

# A continental risk assessment of West Nile virus under climate change

RYAN J. HARRIGAN<sup>1</sup>, HENRI A. THOMASSEN<sup>2</sup>, WOLFGANG BUERMANN<sup>3</sup> and THOMAS B. SMITH<sup>1,4</sup>

<sup>1</sup>Center for Tropical Research, Institute of the Environment and Sustainability, University of California, 619 Charles E. Young Drive East, Los Angeles, California 90095, USA, <sup>2</sup>Comparative Zoology, Institute of Evolution and Ecology, University of Tübingen, Auf der Morgenstelle 28, D-72076 Tübingen, Germany, <sup>3</sup>School of Earth and Environment, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, United Kingdom, <sup>4</sup>Department of Ecology and Evolutionary Biology, University of California, 621 Charles E. Young Drive South, Los Angeles, California 90095, USA

## Abstract

Since first introduced to North America in 1999, West Nile virus (WNV) has spread rapidly across the continent, threatening wildlife populations and posing serious health risks to humans. While WNV incidence has been linked to environmental factors, particularly temperature and rainfall, little is known about how future climate change may affect the spread of the disease. Using available data on WNV infections in vectors and hosts collected from 2003–2011 and using a suite of 10 species distribution models, weighted according to their predictive performance, we modeled the incidence of WNV under current climate conditions at a continental scale. Models were found to accurately predict spatial patterns of WNV that were then used to examine how future climate may affect the spread of the disease. Predictions were accurate for cases of human WNV infection in the following year (2012), with areas reporting infections having significantly higher probability of presence as predicted by our models. Projected geographic distributions of WNV in North America under future climate for 2050 and 2080 show an expansion of suitable climate for the disease, driven by warmer temperatures and lower annual precipitation that will result in the exposure of new and naïve host populations to the virus with potentially serious consequences. Our risk assessment identifies current and future hotspots of West Nile virus where mitigation efforts should be focused and presents an important new approach for monitoring vector-borne disease under climate change.

**Keywords:** climate change, disease ecology, ecological forecasting, ensemble distribution modeling, flavivirus, infectious diseases, maxent, West Nile virus

Received 7 November 2013 and accepted 14 December 2013

## Introduction

West Nile virus (WNV) (Flaviviridae: Flavivirus) is the most prevalent flavivirus in the world. Transmitted by mosquitoes, it was first detected in North America in 1999, and by 2004, the disease had spread across the contiguous United States. Since its introduction, WNV has infected over 30 000 humans and caused nearly 1500 deaths (Centers for Disease Control & Prevention (CDC) 2012, Public Health Agency of Canada 2012), and has had devastating impacts on native host populations (Ladeau *et al.*, 2007). WNV occurrence is influenced by environmental factors, including the composition of vertebrate host communities, climatic conditions, features of the landscape, and anthropogenic stressors (Kilpatrick, 2011). The presence of the virus is also intimately linked to its vector populations of mosquitoes (primarily species in the genus *Culex*), whose distributions are environmentally driven and

may be impacted by climate change (Soverow *et al.*, 2009; Kilpatrick, 2011). Despite intensive study, the links between outbreaks and environmental conditions remain unclear. Of particular interest to health professionals is how future climate conditions may influence the distribution and spread of the disease, and how outbreaks can be predicted.

Particular climatic conditions may facilitate WNV spread and proliferation within a community, either directly by influencing transmission or replication of the virus, or indirectly, by influencing the arthropod vectors' or avian hosts' ecology. For example, severe outbreaks of WNV in Europe and the United States have been associated with droughts (Marra *et al.*, 2004; Shaman *et al.*, 2005). It is believed that drought may facilitate mosquito-breeding conditions by concentrating available water, creating organically rich water bodies that congregate hosts and vectors, and increase the frequency of transmission events (Shaman *et al.*, 2005).

Previous studies linking environmental variables to outbreaks of WNV in either vectors or hosts have been

correspondence: Ryan J. Harrigan, tel. +1 310 206 6234, fax +1 310 825 5446, e-mail: iluvsa@ucla.edu

conducted at local geographic scales with limited temporal coverage. A wide range of correlates have been identified in these studies, including: long-term temperature patterns (Wimberly *et al.*, 2008), precipitation (Shaman *et al.*, 2005), land-use changes and urbanization (Ezenwa *et al.*, 2007), socioeconomic factors (Harrigan *et al.*, 2010), and host biodiversity (Kilpatrick *et al.*, 2006a). The many WNV environmental correlates underscore the complexity of virus-vector-host interactions, especially with different mosquito species, with different biological requirements (Rueda *et al.*, 1990), acting as the predominant vectors of WNV in different geographic regions. While factors other than climate have previously been reported to help predict distributions and prevalence of other vector-borne diseases, climate conditions including temperature and precipitation still help explain much of the variation present (Pérez-Rodríguez *et al.*, 2013). For regional and continental scales, climatic factors provide excellent predictive power in determining risk associated with WNV and other vector-borne diseases (Wimberly *et al.*, 2008; Sehgal *et al.*, 2011; Cornuall *et al.*, 2013). Most importantly, these climate factors can be modeled under future warming conditions so that long-term predictions of shifts in risk can be estimated.

Using a suite of species distribution modeling algorithms, available climate predictors, and WNV incidences in mosquito vectors, primary avian hosts, and secondary human hosts (see Supporting Online Information), we: (i) combine WNV data for the continental United States to map current distributions of the disease in vector and host populations; (ii) identify the most important climatic factors associated with WNV presence under current conditions; (iii) predict spatial and short-term temporal heterogeneity of WNV risk to identify annual hotspots; and (iv) project how these continental risks based on climate may shift under future climate change. By deriving a risk assessment of this devastating disease, we hope to inform mitigation and control strategies in areas that are or will be under threat due to suitable climatic conditions, as well as provide a blueprint for how risk assessments for other emerging infectious diseases may be modeled.

