negatives (areas of actual presence predicted absent) and false positives (areas of actual absence predicted present) (Fielding and Bell, 1997) – rule performance in each of these dimensions is evaluated via the intrinsic testing data set. Change in predictive accuracy from one iteration to the next are used to evaluate whether particular rules should be incorporated into the model or not, and the algorithm runs either 1000 iterations or until convergence (Stockwell and Peters, 1999). The final rule-set is then used to query the environmental data sets to identify areas fitting the rule set predictions to produce a hypothesis of the potential geographic distribution of the species (Soberón and Peterson, 2005).

Since GARP includes several random-walk elements, each replicate model developed produces distinct results, representing alternative solutions to the optimization challenge. Following best-practices approaches (Anderson et al., 2003), we developed 100 replicates of each model. We filtered these replicates based on their error characteristics, retaining the 20 with lowest false negative rates (= percentage of testing occurrence points falling in areas not predicted to be suitable), and then retained the 10 (of the 20) closest to the median of proportional area predicted present, an index of false-positive error rates (Anderson et al., 2003). A consensus of these 'best subset' models was then developed by summing values for each pixel in the map to produce final predictions of potential distributions with 11 thresholds (integers from 0 to 10).

The customary approaches to spatial model validation (e.g. receiver operating characteristic, kappa statistics) are not applicable to situations in which presence-only data are the only information available (Fielding and Bell, 1997; Manel et al., 1999). As such, we validated models using simple calculations of binomial probabilities that coincidence of predictions and independent test data are no better than random, with the probability of k successes in n trials depending on p , the probability of success in any one trial – we estimated p as the proportion of the testing area predicted present, and k as the number of the n testing points that were successfully predicted (Anderson et al., 2003). Binomial probabilities were calculated for each of the 10 thresholds representing predictions of presence (1 = broad, 10 = narrow), in each case testing whether predictivity is better than that expected by chance. In one case, we explored the effects of spatial uncertainty regarding the localization of outbreak sites by calculating success in predicting areas of presence within 100 m of known occurrence sites, adjusting p appropriately to reflect the broader area of potential presence.

2.3. Modelling approach

This study focuses on the question of whether HP-H5N1 occurrences in West Africa follow a consistent and predictable environmental regime. As such, we developed a series of tests of model predictivity, in each case with the models developed and the predictions tested being based on independent suites of occurrence data. Model tests were based on subsets of the 2006 Nigerian occurrence data

described above, as well as on the 12 additional Nigerian occurrences from November 2006 to January 2007 (Figure 1); we also tested our Nigerian models with occurrence data from across West Africa. The basic design of testing was as follows:

2.3.1. Predictivity across training landscape

We divided the 72 Nigerian occurrences from 2006 into two equal groups at random, and used one group for model development and the other for model testing (hereafter 'RND' tests). We also tested the ability of 2006-based ENMs to predict the spatial distribution of cases from November 2006 to January 2007 (hereafter 'YEAR' tests). This scheme assesses the ability of the modelling approach to anticipate the spatial distribution of HP-H5N1 cases were sampling density to be increased, but across a region in which samples are already available.

2.3.2. Predictivity across space (medium scale)

We stratified the 72 Nigerian occurrences from 2006 spatially into quadrants above and below the median longitude and median latitude of the occurrence data. From this spatial stratification, we developed three pairs of quadrants: west versus east of the median longitude (hereafter 'EW' tests), north versus south of the median latitude (hereafter 'NS' tests), and on-diagonal (upper left-hand and lower right-hand quadrants) versus off-diagonal (lower left-hand and upper right-hand quadrants; hereafter 'DIAG' tests). In each case, we developed both reciprocal predictions, testing the ability of ENMs to anticipate the spatial distribution of HP-H5N1 cases in regions for which no sampling is available.

2.3.3. Predictivity across space (broader scale)

We projected ENMs trained using all 2006 Nigerian occurrence points onto the rest of West Africa, and tested their spatial predictions via their coincidence with the 14 cases for which geographic coordinates were available in other West African nations (hereafter 'WA' tests). These tests evaluated the ability of the ENMs to predict into even broader unsampled areas.

3. Results

Almost all tests conducted in this study indicated that independent sets of test points coincided with ENM predictions significantly better than random expectations (Table 1). In other words, models based on known HP-H5N1 occurrences can anticipate spatial distributions of independent samples of HP-H5N1 based on their environmental attributes.

