

Assessing the Influence of Climate on the Spatial Pattern of West Nile Virus Incidence in the United States

Morgan E. Gorris,^{1,2} James T. Randerson,³ Shane R. Coffield,³ Kathleen K. Treseder,⁴ Charles S. Zender,³ Chonggang Xu,⁵ and Carrie A. Manore⁶

¹Information Systems and Modeling, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

²Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

³Department of Earth System Science, University of California, Irvine, Irvine, California, USA

⁴Department of Ecology and Evolutionary Biology, University of California, Irvine, Irvine, California, USA

⁵Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

⁶Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

BACKGROUND: West Nile virus (WNV) is the leading cause of mosquito-borne disease in humans in the United States. Since the introduction of the disease in 1999, incidence levels have stabilized in many regions, allowing for analysis of climate conditions that shape the spatial structure of disease incidence.

OBJECTIVES: Our goal was to identify the seasonal climate variables that influence the spatial extent and magnitude of WNV incidence in humans.

METHODS: We developed a predictive model of contemporary mean annual WNV incidence using U.S. county-level case reports from 2005 to 2019 and seasonally averaged climate variables. We used a random forest model that had an out-of-sample model performance of $R^2 = 0.61$.

RESULTS: Our model accurately captured the V-shaped area of higher WNV incidence that extends from states on the Canadian border south through the middle of the Great Plains. It also captured a region of moderate WNV incidence in the southern Mississippi Valley. The highest levels of WNV incidence were in regions with dry and cold winters and wet and mild summers. The random forest model classified counties with average winter precipitation levels <23.3 mm/month as having incidence levels over 11 times greater than those of counties that are wetter. Among the climate predictors, winter precipitation, fall precipitation, and winter temperature were the three most important predictive variables.

DISCUSSION: We consider which aspects of the WNV transmission cycle climate conditions may benefit the most and argued that dry and cold winters are climate conditions optimal for the mosquito species key to amplifying WNV transmission. Our statistical model may be useful in projecting shifts in WNV risk in response to climate change. <https://doi.org/10.1289/EHP10986>

Introduction

West Nile virus (WNV; Flaviviridae: flavivirus) is the leading cause of mosquito-borne disease in humans in the United States.¹ WNV is primarily transmitted between birds—the viral reservoir—and mosquitoes—the viral vector. Humans are dead-end hosts that can contract WNV when bitten by an infected mosquito; ~20% of humans who are infected by WNV are symptomatic, with a small percent of cases developing into neuroinvasive disease.² WNV was first introduced to the United States in New York in 1999, and by 2005 it was found throughout the contiguous United States.^{3,4} From 2005 to 2019, the annual number of cases ranged between 712 and 5,674 cases per year.³ Although both the number and spatial distribution of WNV cases has varied considerably, over the past 5 y the total number of cases in the contiguous United States has stabilized to ~2,000 cases per year. Many of the cases occur in counties with large population centers in the western United States. However, WNV incidence, or the number of cases normalized by population, has been particularly high in the northern Great Plains region.^{5,6} Because each step of the WNV transmission cycle is sensitive to environmental conditions, climate may be influencing the spatial pattern of WNV incidence.

One important climate control on the WNV transmission cycle may be precipitation. Precipitation regulates the distributions and permanence of landscape water bodies that bring mosquitoes and birds into contact. The *Culex* genus of mosquitoes, some of which frequently take blood meals from humans, is thought to be the most important transmitter of WNV.^{7–9} WNV and *Culex* mosquitoes are often found in temperate climate zones, whereas areas with higher moisture availability in maritime tropical climates favor more tropical mosquito species that are not as effective at transmitting WNV.^{10–12} In temperate regions, one species of mosquitoes that may be an exceptionally important transmitter of WNV is *Culex tarsalis*.⁹ *Cx. tarsalis* thrives in areas with irrigation,^{13,14} which are prevalent in the agricultural regions throughout the Great Plains (i.e., flatlands west of the Mississippi River and east of the Rocky Mountains), where levels of WNV incidence are relatively high. *Cx. tarsalis* is known to change its seasonal preference from feeding primarily on birds in the spring to a more generic feeder in summer and fall.⁹ This cycle would first increase the presence of WNV among birds early in the year and then later, as the mosquitoes switch hosts, increase the likelihood of transmission to humans. An increase in the likelihood of transmission from birds to humans via mosquitoes may be further amplified in areas where mosquitoes, birds, and humans are all co-located around limited water sources, such as sewage drains or artificial standing water in highly urbanized areas or irrigated rural areas.

In addition to precipitation, temperature may be another important climate control on the spatial structure of WNV incidence. Mosquitoes have physiological temperature limits that allow them to successfully overwinter and to survive the summer heat.¹⁵ For example, *Culex quinquefasciatus* survives well at temperatures around 20°C–30°C, but survival drops drastically below 15°C and above 34°C.¹⁵ *Cx. quinquefasciatus* generally lives in regions with mean annual summer temperatures between 24°C and 28°C, and *Culex pipiens* lives in regions between 16°C and 24°C.^{16,17} Moreover, temperatures determine when mosquitoes emerge in spring, their time to adult maturation, reproduction rate, efficacy for

Address correspondence to Morgan E. Gorris, PO Box 1663, Los Alamos, NM 87545 USA. Email: mgorris@lanl.gov
Supplemental Material is available online (<https://doi.org/10.1289/EHP10986>).

The authors declare they have nothing to disclose.

Received 20 January 2022; Revised 15 February 2023; Accepted 28 February 2023; Published 27 April 2023.

Note to readers with disabilities: EHP strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in EHP articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehpsubmissions@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

transmitting WNV, and adult longevity, although these relationships differ among mosquito species.^{16,17} If temperatures become too warm, especially in the hot summer months, mosquito mortality may greatly increase.¹⁶ Hot weather during summer months, when mosquitoes are generally more active, also may cause humans to limit the time they spend outdoors, thus limiting their exposure to mosquitoes.¹⁸ For these reasons, both seasonal temperature thresholds and temperature optimums may be important for structuring the spatial pattern of WNV incidence.

Seasonal temperatures also affect when birds migrate into a region, when they breed, and their geographical distribution.^{19,20} Specific bird species are thought to be important amplifiers of WNV transmission, especially the American robin (*Turdus migratorius*, hereafter “robins”).^{7,21–23} The American crow (*Corvus brachyrhynchos*, hereafter “crows”) may also play an important role in overwintering the virus as it passes from crow to crow within roosts.^{24,25}

So far, studies examining the effects of climate on the spatial pattern of WNV in the United States have been mostly limited to the regional-level or finer and not explored across the contiguous United States as a whole.^{26–29} Our goal was to develop a predictive model of WNV incidence in the United States and explore which seasonal climate conditions constrain the spatial extent of WNV incidence across the United States. Specifically, we aimed to determine which contemporary climate conditions support the highest levels of WNV incidence. First, we used county-level mean annual WNV incidence data to examine the statistical relationships between incidence and seasonal averages of surface air temperature and precipitation. Then, recognizing that the relationships between climate conditions and WNV incidence are highly nonnormal and have multimodal peaks, we employed a random forest model to capture the complexity of the system. In our “Discussion” section below, we hypothesize the biological mechanisms within the WNV transmission cycle that could be related to these climate conditions to ensure our predictive model has a sound biological basis. Our model may provide a means to project WNV incidence in response to climate change, because warming temperatures and changes in precipitation may shift the regions most affected by this disease.

Methods

WNV Data

We obtained human case counts of WNV at the county level by year for the contiguous United States from 1999 to 2019 from the United States Centers for Disease Control and Prevention (U.S. CDC). We converted case counts to WNV incidence (cases per 100,000 population) using annual county-level population estimates from the U.S. Census Bureau.^{30,31} Between 1999 and 2005, WNV was spreading westward across the United States from its introduction point on the East Coast. Thus, we limited our analyses of the contemporary spatial pattern of WNV cases and mean incidence during the period 2005–2019. During that time, there were no known changes in case reporting practices that would have added significant bias into case numbers. In total, our WNV data set included 35,081 disease cases, including both neuroinvasive and non-neuroinvasive cases. We performed our analysis at the county level, which is the finest spatial resolution available from the U.S. CDC for de-identified, aggregated case data; the contiguous United States contains 3,108 counties. No human subjects or institutional review board/ethics review was required for this study because no personal identifiable information was released in conjunction with the human case counts of WNV.

Climate Data

We used monthly surface air temperature and precipitation from the Precipitation elevation Regressions on Independent Slopes

Model (PRISM), available as 4-km gridded products.^{32,33} To compare climate with our county-level WNV incidence data, we spatially averaged the native gridded climate data to the county level using county shapefiles from the 2018 TIGER/Line Shapefiles product by the U.S. Census Bureau.³⁴ We calculated 3-month mean seasonal climate variables from monthly observations during the period December 2004–November 2019 for direct comparison with our WNV data set, starting with winter data prior to the summer WNV season. The different seasons we used for averaging the climate information were winter [December, January, and February (DJF)], spring [March, April, and May (MAM)], summer [June, July, and August (JJA)], and fall [September, October, and November (SON)].