## Materials and methods

### Data collection

Data on West Nile virus 'positives', 'negatives', and 'no collection records' were processed from the Centers for Disease Control and Prevention website ([www.cdc.gov/ncidod/dvbid/westnile/index.htm](http://www.cdc.gov/ncidod/dvbid/westnile/index.htm)) for years 2003–2011. This database is a joint venture between the CDC and Arbonet, with maps of records provided by the United States Geological Service

(USGS). Records are available at the county level within the United States, and each county for each year was assigned to one of three categories: (i) positive for West Nile virus; (ii) negative for West Nile virus; or (iii) no data reported on West Nile virus presence. The centroid of each county was calculated and used as the point location for each record in further modeling analyses.

For each county and each year, three types of positives were recorded: (i) those found in vectors (mosquitoes); (ii) those found in primary hosts (birds); and (iii) those found in secondary hosts (humans). Although total counts were often reported at the county level, only records of presence or absence were used for a number of reasons. First, prevalence measures can be greatly influenced by the way in which samples are collected and/or recorded. Second, prevalence counts in vectors can be largely dependent upon vector control effort within each county, which can in turn be dependent on financial constraints of individual counties. Third, primary host prevalence (in birds) is difficult to measure, as only dead birds are actually tested for West Nile virus. Although some counties do support wildbird and sentinel surveillance programs, the majority of screening occurs within the avian family *Corvidae*, and within this group, dead American crows (*Corvus brachyrhynchos*) are the most often tested bird. Finally, the number of reported cases of West Nile virus in secondary human hosts is likely influenced by a number of factors including human population density, health care reporting practices, and screening and detection methods within hospitals.

### Environmental datasets

Observed climate data for individual years were acquired from the United States Historical Climatology Network, which provides monthly averages of data for 1218 stations across the contiguous United States (USHCN version 2.5.0, 10 October, 2012) (Menne *et al.*, 2009). A continuous data surface was interpolated from these data using ordinary kriging (Oliver & Webster, 1990) with a cell size of  $4 \times 4$  arcmin. Climatic variables were assigned to positive and negative WNV cases at the county level by using the centroid of a county and sampling each variable as an average of the surrounding nine cells ( $3 \times 3$  grid of  $4 \times 4$  arcmin each). Due to the fact that counties (and parishes) within the study area were not equally sized, we note here that climate estimates for larger counties (particularly in environmentally heterogeneous areas such as parts of the Midwest, California, and the Pacific Northwest) must be taken with caution, as errors in estimates may be larger here. Fifty-year mean climatologies from the WorldClim database [[www.worldclim.org](http://www.worldclim.org)] (Hijmans *et al.*, 2005) were used as a baseline for projecting models under future climate change. Initial models were performed using 19 bioclimatic variables, comprising measurements of temperature and precipitation. Due to high correlation between several of these variables (as determined by cross-correlational analyses, Pearson correlation coefficient  $> 0.9$ ) and the fact that some of these variables represent calculated values derived from raw measurements, a subset of eight variables, reflecting both temperature and precipitation measurements, were selected and used for subsequent analyses (See Table S1).

### Distribution models & performance

Several classes of species distribution models have been used to spatially predict diseases across various scales (Harrigan *et al.*, 2010; Sehgal *et al.*, 2011), and the performance of these models under a variety of conditions have been explored in numerous studies (Elith *et al.*, 2006; Siroky, 2009; Thomassen *et al.*, 2010). Often, one or a few of these models are selected based on either predictive skill and/or appropriateness of the model given the input dataset. Recently, however, ensemble modeling methods have been used as a way to avoid selection of any one particular model in favor of a community assemblage of models, either equally weighted or weighted according to the performance of each model. This ensemble modeling technique has been implemented within the R framework in the package BIOMOD (Thuiller *et al.*, 2009). However, these ensemble techniques compare models that incorporate both presence and absence records of species, and unfortunately do not include widely used stand-alone models such as Maxent (Phillips *et al.*, 2006) (Note: the option to run Maxent models in BIOMOD was added in the new package BIOMOD2). To include both presence and presence/absence models (absence data were available for primary vectors and hosts, but not for human cases), we compared the performance of 10 species distribution model techniques (Table S2), using the standard measure of the Receiver Operator Characteristic (ROC) to compare the performance of each model across each dataset. A total of 1620 runs were performed [six runs for each of 10 models across 9 years of data for three species groups (vectors, primary hosts, and secondary hosts)], in the following manner: five runs for which 80% of the data was randomly selected in training of the model, with the remaining 20% used to test the model, followed by a sixth and final run where all of the data were used in the training of the model. These runs incorporating all data points were only used for final visualizations of risk assessments. To measure model and variable performance, cross-validation of the five 80/20 replications was used (using the score obtained from the average of the cross-validations across five repetitions of 80/20 training/testing splits) for 27 separate datasets (vector, primary host, and human presence data for years 2003–2011). Maxent model performance was based on averages of the models for which 20% of the data was held back for testing. We used the following settings of Maxent: 10 000 background points; auto features; regularization multiplier = 1.0; maximum iterations = 500; convergence threshold = 0.00005.

