



## OPEN ACCESS

## EDITED BY

John Hay,  
University at Buffalo, United States

## REVIEWED BY

April Davis,  
New York State Department of Health,  
United States  
Hamish McManus,  
University of New South Wales, Australia

## \*CORRESPONDENCE

Laura Albrecht  
✉ lalbrecht@mines.edu

## SPECIALTY SECTION

This article was submitted to  
Infectious Diseases: Epidemiology and  
Prevention,  
a section of the journal  
Frontiers in Public Health

RECEIVED 16 November 2022

ACCEPTED 31 January 2023

PUBLISHED 17 February 2023

## CITATION

Albrecht L and Kaufeld KA (2023) Investigating the impact of environmental factors on West Nile virus human case prediction in Ontario, Canada. *Front. Public Health* 11:1100543. doi: 10.3389/fpubh.2023.1100543

## COPYRIGHT

© 2023 Albrecht and Kaufeld. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Investigating the impact of environmental factors on West Nile virus human case prediction in Ontario, Canada

Laura Albrecht<sup>1,2\*</sup> and Kimberly A. Kaufeld<sup>1</sup>

<sup>1</sup>Statistical Sciences Group, Los Alamos National Laboratory, Los Alamos, NM, United States,

<sup>2</sup>Department of Applied Mathematics and Statistics, Colorado School of Mines, Golden, CO, United States

West Nile virus is the most common mosquito borne disease in North America and the leading cause of viral encephalitis. West Nile virus is primarily transmitted between birds and mosquitoes while humans are incidental, dead-end hosts. Climate change may increase the risk of human infections as climatic variables have been shown to affect the mosquito life cycle, biting rate, incubation period of the disease in mosquitoes, and bird migration patterns. We develop a zero-inflated Poisson model to investigate how human West Nile virus case counts vary with respect to mosquito abundance and infection rates, bird abundance, and other environmental covariates. We use a Bayesian paradigm to fit our model to data from 2010–2019 in Ontario, Canada. Our results show mosquito infection rate, temperature, precipitation, and crow abundance are positively correlated with human cases while NDVI and robin abundance are negatively correlated with human cases. We find the inclusion of spatial random effects allows for more accurate predictions, particularly in years where cases are higher. Our model is able to accurately predict the magnitude and timing of yearly West Nile virus outbreaks and could be a valuable tool for public health officials to implement prevention strategies to mitigate these outbreaks.

## KEYWORDS

West Nile virus, zero-inflated Poisson, mosquito, human health, spatiotemporal

## 1. Introduction

West Nile virus (WNV) is the most prevalent mosquito-borne disease in North America (1). The virus is transmitted from mosquitoes to humans through biting. WNV can range from no symptoms to severe, where the illness affects the central nervous system. Only 20% of people infected with WNV will develop symptoms while <1% of people develop severe neurological symptoms (2, 3). It was first detected in the United States in 1999 and spread quickly over the next 5 years across the US and Canada and has become endemic (4, 5). Since WNV emerged in the United States, it has become the largest arboviral infection in the contiguous United States (6). The first human case of WNV was detected in Ontario, Canada in 2002 (7). Since then, a WNV surveillance program was implemented in Canada to monitor WNV prevalence in humans, mosquitoes, birds, and other mammals (7). WNV transmission occurs mainly through mosquitoes that can become infectious after biting an infected bird (4). Humans and other mammals are dead end hosts that can become infectious but are unable to transmit the virus back to uninfected mosquitoes (2, 4).

Many environmental factors are known to have an effect on WNV transmission as well as on the mosquito life cycle. In a study by García-Carrasco et al. (8), WNV outbreaks were predicted in Europe using an environmental and spatio-environmental risk model. They found that maximum temperatures, presence of rivers, and density of horses and poultry

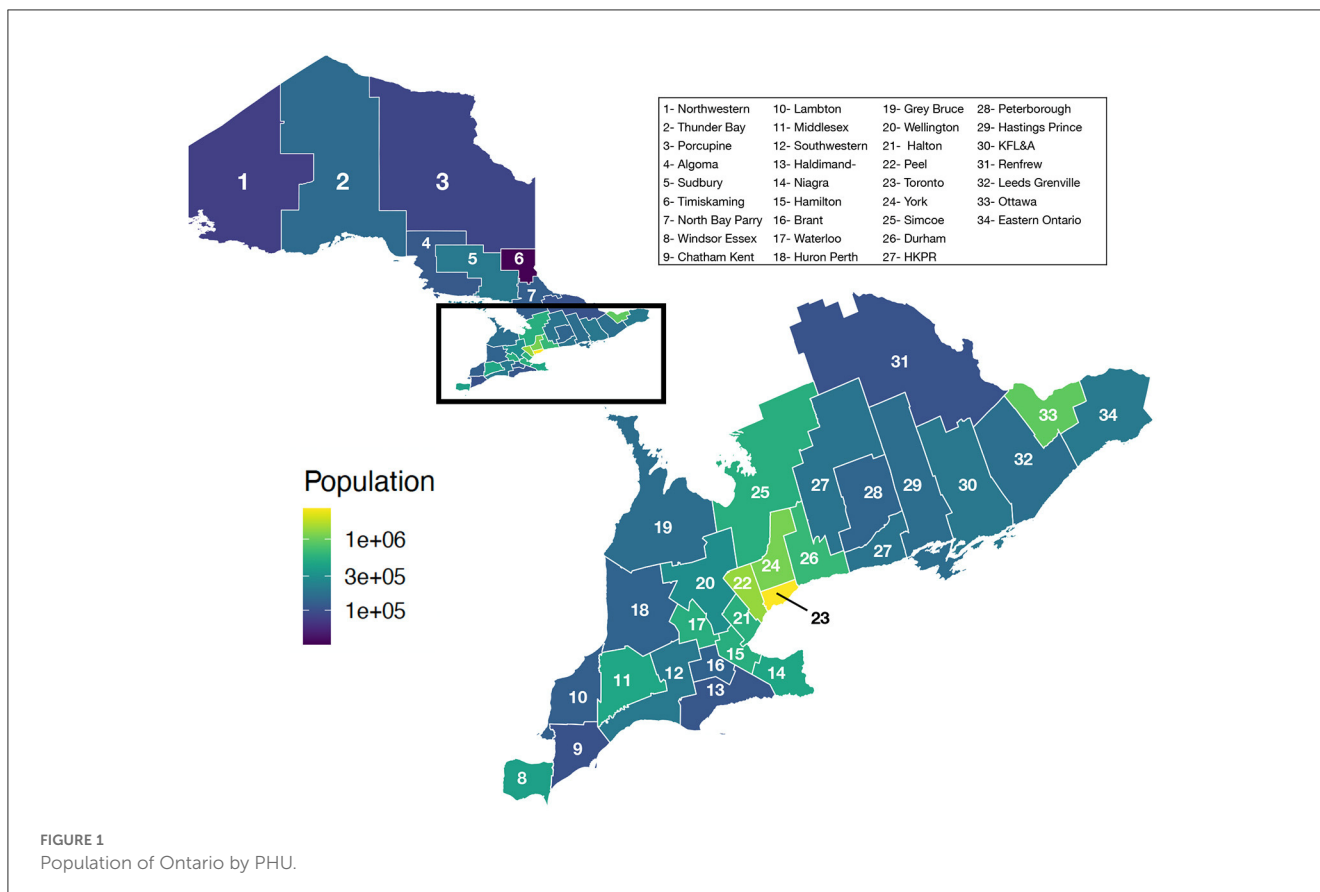
were the best environmental predictors for WNV. In terms of predicting the location of WNV outbreaks, the spatio-environmental model which incorporates latitude and longitude as a proxy for spatial structure predicted the best, suggesting that bird migration rate also plays a role in the geographical pattern of WNV. Another environmental factor that has been correlated with incidence rates for West Nile Virus is the normalized difference vegetation index (NDVI) (9, 10). NDVI provides an index for healthy vegetation and serves as a proxy for suitable conditions for mosquito development (11).