3.1. Predictivity across training landscape

The two reciprocal tests based on random subsamples of 2006 Nigerian occurrence data (RND tests) both indicated significant predictive power of the ENMs, with all thresholds showing predictivity better than random expectations (Table 1). For example, the broadest predictions (threshold >1 of 10 models predict present) predicted 65–72% of Nigeria as present, and predicted the spatial position of >86% of independent test occurrence points correctly ($P < 0.001$); the narrowest predictions (threshold of 10 of 10 models predict present) predicted 17% of Nigeria as present, and predicted the spatial position of 35–37.5% of independent test points correctly ($P < 0.001$).

Table 1 Summary of model predictions and tests in this study, illustrated by information for the threshold ≥ 5 of 10 best subset models predicting potential for presence

	Sample size (train/test)	Prop. area ^a	k (= no. of successes)	Cumulative binomial probability	No. of thresholds significant ^b
RND					
RND1 predicts RND2	37/34	0.466	24	0.008	10
RND2 predicts RND1	34/37	0.487	24	0.003	10
NS					
N predicts S	35/36	0.045	10	<0.001	10
S predicts N	36/35	0.385	23	<0.001	8
EW					
W predicts E	36/35	0.502	23	0.034	9
E predicts W	35/36	0.455	9	0.987	0
DIAG					
On predicts Off	24/47	0.592	21	<0.001	10
Off predicts On	47/24	0.197	17	0.003	10
YEAR					
2006 predicts 2007	72/12	0.610	9	0.094	2
WA					
Nigeria predicts West Africa	72/14	0.272	9	<0.001	8

^a Prop. area indicates the proportion of the test region predicted present at that threshold.

^b Number of thresholds (out of 10) for which model predictions were significantly better than random expectations.

Tests using 12 HP-H5N1 occurrences subsequent to the main sample (YEAR tests) were somewhat successful. Regarding the raw model results, five of the 12 occurrences were predicted by all of the replicate models, and 10 of 12 occurrences were predicted by at least one of the replicate models; two of the 10 thresholds predicted spatial distributions better than randomly ($P < 0.05$; Table 1). However, we noted that several points lay close to predicted areas, so, in view of uncertainties of georeferencing, we explored the effects of introducing a degree of uncertainty into our tests: we traced a 100 m buffer around HP-H5N1 occurrence sites, and noted a marked improvement in model performance. Here, six of 12 occurrences were predicted by all replicate models, and 11 of 12 by at least one model, and six of the 10 thresholds predicted HP-H5N1 occurrences better than random expectations ($P < 0.05$).

3.2. Predictivity across space (medium scale)

We also assessed the ability of ENMs to anticipate spatial distributions of HP-H5N1 occurrences in regions for which no input data were available – these three pairs of tests based on spatially stratified subsets of Nigerian HP-H5N1 occurrence data also indicated, for the most part, significant predictive power of the ENM (Table 1, Figure 2). For both DIAG tests, all thresholds in both reciprocal tests were statistically significantly more predictive than random expectations ($P < 0.05$). For the NS tests, results were similar, except that two of 10 thresholds in the south-predicts-north tests were not significant ($P > 0.05$). Finally, in the EW tests, nine of 10 thresholds were statistically significant in the west-predicts-east tests, but none was significant in the converse test – this ENM dramatically underpredicted HP-H5N1 cases in the southern part of the country (Table 1).

More generally, HP-H5N1 potential presence is predicted more broadly in northern Nigeria than in the southern part of the country; and some areas (e.g. along the southeastern border) are predicted to be largely HP-H5N1 free. The presence of the virus is predicted mostly in the savannah and woodland habitats of the north, whereas absence is predicted in montane areas, coastal mangroves, the freshwater swamps of the Niger Delta, and rainforest areas in the south. Areas of highest predicted HP-H5N1 risk were not the 'greenest' areas, but rather were relatively dry and highly variable seasonally, as can be appreciated from NDVI profiles through the course of the year (Figure 3).