### Weighted ensemble models

Model ensembles typically use either an unweighted mathematical average assembly or a weighted ensemble, for which models are ranked numerically according to their score on a standard measure of importance [i.e. Total Sum of Squares (TSS), Kappa, or ROC curve]. In the program BIOMOD, models are weighted by their ranking. However, weights are determined by a uniform decay index ( $d$ ) that does not consider absolute differences between model performance scores (Thuiller *et al.*, 2009). For instance, a model that is only slightly

worse than another model would still be down-weighted by the same decay index value as a model that performed much worse. In this way, these decay indices may be incorrectly penalizing models that perform well, but not as well as the top models in the ensemble. This may be particularly problematic in cases where several models perform well on a single dataset or when there are large discrepancies in model performances between classes of models.

To address these potential issues, we developed a weighting scheme that uses a decay index directly proportional to the score of the model being weighted (Eqn 1).

$$E = \sum_{i=1}^n M_i w_i \quad (1)$$

where:  $w_i = \frac{M_i}{M_1 + M_2 + \dots + M_i}$ , performance of model  $M_i$  as compared to other models' scores

$E$  = Weighted ensemble model score at each pixel

$i$  = Number of models evaluated in the ensemble

$M$  = Single species distribution model performance score at each pixel

This weighting scheme was applied to each prediction for each of the 27 datasets (vectors, primary hosts, and secondary hosts for years 2003–2011) to arrive at final weighted ensemble predictions. Final pixel values were made spatially continuous by means of ordinary kriging (Oliver & Webster, 1990).

### Individual variable contributions

Individual environmental variables and their comparative performance across models were determined using measures of importance; for all models except Maxent, this was measured as the overall decrease in performance of each model when the variable in question was removed from the model. If the models do not differ greatly in predictive performance, then the variable is deemed not important to the overall model. For Maxent, importance scores were measured slightly differently, as the variables are still included in the model, but randomly permuted to assess the decrease in ROC curve score as a measure of the actual variable importance in determining the model performance. Due to these slightly different means of importance calculations, importance scores were compared across all runs for the eight SDMs (no model is constructed for SREs) within BIOMOD, and then separately for the Maxent models (Table S2). Differences in variable ranking likely reflect these differences in importance calculations, as final spatial predictions for the top models were highly concordant.

### Future ensemble models

Future climate variables (the same suite used for current climate conditions in the study) were downloaded from WorldClim ([www.worldclim.org](http://www.worldclim.org)), in conjunction with the Consultative Group on International Agricultural Research (CGIAR) and the research program on Climate Change, Agriculture and Food Security [CCAFS, [www.ccafs-climate.org](http://www.ccafs-climate.org) (Ramirez, 2008)]. Seven global climate models (GCMs, Table



S5) from the IPCC 4<sup>th</sup> Assessment (Randall *et al.*, 2007) were used at a statistically downscaled resolution of 2.5 arcmin (~4 km). Models were statistically downscaled using the Climgen (Mitchell & Osborn, 2005) generator. Models were mathematically averaged for the projection period 2050–2060 and 2080–2090, and ensemble models used to predict current risk for West Nile virus were then applied to these future conditions. Future models were built under the intermediate A1B scenario (describing a future world of rapid economic growth, global population increases, and a balance of dependency on a variety of energy sources). The A1B scenario was chosen because these models represent the middle of the range for all SRES scenarios and as such are considered more realistic than models based on either extremely high A1 or extremely conservative B2 scenarios.

## Results

### Ensemble distribution model performance

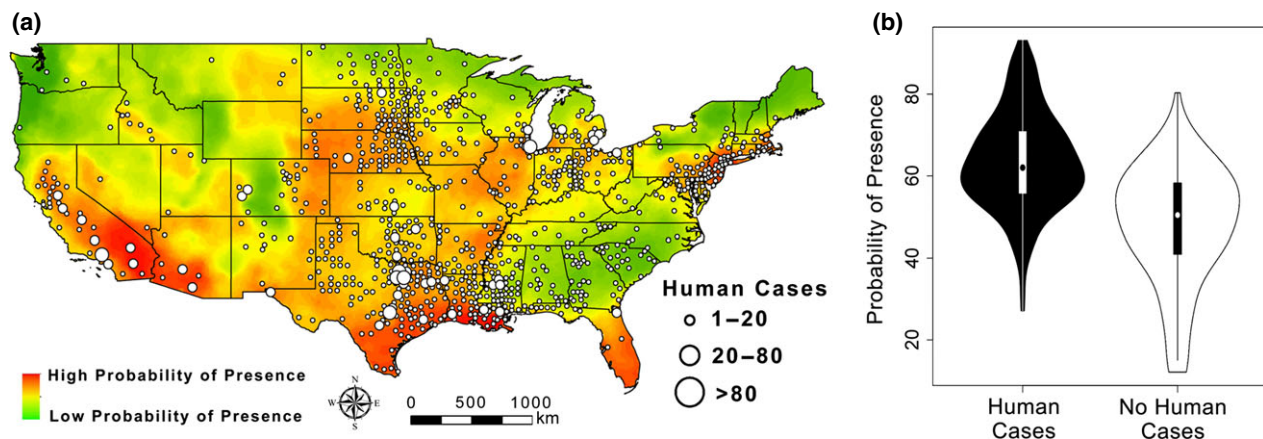
We predicted the spatial distribution of WNV for individual years for which WNV occurrence data were available (2003–2011). Results suggest that climate variables alone accurately predict spatial variation in the probability of presence of WNV, even when using a wide variety of species distribution models (Model ROC curve mean = 0.745, std. = 0.070, range = 0.542–0.935). Final ensemble models incorporating all models weighted by individual performance (see Materials & Methods) accurately predicted human cases of the disease in the following year (2012), including the first human case in Maine, and the particularly severe Texas outbreak in 2012, and assigned significantly higher risk

to counties with human cases ( $t = 12.78$ ,  $P < 0.001$ , Fig. 1). Severity of outbreaks was also well represented by ensemble models, with areas in Texas and Southern California receiving the highest probability of presence in ensemble models and reporting the highest number of human cases in the following year (2012). Conversely, small numbers of cases were reported in areas predicted by our ensemble models to have overall low probabilities of presence (i.e. Southeastern United States and areas of the Pacific Northwest).