Random Forest Model

We used a mean random forest (RF) model [R package “randomForest” (version 4.6-14; <https://cran.microsoft.com/snapshot/2020-04-20/web/packages/randomForest/index.html>)]³⁵ to explore which seasonal climate conditions are important in structuring the spatial pattern of WNV incidence in the United States. RF models have been used in a variety of geoscience applications with general accuracy.^{36–38} Though RF models are primarily considered predictive models, they do offer insight on variable importance.³⁹ The technique of creating an additional summary tree to look inside the so-called “black box” of the RF has also been previously implemented for geoscience.⁴⁰ We chose to use a RF approach to balance the ability to create a predictive model of complex data without losing model interpretability. We created our mean RF model by running 500 iterations of the RF, each time on a different, randomly selected set of training data. We used 80% of the counties for training the model and the remaining 20% for testing. Each iteration of the model generated 1,000 separate regression trees. We randomly selected three predictor variables as candidates at each split in the RF (parameter known as m_{try} , where $m_{try} = \sqrt{p}$ and p is the number of predictor variables).⁴¹ To avoid overfitting, we used a minimum node size of 5, so that each terminal node (i.e., leaves) in our regression trees described at least five counties. We used a maximum number of nodes of 20, so that we did not lose interpretability of our model results. RF models are relatively insensitive to tuning parameters, and the defaults perform well on most data sets.^{41–43}

RF models do not inherently account for spatial autocorrelation. To determine whether we needed to incorporate spatial information into our model, we first tested the WNV case data for spatial autocorrelation using Moran’s I test. We created a binary, spatial adjacency matrix using the county shapefile in QGIS (version 3.10; QGIS Development Team) where counties were defined as adjacent if they had touching geopolitical boundaries. With a Moran’s I observed value = 0.6901, expected value = –0.0003, at $p < 0.001$, we rejected the null hypothesis that there is no spatial autocorrelation present in the mean WNV incidence data. To account for this spatial autocorrelation, we applied eigenvector spatial filtering (ESF) to our RF.^{44,45} We computed the eigenvalue spatial decomposition of the spatial adjacency matrix, creating the same number of eigenvectors as counties ($n = 3,108$). Adding these eigenvectors as predictors into the RF acts as control variables to identify and isolate stochastic spatial dependencies among the counties, so the model can proceed as if the observations are independent.⁴⁴ Because there is currently no set standard for the number of eigenvectors to include, we chose to include the first three as predictor variables in the RF. This choice was made based on exploring the number of eigenvectors, and therefore additional variables, in the RF and the observed diminishing increase in model performance.

We estimated the spread in our RF predictions by calculating the standard deviation (SD) across the 500 iterations. We also averaged the variable importance and node purities across the

iterations. We measured the importance of each climate variable in the model by calculating the decrease in modeled mean-square error from randomly permuting each predictor variable. We averaged the out-of-sample county-level estimations of WNV incidence from the 500 iterations to create our mean RF and for reporting our final estimation of county-level mean annual WNV incidence.

To understand important splitting points for each of the seasonal climate predictors, we ran an additional regression tree using the out-of-sample predicted county-level WNV incidence from our mean RF as input training data [R package “rpart” (version 4.1-15; <https://cran.r-project.org/web/packages/rpart/rpart.pdf>)].⁴⁶ This created a summary tree similar in performance to the mean RF [$R^2 = 0.58$, root-mean-square error (RMSE) = 3.5 cases per 100,000 population per year]. This tree allowed us to explore important climate thresholds that structure areas of higher and lower WNV incidence.

Our analyses were performed in R (version 3.6.1; R Development Core Team). We created our figures and maps in Python (version 3.7.6), using package “Basemap” (version 1.2.0) and package “matplotlib” (version 3.1.3). We edited our figures for layout in Adobe Illustrator 2022.

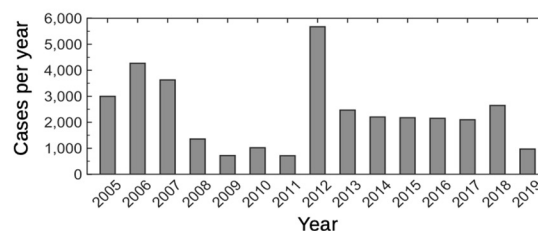
Results

The Mean Spatial Pattern and Magnitude of WNV in the United States

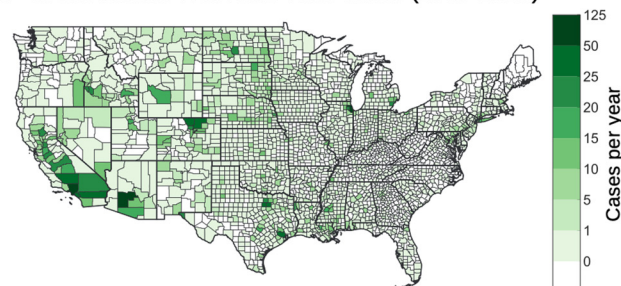
Despite interannual variability in the location and number of WNV cases (Figure 1A), the maps of mean annual WNV cases (Figure 1B) and mean annual WNV incidence (Figure 1C) showed distinct spatial patterns. Many counties had low levels of WNV incidence, and very few had high levels: 1,008 counties (33%) have never had a case of WNV; 1,749 counties (56%) have mean annual incidence averaged from 2005 to 2019 fewer than 5 cases per 100,000 population per year; 224 counties (7%) have mean annual incidence from 5 to <15 cases per 100,000 population per year; and 127 counties (4%) have mean annual incidence of at least 15 cases per 100,000 population per year (Figure 1D). The highest numbers of mean annual WNV cases occurred in counties with large cities. The 10 counties with the highest mean annual number of cases from 2005 to 2019 included the cities of Los Angeles (Los Angeles County, California), Phoenix (Maricopa County, Arizona), Chicago (Cook County, Illinois), Dallas (Dallas County, Texas), Anaheim/Santa Ana/Irvine (Orange County, California), Houston (Harris County, Texas), Fort Worth (Tarrant County, Texas), Bakersfield (Kern County, California), Riverside (Riverside County, California), and Fort Collins (Larimer County, Colorado); together, this set accounted for 21% (492 cases) of the mean annual number of cases during this period. Half of the mean annual number of WNV cases (1,172 cases) can be attributed to just 70 of the 3,108 counties in the United States. Apart from Chicago, Illinois, and Detroit, Michigan, it was more challenging to identify large population centers by the number of WNV cases across the eastern half of the United States.

Normalizing by population, the WNV incidence map revealed a distinct V-shaped pattern of elevated incidence throughout the Great Plains of the United States (Figure 1C). The high-incidence region spanned several northern states on the Canadian border (Montana, North Dakota, and western Minnesota) and narrowed to the south through Wyoming, South Dakota, Nebraska, eastern Colorado, western Kansas, western Oklahoma, and northwestern Texas. There was also a hot spot of elevated WNV incidence in southern Idaho and along the Oregon–Idaho border on the southern Columbia Plateau. WNV incidence was low in the eastern United States, except for a pocket of moderate incidence throughout Louisiana and Mississippi in the southern Mississippi

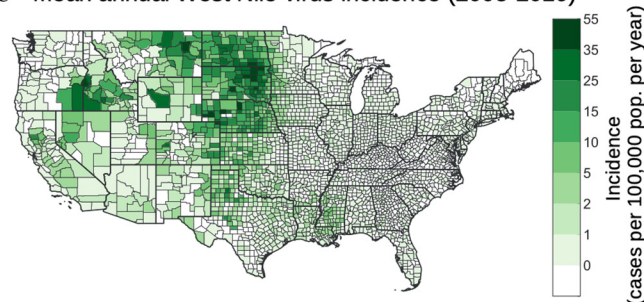
A Annual US West Nile virus cases



B Mean annual West Nile virus cases (2005-2019)



C Mean annual West Nile virus incidence (2005-2019)



D Histogram of mean annual incidence (2005-2019)

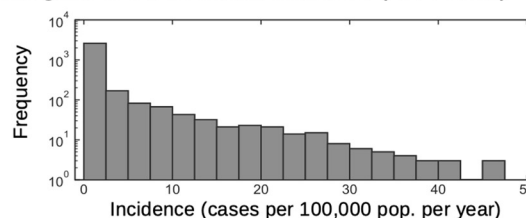


Figure 1. Total number of cases and mean spatial pattern of WNV in the United States during the period 2005–2019, including (A) a time series of WNV cases for the contiguous United States derived from data compiled by the U.S. Centers for Disease Control and Prevention for the sum of neuroinvasive and non-neuroinvasive cases; (B) a county-level map of the number of mean annual WNV cases averaged from 2005–2019; (C) a county-level map of mean annual incidence (cases per 100,000 population per year) derived from the case data shown in panel (B) and annual, county-level U.S. Census population estimates; and (D) the frequency distribution of mean annual incidence across the 3,108 counties in the United States. See Excel Table S1, S2, and S3 for corresponding numerical data. Graphs and maps were created in Python (version 3.7.6), using package “Basemap” (version 1.2.0), and “matplotlib” (version 3.1.3). We edited our figures for layout in Adobe Illustrator 2022. Note: WNV, West Nile virus.

River Valley. In subsequent analyses, our goal was to identify the climate drivers that help explain the V-shaped zone of elevated incidence and, more generally, the full spatial pattern of incidence.