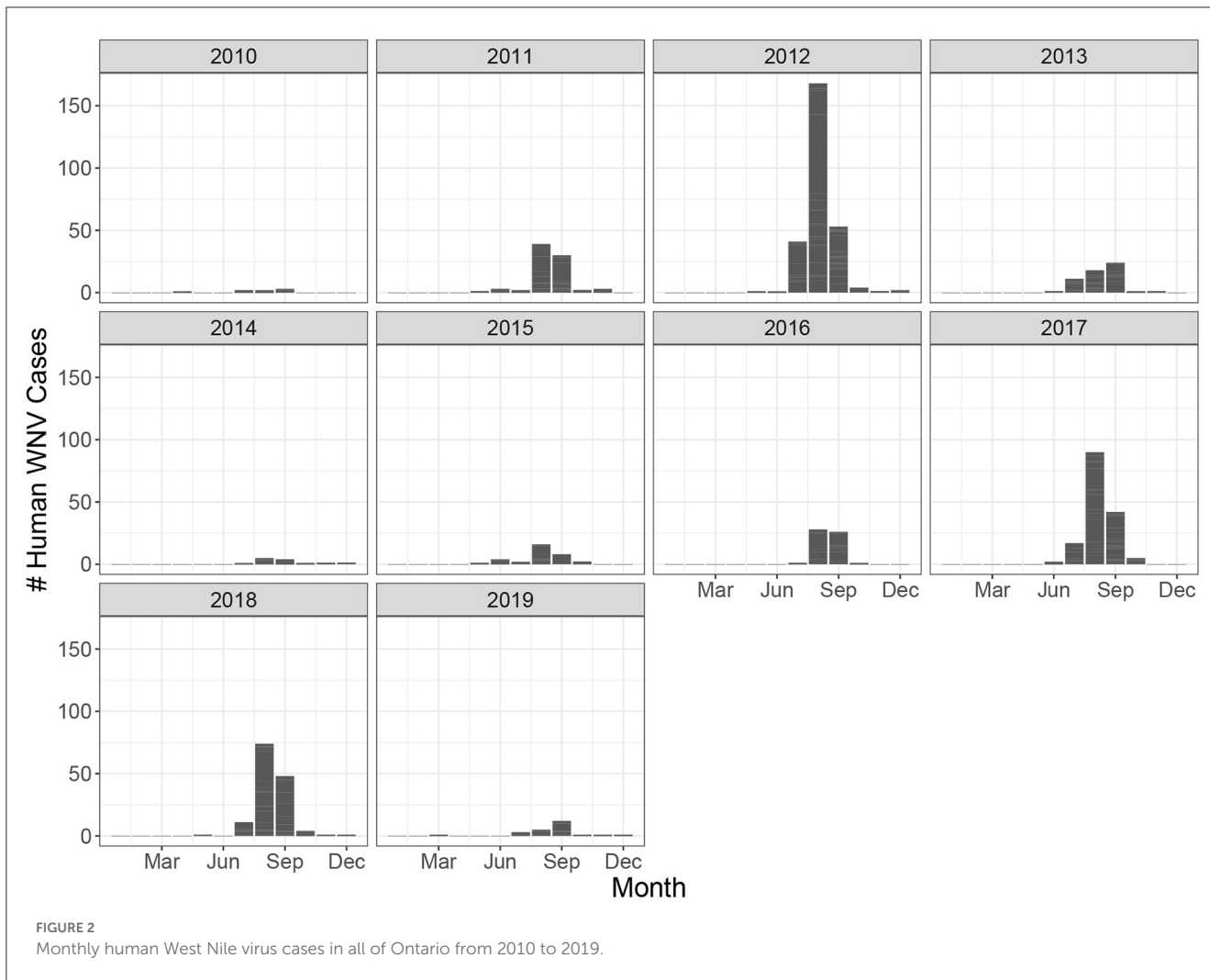
Mosquitoes carry the potential to transmit pathogens from reservoir hosts to humans. In particular, *Culex* mosquitoes are the main vectors involved in the spread of WNV in North America due to their preference for avian blood meals, especially the American Robin (4, 12–14). Giordano et al. (13) found the abundance of bird species can have unpredictable impacts on transmission of the virus due to the feeding preferences of mosquitoes on certain vectors. In late summer, *Culex* sp feeding behavior shifts away from birds to human blood meals (15). This coincides with the peak of human cases in most years.

Climate warming may also change the expansion of vector hosts, leading to WNV spread and become more severe in new parts of the world (16, 17). It can lead to an increase in the frequency and intensity of extreme weather events, such as heatwaves, floods or droughts, which could intensify the interaction between vectors, viruses and hosts (18, 19). Drought, for example, can lead to an increase in the abundance of mosquitoes (20, 21). In regards to

Canada, Ludwig et al. (22) showed that as temperatures continue to rise due to climate change, there is an increased risk of vector borne diseases in Canada.

Public health agencies across the US and Canada collect records on mosquito abundance and viral testing for WNV through mosquito trap data as well as human cases (2, 7). Through these programs, the mosquitoes are classified and counted by species based upon geographic locations. As WNV is of great concern, Temple et al. (23) constructed a presence/absence model to account for viral presence in mosquitoes, while accounting for underlying environmental variables for the presence of West Nile virus mosquito presence. They found a higher rate of mosquitoes in the summer, with the leading predictors found to be the number of freezing weeks, urban landscapes and the proportion of *Culex* species. However, this study was tied to mosquitoes instead of human cases. Giordano et al. (13) conducted the first epidemiological study of human WNV cases in Ontario using data from 2002 to 2013 in seven of the most populous PHUs. They found a positive correlations for temperature, cumulative precipitation, mosquito abundance, and the minimum infection rate. This study also found strong spatial autocorrelation of the positive mosquito pools but did not explore any spatial structure in the human WNV case data. We develop a statistical model for human WNV cases to better understand the relationship between human cases and environmental variables across the 34 PHUs in lower Ontario. We use a zero-inflated Poisson model, which is used to model data that has an excess of zeros. We compare a spatial and non-spatial version





of the model to assess if there is an underlying spatial structure to the WNV counts. To our knowledge, this has not been done before.