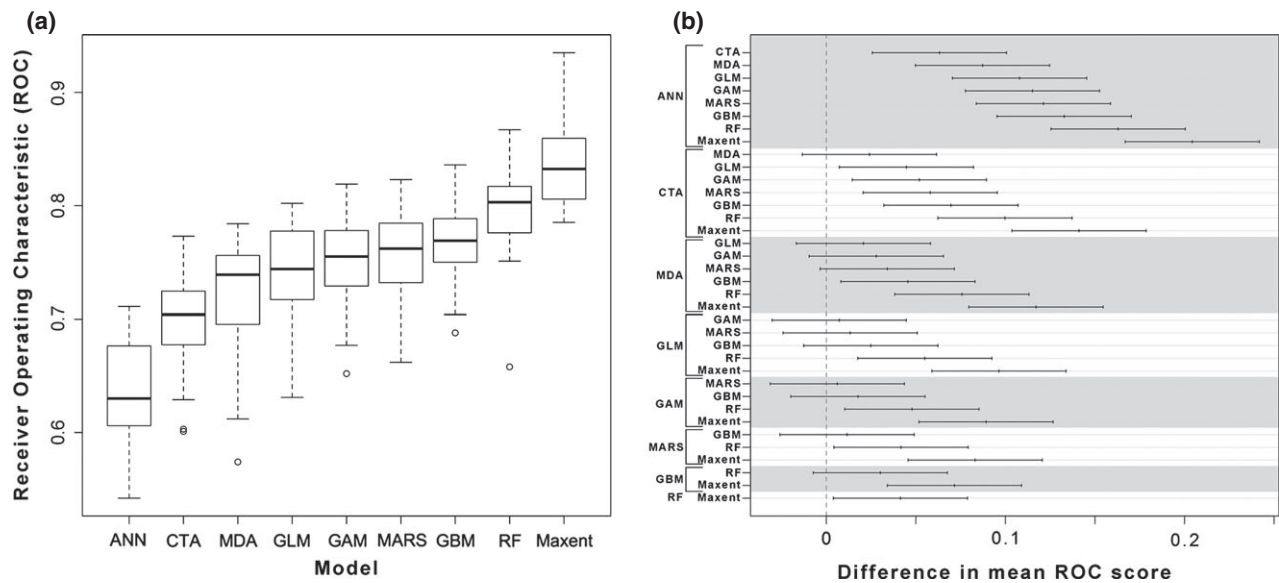
Model performance varied significantly among individual models (Fig. 2a and b). Among the 10 models evaluated, Maxent, and two of the nonlinear bifurcating models, Random Forests and Generalized Boosting Models, consistently ranked among the best performers across datasets (Fig. 2b; Table S4). While every dataset is likely to provide unique variation that particular models may perform well on, the consistency of our top-ranked models' performance across vectors and hosts, and across years using varying environmental predictors, suggests these nonlinear methods (Maxent, Random Forests, and Generalized Boosting Models) should be considered valuable tools in predicting the distribution of diseases. Results also suggest that distributions of diseases based on single algorithms may not always accurately capture the variation necessary for forecasting current and emerging diseases.

### Climate predictors of West Nile virus

The most important climate variables predicting WNV occurrence were maximum temperature of the warmest



**Fig. 1** Predicting the risk of West Nile virus (WNV) across the contiguous United States to identify annual outbreaks. (a) Predictions represent the mathematical average of ensemble models of probability of presence in vectors from years 2003–2011. Overlaid on these predictions are locations of human cases of WNV in 2012; circles are proportional to the number of cases reported. (b) Prediction models assigned significantly higher risk ( $t = 12.78$ ,  $P < 0.001$ ) to counties in the United States that had human cases of WNV in 2012 as compared to those that reported no human cases. Violin plots show kernel density distributions of predicted probabilities of presence for each county in the United States. Bars represent interquartile ranges.



**Fig. 2** Evaluation of species distribution models according to their predictive performance on final models. (a) Receiver Operating Characteristic (ROC) curve mean for each model, standard deviation is indicated by dashed lines, (b) Individual pairwise comparisons of model performance as assessed by ROC score. The dashed vertical line indicates no significant difference (using a 95% confidence interval) between mean scores. Comparisons that do not cross this vertical line are those where the model indicated to the left had significantly lower ROC scores than the model indicated within bracket.