Statistical Relationships between Seasonal Climate and WNV Incidence

By comparing county-level mean annual WNV incidence with seasonal precipitation and temperature variables, we found that

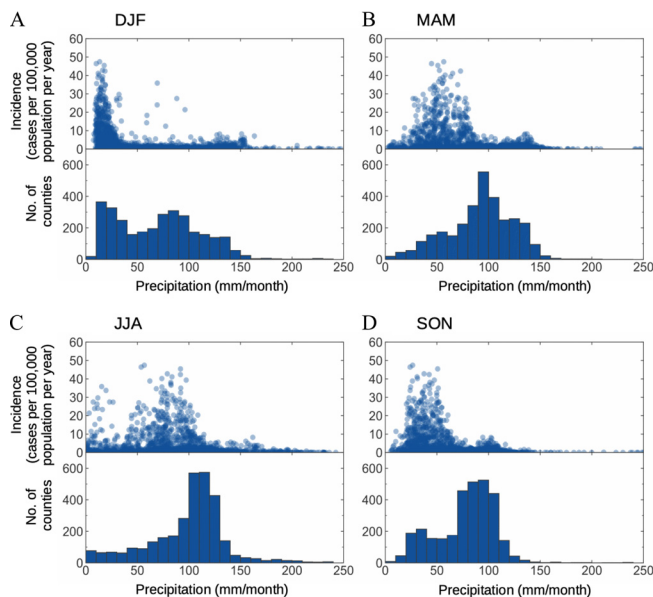


Figure 2. Incidence and histograms of the county number of WNV cases as a function of precipitation. Four seasonal intervals are shown: (A) winter (DJF), (B) spring (MAM), (C) summer (JJA), and (D) fall (SON). In each histogram, the number of counties was aggregated in 10-mm/month precipitation bins. Mean annual incidence and mean monthly precipitation variables during the period 2005–2019 were used to create the plots. See Excel Table S3 for corresponding numerical data. Surface air temperature and precipitation data are from the Precipitation elevation Regressions on Independent Slopes Model (PRISM). West Nile virus case data is provided by the U.S. Centers for Disease Control and Prevention. Note: DJF, December, January, February; JJA, June, July, August; MAM, March, April, May; SON, September, October, November; WNV, West Nile virus.

areas with dry and cold winters supported higher levels of incidence (Figures 2 and 3). From visual inspection, an important threshold occurred for winter (DJF) precipitation at a level of about 25 mm/month (Figure 2A). A total of 560 counties had precipitation levels below this threshold; this set of counties had a mean annual incidence of 8.8 cases per 100,000 population per year ($SD = 9.5$). In contrast, the 2,548 counties above this threshold had a mean annual incidence of 0.7 cases per 100,000 population per year ($SD = 2.1$). This factor of 13 difference in incidence for these two sets was highly significant when evaluated using a one-tailed independent t -test [$t(3,106) = 3.4$, $p = 0.0003$]. Low precipitation during other seasonal periods was also associated with higher incidence but with less-defined breakpoints.

Similarly, counties that had mean winter temperatures below 0°C ($n = 1,298$) had higher levels of WNV incidence [~ 4.0 cases per 100,000 population per year (Figure 3A)]. In contrast, counties with winter temperature above 0°C had lower incidence (0.9 cases per 100,000 population per year). Cooler temperatures during other seasons also favored higher levels of WNV incidence (Figure 3B and D). WNV incidence was relatively low in counties with summer temperatures above 24°C (Figure 3C). The strong nonlinearities and apparent threshold-type relationships between seasonal temperature and precipitation and WNV incidence provided motivation for the use of a RF model (Figures 2 and 3).

To further explore the relationship between seasonal climate drivers and mean annual WNV incidence, we calculated the county-level univariate linear correlation between each seasonal climate driver and mean annual WNV incidence (Table 1). Among all seasonal climate drivers, fall precipitation had the strongest negative correlation with WNV incidence ($r = -0.43$), so counties with lower fall precipitation had higher WNV incidence. There was also a relatively strong negative correlation

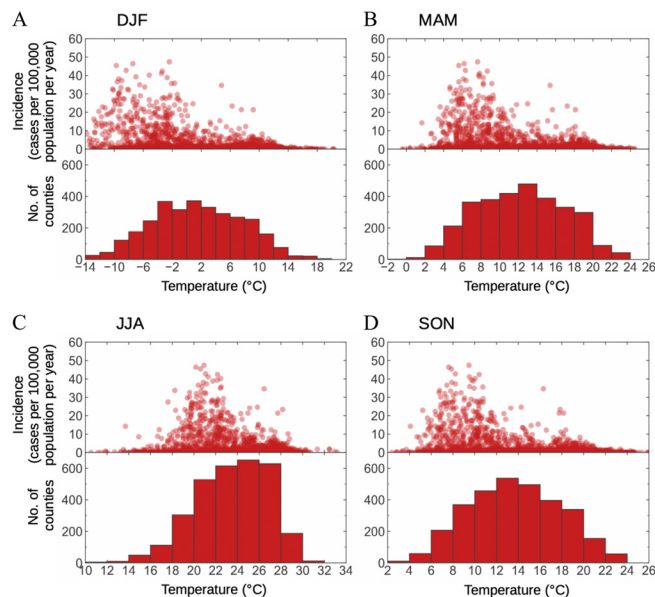


Figure 3. Incidence and histograms of county number as a function of air temperature. Four seasonal intervals are shown: (A) winter (DJF), (B) spring (MAM), (C) summer (JJA), and (D) fall (SON). In each histogram, the number of counties was aggregated in 2°C temperature bins. Mean annual incidence and mean surface air temperature during the period 2005–2019 were used to create these plots. See Excel Table S3 for corresponding numerical data. Surface air temperature and precipitation data are from the Precipitation elevation Regressions on Independent Slopes Model (PRISM). WNV case data is provided by the U.S. Centers for Disease Control and Prevention. Note: DJF, December, January, February; JJA, June, July, August; MAM, March, April, May; SON, September, October, November; WNV, West Nile virus.

between precipitation and WNV incidence in winter ($r = -0.37$) and spring ($r = -0.35$). Among the seasonal temperature variables, winter temperature had the strongest negative correlation with WNV incidence ($r = -0.29$), but with a magnitude that was lower than the fall, winter, and spring precipitation variables. The levels of variance explained by these single variables provided a baseline for evaluating the success of our more complex RF model.

When building and interpreting our model of the spatial pattern of WNV incidence, it was also important to consider the collinearity among the different climate drivers. The spatial structure of temperature was highly correlated among all the different seasons (Figure 4; 0.80 – 0.99), whereas seasonal precipitation exhibited relatively higher levels of correlation among fall, winter, and spring seasons (0.76 – 0.83) but not between summer and other seasons (0.23 – 0.53). Across temperature and precipitation variables,

Table 1. The county-level univariate Pearson correlation coefficient (r), associated p -values, and coefficient of determination (R^2) between each county-level seasonal climate variable and WNV incidence in the United States from 2005 to 2019.

Season	Explanatory variables					
	Precipitation			Temperature		
	Pearson r	p -Value	R^2	Pearson r	p -Value	R^2
DJF	−0.37	5.43×10^{-100}	0.13	−0.29	4.00×10^{-62}	0.09
MAM	−0.35	5.52×10^{-92}	0.12	−0.25	4.18×10^{-44}	0.06
JJA	−0.23	8.17×10^{-40}	0.05	−0.14	6.69×10^{-15}	0.02
SON	−0.43	5.53×10^{-140}	0.19	−0.27	5.50×10^{-52}	0.07

Note: Surface air temperature and precipitation data are from the Precipitation elevation Regressions on Independent Slopes Model (PRISM). DJF, December, January, February; JJA, June, July, August; MAM, March, April, May; SON, September, October, November; WNV, West Nile virus.

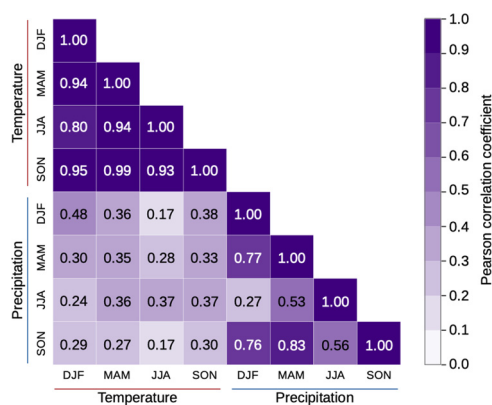


Figure 4. A spatial correlation matrix between the eight seasonal climate variables across all counties in the contiguous United States. The matrix displays the Pearson correlation coefficients between each pair of predictor variables. Each predictor variable consisted of a vector with 3,108 elements, with each element representing the mean climate condition for a single county. Surface air temperature and precipitation data are from the Precipitation elevation Regressions on Independent Slopes Model (PRISM). Note: DJF, December, January, February; JJA, June, July, August; MAM, March, April, May; SON, September, October, November.

correlations were considerably lower (0.17–0.48), with the largest positive correlation occurring between precipitation and temperature during winter. The greater degree of independence among precipitation and temperature variables suggested that more complex models may be able to capture the spatial structure of disease incidence by combining information from both variables.

RF Model of WNV Incidence

Our RF model successfully captured the V-shaped area of increased WNV incidence throughout the Great Plains (Figure 5A). The R^2 performance between predicted and observed incidence of our RF model averaged across the 500 iterations was 0.70 (0.65–0.74) for our in-sample training data and 0.61 (0.43–0.74) for our out-of-sample testing data. Relative to the 2005–2019 mean county incidence value of 2.2 cases per 100,000 population per year, the RMSE for our in-sample training data was 3.0 cases per 100,000 population per year (2.7–3.2). For our out-of-sample testing data, the RMSE was 3.4 cases per 100,000 population per year (2.5–4.3).

The model produced a smoothed pattern of WNV incidence levels between counties even though there was county-level heterogeneity in the observed mean annual incidence data (Figure 1C). The highest levels of SD between the 500 iterations were in the regions with the highest levels of incidence, whereas counties with low incidence had correspondingly low SDs (Figures 1C, 5B, 6A). Overall, our RF model had a slight high-incidence bias throughout the western United States, yet still underestimated the incidence in counties with the highest levels, including those in the central Great Plains (Colorado, Kansas, Montana, Nebraska, North Dakota, South Dakota, Texas), Columbia Plateau (Idaho, Nevada, Oregon), and Mississippi Valley (Arkansas, Louisiana, Mississippi; Figure 5C). The model underestimated incidence for any county above a mean annual incidence of 20 cases per 100,000 population per year (Figure 6B).