The paper is presented as follows. In Section 2, we introduce the WNV data in Ontario, Canada and describe the environmental variables used in the analysis. We also present the spatial and non-spatial zero-inflated Poisson model for modeling WNV human cases. In Section 3 we present and compare model results for WNV. We conclude with a short discussion on the zero-inflated Poisson model and future use in human case disease modeling.

## 2. Materials and methods

### 2.1. Data sources

Ontario is the most populous province in Canada with just over 15 million residents and the third largest in size spanning 1.076 million km<sup>2</sup> (Figure 1). Ontario is divided into 34 Public Health Units (PHUs). We collected monthly human case data from 2010 to 2019 for each PHU from the Public Health Ontario website (24). We focus our analysis on the 27 PHUs located in lower Ontario as upper Ontario is sparsely populated and has

had only 8 total cases reported across all 10 years in our data set. Most human cases are reported between May and October with a peak in August while very few cases are observed in the winter months (Figure 2). We observe high variability in human cases from year to year with some years having very few cases, for example 2010 and 2014, and others having large outbreaks i.e., 2012, 2017. It appears some spatial correlation may be present in the case data as the majority of cases are clustered around Toronto (Figure 3). While severe cases are likely to be detected, most cases will not be reported. Hence, the WNV cases reported are only a small subset of the true number of infections.

We collected mosquito trap data, bird abundance, temperature and precipitation, land cover, population, and Normalized Difference Vegetation Index (NDVI) data to be used as potential covariates to model the human WNV cases.

Since 2002, each PHU has used mosquito traps to obtain information on mosquito abundance and viral testing. Officials place light traps to attract mosquitoes and return to collect the traps on a weekly or biweekly basis from May through October. They record the total abundance of mosquitoes collected in the trap. A subset of this total is then further identified by species. Then,

mosquitoes are pooled into groups and tested for WNV in a lab. Each pooled sample is recorded as either positive or negative. A positive pool indicates that at least one mosquito in the pool is positive but does not provide insight into the number or fraction of mosquitoes in the pool that are positive. Pool sizes range from 1–60 with the majority of pools containing less than 10 mosquitoes. To account for the differences in pool sizes, the mosquito infection rate is often used as a proxy.

A common correction is to use the Minimum Infection Rate (MIR) which assumes that only one individual is positive for every positive pool. This effectively ignores the pooling information and provides a lower bound on the infection rate. This method is appropriate only when infections are considered to be low (25). Otherwise, the rate tends to be too narrow as it does not reflect the information lost in pooling. Instead, since our data contains varying pool sizes, we use the MLE method as described in Biggerstaff (26) which incorporates the different pool sizes in the infection rate calculation. Using this method, we assume the number of positive mosquito pools,  $X_i$  for  $i = 1, \dots, M$ , where  $M$  is the number of distinct pool sizes, is binomially distributed as shown in Equation 1 where  $n_i$  is the number of pools of size  $m_i$ , and  $X_i$  is the number of  $n_i$  pools observed that are positive.

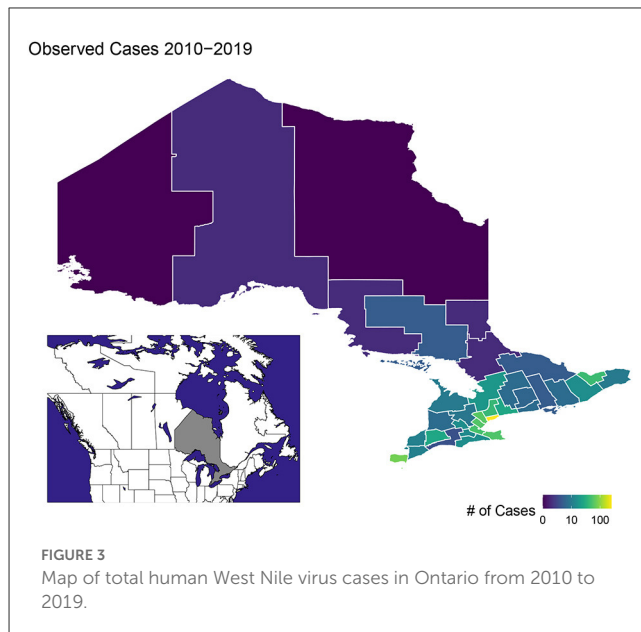
$$X_i \sim \text{Binomial}(n_i, 1 - (1 - p)^{m_i}) \tag{1}$$

We gathered observational bird data from Cornell’s citizen science eBird project (27). We calculated the relative abundance for several bird species known to be good carriers of WNV including the American Robin, House Sparrow, and American Crow (28). Relative abundance is used to correct for the inherent sampling bias present in the eBird data. Kilpatrick (29) found abundance of bird species can have unpredictable impacts on transmission of the virus due to the feeding preferences of mosquitoes on certain hosts. Feeding preferences of *Culex* mosquitoes have also been observed to change throughout the season with mosquitoes preferring human blood meals more later in the year (12).

We collected temperature and precipitation data from weather stations in Ontario using the weathercan package in R (30, 31). We obtained 12 land classification proportions that had been inferred from satellite imagery (32). Census data was collected from Statistics Canada and includes population, population density, number of dwellings, dwell density, and area (33). We collected monthly NDVI data at 1km spatial resolution from the MODISstp package (34). All covariates were aggregated to obtain monthly averages across each PHU.

## 2.2. Model

The West Nile Virus case data contains many months in which zero cases are reported in a given PHU. Due to the excess of zeros, we use Bayesian zero-inflated Poisson (ZIP) model. The ZIP model allows us to model zero observations from two distinct processes, the one associated with the number of cases, and a model for the excess zeros. The model has two parts, a Poisson count model and a logit model for the zeros. We fit two different ZIP models, a spatial and non-spatial model. We will describe the non-spatial version and extrapolate to the spatial version.



