Stochastic Environmental Research and Risk Assessment (2022) 36:2337–2352
https://doi.org/10.1007/s00477-022-02257-4

ORIGINAL PAPER

Bayesian time-varying occupancy model for West Nile virus in Ontario, Canada

Seth D. Temple[1,2] · Carrie A. Manore[3] · Kimberly A. Kaufeld[1]

Accepted: 19 May 2022 / Published online: 28 June 2022

Abstract

Occupancy models determine the true presence or absence of a species by adjusting for imperfect detection in surveys. They often assume that species presences can be detected only if sites are occupied during a sampling season. We extended these models to estimate occupancy rates that vary throughout a sampling season as well as account for spatial dependence among sites. For these methods, we constructed a fast Gibbs sampler with the Pólya-Gamma augmentation strategy to conduct inference on covariate effects. We applied these methods to evaluate how environmental conditions and surveillance practices are associated with the presence of West Nile virus in mosquito traps across Ontario, Canada from 2002 to 2017. We found that urban land cover and warm temperatures drove viral occupancy, whereas viral testing on pools with higher proportions of *Culex* mosquitoes was more likely to result in a positive test for West Nile virus. Models with time-varying occupancy effects achieved much lower Watanabe-Akaike information criteria than models without such effects. Our final model had strong predictive performance on test data that included some of the most extreme seasons, demonstrating the promise of these methods in the study of pathogens spread by mosquito vectors.

Keywords Occupancy models · Spatio-temporal · Bayesian logistic regression · Vector-borne pathogens

1 Introduction

Vector-borne pathogens are becoming more prevalent and colonizing new spatial regions as a result of climate change and other human impacts, posing serious threats to human and animal populations. This trend is true for mosquito-borne pathogens because the mosquito life cycle depends on local weather and land use conditions (Bartlow et al. 2019). For instance, Gorris et al. (2021) inferred northward range expansions for several *Culex* mosquito species in updating the original maps of Darsie and Ward (1981). They found that abiotic factors like warm temperatures, water availability, terrain, and land cover characterized ecological niches for these mosquitoes. Relating environmental conditions to vector disease ecology under a statistical modeling framework can reveal key relationships between vectors and their environments and facilitate projections of disease spread.

Mosquitoes carry the potential to transmit pathogens maintained in reservoir hosts to humans and other animals. *Culex* mosquitoes, in particular, have been documented as potent disease vectors for West Nile virus (WNV), Eastern equine encephalitis virus (EEV), Zika virus, and other pathogens (Gorris et al. 2021). Turell et al. (2005) confirmed through laboratory experiments that *Culex* species could spread WNV efficiently as enzootic or bridge vectors. Since its arrival in North America in the late 1990s, WNV has spread across the United States and Canada and become endemic (Hadfield et al. 2019). It is a major public health concern, costing the American and Canadian economies hundreds of millions of dollars (Giordano et al. 2018).

✉ Seth D. Temple
sdtemple@uw.edu

Carrie A. Manore
cmanore@lanl.gov

Kimberly A. Kaufeld
kkaufeld@lanl.gov

[1] Statistical Sciences, Los Alamos National Laboratory, Los Alamos, NM, USA

[2] Department of Statistics, University of Washington, Seattle, WA, USA

[3] Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, NM, USA

Various public health agencies across the US and Canada maintain records on mosquito abundance and viral testing. These programs involve classifying and counting mosquito species found at trap sites across a geographic region. In many instances, pools of mosquitoes were tested for WNV and/or other pathogens. Although the dataset we are using contains direct information about vector-borne pathogens, most studies have only explored the patterns of vector abundance. There are established methods to model abundance in ecology (Royle and Nichols 2003; Royle and Dorazio 2006), whereas analyzing viral data requires another approach. Considering that WNV is a serious public health concern, we developed a new statistical method to make use of presence/absence indicators from viral tests to create an occupancy model specifically for viral presence in mosquitoes.