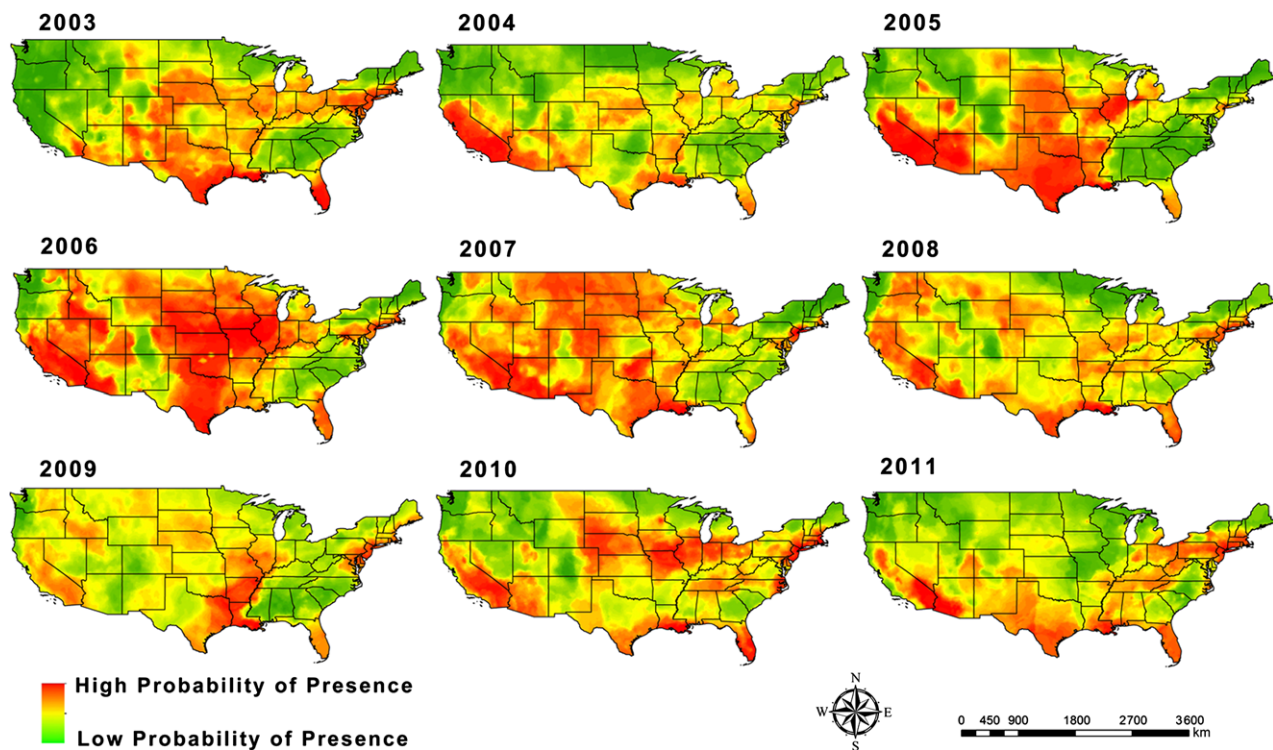
month and measures of annual and seasonal precipitation (Table S2; Figure S1). Higher maximum temperatures in the warmest month were associated with higher probability of presence in our analyses. Further, both annual total precipitation and precipitation seasonality were found to have an inverse relationship with the probability of WNV presence in vectors, primary hosts, and secondary hosts. Our findings suggest regions with lower annual total precipitation have higher probability of WNV cases. Regions with high annual precipitation and lower seasonal variation, such as the southeastern United States and the Pacific Northwest, were predicted to have lower probabilities of presence than drier more seasonal regions (Fig. 3, Figures S1–S3).

Comparisons of interannual variation in WNV probability of presence reveal that some regions are consistently at risk for WNV infections during summer months, whereas others appear more stochastic in their risk to WNV (Fig. 4). Regions such as the Pacific Northwest and Southeast have experienced consistently low probabilities of WNV presence. The Midwestern United States, a region often recording high numbers of WNV cases (CDC, 2012), shows high year-to-year variation in the probability of WNV presence, likely caused by a decrease in the number of cases in recent years (Figure S4). In contrast, the northeast regions and parts of the southern United States have experienced consistently higher probabilities of WNV presence, an indication that these regions are rarely spared annual WNV

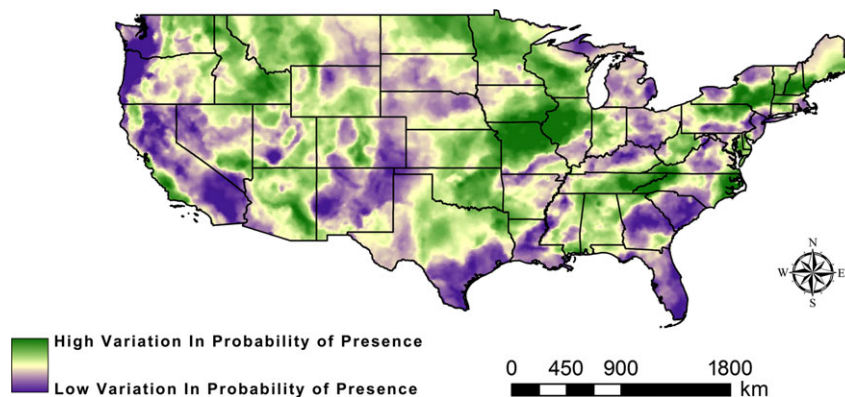
outbreaks (Fig. 4). These results suggest that mitigation efforts should be focused in these regions.

#### Future predictions

The projected geographic distribution of WNV in North America under future climate conditions for mid-century shows a northward and altitudinal expansion of the suitability of WNV under climate change (Fig. 5). Because current WNV distributions are highly dependent on temperature variables (e.g., maximum temperature of the warmest month), the ensemble model predicts that as temperatures in northern latitudes and western mountain regions continue to increase (Figures S5 and S6), conditions favorable for early amplification cycles and expanding vector distributions will enhance ecological conditions for WNV. Although projections of future precipitation across North America are more uncertain and seasonally dependent than those developed for temperature, annual mean precipitation is predicted to decrease across much of southern North America (Randall *et al.*, 2007) (Figures S5 and S6). Lower rainfall, particularly in the spring seasons in the southeast and southern Midwest United States, is predicted to increase WNV risk. Increases in rainfall across parts of central United States may lead to decreases in the probability of WNV, although regions of decreased risk are relatively low compared to those regions that will be at higher risk under climate change (Fig. 5).