For public health unit  $i$ , ( $i = 1, \dots, 27$ ), and month  $t$ , ( $t = 1, \dots, 120$ ), define  $Y_{it}$  to be the number of human West Nile Virus cases where cases start in January 2010. We model the West Nile cases as a Poisson generalized linear model,

$$Y_{it}|X_{it} \sim \text{Poisson}(E_{it}e^{X_{it}^T\beta}) \tag{2}$$

Where  $E_{it}$  is the expected number of disease counts in the absence of covariate effects. This is the number of cases expected in each PHU if each person is equally likely to get West Nile virus. However, when there are no cases observed in PHU  $i$  and time  $t$  we assume that these zeroes occur as structural zeroes with probability  $\omega_{it}$ , and model the zero case counts as Poisson variables with probability  $(1 - \omega_{it})$ . We assume the following Bayesian zero-inflated Poisson model:

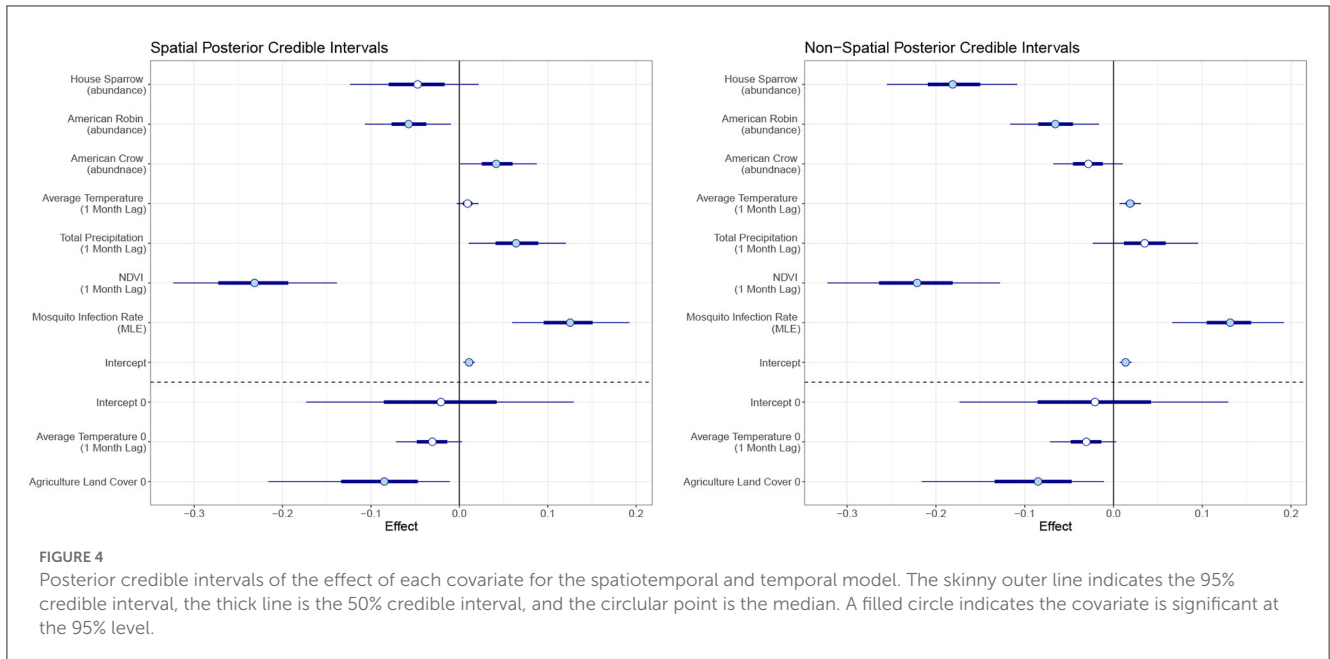
$$P(Y_{it} = y_{it}|\omega_{it}, \lambda_{it}) = \begin{cases} \omega_{it} + (1 - \omega_{it})\exp(-\lambda_{it}) & \text{for } y_{it} = 0 \\ (1 - \omega_{it})\exp(-\lambda_{it})\lambda_{it}^{y_{it}}/y_{it}! & \text{for } y_{it} > 0 \end{cases} \tag{3}$$

We assume a logit link for each probability,  $\omega_{it}$ , and relate these quantities to environmental variables, temporal and, in the spatial version, spatial random effects. More specifically, we assume a Poisson distribution of the cases as

$$\log(\lambda_{it}) = X_{it}\beta + \phi y_{i,t-1} \tag{4}$$

Where  $X_{it}$  is a vector of covariates related to exposure risk,  $\beta$  is the vector of regression coefficients. We include a temporal random effect,  $\phi$ , which is an autoregressive process. It is used to account for the idea that the cases within PHU  $i$  during month  $t$  are closely related to the cases the month before, i.e., PHU  $i$  at month  $t - 1$ . We assume that the number of cases relate to mosquito infection rate, bird abundance, NDVI, temperature, and precipitation. The model for the zero cases also includes a random effect to account for temporal autocorrelation as in the Poisson model. That is,

$$\text{logit}(\omega_{it}) = Z_{it}\gamma + \phi y_{i,t-1} \tag{5}$$



**FIGURE 4** Posterior credible intervals of the effect of each covariate for the spatiotemporal and temporal model. The skinny outer line indicates the 95% credible interval, the thick line is the 50% credible interval, and the circular point is the median. A filled circle indicates the covariate is significant at the 95% level.

**TABLE 1** Posterior results for the spatial and non-spatial model.

	Parameter	Spatial			Non-Spatial		
		Mean	SD	95% CI	Mean	SD	95% CI
Poisson	Average Temperature (°C)	0.009	0.008	(-0.006, 0.024)	0.019	0.007	<b>(0.005, 0.034)</b>
	Total Precipitation (cm)	0.065	0.034	<b>(0.011, 0.121)</b>	0.035	0.035	(-0.033, 0.103)
	Mosquito infection rate (MLE)	0.125	0.042	<b>(0.043, 0.206)</b>	0.130	0.038	<b>(0.052, 0.204)</b>
	NDVI	-0.231	0.058	<b>(-0.341, -0.122)</b>	-0.022	0.006	<b>(-0.034, -0.011)</b>
	American robin	-0.058	0.029	<b>(-0.119, -0.001)</b>	-0.065	0.029	<b>(-0.125, -0.01)</b>
	House sparrow	-0.049	0.045	(-0.137, 0.033)	-0.181	0.045	<b>(-0.269, -0.096)</b>
	American crow	0.043	0.026	<b>(0.001, 0.097)</b>	-0.029	0.024	(-0.076, 0.016)
	Intercept	0.011	0.004	<b>(0.003, 0.019)</b>	0.014	0.004	<b>(0.006, 0.022)</b>
Zero- Inflated	Agriculture land cover	-0.096	63.731	<b>(-0.238, -0.006)</b>	-0.097	0.063	<b>(-0.241, -0.005)</b>
	Average temperature (°C)	-0.032	23.100	(-0.077, 0.006)	-0.031	0.024	(-0.081, 0.006)
	Intercept	-0.022	94.348	(-0.206, 0.16)	-0.014	0.097	(-0.213, 0.17)

This shows the mean, standard deviation, and 95% Credible Interval (CI) of the effect size of each parameter from the posterior draws. Credible intervals in bold indicate the parameter is significant at the 95% level.

Where  $Z_{it}$  is a vector of covariates related to the structural zeros,  $\gamma$  is the vector of regression coefficients.