3.3. Predictivity across space (broader scale)

Given the initial signal of predictive power regarding HP-H5N1 within Nigeria, we projected a Nigeria 2006-based ENM across West Africa to develop a broader-scale test of ENM predictive ability (Table 1, Figure 4). As within Nigeria, HP-H5N1 potential for occurrence is predicted in the savannah and woodland belt and across the southern portion of the Sahel. The virus is not predicted to occur in the Sahara, montane zones, coastal mangroves or rainforest. In particular, we note that the Upper Guinean rainforest block (Guinea, Ivory Coast, Liberia, Sierra Leone, western Ghana) and the Lower Guinea rainforest block (Cameroon, southern Nigeria) are not predicted to be suitable for HP-H5N1,

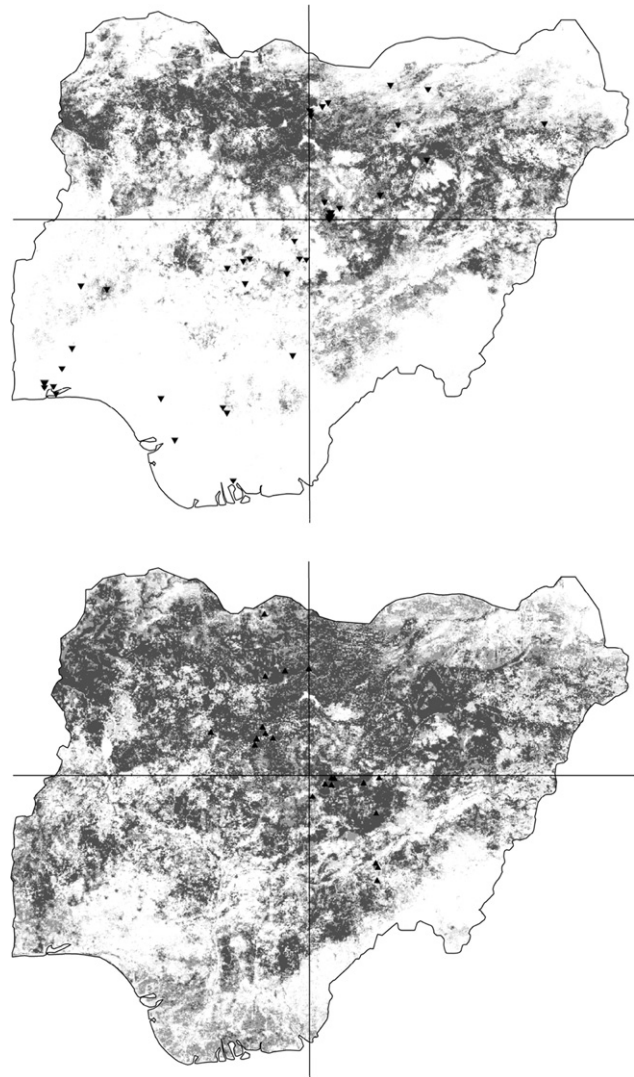


Figure 2 Example of spatially stratified tests of ENM predictions of highly pathogenic H5N1 distributions within Nigeria. Here, occurrences in on-diagonal quadrants were used to predict distributions of cases in off-diagonal quadrants, and vice versa. Model predictions are shown as ramps of model agreement in predictions: white = 10 of 10 models predict absence; light grey = 1–5 of 10 models predict potential presence; dark grey = 6–9 of 10 models predict potential presence; and darkest grey = all 10 models agree in predicting potential presence. Only independent test points are plotted on each map.

but the savannah areas of the Dahomey Gap (Benin, eastern Ghana, Togo) are strongly predicted as suitable. The coincidence of the projection of the Nigerian ENM rule sets across West Africa coincided with the 14 independent test points better than random expectations at eight of 10 thresholds (all $P < 0.01$), except for the two most restrictive thresholds (both $P > 0.06$).

4. Discussion

ENM applications to transmission are still preliminary, and are certainly relatively new to the field (Peterson, 2007).

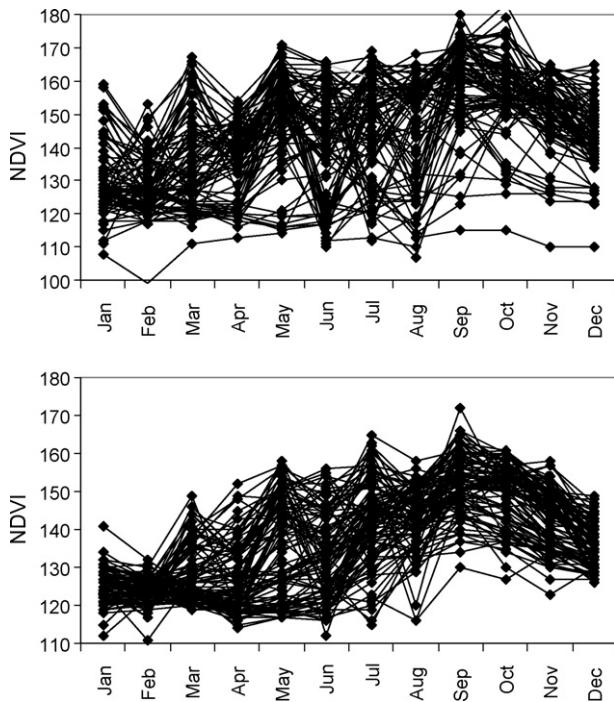


Figure 3 Summary of Normalized Difference Vegetation Index 'greenness' profiles of Nigeria through the year for (top) 85 randomly selected points of predicted absence, and (bottom) 96 randomly selected points of predicted presence. Note considerable reduction of variance and accentuated seasonality in the bottom graph.