**Fig. 3** Probability of presence of West Nile virus (WNV) in vectors from years 2003–2011, as determined by ensemble modeling of 60 runs across 10 separate models. Areas in red indicate high probability of presence (maximum probability = 1), while those in green represent areas of low predicted presence (minimum probability = 0).



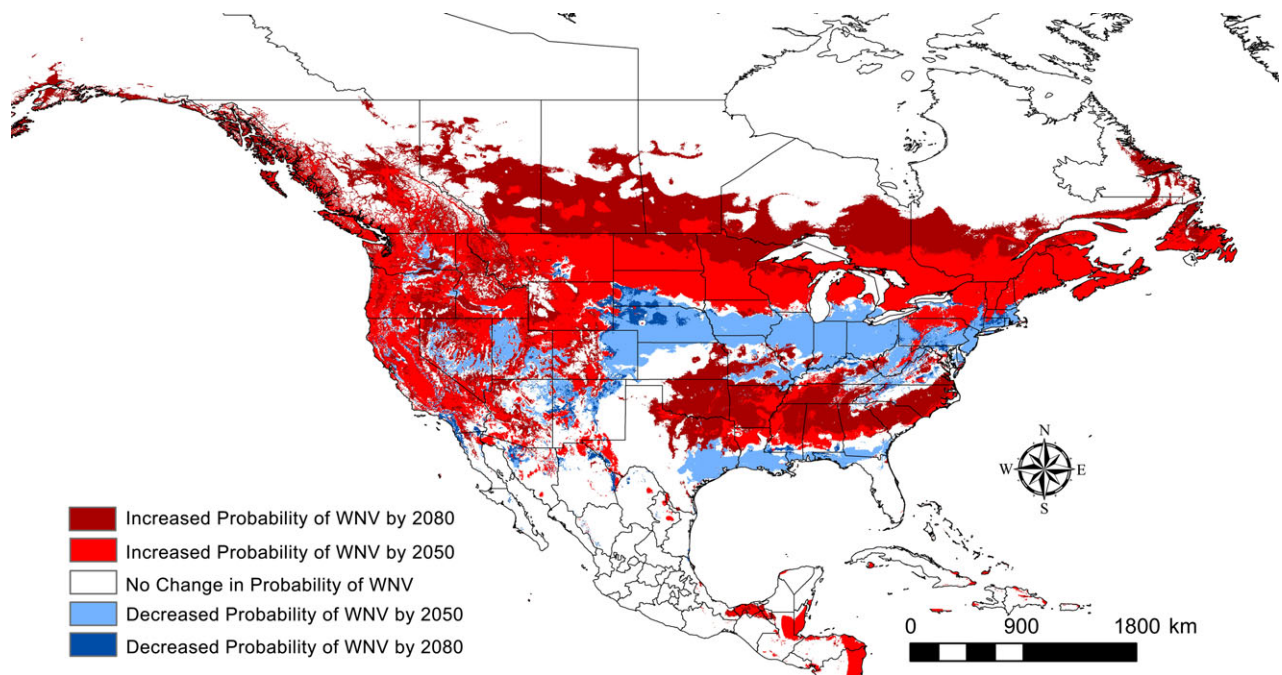
**Fig. 4** Variation in probability of presence of West Nile virus (WNV) in vectors, primary hosts, and secondary hosts from the year 2005 to the year 2011, as determined by ensemble models across years. While the Midwestern United States has historically had high numbers of WNV cases, recent years (2009–2011) saw a decrease in the number of cases in this region. Consistently low probabilities of presence in the Pacific Northwest and parts of the southwest region of the United States are contrasted with consistently high probabilities of presence in portions of the northeast and the Gulf Coast region. 2003 and 2004 were excluded from variation measures as WNV had not reached its current full continental range until 2005.

## Discussion

Climate has previously been shown to be a strong predictor of the distribution of vector-borne diseases (Wimberly *et al.*, 2008; Sehgal *et al.*, 2011), and we find that at continental scale, both temperature and

precipitation can be used to accurately assess risk of West Nile virus for both primary host and human populations. Maximum temperatures of the warmest months were shown to greatly influence the probability of presence, perhaps not surprisingly, given that variation in temperature has been shown to play an





**Fig. 5** Probability of presence of West Nile virus (WNV) in vectors projected across North America, for the year 2050 and 2080 under the A1B climate scenario. Areas in red indicate increased probability of presence (by at least 10% compared to current), while those in blue represent areas of decreased predicted prevalence (by at least 10% compared to current). Northern latitude increases are driven primarily by a positive relationship between temperature variables (BIO5, the maximum temperature of the warmest month), where as increases in WNV probability in the middle of the continent are driven by the counterintuitive inverse relationship between lower precipitation and higher WNV probability of presence.

important role in vector life cycles, and vectorial capacity, particularly for members of the *Culex* sp. complex (Rueda *et al.*, 1990; Kilpatrick *et al.*, 2008; Kilpatrick, 2011). Differences in vector life cycles may also lead to variation in how effectively WNV is transmitted during the course of typical spring and summer seasons (Miramontes *et al.*, 2006). Precipitation, too, is likely an important limiting factor for mosquito species, and several mechanistic hypotheses between precipitation and disease occurrence have been previously suggested (Epstein, 2001; Shaman *et al.*, 2005; Wimberly *et al.*, 2008). Somewhat counterintuitive is our result that lower annual precipitation is correlated with higher numbers of WNV cases. This could be explained either by reducing the number of predators and competitors for mosquito populations (Chase & Knight, 2003; Wang *et al.*, 2010) or by forcing closer congregations of vectors and hosts to areas with available standing water when rainfall is scarce or during droughts (Shaman *et al.*, 2005).