The spatial version of the zero-inflated Poisson model borrows information from neighboring sites. For example, for one public health unit,  $i$ , if the neighboring public health units have cases, it is more likely that the public health unit,  $i$ , will also have WNV cases. The spatial model is modified slightly from Equation (4), as a spatial random effect,  $u_i$  is incorporated for each PHU  $i$ . That is,

$$\log(\lambda_{it}) = X_{it}\beta + u_i + \phi y_{i,t-1} \tag{6}$$

$$\text{logit}(w_{it}) = Z_{it}\gamma + \phi y_{i,t-1} \tag{7}$$

Similarly to Equation (4),  $X_{it}$  and  $Z_{it}$  are vectors of covariates associated with the non-zero and zero process respectively. The  $\beta$

and  $\gamma$  are vectors of coefficients associated with these covariates and  $u_i$  are the spatially correlated random effects for the Poisson process.

We utilize a Bayesian paradigm and put multivariate normal priors on the regression coefficients. We put an intrinsic conditional autoregressive prior on  $u_i$  (see [Supplementary material](#) for details). Posterior estimates are found using a random walk Metropolis-Hastings Markov Chain Monte Carlo (MCMC) in NIMBLE (35–37). The MCMC simulation was run for three independent chains for 500,000 iterations with a burn-in period of 100,000 iterations and a thinning level of 100. Posterior trace plots were inspected to ensure convergence of each chain.

For model selection, we utilized the Watanabe-Akaike information criteria (WAIC) as a measure of out-of-sample

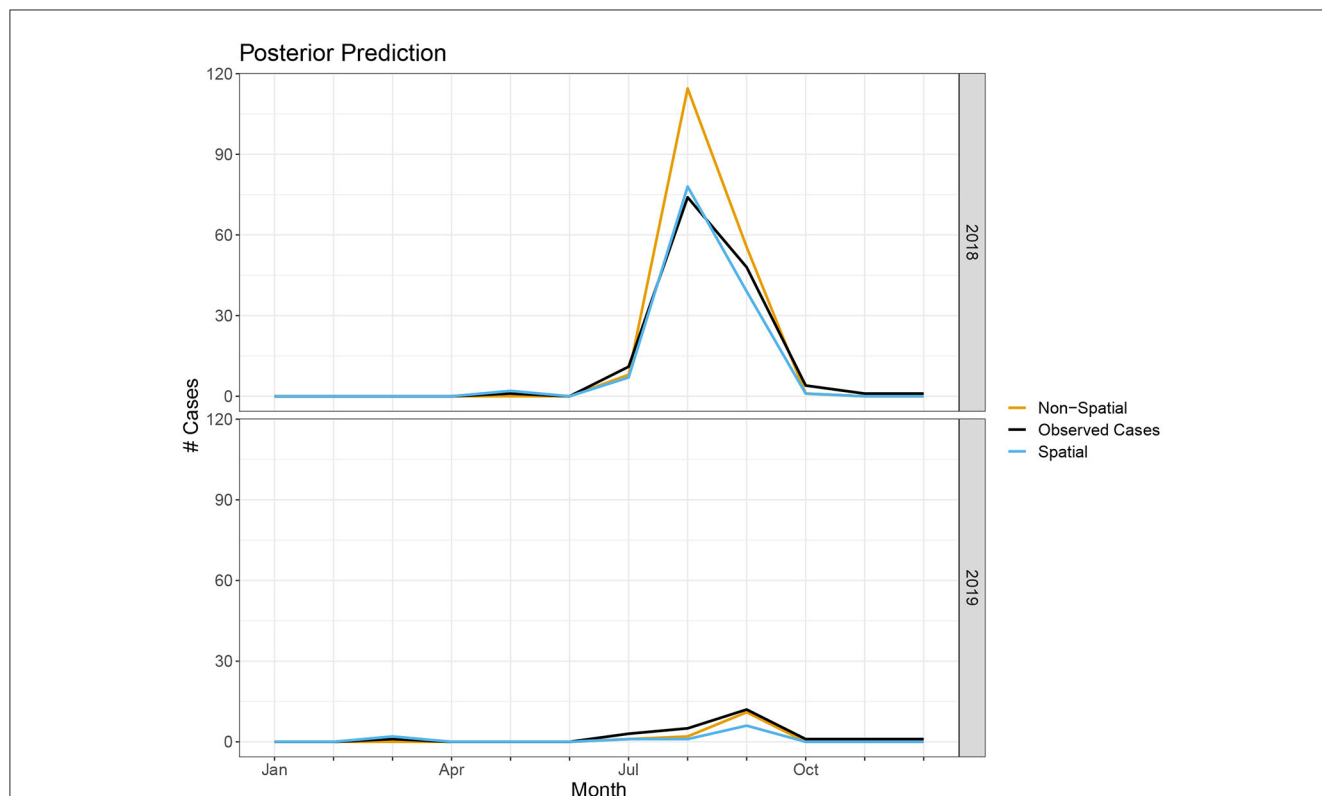


FIGURE 5 Number of cases predicted in spatial and non-spatial model compared to true observed cases in 2018 and 2019.

predictive accuracy which adds a correction for the effective number of parameters to adjust for overfitting (38, 39). A lower WAIC values indicates a better model fit. NIMBLE calculates the WAIC from the posterior samples produced by the MCMC algorithm. All temporal variables were lagged by 0, 1, and 2 months and the WAIC values of each model were compared to select the most appropriate lags.

In order to test the predictability of our model, we held out data from 2018 to 2019 for posterior prediction. We drew 1,000 samples from the posterior distributions of each parameter to obtain 1,000 predictions of the number of cases in each month and PHU ( $Y_{it}^{pred}$ ). To evaluate our predictions, we calculate the mean squared error between the predicted number of cases and the true number of cases as in Equation 8.

$$MSE = \frac{1}{n} \sum_{i,t} (Y_{it}^{pred} - Y_{it})^2 \tag{8}$$

We also compare visually the median number of predicted cases in each year to the observed data.