The measure of variable importance indicated winter precipitation, fall precipitation, and winter temperature were the top three most important seasonal climate drivers to determine the county-level WNV incidence (Table 2). Additionally, our summary tree identified summer temperature and precipitation as another important conditional split to determine levels of incidence (Figure 7). Two splits on the second eigenvector were also

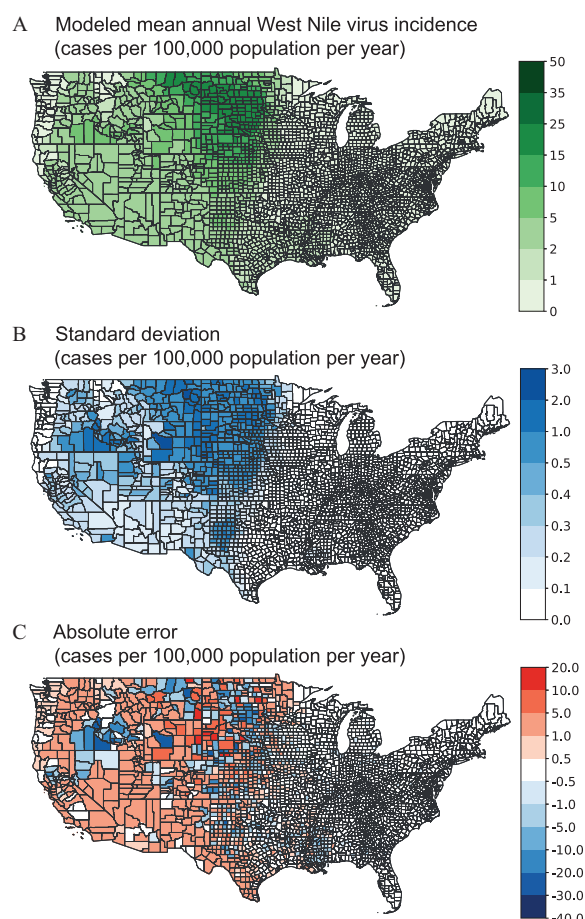


Figure 5. A map of (A) modeled mean annual WNV incidence predicted by the RF model where each county is the average of the 500 iterations of out-of-sample testing data, (B) the standard deviation across those 500 iterations, and (C) the absolute error from observed WNV incidence levels (2005–2019). All are in units of cases per 100,000 population per year. See Table S3 for corresponding numerical data. Analyses were performed in R (version 3.6.1; R Development Core Team), maps were created in Python (version 3.7.6), using package “Basemap” (version 1.2.0) and “matplotlib” (version 3.1.3). We edited our figures for layout in Adobe Illustrator 2022. Note: RF, random forest; WNV, West Nile virus.

included in the summary tree, both of which mediated latitudinal splits in the county-level incidence. Counties that had dry and cold winters and wetter and warm summers were assigned the highest level of WNV incidence. On average, counties with winter precipitation levels <23.3 mm/month had incidences nearly 11 times greater than counties that were wetter.

Our summary tree was comprised of one long and one short branch (Figure 7). The initial conditional split for winter precipitation produced a similar V-shaped area throughout the Great Plains for mean annual WNV incidence. Along the short branch, the conditional pattern of fall precipitation was an important factor for limiting WNV incidence levels in the eastern United States. Along the long branch, temperature and eigenvector conditional breaks were staggered latitudinally in comparison with the precipitation breaks.

Discussion

The Relationships between Climate and WNV Incidence

Climate conditions are likely structuring both the mean spatial extent of WNV incidence and the interannual variability in the number and location of disease cases. Our study was the first to

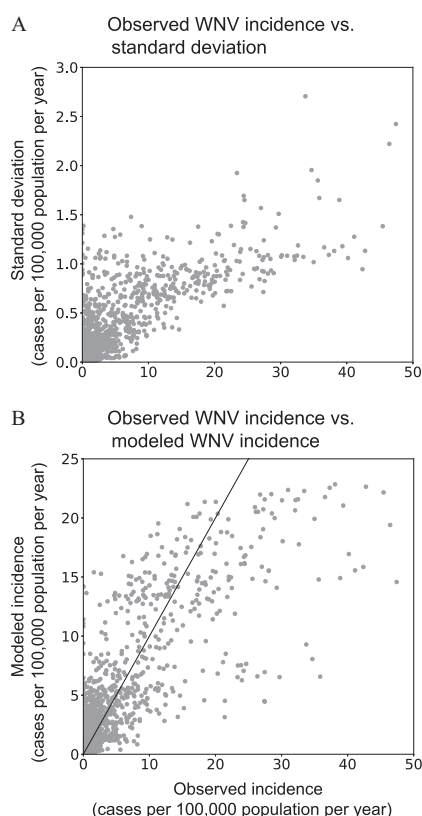


Figure 6. RF model performance plots showing county-level values for (A) observed WNV incidence vs. the standard deviation among the 500 model iterations and (B) observed WNV incidence vs. the modeled WNV incidence (2005–2019). Each dot represents one of the 3,108 counties in the contiguous United States. A 1-to-1 line was added as a guide to (B). See Table S3 for corresponding numerical data. Note: RF, random forest; WNV, West Nile virus.

our knowledge to use climate conditions to constrain the mean geospatial pattern and levels of WNV incidence across the contiguous United States as a whole, in comparison with previous studies that focused on a subregion or the interannual relationships between climate and WNV at subregional levels.^{26–29} Our

Table 2. Variable importance measured by the percentage increase in MSE if that variable were to be excluded from the analysis and total decrease in node impurity measured by the residual sum of squares. We show results for the climate variables in winter (DJF), spring (MAM), summer (JJA), and fall (SON), as well as the three eigenvectors included to account for spatial autocorrelation (E1, E2, E3).

Explanatory variable	Percentage increase in MSE	Decrease in node impurity
DJF precipitation	12.9	14,956
MAM precipitation	2.0	2,065
JJA precipitation	1.8	2,087
SON precipitation	9.1	7,344
DJF temperature	4.2	4,340
MAM temperature	0.7	716
JJA temperature	1.7	1,851
SON temperature	0.9	963
E1	3.9	4,661
E2	6.2	5,200
E3	4.3	4,843

Note: Surface air temperature and precipitation data are from the Precipitation elevation Regressions on Independent Slopes Model (PRISM). DJF, December, January, February; JJA, June, July, August; MAM, March, April, May; MSE, mean-square error; SON, September, October, November.

model suggested that although most all seasonal climate drivers may contribute information on the spatial pattern of WNV incidence in the United States, precipitation levels in winter and fall and temperatures in winter may be the most important. We found evidence for this importance in the variables that contributed the most to the RF model. From the summary tree, we found dry and cold winters constrained the V-shaped pattern of higher WNV incidence throughout the Great Plains. Within that region, areas with cold winters and warm summer temperatures were assigned the highest levels of WNV incidence.

Our model produced a more smoothed map of WNV incidence in comparison with observations. This smoothed estimate may be a better proxy of WNV risk, because nonclimate factors may also contribute to the county-level heterogeneity in incidence. For example, the availability of county-level health care services and variability in the accuracy of diagnoses and reporting could affect reported incidence rates.

Our model may provide a means to predict the future spatial extent of WNV risk in response to climate change. However, sub-regional patterns of WNV incidence are likely structured by a series of environmental variables, including vegetation type,^{47–50} land use,^{51,52} and hydrology.⁵³ Dry agricultural areas supported by irrigation may be a key component to bringing birds, mosquitoes, and humans into close proximity, increasing the speed and likelihood of WNV transmission. Our model was limited to only incorporating seasonal temperature and precipitation, though an important next step in future work is to incorporate additional information on land cover, land use, and water resources and irrigation.

The Spatial Pattern of WNV Incidence in Relation to Mosquitoes

Climate conditions may be influencing the spatial pattern of WNV incidence by controlling the habitat range of important mosquito vector species. Although many species can carry WNV, some species are more easily infected, better at amplifying the virus, and more successful at transmitting the disease to a host than others.^{9,54–57} As of 2019, WNV had been found throughout the United States in 65 different mosquito species.⁵⁸ *Cx. tarsalis* is likely one of the most important mosquito species for WNV transmission within the northern Great Plains, because it is an effective vector and primarily breeds in rural habitats.^{9,29,59} It is possible for this species to acquire WNV through vertical transmission (parent to offspring) and carry the virus through diapause into the following year.⁶⁰

A spatial analysis of the 2003 WNV levels in the northern Great Plains found that climatically averaged temperature and precipitation best described the spatial pattern of disease.²⁹ This study by Wimberly et al. hypothesized these climate conditions may be capturing the geographical range of *Cx. tarsalis*.²⁹ Wimberly et al. found average monthly precipitation between May and July at ~67 mm/month was optimal for WNV amplification and transmission in the northern Great Plains.²⁹ In our study across the contiguous United States, we found a similar optimal result, in which counties with higher levels of WNV incidence also occur in areas that have ~50–80 mm/month of rainfall in the spring and summer months (Figure 2B,C). East of the northern Great Plains, precipitation increases, and *Aedes vexans* and *Cx. pipiens* may assume the role of the most important vector.^{29,61} This possibly supports the success of *Cx. tarsalis* as a vector for WNV under drier conditions. Wimberly et al. also found a positive linear relationship between WNV incidence and summer temperatures when analyzing mean temperatures between 17.1°C and 28.8°C²⁹; this finding may describe the effects of temperature on the amplification of mosquitoes as WNV vectors by potentially influencing the mosquito