Although total number of WNV cases is difficult to predict among years, regions of highest risk appear to be consistent across years. These 'hotspots' include much of the California Central Valley, Southwest Arizona, southern Texas, and Louisiana. In 2012, one-third of all reported cases of WNV were from the

state of Texas (Murray *et al.*, 2013), centered around the city of Dallas and its suburbs, a region that was identified by our models as being of high risk for WNV. While total number of WNV cases was lower in this region in 2013, the relative number of cases was still highest in these climatically suitable regions.

Despite not including socioeconomic variables in our models, highly urbanized areas of the Northeast United States were also identified at risk for higher probabilities of WNV under present climate conditions. In 2012, the first human case of the disease was reported from Maine (CDC, 2012), and this county was ranked highest for probability of presence (0.64) among counties in Maine by our risk assessment model. This example may suggest that a subtle northward movement of WNV has already started, with shifting climate conditions that cause previously unsuitable, colder northern latitudes to become suitable for vector-borne diseases during summer months. Our results on future projections of WNV suggest that this temperature-driven northward expansion substantially accelerates during the course of a century, exposing many primary and secondary hosts to infection for the first time, which could have as devastating an impact on host populations as the initial emergence of WNV did in North America (Ladeau *et al.*, 2007; Kilpatrick, 2011). The

inverse relationship between precipitation and WNV presence can also contribute to future shifts in the distribution of the disease, because temperate regions are likely to experience new precipitation regimes under most climate change scenarios (Christensen *et al.*, 2007; Min *et al.*, 2011). Lower precipitation estimates for much of the southeastern United States suggest that these areas, previously experiencing only sporadic outbreaks of WNV, will have increased risk of the disease under climate change. Warmer temperatures are also likely to shift vector distributions, such that mosquitoes typically associated with more tropical distributions, may increase their northern ranges. Other vector-borne diseases, previously only found at more tropical latitudes, have been shown to have increasingly temperate distributions under climate change (Hales *et al.*, 2002; Loiseau *et al.*, 2012), underlining the importance of new, routine surveillance in areas previously thought to be immune to transmission of disease by vectors once considered tropical in nature.

We present the first predictions of continental WNV risk under present and future climate conditions for North America. However, there are important factors other than climate that influence infection risk. Local land-use changes, socioeconomic conditions, host switching by vectors, and host community diversity may all contribute to explaining presence and increased transmission of the disease (Kilpatrick, 2011; Brown, 2008; Ezenwa *et al.*, 2007; Harrigan *et al.*, 2010; Kilpatrick *et al.*, 2006a,b). Predictions of future land-use change or socioeconomic status carry with them even more uncertainty than climate predictions, and so these potentially important variables were excluded from future risk assessments. Nevertheless, our results show that the distribution of WNV across the continent is significantly influenced by large-scale climatic conditions that are changing at an unprecedented rate (Coumou & Rahmstorf, 2012). Most importantly, the distribution of WNV is predicted to shift northward and to higher elevations under future climate conditions. This will likely have significant negative consequences for naïve host populations not currently exposed to the disease. Novel host exposure to WNV (Ladeau *et al.*, 2007; Lindsey *et al.*, 2008) and other diseases (Ouellet *et al.*, 2005; Beadell *et al.*, 2006) has already been shown to have devastating impacts on populations. Accurate forecasting and risk assessment will be an important first step in mitigating threats to both wildlife and human populations.

## Acknowledgements

We thank Jennifer Ward, Esther Ann, and Allen Lee for assistance with data entry, and the Lloyd-Smith laboratory, Zhenxin

Fan, and Tony Chasar for suggestions on methods and presentation of results. This work was supported by a US Environmental Protection Agency grant (R 833778) to TBS. Although the research described in this article has been funded wholly or partially by the United States Environmental Protection Agency through grant/cooperative agreement R 833778, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## References