### 3. Results

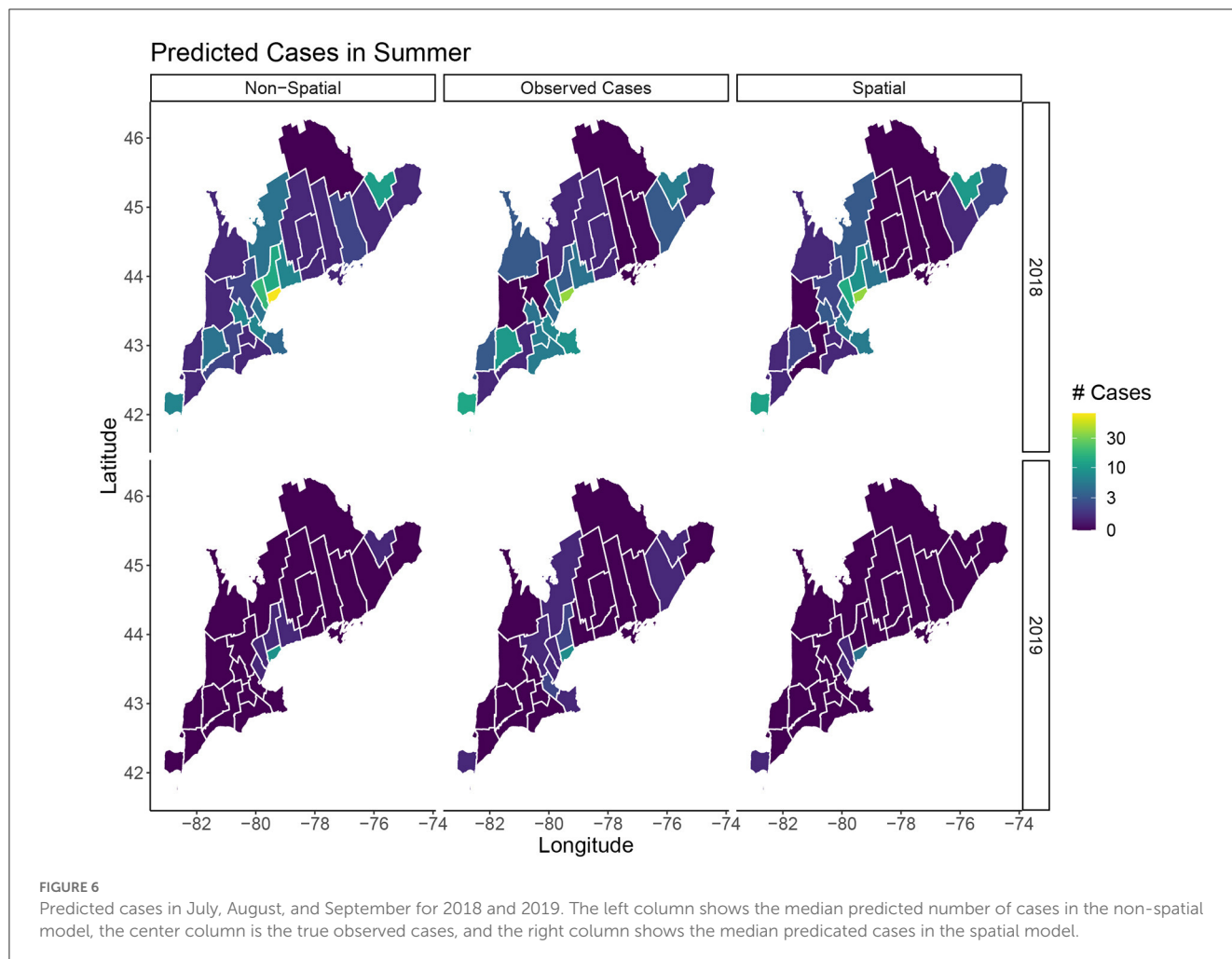
The temporal and spatiotemporal models were compared to see if the spatial dependence from the neighboring sites adds more value for predicting WNV cases. As part of the modeling, we explored different lagged windows, i.e., 1 or 2

month temperature or precipitation lag. This was based on the knowledge that it takes a month or two for a WNV cases to be reported (40).

In our final model, we used mosquito infection rate MLE, NDVI (1 month lag), average temperature (1 month lag), total precipitation (1 month lag), and crow, robin, and sparrow abundance for the Poisson covariates. For the structural zero covariates, we used percent of agriculture and average temperature lagged by 1 month. Posterior effect estimates for the spatial and non-spatial model are shown in Figure 4. A filled circle indicates the predictor is significant at the 95% level. The thicker line shows the 50% credible interval and the thin line shows the 95% credible interval. Point estimates and credible interval values are also included in Table 1.

In the temporal model, we found MLE and average temperature had a significant positive effect on human WNV cases at the 95% level with mosquito MLE having the largest effect. NDVI, sparrow and robin abundance had negative effects. Agriculture land cover showed a significant negative effect on the log-odds for the structural zeroes. That is, an increase in the percent of land cover that is agricultural leads to a reduction in the log-odds for structural zeroes.

For the spatiotemporal model, total precipitation and crow abundance also have a significant positive effect at the 95% level. All other estimates are similar though the magnitudes differ slightly from the temporal model. The WAIC values for the non-spatial and spatial models are 1,404 and 1,328, respectively, which indicates the spatial model performs slightly better than the non-spatial version.



In terms of posterior predictions, the spatiotemporal model also performs slightly better with an MSE of 1.17 compared to the temporal model MSE of 2.02. Median predictions for the spatial and non-spatial model are shown in Figure 5. We see the spatial model more accurately captures the peak infections in the summer of 2018 when cases are highest. Maps of the median posterior predictions by PHU for the summer months (July, August, and September) are shown in Figure 6. The non-spatial model in 2018 highly overpredicts the number of cases in the greater Toronto region while the spatial model predictions are close to the observed data. Very few cases were observed in 2019 and both model predictions are similar to the observed data. We calculated the difference between the number of cases observed in each month with the median number of cases predicted. Spatial maps of the differences for summer months in 2018 and 2019 are shown in Supplementary Figure S1. The number of cases predicted in all other seasons are very low and only contain small deviations from the observed data.

## 4. Discussion

Human case data of WNV is complicated to study and has many limitations. Case data is only reported at the monthly timescale while many of our covariates are measured at hourly,

daily, or weekly increments. Aggregating these covariates to the monthly scale may lead to a loss of information and may mask the underlying relationship between these covariates and human cases. Additionally, cases are greatly under reported due to a high number of asymptomatic cases and limited surveillance testing. We develop a zero-inflated Poisson model with lagged temporal covariates to help elucidate the relationship between human cases of WNV and environmental variables.

We find the mosquito infection rate correlates strongly with human WNV cases. Average temperature and total precipitation lagged by 1 month also have slight positive correlations with human cases. Higher temperatures result in an increasing abundance of mosquitoes, a shorter extrinsic incubation period, and increased biting rates, thus, leading to a higher probability of transmission of WNV (41). Our results show a seemingly counter-intuitive negative correlation between NDVI and cases since an increase in NDVI should lead to an increase in mosquitoes (42, 43). However, this relationship was previously observed in several other studies of WNV (20, 44). This may be related to the tendency of *Culex* mosquitoes to thrive better in artificial pools of water where fewer predators are present (45). Low NDVI can also be indicative of a drought which may cause mosquitoes and birds to congregate near the remaining bodies of water thus increasing transmission (21). We decided to only include the percent of agriculture among

the land cover variables due to the high correlation between the different land cover variables. Miramontes et al. (46) found an association between agricultural activity and human incidence of West Nile virus.

The relationship between birds and WNV is complex due to the shifting seasonal biting preferences of mosquitoes. While robins are known to be good carriers of WNV, we find an increase in the abundance of robins is associated with a decrease in human cases. This could be due to the feeding preference of mosquitoes being skewed toward birds at the times where the abundance of this species is higher. While using relative abundance helps standardize the eBird observations, it is still impacted by factors such as bias in the reporting rate of rare and common species (47).