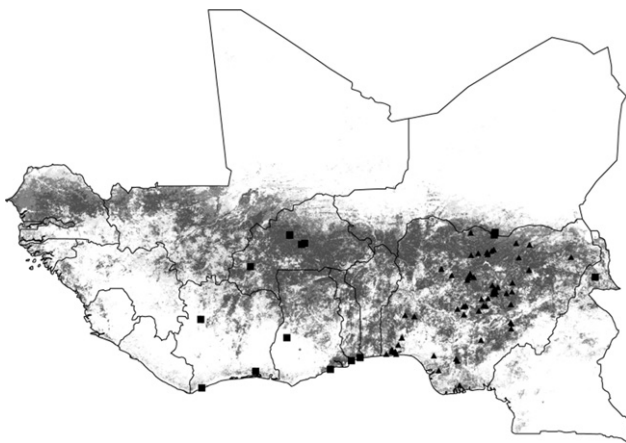


Figure 4 Regional projection across all of West Africa of highly pathogenic H5N1 ecological niche model results based on 2006 Nigeria occurrences (black triangles). Model predictions are shown as ramps of model agreement in predictions: white = 10 of 10 models predict absence; light grey = 1–5 of 10 models predict potential presence; dark grey = 6–9 of 10 models predict potential presence; and darkest grey = all 10 models agree in predicting potential presence. Solid triangles indicate independent test occurrence data from Nigeria; solid squares indicate independent test data from other West African countries.

Given this novelty, two features of ENMs should be emphasized at this point to facilitate interpretation of model results. Firstly, ENMs are frequently coarse-resolution, distribution-wide views of biological phenomena that outline broad potential for disease occurrence; particular landscape features, management regimes (e.g. biosafety measures), and chance events (e.g. introduction of the pathogen) may prevent this potential from being realized, but the model results indeed indicate the coarser-scale potential for such occurrences. Secondly, given the potential nature of predictions of presence (compared to ENM predictions of absence), false-positive error is much more serious in model evaluation than false-negative error. These features of ENM must be taken into account in any consideration of model predictions and their utility.

The ecological niche model predictions that we have developed for Nigeria and West Africa are exploratory, designed to test the basic hypothesis that environmental correlates exist. Although ENMs have been applied broadly to biodiversity questions (Araujo et al., 2005; Guisan and Zimmermann, 2000; Peterson, 2006; Thuiller, 2003), their application to disease systems remains preliminary (Peterson, 2007). Although several initial tests have been published (Costa et al., 2002; Levine et al., 2004; Peterson and Shaw, 2003; Peterson et al., 2002a), the failings and biases of the technique in a disease transmission context are still being discovered and understood.

We recognize several limitations in our analyses. First, imprecision inherent in georeferencing infection sites sets a base level of error, and guarantees some predictive failures. Given that poultry is frequently traded and moved to markets, HP-H5N1 infections may frequently appear at sites not coincident with transmission sites – a number of Nigerian HP-H5N1 cases were detected in poultry markets, to which infected birds were presumably transported over unknown distances from actual transmission sites. These factors – movement, transportation, trade and biosecurity measures on poultry farms – may impact the epidemiology of the disease, but we focus explicitly on ecological and geographical factors with the aim of developing a model of the ecological niche of the virus.

Another important challenge for these analyses is that of distinguishing true spatial and ecological biases in case distributions (i.e. the ecological niche!) from the spatial and ecological biases in distributions of the major known HP-H5N1 host in Nigeria (chickens). The total Nigerian chicken population is >140 million birds, including 'backyard chickens', raised without biosecurity measures (~60%), commercially farmed chickens under high biosecurity (~25%), and semi-commercial chickens, raised with some biosecurity measures (~15%) (Adene and Oguntade, 2006). Most commercial birds (65%) are raised in the southwestern part of the country, near Lagos (Adene and Oguntade, 2006).