One of the challenges in developing an accurate viral occupancy model is that mosquito populations, viral transmission, and human sampling effort vary across time and space. Our method generalizes the occupancy model to allow occupancy to change during a sampling season (MacKenzie et al. 2002, 2003, 2017). Occupancy modeling posits a data-generating process by which a presence arises when a species is both occupying a sampling site and detected during a sampling period. The forces affecting occupancy and detection differ; occupancy may depend on local environmental conditions whereas detection may depend on sampling procedures and effort. Historically, these models have assumed that fauna either occupied or did not occupy a site throughout a sampling season (MacKenzie et al. 2002, 2003; Johnson et al. 2013; Hooten and Hobbs 2015). We relaxed this assumption to adjust for seasonality in vector-borne pathogens. Occupancy and detection probabilities can be construed as the outputs of generalized linear models (GLMs) with predictors belonging to three different dimensions of heterogeneity: site, season, and period. In our example, site is the areal region, season is the yearly pattern, and period is the epidemiological week. Hence, time-varying occupancy models can be implemented by introducing occupancy covariate effects that vary over periods.

The rest of this paper is presented as follows. Section 2 motivates the method development by introducing empirical data on West Nile virus in Ontario, Canada. Section 3 formalizes the time-varying occupancy model and places it alongside its historical counterparts. Monte Carlo Markov Chain (MCMC) methods for Bayesian binary regression are laid out in Section 3.3. Models with spatial random effects are discussed in Section 3.4. In Section 4, we present a model to study occupancy and detection patterns for WNV in Ontario. Besides serving as an example for time-varying occupancy, this case study uncovered strong statistical signals that were in line with contemporary scientific knowledge. We conclude with commentary on the time-varying occupancy model and its future use in vector disease ecology.

2 West Nile virus mosquito data

The province of Ontario is 1.076 million km^2 and holds about 20% of Canada's population. It is apportioned into 34 public health units (PHUs), ranging from rural, large, and sparsely populated PHUs in the northwest to urban, small, and densely populated PHUs in the southeast. Between 2002 and 2017, these PHUs trapped mosquitoes and tested them for WNV. Officials baited miniature light traps and returned 24 hours later to collect mosquitoes and diagnose WNV status (Giordano et al. 2018). Surveillance was conducted at hundreds of trap sites each week from May to October. Most PHUs surveyed traps at least once a week, so weekly aggregation resulted in fewer missing observations. Using the MMWRweek R package (Niemi 2020), we defined epidemiological weeks (epiweeks) according to the Morbidity and Mortality Weekly Report standard of the Centers for Disease Control and Prevention (CDC) and other public health agencies throughout the world. We chose this definition because human cases for infectious diseases like WNV are commonly reported this way.

2.1 Mosquito traps

Trap data included species classifications, abundance counts, and test results for WNV from mosquito seasons in 2002-2017. Pools of mosquitoes were blended together and then evaluated in aggregate for WNV, obscuring true counts for how many mosquitoes harbored the virus (Kesavaraju et al. 2012). Agency protocols and local mosquito abundance and diversity also impacted this detection process. Some PHUs collected and assessed more specimens than others. On the other hand, baiting for potential WNV vectors could have affected the viral testing. For example, the *Culex* genus is especially relevant to WNV transmission (Turell et al. 2005) and widespread throughout Ontario (Gorris et al. 2021). Zero inflation was a concern as well at finer spatial resolutions. These nuances and limitations in the viral testing informed our decision to model the binary response positive test versus negative test(s). We converted counts of WNV tests to presence-absence observations; namely, 1 corresponded to a positive test and 0 corresponded to no positive tests.