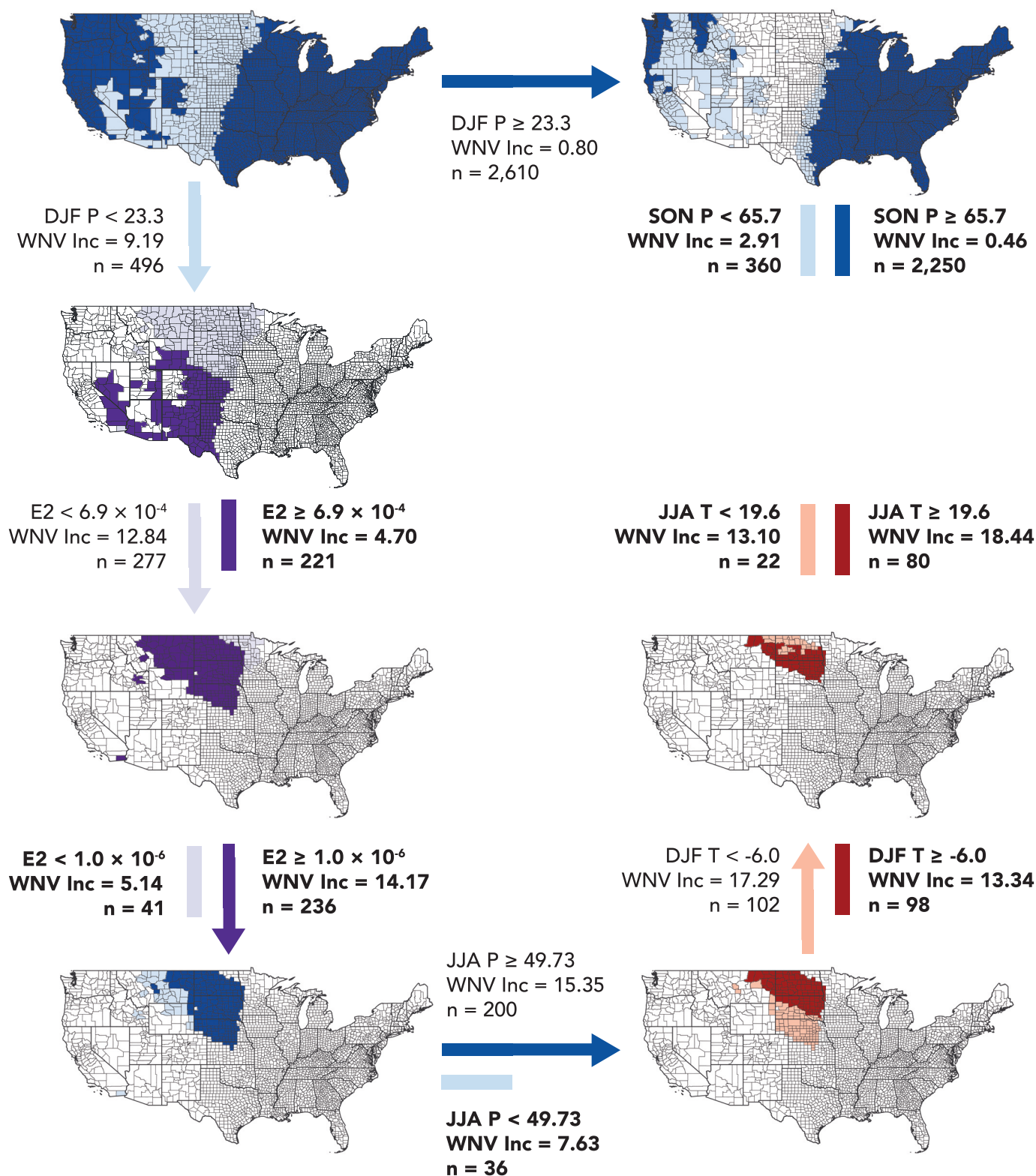


Figure 7. The summary regression tree showing county-level conditional climate splits, created from the mean RF predictions. Precipitation (P) has units of millimeters/month and air temperature (T) has units of °C. Eigenvectors are denoted by E, where 1, 2, or 3 is the corresponding column number in the eigenvector matrix. The seasons are partitioned into winter (DJF), spring (MAM), summer (JJA), and fall (SON). WNV incidence (WNV Inc) is in units of cases per 100,000 population per year. The number of counties that fall into each division is denoted by n . Terminal nodes are in bold font. Note: Maps were created in Python (version 3.7.6), using package “Basemap” (version 1.2.0) and “matplotlib” (version 3.1.3). We edited our figures for layout in Adobe Illustrator 2022. DJF, December, January, February; JJA, June, July, August; MAM, March, April, May; RF, random forest; SON, September, October, November; WNV, West Nile virus.

incubation rate.⁶² In contrast, our analysis across the contiguous United States examined a broader range of summer temperatures between 10.5°C and 33.2°C and found there was an optimal summer temperature for increased WNV incidence around 21°C (Figure 3C). Similar to the findings from Wimberly et al.,²⁹ summer temperatures above or below 19.5°C were identified as a conditional split in our summary tree, where higher summer temperatures were assigned higher values of WNV incidence.

The important seasonal climate drivers of the spatial structure of WNV incidence are likely describing the ecological niche of key mosquito species driving the transmission of WNV. Recently updated habitat suitability maps for *Culex* mosquitoes highlight the Great Plains of the United States as highly suitable for both *Cx. salinarius* and *Cx. tarsalis*.¹¹ The suitability maps for these two species were largely structured by the cultivated and managed vegetation land cover variable, which further suggests that agricultural areas may be a key ecosystem for WNV transmission. The absence or reduced population sizes of these key mosquito species in other regions of the United States due to unsuitable habitats or competition among other mosquito species may explain the lower levels of WNV incidence, especially in the eastern United States, where there is greater human population density and the presence of important bird species. Future work examining the connection between the geographical distributions (and population density) of important mosquito species and WNV incidence in the United States could help evaluate whether mosquitoes are indeed the connection between climate conditions and WNV incidence.

The Spatial Pattern of WNV Incidence in Relation to Birds

Birds play an important role in the transmission of WNV by acting as viral host reservoirs. For this reason, the presence of birds is known to modulate the rate of WNV transmission.^{63–65} One of the most important bird species for WNV transmission, considered the key host reservoir, are robins.⁶⁵ Even in ecosystems where robins are not abundant, they can account for most mosquito feedings.⁶⁵ Unlike other species, robin population levels are not affected by WNV,⁶⁶ but removing robins from a community does decrease WNV transmission.⁶⁵ Robins are ubiquitously present throughout the contiguous United States and are present year-round in the Great Plains.⁶⁷ There is a relatively high abundance of robins in the Great Plains in comparison with other parts of the United States during the nonbreeding season (DJF), prebreeding migration (FMAM), and postbreeding migration (SOND),⁶⁸ which could help WNV persistence in this region. The large number of robins that migrate through this area may play a role in the WNV transmission cycle because there are more potential viral hosts. However, robins do not seem to be the key determinant of the spatial pattern of elevated WNV incidence because they are also present outside of the areas of highest WNV incidence.

Crows are another bird of particular interest in the WNV transmission cycle. Although crow population levels have suffered from WNV,^{69,70} crows may be relatively unimportant for the amplification of WNV⁶⁵ because they make up a very small percent of mosquito feedings.²¹ However, crows are still a highly competent host for WNV,⁷¹ and new findings suggest they may play an important role in overwintering the virus.²⁵ Crows live in large winter roosts and may allow WNV to persist throughout winter by continually spreading the disease from bird to bird.^{24,25} Crows are also present year-round in the Great Plains,⁶⁸ so the presence of WNV in the crow populations may accelerate bird-to-mosquito transmission in spring when mosquitoes emerge.²⁵ However, crows also do not seem to be the key determinant of the spatial pattern of WNV because they are found throughout

much of the eastern United States, too, where there are lower levels of WNV incidence.

The Spatial Pattern of WNV Incidence in Relation to Humans

To study the relative levels of WNV in the United States, we chose to create our model based on disease incidence rather than case counts. Because the highest levels of WNV incidence occur in generally less-populated areas throughout the northern Great Plains but not across all less-populated regions in the western United States, the abundance of humans does not appear to be an important factor in structuring the spatial extent of WNV.

Because the ratio of mosquitoes to humans may be important for WNV transmission, previous studies have incorporated population density as a predictor of incidence. However, these studies found mixed results where both suburbs and rural areas were associated with higher and lower levels of incidence.^{49,50,52,72} Higher levels of human immunity in a population can reduce the number of WNV cases,⁷³ which may be an important factor when predicting the interannual levels of WNV cases. However, human immunity likely does not structure the long-term spatial extent of the disease. Estimated population seroprevalence in the United States is low (<14%), so most of the population is still susceptible to WNV.^{6,73,74}

Socioeconomic status and population demographics may also influence the structure of WNV incidence.^{4,47,75,76} At the county level, differences in health care may cause some of the heterogeneity in case counts, especially if people need to travel to population centers to seek medical services. In addition, we hypothesize that differences in efficacy of county-level case reporting contribute to some of the relatively sharp differences in incidence in the northern Great Plains (Figure 2C). If so, this measurement uncertainty is likely one important factor that limits the model performance for both training and out-of-sample metrics (Table 2).

Human behavior and the amount of time spent outdoors are also affected by weather and climate conditions.⁷⁷ People may choose to remain indoors during the hot and humid summer months, causing less human exposure to infected mosquitoes.¹⁸ However, cooler temperatures in the evenings and nights may draw people outside, which coincides with the timing of when *Culex* mosquitoes preferentially blood-feed.^{78,79}

Agricultural Areas Inadvertently Supporting Higher Levels of WNV Incidence

The climate conditions within the Great Plains, paired with favorable soils, make it a primary area for agriculture.⁸⁰ We suggest irrigated agricultural areas might be a key element that concentrates mosquitoes, birds, and humans around limited water sources in dry regions and inadvertently amplifies the transmission and higher incidence levels of WNV. People in the Great Plains are more likely to be outdoors for their occupations. The nectar and sugars from crops may provide food to male and nonbreeding female mosquitoes, whereas irrigated croplands may provide water resources necessary for mosquito breeding habitat.