Our model is able to accurately capture the trends and peak timing of WNV outbreaks. While the posterior estimates for the non-spatial and spatial models are similar, the WAIC values indicate a modest improvement when using the spatial model. Our posterior prediction for 2018 shows the spatial model more accurately predicts the peak number of cases in the summer (Figure 5). The number of cases in 2019 is much lower and the difference between the two model predictions is negligible. The improvements with the spatiotemporal model are modest but, especially when cases are high, it does appear there is a spatial structure to the human cases and including spatial random effects leads to better predictions. Our model's ability to accurately predict the magnitude and timing of yearly WNV outbreaks could be valuable to public health officials to implement prevention strategies and public health announcements to mitigate these outbreaks.

One limitation of this study is the narrow geographic scope. Currently our results are specific to the province of Ontario and cannot be generalized to other locations. Future work applying this model to locations across North America is necessary to determine the reliability of our results in other locations. Expanding this analysis is complicated due to varying levels of data availability in different public health agencies. In the United States, for instance, county level data for human cases is reported on a yearly temporal scale and monthly case data is only available at the state level. Investigating seasonal trends will be more challenging at these larger spatial and temporal scales.

The correlations observed between these environmental factors with human cases indicates climate change could lead to an increase of WNV cases across Ontario. Currently PHUs in upper Ontario have reported very few cases of WNV but increasing temperatures and an increase in drought conditions may lead to more cases in this region in the coming years. Also, an increase in temperature could lead to changes in bird migration patterns and locations. Some species of birds have reportedly been migrating earlier in the spring as a result of warmer spring mean temperatures (48). This could impact the peak timing and length of the WNV disease season. The accurate predictive power of our model could allow us to simulate WNV outbreaks under different climate change scenarios. Knowledge of the intricate relationships between these

variables and human cases is vital for planning and prevention of future outbreaks of WNV.

## Data availability statement

Publicly available datasets were analyzed in this study. This data can be found at: <https://www.publichealthontario.ca/en/data-and-analysis/infectious-disease/west-nile-virus>.

## Author contributions

LA and KAK: conceptualization, methodology, statistical analysis, and writing. LA: computer programming. All authors contributed to the article and approved the submitted version.

## Funding

This work was funded by a Laboratory Directed Research and Development (LDRD) Directed Research grant 20210062DR. This work is approved for distribution under LA-UR-22-32109. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of Los Alamos National Laboratory. Los Alamos National Laboratory, 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.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2023.1100543/full#supplementary-material>