Free mingling of backyard poultry and wild birds has been identified as a risk factor for HP-H5N1 transmission (Chotpitayasunondh et al., 2005; de Benedictis et al., 2007; WHO, 2005). In Nigeria, however, at least at the coarse scales examined herein, backyard chicken distributions and our reconstructed risk areas are virtually inverse: backyard chicken populations are highest and HP-H5N1 predictions lowest in southeastern Nigeria, and backyard chicken

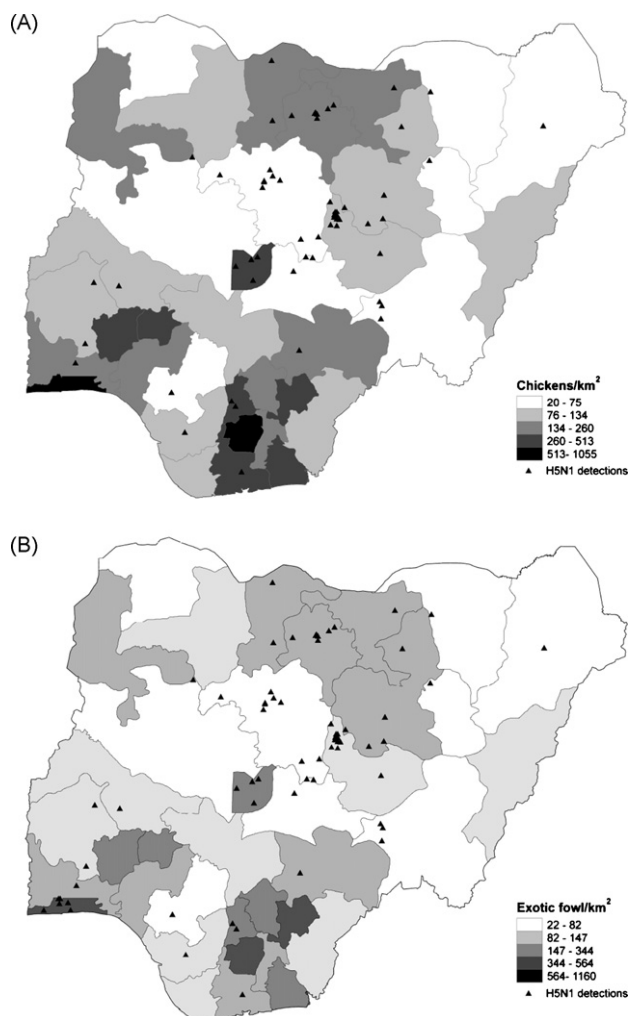


Figure 5 Summary of population density (individuals/km²) in Nigeria for (A) backyard chickens and (B) exotic fowl (ducks, guinea fowls, ostriches, and pigeons). Solid triangles show the distribution of highly pathogenic H5N1 cases. *Source:* Federal Ministry of Agriculture and Rural Development, Federal Department of Livestock and Pest Control Services in Adene and Oguntade (2006).

populations are lowest and HP-H5N1 predictions highest in northern Nigeria (Figure 5). Moreover, HP-H5N1 outbreak localities do not necessarily coincide with areas of high backyard chicken population – for example, the state with the highest backyard chicken populations (Imo, southwest Nigeria) has had no cases of HP-H5N1, despite having roughly ten-fold higher density of backyard chickens as Plateau, the state with the highest number of HP-H5N1 outbreak sites. Similarly, we observed little coincidence between HP-H5N1 outbreaks and areas of high density of commercially farmed birds in the southeast (Adene and Oguntade, 2006). This result coincides with experience in Thailand, where chickens are the most frequent victims of poultry HP-H5N1 outbreaks, but outbreaks do not correspond to the distribution of backyard chickens (Gilbert et al., 2006).

Recent studies in Southeast Asia (Gilbert et al., 2006, 2007) identified predictable foci of HP-H5N1 activity based on free-range duck farming and rice-paddy cultivation.

Although that association has clearly and easily interpretable foundations, our results suggest that predictable ecology may be more pervasive in HP-H5N1 geography than might have been expected. Several elements in the HP-H5N1 transmission cycle could be responsible for this predictivity: ecological biases associated with initial arrival of virus propagules in a poultry population via migratory birds (Chen et al., 2006; Olsen et al., 2006), transmission among Nigerian poultry flocks (Gilbert et al., 2006; Yasue et al., 2006), or even with transportation routes within Nigeria that might be responsible for communicating infections – most likely, the truth lies in a combination of such factors (Kilpatrick et al., 2006). The precise basis for this predictivity has yet to be identified, but the existence of an environmental signal in HP-H5N1 transmission may offer valuable clues as to its nature.