Early estimates from 2004–2006 examining the number of human and veterinary WNV cases across the United States and levels of irrigation found a statistically significant positive association between the two.⁸¹ It has also been noted that county-level WNV incidence in the northern Great Plains is positively correlated with the amount of irrigated cropland.²⁹ Irrigated croplands were identified as an important risk factor for WNV in areas of higher incidence levels in Colorado⁸² and throughout regions with WNV outbreaks in Europe.⁸³

Climate Controls on the Spatial Distribution of WNV vs. Interannual Variability

Like our findings regarding the spatial structure of WNV incidence in the United States, previous work examining the interannual variability in WNV cases also found temperature and precipitation to be important drivers of WNV. However, the relationships between climate variables and WNV incidence are fundamentally different between our spatial and their temporal analyses. From year to year, warmer winters and springs led to an increase in WNV cases throughout the United States.^{84–86} These warmer temperatures may allow more mosquitoes to overwinter and begin an earlier breeding season in spring. The large number of WNV cases in 2012 in the southern Great Plains followed an anomalously warm spring but mild summer, possibly allowing the mosquito breeding season to both begin early and sustain itself throughout summer by not reaching the physiological temperature limit of the mosquitoes.⁸⁴ Although our model suggested that regions with colder winter temperatures support higher mean levels of WNV incidence, areas that have anomalously warmer winters within this climate space may lead to years with above-average levels of WNV cases.

The role of precipitation on interannual WNV levels varies spatially.^{53,61,86–88} Generally, lower levels of precipitation cause places that are already dry (e.g., western United States) to have lower levels of WNV incidence and cause places that are typically wet (e.g., eastern United States) to have higher WNV incidence.^{84,87,89} Drought conditions have also been linked to increases in WNV.^{53,90} Our RF model indicated areas with dry and cold winters and mild-to-warm summers supported higher levels of WNV incidence.

We hypothesize climate plays different roles in structuring the mean spatial pattern of WNV incidence in the United States versus the interannual variability in the number of disease cases. Spatially, climate conditions likely structure the range of the important vector species for WNV. Mosquitoes able to flourish in cold and dry winters are also efficient transmitters of WNV, increasing regional levels of WNV incidence. However, interannual variability in climate may influence several facets of the mosquito lifecycle and local resources available to the existing mosquito populations. These differences may explain why some climate conditions have opposing relationships to levels of WNV incidence spatially versus temporally.

Model Limitations and Future Research

Although our RF model could explain up to 61% of the variance in the spatial structure of WNV incidence using only seasonal temperature and precipitation, additional predictive variables may improve the model. Irrigated agricultural areas may be a key factor in bringing humans, birds, and mosquitoes around limited water supplies. Though climate conditions generally structure where areas of agriculture are, future models should consider incorporating agricultural variables and other information about land use. Our model did not incorporate any information regarding the dominant regional mosquito vector species, which could enhance model performance if different vector species have orthogonal responses to climate drivers. Our model also did not include any specific measure of bird abundance, species diversity, or the presence of a species. In future work, our RF model may be useful for predicting the response of WNV incidence to climate change, because shifts in precipitation and increasing temperatures may cause mosquitoes to change their geographical range.

Acknowledgments

The authors thank J. Lehman and the Arboviral Diseases Branch team at the National Center of Emerging Zoonotic Infectious Diseases of the U.S. CDC for providing us with WNV case data.

M.E.G. received support from a Department of Defense National Defense Science and Engineering Graduate Fellowship (32 CFR 168a) and Laboratory Directed Research and Development (LDRD) Director's Postdoc Fellowship through Los Alamos National Laboratory (LANL). M.E.G. also thanks the LANL Center for Nonlinear Studies. M.E.G., C.A.M., and C.X. were supported by the LDRD program of LANL under project number 20210062DR. J.T.R. received support from the U.S. Department of Energy Office of Science Division Biological and Environmental Research RUBISCO science focus area project. S.R.C. was supported by a National Science Foundation (NSF) Graduate Research Fellowship (DGE-1839285). K.K.T. was supported by grants from NSF (DEB 1912525) and the Department of Energy Office of Biological and Environmental Research (DE-SC0020382). C.S.Z. received support from the Borrego Valley Endowment Fund (BVEF-204987). C.A.M. was also supported by the LDRD program of LANL under project numbers 20190581ECR and 20210062DR. J.T.R. and C.X. also acknowledge support from a University of California Laboratory Fees Research Program grant investigating California drought, wildfire, and tree mortality. This work is approved for distribution under LA-UR-22-20129. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of their affiliations, including LANL. LANL, an affirmative action/equal opportunity employer, is managed by Triad National Security, LLC, for the National Nuclear Security Administration of the U.S. Department of Energy under contract 89233218CNA000001.

References

1. U.S. CDC (U.S. Centers for Disease Control and Prevention). 2022. West Nile Virus. <https://www.cdc.gov/westnile/index.html> [accessed 15 January 2022].
2. U.S. CDC. 2018. West Nile Virus, Symptoms, Diagnosis, & Treatment. <https://www.cdc.gov/westnile/symptoms/index.html> [accessed 15 January 2022].
3. U.S. CDC. 2021. Final Cumulative Maps & Data for 1999–2021. <https://www.cdc.gov/westnile/statsmaps/cumMapsData.html> [accessed 15 January 2022].
4. Hayes EB, Komar N, Nasci RS, Montgomery SP, O'Leary DR, Campbell GL. 2005. Epidemiology and transmission dynamics of West Nile virus disease. *Emerg Infect Dis* 11(8):1167–1173, PMID: 16102302, <https://doi.org/10.3201/eid1108.050289a>.
5. Chuang TW, Wimberly MC. 2012. Remote sensing of climatic anomalies and West Nile virus incidence in the northern Great Plains of the United States. *PLoS One* 7(10):e46882, PMID: 23071656, <https://doi.org/10.1371/journal.pone.0046882>.
6. Petersen LR, Carson PJ, Biggerstaff BJ, Custer B, Borchardt SM, Busch MP. 2013. Estimated cumulative incidence of West Nile virus infection in US adults, 1999–2010. *Epidemiol Infect* 141(3):591–595, PMID: 22640592, <https://doi.org/10.1017/S0950268812001070>.
7. Hamer GL, Kitron UD, Goldberg TL, Brawn JD, Loss SR, Ruiz MO, et al. 2009. Host selection by *Culex pipiens* mosquitoes and West Nile virus amplification. *Am J Trop Med Hyg* 80(2):268–278, PMID: 19190226, <https://doi.org/10.4269/ajtmh.2009.80.268>.
8. Kilpatrick AM, Kramer LD, Campbell SR, Alleyne EO, Dobson AP, Daszak P. 2005. West Nile virus risk assessment and the bridge vector paradigm. *Emerg Infect Dis* 11(3):425–429, <https://doi.org/10.3201/eid1103.040364>.
9. Turell MJ, Dohm DJ, Sardelis MR, Oguinn ML, Andreadis TG, Blow JA. 2005. An update on the potential of North American mosquitoes (Diptera: Culicidae) to transmit West Nile virus. *J Med Entomol* 42(1):57–62, PMID: 15691009, <https://doi.org/10.1093/jmedent/42.1.57>.
10. Darsie RF Jr, Ward RA. 2005. *Identification and Geographical Distribution of the Mosquitoes of North America, North of Mexico*. Mount Laurel, NJ: American Mosquito Control Association.
11. Gorris ME, Bartlow AW, Temple SD, Romero-Alvarez D, Shutt DP, Fair JM, et al. 2021. Updated distribution maps of predominant *Culex* mosquitoes across the Americas. *Parasit Vectors* 14(1):547, PMID: 34688314, <https://doi.org/10.1186/s13071-021-05051-3>.
12. Reisen WK. 2013. Ecology of West Nile virus in North America. *Viruses* 5(9):2079–2105, PMID: 24008376, <https://doi.org/10.3390/v5092079>.
13. Lungstrom L. 1954. Biological studies on *Culex tarsalis* (Diptera Culicidae) in Kansas. *Trans Kans Acad Sci* 1903 57(1):86–96, <https://doi.org/10.2307/3625649>.
14. Brown D, Armijos MV, Scott TW, Nielsen CF, Boyce WM, Kelley K, et al. 2008. Risk factors associated with human infection during the 2006 West Nile virus