## References

- Lanciotti R, Roehrig J, Deubel V, Smith J, Parker M, Steele K, et al. Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. *Science*. (1999) 286:2333–7. doi: 10.1126/science.286.5448.2333
- West Nile virus. *Centers for Disease Control and Prevention* (2022). Available online at: <https://www.cdc.gov/westnile/index.html> (accessed November 16, 2022).
- Mostashari F, Bunning ML, Kitsutani PT, Singer DA, Nash D, Cooper MJ, et al. Epidemic West Nile encephalitis, New York, 1999: results of a household-based seroepidemiological survey. *Lancet*. (2001) 358:261–4. doi: 10.1016/S0140-6736(01)05480-0
- Hayes EB, Komar N, Nasci RS, Montgomery SP, O'Leary DR, Campbell GL. Epidemiology and transmission dynamics of West Nile Virus Disease. *Emerg Infect Dis*. (2005) 11:1167–73. doi: 10.3201/eid1108.050289a
- Hadfield J, Brito AF, Swetnam DM, Vogels CB, Tokarz RE, Andersen KG, et al. Twenty years of West Nile virus spread and evolution in the Americas visualized by Nextstrain. *PLoS Pathogens*. (2019) 15:e1008042. doi: 10.1371/journal.ppat.1008042
- Curren EJ, Lehman J, Kolsin J, Walker WL, Martin SW, Staples JE, et al. West Nile virus and other nationally notifiable arboviral diseases—United States, 2017. *Morbidity and Mortality Weekly Report*. (2018) 67:1137. doi: 10.15585/mmwr.mm6741a1
- Surveillance of West Nile virus. *Public Health Agency of Canada* (2015). Available online at: <https://www.canada.ca/en/public-health/services/diseases/west-nile-virus/surveillance-west-nile-virus.html> (accessed November 16, 2022).
- García-Carrasco JM, Muñoz AR, Olivero J, Segura M, Real R. Predicting the spatio-temporal spread of West Nile virus in Europe. *PLoS Neglect Trop Dis*. (2021) 15:e0009022. doi: 10.1371/journal.pntd.0009022
- Brown H, Duik-Wasser M, Andreadis T, Fish D. Remotely-sensed vegetation indices identify mosquito clusters of West Nile virus vectors in an urban landscape in the northeastern United States. *Vector Borne Zoonotic Dis*. (2008) 8:197–206. doi: 10.1089/vbz.2007.0154
- Ward MP. Equine West Nile virus disease occurrence and the normalized difference vegetation index. *Prevent Vet Med*. (2009) 88:205–12. doi: 10.1016/j.prevetmed.2008.10.003
- Fuller DO, Troyo A, Beier JC. El Niño Southern Oscillation and vegetation dynamics as predictors of dengue fever cases in Costa Rica. *Environ Res Lett*. (2009) 4:014011. doi: 10.1088/1748-9326/4/1/014011
- Kilpatrick AM, Kramer LD, Jones MJ, Marra PP, Daszak P. West Nile virus epidemics in north america are driven by shifts in mosquito feeding behavior. *PLoS Biol*. (2006) 4:e82. doi: 10.1371/journal.pbio.0040082
- Giordano BV, Kaur S, Hunter FF. West Nile virus in Ontario, Canada: a twelve-year analysis of human case prevalence, mosquito surveillance, and climate data. *PLoS ONE*. (2017) 12:e0183568. doi: 10.1371/journal.pone.0183568
- Gorris ME, Bartlow AW, Temple SD, Romero-Alvarez D, Shutt DP, Fair JM, et al. Updated distribution maps of predominant Culex mosquitoes across the Americas. *Parasites Vectors*. (2021) 14:1–13. doi: 10.1186/s13071-021-05051-3
- Kilpatrick AM, Wheeler SS. Impact of West Nile virus on bird populations: limited lasting effects, evidence for recovery, and gaps in our understanding of impacts on ecosystems. *J Med Entomol*. (2019) 56:1491–7. doi: 10.1093/jme/tjz149
- Cotton PA. Avian migration phenology and global climate change. *Proc Natl Acad Sci USA*. (2003) 100:12219–22. doi: 10.1073/pnas.1930548100
- Paz S. Climate change impacts on West Nile virus transmission in a global context. *Philos Trans R Soc B Biol Sci*. (2015) 370:20130561. doi: 10.1098/rstb.2013.0561
- Chen CC, Epp T, Jenkins E, Waldner C, Curry PS, Soos C. Modeling monthly variation of culex tarsalis (Diptera: Culicidae) abundance and West Nile virus infection rate in the canadian prairies. *Int J Environ Res Public Health*. (2013) 10:3033–51. doi: 10.3390/ijerph10073033
- Semenza JC, Suk JE. Vector-borne diseases and climate change: a European perspective. *FEMS Microbiol Lett*. (2018) 365:fnx244. doi: 10.1093/femsle/fnx244
- Wang G, Minnis RB, Belant JL, Wax CL. Dry weather induces outbreaks of human West Nile virus infections. *BMC Infect Dis*. (2010) 10:38. doi: 10.1186/1471-2334-10-38
- Paull SH, Horton DE, Ashfaq M, Rastogi D, Kramer LD, Diffenbaugh NS, et al. Drought and immunity determine the intensity of West Nile virus epidemics and climate change impacts. *Proc R Soc B Biol Sci*. (2017) 284:20162078. doi: 10.1098/rspb.2016.2078
- Ludwig A, Zheng H, Vrbova L, Drobot M, Iranpour M, Lindsay L. Increased risk of endemic mosquito-borne diseases in Canada due to climate change. *Can Commun Dis Rep*. (2019) 45:91–7. doi: 10.14745/ccdr.v45i04a03
- Temple SD, Manore CA, Kaufeld KA. Bayesian time-varying occupancy model for West Nile virus in Ontario, Canada. *Stochastic Environ Res Risk Assess*. (2022) 36:2337–52. doi: 10.1007/s00477-022-02257-4
- PHO. *Public Health Ontario* (2022). Available online at: <https://www.publichealthontario.ca/en/data-and-analysis/infectious-disease/west-nile-virus> (accessed November 16, 2022).
- Biggerstaff BJ, Petersen LR. A modeling framework for evaluation and comparison of trigger strategies for switching from minipool to individual-donation testing for West Nile virus. *Transfusion*. (2009) 49:1151–9. doi: 10.1111/j.1537-2995.2009.02112.x
- Biggerstaff BJ. Confidence intervals for the difference of two proportions estimated from pooled samples. *J Agric Biol Environ Stat*. (2008) 13:478–96. doi: 10.1198/108571108X379055
- Sullivan BL, Wood CL, Iliff MJ, Bonney RE, Fink D, Kelling S. eBird: a citizen-based bird observation network in the biological sciences. *Biol Conservat*. (2009) 142:2282–92. doi: 10.1016/j.biocon.2009.05.006
- Kain MP, Bolker BM. Predicting West Nile virus transmission in North American bird communities using phylogenetic mixed effects models and eBird citizen science data. *Parasit Vectors*. (2019) 12:1–22. doi: 10.1186/s13071-019-3656-8
- Kilpatrick AM. Globalization, Land Use and the Invasion of West Nile Virus. *Science*. (2011) 334:323–7. doi: 10.1126/science.1201010
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna (2022). Available online at: <https://www.R-project.org/>.
- LaZerte SE, Albers S. Weathercan: download and format weather data from Canada. *J Open Source Software*. (2018) 3:571. doi: 10.21105/joss.00571
- Tuanmu MN, Jetz W. A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. *Glob Ecol Biogeogr*. (2014) 23:1031–1045. doi: 10.1111/geb.12182
- LLC M. *Statistics Canada* (2016). Available online at: <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/index-eng.cfm> (accessed November 16, 2022).
- Busetto L, Ranghetti L. MODISr: an R package for preprocessing of MODIS Land Products time series. *Comput Geosci*. (2016) 97:40–8. doi: 10.1016/j.cageo.2016.08.020
- de Valpine P, Turek D, Paciorek C, Anderson-Bergman C, Temple Lang D, Bodik R. Programming with models: writing statistical algorithms for general model structures with NIMBLE. *J Comput Graph Stat*. (2017) 26:403–13. doi: 10.1080/10618600.2016.1172487
- de Valpine P, Paciorek C, Turek D, Michaud N, Anderson-Bergman C, Obermeyer F, et al. *NIMBLE: MCMC, Particle Filtering, and Programmable Hierarchical Modeling*. R package version 0.12.2 (2022). Available online at: <https://cran.r-project.org/package=nimble> (accessed November 16, 2022).
- de Valpine P, Paciorek C, Turek D, Michaud N, Anderson-Bergman C, Obermeyer F, et al. *NIMBLE User Manual*. R package manual version 0.12.2 (2022). Available online at: <https://r-nimble.org> (accessed November 16, 2022).
- Watanabe S. A widely applicable bayesian information criterion. *J Mach Learn Res*. (2013) 14:867–97.
- Gelman A, Hwang J, Vehtari A. Understanding predictive information criteria for Bayesian models. *Stat Comput*. (2014) 24:997–1016. doi: 10.1007/s11222-013-9416-2
- DeFelice NB, Birger R, DeFelice N, Gagner A, Campbell SR, Romano C, et al. Modeling and surveillance of reporting delays of mosquitoes and humans infected with west Nile virus and associations with accuracy of West Nile Virus forecasts. *JAMA Netw Open*. (2019) 2:e191375. doi: 10.1001/jamanetworkopen.2019.3175
- Dohm DJ, O'Guinn ML, Turell MJ. Effect of environmental temperature on the ability of culex pipiens (Diptera: Culicidae) to transmit west Nile Virus. *J Med Entomol*. (2002) 39:221–25. doi: 10.1603/0022-2585-39.1.221
- Poh KC, Chaves LF, Reyna-Nava M, Roberts CM, Fredregill C, Bueno Jr R, et al. The influence of weather and weather variability on mosquito abundance and infection with West Nile virus in Harris County, Texas, USA. *Sci Total Environ*. (2019) 675:260–72. doi: 10.1016/j.scitotenv.2019.04.109
- Bisanzio D, Giacobini M, Bertolotti L, Mosca A, Balbo L, Kitron U, et al. Spatio-temporal patterns of distribution of West Nile virus vectors in eastern Piedmont Region, Italy. *Parasit Vectors*. (2011) 4:1–11. doi: 10.1186/1756-3305-4-230
- Myer MH, Johnston JM. Spatiotemporal bayesian modeling of West Nile Virus: identifying risk of infection in mosquitoes with local-scale predictors. *Sci Total Environ*. (2019) 650:2818–29. doi: 10.1016/j.scitotenv.2018.09.397
- Horsfall WE. Mosquitoes. Their bionomics and relation to disease. *JAMA*. (1955) 157:1457.
- Miramontes Jr R, Lafferty WE, Lind BK, Oberle MW. Is agricultural activity linked to the incidence of human West Nile virus? *Am J Prevent Med*. (2006) 30:160–3. doi: 10.1016/j.amepre.2005.10.008
- Isaac NJB, van Strien AJ, August TA, de Zeeuw MP, Roy DB. Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods Ecol Evol*. (2014) 5:1052–60. doi: 10.1111/2041-210X.12254
- Horton KG, La Sorte FA, Sheldon D, Lin TY, Winner K, Bernstein G, et al. Phenology of nocturnal avian migration has shifted at the continental scale. *Nat Climate Change*. (2020) 10:63–8. doi: 10.1038/s41558-019-0648-9