Interpretation of the nature of the environmental signal associated with high HP-H5N1 transmission is complex. The NDVI data used in this study are correlated with photosynthetic mass (Tucker, 1979), and our time series of NDVI images thus summarize patterns of vegetation phenology across landscapes. In the crudest sense, our NDVI profiles identify areas of drought-sensitive vegetation phenology as particularly key in HP-H5N1 transmission (Figure 3), but the details are still under study and exploration (Peterson et al., 2007a).

Perhaps most importantly, projecting the Nigerian ENMs across the entire region yielded a view of West African HP-H5N1 distributions that was highly predictive of what independent test data could be assembled. Such validated model predictions offer the possibility of public health applications, providing information that may be used to prioritize surveillance and remediation activities. Similarly, such predictions may be helpful to policy makers planning expansions to and investment in the Nigerian chicken industry, particularly as regards biosecurity measures. The spatial limits of the predictivity we have documented remain an open question – our initial demonstration of predictable HP-H5N1 geography across West Africa awaits further testing and comparison with HP-H5N1 occurrence information from other regions.

Authors' contributions: ATP conceived and designed the experiments; RW and FOF analyzed the data; RW and ATP drafted the manuscript. All authors read and approved the final manuscript. RW and ATP are guarantors of the paper.

Acknowledgements: Thanks to Monica Papeş, Ryan Lash, and Yoshinori Nakazawa for their expert assistance with GIS and satellite imagery, and to the National Biological Information Infrastructure of the U.S. Geological Survey for financial support.

Funding: U.S. Geological Survey.

Conflicts of Interest: None declared.

Ethical approval: Not required.