- outbreak in Davis, a residential community in northern California. *Am J Trop Med Hyg* 78(1):53–62, PMID: 18187785, <https://doi.org/10.4269/ajtmh.2008.78.53>.
15. Rueda LM, Patel KJ, Axtell RC, Stinner RE. 1990. Temperature-dependent development and survival rates of *Culex quinquefasciatus* and *Aedes aegypti* (Diptera: Culicidae). *J Med Entomol* 27(5):892–898, PMID: 2231624, <https://doi.org/10.1093/jmedent/27.5.892>.
 16. Ciota AT, Mataracchiero AC, Kilpatrick AM, Kramer LD. 2014. The effect of temperature on life history traits of culex mosquitoes. *J Med Entomol* 51(1):55–62, PMID: 24605453, <https://doi.org/10.1603/me13003>.
 17. Dohm DJ, O'Guinn ML, Turell MJ. 2002. Effect of environmental temperature on the ability of *Culex pipiens* (Diptera: Culicidae) to transmit West Nile virus. *J Med Entomol* 39(1):221–225, PMID: 11931261, <https://doi.org/10.1603/0022-2585-39.1.221>.
 18. Hedquist BC, Brazel AJ. 2014. Seasonal variability of temperatures and outdoor human comfort in Phoenix, Arizona, U.S.A. *Build Environ* 72:377–388, <https://doi.org/10.1016/j.buildenv.2013.11.018>.
 19. Dunn PO, Winkler DW. 1999. Climate change has affected the breeding date of tree swallows throughout North America. *Proc Biol Sci* 266(1437):2487–2490, PMID: 10693819, <https://doi.org/10.1098/rspb.1999.0950>.
 20. Marra PP, Francis CM, Mulvihill RS, Moore FR. 2005. The influence of climate on the timing and rate of spring bird migration. *Oecologia* 142(2):307–315, PMID: 15480801, <https://doi.org/10.1007/s00442-004-1725-x>.
 21. Molaei G, Andreadis TG, Armstrong PM, Anderson JF, Vossbrinck CR. 2006. Host feeding patterns of *Culex* mosquitoes and West Nile virus transmission, northeastern United States. *Emerg Infect Dis* 12(3):468–474, PMID: 16704786, <https://doi.org/10.3201/eid1203.051004>.
 22. Kilpatrick AM. 2011. Globalization, land use, and the invasion of West Nile virus. *Science* 334(6054):323–327, PMID: 22021850, <https://doi.org/10.1126/science.1201010>.
 23. Diuk-Wasser MA, Molaei G, Simpson JE, Folsom-O'Keefe CM, Armstrong PM, Andreadis TG. 2010. Avian communal roosts as amplification foci for West Nile virus in urban areas in northeastern United States. *Am J Trop Med Hyg* 82(2):337–343, PMID: 20134014, <https://doi.org/10.4269/ajtmh.2010.09-0506>.
 24. Hinton MG, Reisen WK, Wheeler SS, Townsend AK. 2015. West Nile virus activity in a winter roost of American crows (*Corvus brachyrhynchos*): is bird-to-bird transmission important in persistence and amplification? *J Med Entomol* 52(4):683–692, PMID: 26335475, <https://doi.org/10.1093/jme/tjv040>.
 25. Montecino-Latorre D, Barker CM. 2018. Overwintering of West Nile virus in a bird community with a communal crow roost. *Sci Rep* 8(1):6088, PMID: 29666401, <https://doi.org/10.1038/s41598-018-24133-4>.
 26. Beeman SP, Morrison AM, Unnasch TR, Unnasch RS. 2021. Ensemble ecological niche modeling of West Nile virus probability in Florida. *PLoS One* 16(10): e0256868, PMID: 34624026, <https://doi.org/10.1371/journal.pone.0256868>.
 27. Hernandez E, Torres R, Joyce AL. 2019. Environmental and sociological factors associated with the incidence of West Nile virus cases in the Northern San Joaquin Valley of California, 2011–2015. *Vector Borne Zoonotic Dis* 19(11):851–858, PMID: 31211639, <https://doi.org/10.1089/vbz.2019.2437>.
 28. Hess A, Davis JK, Wimberly MC. 2018. Identifying environmental risk factors and mapping the distribution of West Nile virus in an endemic region of North America. *GeoHealth* 2(12):395–409, PMID: 32159009, <https://doi.org/10.1029/2018GH000161>.
 29. Wimberly MC, Hildreth MB, Boyte SP, Lindquist E, Kightlinger L. 2008. Ecological niche of the 2003 West Nile epidemic in the northern Great Plains of the United States. *PLoS One* 3(12):e3744, PMID: 19057643, <https://doi.org/10.1371/journal.pone.0003744>.
 30. U.S. Census Bureau. 2021. County Intercensal Datasets: 2000–2010. <https://www.census.gov/data/datasets/time-series/demo/popest/intercensal-2000-2010-counties.html> [accessed 1 July 2021].
 31. U.S. Census Bureau. 2021. County Population Totals: 2010–2019. <https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-total.html> [accessed 1 July 2021].
 32. Daly C, Neilson RP, Phillips DL. 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J Appl Meteor* 33(2):140–158, [https://doi.org/10.1175/1520-0450\(1994\)033<0140:ASTMFV>2.0.CO;2](https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFV>2.0.CO;2).
 33. Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, et al. 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int J Climatol* 28(15):2031–2064, <https://doi.org/10.1002/joc.1688>.
 34. U.S. Census Bureau. 2021. TIGER/Line Shapefiles 2018. <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2018.html> [accessed 1 July 2021].
 35. Breiman L. 2001. Random Forests. *Mach Learn* 45:5–32, <https://doi.org/10.1023/A:1010933404324>.
 36. Coffield SR, Hemes KS, Koven CD, Goulden ML, Randerson JT. 2021. Climate-driven limits to future carbon storage in California's wildland ecosystems. *AGU Adv* 2(3):e2021AV000384, <https://doi.org/10.1029/2021AV000384>.
 37. Keyel AC, Elison Timm O, Backenson PB, Prussing C, Quinones S, McDonough KA, et al. 2019. Seasonal temperatures and hydrological conditions improve the prediction of West Nile virus infection rates in *Culex* mosquitoes and human case counts in New York and Connecticut. *PLoS One* 14(6):e0217854, PMID: 31158250, <https://doi.org/10.1371/journal.pone.0217854>.
 38. Lee S, Kim JC, Jung HS, Lee MJ, Lee S. 2017. Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. *Geomat Nat Hazards Risk* 8(2):1185–1203, <https://doi.org/10.1080/19475705.2017.1308971>.
 39. Breiman L. 2001. Random forests. *Mach Learn* 45(1):5–32, <https://doi.org/10.1023/A:1010933404324>.
 40. Faivre NR, Jin Y, Goulden ML, Randerson JT. 2016. Spatial patterns and controls on burned area for two contrasting fire regimes in Southern California. *Ecosphere* 7(5):e01210, <https://doi.org/10.1002/ecs2.1210>.
 41. Liaw A, Wiener M. 2022. Classification and regression by randomForest. *R News* 2(3):18–22.
 42. Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J, et al. 2007. Random forests for classification in ecology. *Ecology* 88(11):2783–2792, PMID: 18051647, <https://doi.org/10.1890/07-0539.1>.
 43. Freeman EA, Moisen GG, Coulston JW, Wilson BT. 2016. Random forests and stochastic gradient boosting for predicting tree canopy cover: comparing tuning processes and model performance. *Can J for Res* 46(3):323–339, <https://doi.org/10.1139/cjfr-2014-0562>.
 44. Griffith D, Chun Y. 2014. Spatial Autocorrelation and Spatial Filtering. In: *Handbook of Regional Science*. Fischer M, Nijkamp P, eds. Berlin, Germany: Springer, 1477–1507. https://doi.org/10.1007/978-3-642-23430-9_72 [accessed 8 September 2022].
 45. Liu X, Kounadi O, Zurita-Milla R. 2022. Incorporating spatial autocorrelation in machine learning models using spatial lag and eigenvector spatial filtering features. *ISPRS Int J Geo-Inf* 11(4):242, <https://doi.org/10.3390/ijgi11040242>.
 46. Therneau T, Atkinson B, Ripley B. 2022. Package “rpart.” <https://cran.r-project.org/web/packages/rpart/rpart.pdf> [accessed 1 July 2021].
 47. Brownstein JS, Rosen H, Purdy D, Miller JR, Merlino M, Mostashari F, et al. 2002. Spatial analysis of West Nile virus: rapid risk assessment of an introduced vector-borne zoonosis. *Vector Borne Zoonotic Dis* 2(3):157–164, PMID: 12737545, <https://doi.org/10.1089/15303660260613729>.
 48. Ezenwa VO, Milheim LE, Coffey MF, Godsey MS, King RJ, Gupta SC. 2007. Land cover variation and West Nile virus prevalence: patterns, processes, and implications for disease control. *Vector Borne Zoonotic Dis* 7(2):173–180, PMID: 17627435, <https://doi.org/10.1089/vbz.2006.0584>.
 49. Gibbs SEJ, Wimberly MC, Madden M, Masour J, Yabsley MJ, Stallknecht DE. 2006. Factors affecting the geographic distribution of West Nile virus in Georgia, USA: 2002–2004. *Vector Borne Zoonotic Dis* 6(1):73–82, PMID: 16584329, <https://doi.org/10.1089/vbz.2006.6.73>.
 50. Ruiz MO, Tedesco C, McTighe TJ, Austin C, Kitron U. 2004. Environmental and social determinants of human risk during a West Nile virus outbreak in the greater Chicago area, 2002. *Int J Health Geogr* 3(1):8, PMID: 15099399, <https://doi.org/10.1186/1476-072X-3-8>.
 51. Bowden SE, Magori K, Drake JM. 2011. Regional differences in the association between land cover and West Nile virus disease incidence in humans in the United States. *Am J Trop Med Hyg* 84(2):234–238, PMID: 21292890, <https://doi.org/10.4269/ajtmh.2011.10-0134>.
 52. Ruiz MO, Walker ED, Foster ES, Haramis LD, Kitron UD. 2007. Association of West Nile virus illness and urban landscapes in Chicago and Detroit. *Int J Health Geogr* 6(1):10, PMID: 17352825, <https://doi.org/10.1186/1476-072X-6-10>.
 53. Shaman J, Day JF, Komar N. 2010. Hydrologic conditions describe West Nile virus risk in Colorado. *Int J Environ Res Public Health* 7(2):494–508, PMID: 20616987, <https://doi.org/10.3390/ijerph7020494>.
 54. Sardelis MR, Turell MJ, Dohm DJ, O'Guinn ML. 2001. Vector competence of selected North American *Culex* and *Coquillettia* mosquitoes for West Nile virus. *Emerg Infect Dis* 7(6):1018–1022, PMID: 11747732, <https://doi.org/10.3201/eid0706.010617>.
 55. Turell MJ, O'Guinn ML, Dohm DJ, Jones JW. 2001. Vector competence of North American mosquitoes (Diptera: Culicidae) for West Nile virus. *J Med Entomol* 38(2):130–134, PMID: 11296813, <https://doi.org/10.1603/0022-2585-38.2.130>.
 56. Turell MJ, O'Guinn M, Oliver J. 2000. Potential for New York mosquitoes to transmit West Nile virus. *Am J Trop Med Hyg* 62(3):413–414, PMID: 11037788, <https://doi.org/10.4269/ajtmh.2000.62.413>.
 57. Goddard LB, Roth AE, Reisen WK, Scott TW. 2002. Vector competence of California mosquitoes for West Nile virus. *Emerg Infect Dis* 8(12):1385–1391, PMID: 12498652, <https://doi.org/10.3201/eid0812.020536>.
 58. Rochlin I, Faraji A, Healy K, Andreadis TG. 2019. West Nile virus mosquito vectors in North America. *J Med Entomol* 56(6):1475–1490, PMID: 31549725, <https://doi.org/10.1093/jme/tjz146>.