- Beadell JS, Ishtiaq F, Covas R *et al.* (2006) Global phylogeographic limits of Hawaii's avian malaria. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 2935–2944.
- Brown HE (2008) Ecological factors associated with West Nile virus transmission, Northeastern United States. *Emerging Infectious Diseases*, **14**, 1539–1545.
- Chase JM, Knight TM (2003) Drought-induced mosquito outbreaks in wetlands. *Ecology Letters*, **6**, 1017–1024.
- Christensen JH, Hewitson B, Busuioac A *et al.* (2007) Regional climate projections. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL), pp. 847–940. Chapter 11. Cambridge University Press, Cambridge.
- Cornuall J, Khimoun A, Harrigan RJ *et al.* (2013) The role of ecology in the geographical separation of blood parasites infecting an insular bird. *Journal of Biogeography*, **40**, 1313–1323.
- Coumou D, Rahmstorf S (2012) A decade of weather extremes. *Nature Climate Change*, **2**, 491–496.
- Elith J, Graham CH, Anderson RP *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.
- Epstein PDC (2001) West Nile virus and drought. *Global Change & Human Health*, **3**, 105–107.
- Ezenwa VO, Milheim LE, Coffey MF *et al.* (2007) Land cover variation and West Nile virus prevalence: patterns, processes, and implications for disease control. *Vector Borne Zoonotic Diseases*, **7**, 173–180.
- Hales SH, de Wet N, Maindonald J (2002) Potential effect of population and climate changes on global distribution of dengue fever: an empirical model. *The Lancet*, **360**, 830–834.
- Harrigan RJ, Thomassen HA, Buermann W *et al.* (2010) Economic conditions predict prevalence of West Nile virus. *Plos One*, **5**, e15437 doi: 10.1371/journal.pone.0015437.
- Hijmans RJ, Cameron SE, Parra JL *et al.* (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
- Kilpatrick AM (2011) Globalization, land use, and the invasion of West Nile virus. *Science*, **334**, 323–327.
- Kilpatrick AM, Daszak P, Jones MJ *et al.* (2006a) Host heterogeneity dominates West Nile virus transmission. *Proceedings of the Royal Society B-Biological Sciences*, **273**, 2327–2333.
- Kilpatrick AM, Kramer LD, Jones MJ *et al.* (2006b) West Nile virus epidemics in North America are driven by shifts in mosquito feeding behavior. *Plos Biology*, **4**, 606–610.
- Kilpatrick AM, Meola MA, Moudy RM *et al.* (2008) Temperature, viral genetics, and the transmission of West Nile virus by *Culex pipiens* mosquitoes. *Plos Pathogens*, **4**, e1000092.
- Ladeau SL, Kilpatrick AM, Marra PP (2007) West Nile virus emergence and large-scale declines of North American bird populations. *Nature*, **447**, 710–713.
- Lindsey NP, Kuhn S, Campbell GL *et al.* (2008) West Nile virus neuroinvasive disease incidence in the United States, 2002–2006. *Vector Borne Zoonotic Diseases*, **8**, 35–39.
- Loiseau C, Harrigan RJ, Cornel AJ *et al.* (2012) First evidence and predictions of *Plasmodium* transmission in Alaskan bird populations. *Plos One*, **7**, e44729.
- Marra PP, Griffing S, Caffrey C *et al.* (2004) West Nile virus and wildlife. *BioScience*, **54**, 393–402.
- Menne MJ, Williams CN, Vose RS (2009) The United States historical climatology network monthly temperature data. *Bulletin of the American Meteorological Society*, **90**, 993–1007.



- Min S-K, Zhang X, Zwiers FW *et al.* (2011) Human contribution to more-intense precipitation extremes. *Nature*, **470**, 378–381.
- Miramontes R, Laffery WE, Lind BK *et al.* (2006) Is agricultural activity linked to the incidence of human West Nile virus? *American Journal of Preventive Medicine*, **30**, 160–163.
- Mitchell TG, Osborn TJ (2005) ClimGen: a flexible tool for generating monthly climate data and scenarios. *Tyndall Centre for Climate Change Research Working Paper*.
- Murray KO, Ruktanonchai D, Hesalroad D, Fonken E, Nolan MS (2013) West Nile virus, Texas, USA, 2012. *Emerging Infectious Diseases*, **19**, doi:10.3201/eid1911.130768.
- Oliver MA, Webster R (1990) Kriging: a method of interpolation for geographical information systems. *International Journal of Geographical Information Systems*, **4**, 313–332.
- Ouellet M, Mikaelian I, Pauli BD *et al.* (2005) Historical evidence of widespread chytrid infection in North American amphibian populations. *Conservation Biology*, **19**, 1431–1440.
- Pérez-Rodríguez A, Fernández-González S, De La Hera I *et al.* (2013) Finding the appropriate variables to model the distribution of vector-borne parasites with different environmental pressures: climate is not enough. *Global Change Biology*, **19**, 3245–3253.
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- Ramirez JJA (2008) *High Resolution Statistically Downscaled Future Climate Surfaces* (ed. Agriculture ICFT). CGIAR Research Program on Climate Change, Agriculture and Food Security, Cali, Colombia.
- Randall DA, Wood RA, Bony S *et al.* (2007) Climate models and their evaluation. In: *Climate Change 2007: The Physical Science Basis; Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL), pp. 589–662. Chapter 8. Cambridge University Press, Cambridge.
- Rueda LM, Patel KJ, Axtell RC *et al.* (1990) Temperature-dependent development and survival rates of *Culex quinquefasciatus* and *Aedes aegypti* (Diptera: Culicidae). *Journal of Medical Entomology*, **5**, 892–898.
- Sehgal RN, Buermann W, Harrigan RJ *et al.* (2011) Spatially explicit predictions of blood parasites in a widely distributed African rainforest bird. *Proceedings of the Biological Sciences*, **278**, 1025–1033.
- Shaman J, Day JF, Stieglitz M (2005) Drought-induced amplification and epidemic transmission of West Nile Virus in southern Florida. *Journal of Medical Entomology*, **42**, 134–141.
- Siroky DS (2009) Navigating Random Forests and related advances in algorithmic modeling. *Statistics Surveys*, **3**, 147–163.
- Soverow JE, Wellenius GA, Fisman DN *et al.* (2009) Infectious disease in a warming world: how weather influenced West Nile virus in the United States (2001–2005). *Environmental Health Perspectives*, **117**, 1049–1052.
- Thomassen HA, Cheviron ZA, Freedman AH *et al.* (2010) Spatial modelling and landscape-level approaches for visualizing intra-specific variation. *Molecular Ecology*, **19**, 3532–3548.
- Thuiller W, Lafourcade B, Engler R *et al.* (2009) BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369–373.
- Wang G, Minnis RB, Belant JL *et al.* (2010) Dry weather induces outbreaks of human West Nile virus infections. *Bmc Infectious Diseases*, **10**, 38.
- Wimberly MC, Hildreth MB, Boyte SP *et al.* (2008) Ecological niche of the 2003 West Nile virus epidemic in the Northern Great Plains of the United States. *Plos One*, **3**, e3744. doi: 10.1371/journal.pone.0003744.

### Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Data S1.** Supplementary methods and results.