References

- Adene, D.F., Oguntade, A.E., 2006. The structure and importance of the commercial and village based poultry industry in Nigeria. FAO, Rome, Italy.
- Anderson, R.P., Lew, D., Peterson, A.T., 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecol. Modell.* 162, 211–232.
- Araujo, M.B., Pearson, R.G., Thuiller, W., Erhard, M., 2005. Validation of species-climate impact models under climate change. *Global Change Biol.* 11, 1504–1513.
- Beveridge, W.I.B., 1991. The chronicle of influenza epidemics. *Hist. Philos. Life Sci.* 13, 223–235.
- Chen, H.L., Li, Y.B., Li, Z.J., Shi, J.Z., Shinya, K., Deng, G.H., Qi, Q.L., Tian, G.B., Fan, S.F., Zhao, H.D., Sun, Y.X., Kawaoka, Y., 2006. Properties and dissemination of H5N1 viruses isolated during an influenza outbreak in migratory waterfowl in western China. *J. Virol.* 80, 5976–5983.
- Chotpitayasunondh, T., Ungchusak, K., Hanshaoworakul, W., Chunsuthiwat, S., Sawanpanyalert, P., Kijphati, R., Lochindarat, S., Srisan, P., Suwan, P., Osotthanakorn, Y., Anantasetagoon, T., Kanjanawasri, S., Tanupattarachai, S., Weerakul, J., Chaiwirattana, R., Maneerattanaporn, M., Poolsavathitkool, R., Chokephaibulkit, K., Apisarnthanarak, A., Dowell, S.F., 2005. Human disease from influenza A (H5N1), Thailand, 2004. *Emerg. Infect. Dis.* 11, 201–209.
- Costa, J., Peterson, A.T., Beard, C.B., 2002. Ecologic 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.* 67, 516–520.
- de Benedictis, P., Joannis, T.M., Lombin, L.H., Shittu, I., Beato, M.S., Rebonato, V., Cattoli, G., Capua, I., 2007. Field and laboratory findings of the first incursion of the Asian H5N1 highly pathogenic avian influenza virus in Africa. *Avian. Pathol.* 36, 115–117.
- Ducatez, M.F., Olinger, C.M., Owoade, A.A., De Landtsheer, S., Ammerlaan, W., Niesters, H.G., Osterhaus, A.D., Fouchier, R.A., Muller, C.P., 2006. Avian flu: multiple introductions of H5N1 in Nigeria. *Nature* 442, 37.
- Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, 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., Soberon, J., Williams, S., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29, 129–151.
- Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* 24, 38–49.
- Gilbert, M., Chaitaweesub, P., Parakarnawongsa, T., Premashtira, S., Tiensin, T., Kalpravidh, W., Wagner, H., Slingenbergh, J., 2006. Free-grazing ducks and highly pathogenic avian influenza, Thailand. *Emerg. Infect. Dis.* 12, 227–234.
- Gilbert, M., Xiao, X.M., Chaitaweesub, P., Kalpravidh, W., Premashtira, S., Boles, S., Slingenbergh, J., 2007. Avian influenza, domestic ducks and rice agriculture in Thailand. *Agric. Ecosyst. Environ.* 119, 409–415.
- Grinnell, J., 1917. The niche relationships of the Californian Thrasher. *Auk* 34, 427–433.
- Grinnell, J., 1924. Geography and evolution. *Ecology* 5, 225–229.
- Guan, Y., Poon, L.L.M., Cheung, C.Y., Ellis, T.M., Lim, W., Lipatov, A.S., Chan, K.H., Sturm-Ramirez, K.M., Cheung, C.L., Leung, Y.H.C., Yuen, K.Y., Webster, R.G., Peiris, J.S.M., 2004. H5N1 influenza: a protean pandemic threat. *Proc. Natl. Acad. Sci. USA* 101, 8156–8161.
- Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. *Ecol. Modell.* 135, 147–186.
- Johnson, N., Mueller, J., 2002. Updating the accounts: global mortality of the 1918–1920 'Spanish' influenza pandemic. *Bull. Hist. Med.* 76, 105–115.
- Kilpatrick, A.M., Chmura, A.A., Gibbons, D.W., Fleischer, R.C., Marra, P.P., Daszak, P., 2006. Predicting the global spread of H5N1 avian influenza. *Proc. Natl. Acad. Sci. USA* 103, 19368–19373.
- Kou, Z., Lei, F.M., Yu, J., Fan, Z.J., Yin, Z.H., Jia, C.X., Xiong, K.J., Sun, Y.H., Zhang, X.W., Wu, X.M., Gao, X.B., Li, T.X., 2005. New genotype of avian influenza H5N1 viruses isolated from tree sparrows in China. *J. Virol.* 79, 15460–15466.
- 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.
- 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.
- McNyset, K., Blackburn, J., 2006. Does GARP really fail miserably? A response to Stockman et al. (2006). *Divers. Distrib.* 12, 782–786.
- OIE, 2005. Rapport de mission: Mission to Russia to assess the avian influenza situation in wildlife and the national measures being taken to minimize the risk of international spread. World Organisation for Animal Health, Paris. <http://www.oie.int/download/Missions/2005/ReportRussia2005Final2.pdf> [accessed 11 December 2007].
- Olsen, B., Munster, V.J., Wallensten, A., Waldenstrom, J., Osterhaus, A., Fouchier, R.A.M., 2006. Global patterns of influenza A virus in wild birds. *Science* 312, 384–388.
- Owoade, A.A., Ducatez, M.F., Muller, C.P., 2006. Seroprevalence of avian influenza virus, infectious bronchitis virus, reovirus, avian pneumovirus, infectious laryngotracheitis virus, and avian leukosis virus in Nigerian poultry. *Avian Dis.* 50, 222–227.
- Peterson, A.T., 2006. Uses and requirements of ecological niche models and related distributional models. *Biodivers. Inform.* 3, 59–72.
- Peterson, A.T., 2007. Ecological niche modeling and spatial patterns of disease transmission. *Emerg. Infect. Dis.* 12, 1822–1826.
- Peterson, A.T., Shaw, 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.
- Peterson, A.T., Sanchez-Cordero, V., Ben Beard, C., Ramsey, J.M., 2002a. Ecologic niche modeling and potential reservoirs for Chagas disease, Mexico. *Emerg. Infect. Dis.* 8, 662–667.
- Peterson, A.T., Ortega-Huerta, M.A., Bartley, J., Sanchez-Cordero, V., Soberon, J., Buddemeier, R.H., Stockwell, D.R.B., 2002b. Future projections for Mexican faunas under global climate change scenarios. *Nature* 416, 626–629.
- Peterson, A.T., Vieglais, D.A., Andreasen, J.K., 2003. Migratory birds modeled as critical transport agents for West Nile virus in North America. *Vector Borne Zoonotic Dis.* 3, 27–37.
- Peterson, A.T., Bauer, J.T., Mills, J.N., 2004. Ecologic and geographic distribution of filovirus disease. *Emerg. Infect. Dis.* 10, 40–47.
- Peterson, A.T., Benz, B.W., Papes, M., 2007a. Highly pathogenic H5N1 avian influenza: entry pathways into North America via bird migration. *PLOS One* 2, e261.
- Peterson, A.T., Papes, M., Eaton, M., 2007b. Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography* 30, 550–560.
- ProMED, 2006. Avian influenza - worldwide: Nigeria, OIE. ProMED archive number 20060208.0409. <http://www.promedmail.org> [accessed 11 December 2007].