59. Chuang TW, Hildreth MB, Vanroekel DL, Wimberly MC. 2011. Weather and land cover influences on mosquito populations in Sioux Falls, South Dakota. *J Med Entomol* 48(3):669–679, PMID: [21661329](#), <https://doi.org/10.1603/me10246>.
60. Nelms BM, Fechter-Leggett E, Carroll BD, Macedo P, Klueh S, Reisen WK. 2013. Experimental and natural vertical transmission of West Nile virus by California *Culex* (Diptera: Culicidae) mosquitoes. *J Med Entomol* 50(2):371–378, PMID: [23540126](#), <https://doi.org/10.1603/me12264>.
61. DeGroot JP, Sugumaran R, Brend SM, Tucker BJ, Bartholomay LC. 2008. Landscape, demographic, entomological, and climatic associations with human disease incidence of West Nile virus in the state of Iowa, USA. *Int J Health Geogr* 7(1):19, PMID: [18452604](#), <https://doi.org/10.1186/1476-072X-7-19>.
62. Reisen WK, Fang Y, Martinez VM. 2006. Effects of temperature on the transmission of West Nile virus by *Culex tarsalis* (Diptera: Culicidae). *J Med Entomol* 43(2):309–317, PMID: [16619616](#), [https://doi.org/10.1603/0022-2585\(2006\)043\[0309:EOTOTT\]2.0.CO;2](https://doi.org/10.1603/0022-2585(2006)043[0309:EOTOTT]2.0.CO;2).
63. Bergsman LD, Hyman JM, Manore CA. 2016. A mathematical model for the spread of West Nile virus in migratory and resident birds. <https://www.aims-science.org/article/id/18c94254-45ae-4773-9c81-13636583f480> [accessed 6 September 2022].
64. Ezenwa VO, Godsey MS, King RJ, Guptill SC. 2006. Avian diversity and West Nile virus: testing associations between biodiversity and infectious disease risk. *Proc Biol Sci* 273(1582):109–117, PMID: [16519242](#), <https://doi.org/10.1098/rspb.2005.3284>.
65. Kilpatrick AM, Daszak P, Jones MJ, Marra PP, Kramer LD. 2006. Host heterogeneity dominates West Nile virus transmission. *Proc Biol Sci* 273(1599):2327–2333, PMID: [16928635](#), <https://doi.org/10.1098/rspb.2006.3575>.
66. George TL, Harrigan RJ, LaManna JA, DeSante DF, Saracco JF, Smith TB. 2015. Persistent impacts of West Nile virus on North American bird populations. *Proc Natl Acad Sci USA* 112(46):14290–14294, PMID: [26578774](#), <https://doi.org/10.1073/pnas.1507747112>.
67. Vanderhoff N, Pyle P, Patten MA, Sallabanks R, James FC. 2020. American Robin: *Turdus migratorius*. In: *Birds of the World*. Rodewald PG, ed. Ithaca, NY: Cornell Lab of Ornithology. <https://doi.org/10.2173/bow.amerob.01>.
68. Fink D, Auer T, Johnston A. eBird Status and Trends: Data Access. Ithaca, New York: Cornell Lab of Ornithology. <https://doi.org/https://doi.org/10.2173/ebirdst.2020> [accessed 8 September 2022].
69. LaDeau SL, Calder CA, Doran PJ, Marra PP. 2011. West Nile virus impacts in American crow populations are associated with human land use and climate. *Ecol Res* 26(5):909–916, PMID: [32214652](#), <https://doi.org/10.1007/s11284-010-0725-z>.
70. Yaremych SA, Warner RE, Mankin PC, Brawn JD, Raim A, Novak R. 2004. West Nile virus and high death rate in American crows. *Emerg Infect Dis* 10(4):709–711, <https://doi.org/10.3201/eid1004.030499>.
71. McLean RG, Ubico SR, Docherty DE, Hansen WR, Sileo L, McNamara TS. 2001. West Nile virus transmission and ecology in birds. *Ann NY Acad Sci* 951(1):54–57, PMID: [11797804](#), <https://doi.org/10.1111/j.1749-6632.2001.tb02684.x>.
72. Schweitzer BK, Kramer WL, Sambol AR, Meza JL, Hinrichs SH, Iwen PC. 2006. Geographic factors contributing to a high seroprevalence of West Nile virus-specific antibodies in humans following an epidemic. *Clin Vaccine Immunol* 13(3):314–318, PMID: [16522771](#), <https://doi.org/10.1128/CVI.13.3.314-318.2006>.
73. Paull SH, Horton DE, Ashfaq M. 2017. Drought and immunity determine the intensity of West Nile virus epidemics and climate change impacts. *Proc Biol Sci* 284(1848):20162078, PMID: [28179512](#), <https://doi.org/10.1098/rspb.2016.2078>.
74. Murphy TD, Grandpre J, Novick SL, Seys SA, Harris RW, Musgrave K. 2005. West Nile virus infection among health-fair participants, Wyoming 2003: assessment of symptoms and risk factors. *Vector Borne Zoonotic Dis* 5(3):246–251, PMID: [16187893](#), <https://doi.org/10.1089/vbz.2005.5.246>.
75. Harrigan RJ, Thomassen HA, Buermann W, Cummings RF, Kahn ME, Smith TB. 2010. Economic conditions predict prevalence of West Nile virus. *PLoS One* 5(11):e15437, PMID: [21103053](#), <https://doi.org/10.1371/journal.pone.0015437>.
76. Liu A, Lee V, Galusha D, Slade MD, Diuk-Wasser M, Andreadis T, et al. 2009. Risk factors for human infection with West Nile virus in Connecticut: a multi-year analysis. *Int J Health Geogr* 8(1):67, PMID: [19943935](#), <https://doi.org/10.1186/1476-072X-8-67>.
77. Chen L, Ng E. 2012. Outdoor thermal comfort and outdoor activities: a review of research in the past decade. *Cities* 29(2):118–125, <https://doi.org/10.1016/j.cities.2011.08.006>.
78. Fritz M, Walker E, Yunker A, Dworkin I. 2014. Daily blood feeding rhythms of laboratory-reared North American *Culex pipiens*. *J Circadian Rhythms* 12(1):1, PMID: [24450879](#), <https://doi.org/10.1186/1740-3391-12-1>.
79. Yee WL, Foster WA. 1992. Diel sugar-feeding and host-seeking rhythms in mosquitoes (Diptera: Culicidae) under laboratory conditions. *J Med Entomol* 29(5):784–791, PMID: [1357175](#), <https://doi.org/10.1093/jmedent/29.5.784>.
80. U.S. Department of Agriculture National Agricultural Statistics Service. CropScape Cropland Data Layers. <https://nassgeodata.gmu.edu/CropScape/> [accessed 1 July 2021].
81. Gates MC, Boston RC. 2009. Irrigation linked to a greater incidence of human and veterinary West Nile virus cases in the United States from 2004 to 2006. *Prev Vet Med* 89(1–2):134–137, PMID: [19185941](#), <https://doi.org/10.1016/j.prevetmed.2008.12.004>.
82. Eisen L, Barker CM, Moore CG, Pape WJ, Winters AM, Cheronis N. 2010. Irrigated agriculture is an important risk factor for West Nile virus disease in the hyperendemic Larimer-Boulder-Weld area of north central Colorado. *J Med Entomol* 47(5):939–951, PMID: [20939393](#), <https://doi.org/10.1093/jmedent/47.5.939>.
83. Marcantonio M, Rizzoli A, Metz M, Rosà R, Marini G, Chadwick E, et al. 2015. Identifying the environmental conditions favouring West Nile virus outbreaks in Europe. *PLoS One* 10(3):e0121158, PMID: [25803814](#), <https://doi.org/10.1371/journal.pone.0121158>.
84. Hahn MB, Monaghan AJ, Hayden MH, Eisen RJ, Delorey MJ, Lindsey NP, et al. 2015. Meteorological conditions associated with increased incidence of West Nile virus disease in the United States, 2004–2012. *Am J Trop Med Hyg* 92(5):1013–1022, PMID: [25802435](#), <https://doi.org/10.4269/ajtmh.14-0737>.
85. Manore CA, Davis J, Christofferson RC, Wesson D, Hyman JM, Mores CN. 2014. Towards an early warning system for forecasting human West Nile virus incidence. *PLoS Curr* 6:eurrents.outbreaks.ed6f0f8a61d20ae5f32aaa5c2b8d3c23, PMID: [25914857](#), <https://doi.org/10.1371/currents.outbreaks.f0b3978230599a56830ce30cb9ce0500>.
86. Wimberly MC, Lamsal A, Giacomo P, Chuang TW. 2014. Regional variation of climatic influences on West Nile virus outbreaks in the United States. *Am J Trop Med Hyg* 91(4):677–684, PMID: [25092814](#), <https://doi.org/10.4269/ajtmh.14-0239>.
87. Landesman WJ, Allan BF, Langerhans RB, Knight TM, Chase JM. 2007. Inter-annual associations between precipitation and human incidence of West Nile virus in the United States. *Vector Borne Zoonotic Dis* 7(3):337–343, PMID: [17867908](#), <https://doi.org/10.1089/vbz.2006.0590>.
88. Wang G, Minnis RB, Belant JL, Wax CL. 2010. Dry weather induces outbreaks of human West Nile virus infections. *BMC Infect Dis* 10(1):38, PMID: [20181272](#), <https://doi.org/10.1186/1471-2334-10-38>.
89. Little E, Campbell SR, Shaman J. 2016. Development and validation of a climate-based ensemble prediction model for West Nile virus infection rates in *Culex* mosquitoes, Suffolk County, New York. *Parasit Vectors* 9(1):443, PMID: [27507279](#), <https://doi.org/10.1186/s13071-016-1720-1>.
90. Smith KH, Tyre AJ, Hamik J, Hayes MJ, Zhou Y, Dai L. 2020. Using climate to explain and predict West Nile virus risk in Nebraska. *GeoHealth* 4(9):e2020GH000244, PMID: [32885112](#), <https://doi.org/10.1029/2020GH000244>.