- ProMED, 2007. Avian influenza, human (29): Nigeria, Confirmed. ProMED archive number 20070203.0439. <http://www.promedmail.org> [accessed 11 December 2007].
- Rand McNally, 1998. New Millenium World Atlas Deluxe. Rand McNally & Co., Skokie, IL.
- Salzberg, S.L., 2007. Genome analysis linking recent European and African influenza (H5N1) viruses. *Emerg. Infect. Dis.* 13, 713–718.
- Soberón, J., Peterson, A.T., 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodivers. Inform.* 2, 1–10.
- Stockman, A.K., Beamer, D.A., Bond, D.A., 2006. An evaluation of a GARP model as an approach to predicting the spatial distribution of non-vagile invertebrate species. *Divers. Distrib.* 12, 81–89.
- 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.
- Suarez, D.L., 2000. Evolution of avian influenza viruses. *Vet. Microbiol.* 74, 15–27.
- Swayne, D.E., Suarez, D.L., 2000. Highly pathogenic avian influenza. *Rev. Sci. Tech. Off. Int. Epiz.* 19, 463–482.
- Tanimura, N., Tsukamoto, K., Okamatsu, M., Mase, M., Imada, T., Nakamura, K., Kubo, M., Yamaguchi, S., Irishio, W., Hayashi, M., Nakai, T., Yamauchi, A., Nishimura, M., Imai, K., 2006. Pathology of fatal highly pathogenic H5N1 avian influenza virus infection in large-billed crows (*Corvus macrorhynchos*) during the 2004 outbreak in Japan. *Vet. Pathol.* 43, 500–509.
- Taubenberger, J.K., 2006. The origin and virulence of the 1918 "Spanish" influenza virus. *Proc. Am. Philos. Soc.* 150, 86–112.
- Thuiller, W., 2003. BIOMOD - optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Chang. Biol.* 9, 1353–1362.
- Tsoar, A., Allouche, O., Steinitz, O., Rotem, D., Kadmon, R., 2007. A comparative evaluation of presence-only methods for modelling species distribution. *Divers. Distrib.* 13, 397–405.
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150.
- Webby, R.J., Webster, R.G., 2003. Are we ready for pandemic influenza? *Science* 302, 1519–1522.
- Webster, R.G., Bean, W.J., Gorman, O.T., Chambers, T.M., Kawaoka, Y., 1992. Evolution and ecology of influenza-A viruses. *Microbiol. Rev.* 56, 152–179.
- Webster, R.G., Peiris, M., Chen, H., Guan, Y., 2006. H5N1 outbreaks and enzootic influenza. *Emerg. Infect. Dis.* 12, 3–8.
- WHO, 2003. Influenza: Fact sheet No. 211, Revised March 2003. World Health Organization, Geneva. <http://www.who.int/mediacentre/factsheets/fs211/en/print.html> [accessed 11 December 2007].
- WHO, 2005. Influenza pandemic risk assessment and preparedness in Africa. WHO Regional Office for Africa, Brazzaville. http://www.afro.who.int/csr/epr/avian_flu/afr_avian_flu_31_10_05.pdf [accessed 11 December 2007].
- WHO, 2007a. H5N1 avian influenza: Timeline of major events, Updated 5 December 2007. World Health Organization, Geneva. http://www.who.int/csr/disease/avian_influenza/Timeline_5_Dec_07.pdf [accessed 11 December 2007].
- WHO, 2007b. Cumulative number of confirmed human cases of avian influenza A/(H5N1) reported to WHO. World Health Organization, Geneva. http://www.who.int/csr/disease/avian_influenza/country/cases_table_2007_07_11/en/index.html [accessed 11 December, 2007].
- Xu, X.Y., Subbarao, K., Cox, N.J., Guo, Y.J., 1999. Genetic characterization of the pathogenic influenza A/Goose/Guangdong/1/96 (H5N1) virus: similarity of its hemagglutinin gene to those of H5N1 viruses from the 1997 outbreaks in Hong Kong. *Virology* 261, 15–19.
- Yasue, M., Feare, C.J., Bennun, L., Fiedler, W., 2006. The epidemiology of H5N1 avian influenza in wild birds: why we need better ecological data. *Bioscience* 56, 923–929.