After data aggregation, there were 7396 trap observations, 1054 of which had a positive WNV test. Exploratory data analysis highlighted some interesting trends: (1) the majority of positive cases belonged to the *Culex pipiens* morphological group, (2) PHUs in the Greater Toronto Area (GTA) tested the most, (3) testing was generally consistent between years, and (4) most positive tests occurred between mid to late summer. Presences were mainly in southern Ontario and unequally distributed about the metro areas of Toronto, Ottawa, and Detroit. 2002, 2012, and 2017 had more cases than the average season whereas 2009 had fewer cases (Figure 1).

2.2 Environmental factors

Prior studies of this dataset have reported temperature, water resources, and land cover types that covary strongly with mosquito abundance (DeMets et al. 2020a, b). Using Poisson GLMs for the GTA data, Yoo (2014) and Yoo et al. (2016) discovered positive associations between temperature, precipitation, and population density and *Culex pipiens-restuans* abundance. Wang et al. (2011) found that temperature provided a stronger leading indication than precipitation in a Gamma GLM for *Culex pipiens-restuans* abundance in the Peel region (PEE). In general, mosquito species important to WNV transmission thrive in human-occupied land with standing water available and warm weather conditions. Birds also play a part in the enzootic cycle of WNV. Their competence as reservoir hosts impacts transmission in dynamic ways (Allan et al. 2009; Ciota and Kramer 2013).

Based on this literature review, we gathered environmental predictors to consider in our models. For climate and land type, we collected 19 bioclimatic variables (Vega et al. 2017) and 12 land classification proportions (Tuanmu and Jetz 2014) that had been inferred from satellite imagery. For weather trends and water availability, we assembled temperature, precipitation, and water level statistics from the daily reports of weather (Dunnington 2017) and hydrological stations (Albers 2017). Lastly, we computed Shannon and Gini-Simpson diversity indices (Simpson 1949; Willis and Martin 2020) based on observational data from the citizen-science project eBird (Sullivan et al. 2009). We chose these indices because they were less affected by the sampling effort biases of the eBird user base.

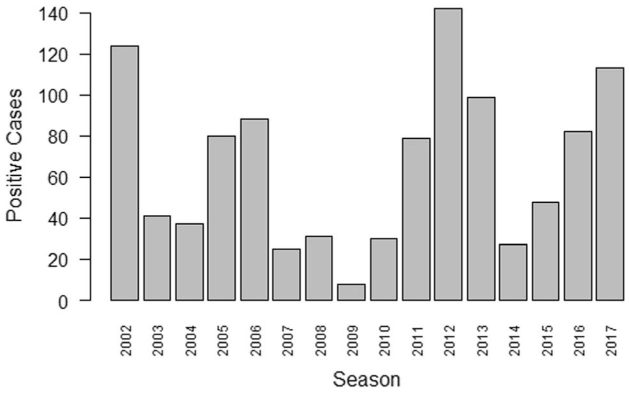

Fig. 1 Total number of positive WNV cases by sampling season

We aggregated these covariates to align with the PHUs and weekly reporting. In some time periods, some PHUs lacked weather, hydrological, or bird diversity measurements. We imputed values for this small fraction of the covariate data. Missing temperature, precipitation, and water level values were replaced with distance-weighted averages, whereas missing bird diversity indices were filled in with medians.

These environmental covariates varied over space and time. Regions at lower latitudes experienced warmer temperatures, and epiweeks 30 to 40 aligned with the summer heat. Bird diversity was typically strong throughout our study period, but dips did occur and may be predictive. We found water level to be relatively stable for each PHU, which may indicate that its data source consists of managed water resources. Plotting precipitation over time did not uncover any noteworthy trends. Agricultural land was prevalent in southern Ontario outside of downtown Toronto. While land cover and climate characterized sites only, temperature, precipitation, hydrological resources, and bird diversity changed over time and may inform time-varying occupancy.

3 Bayesian time-varying occupancy model

Study designs for occupancy modeling encompass heterogeneity along three different dimensions: site, season, and period. Responses and covariates comprise a fourth dimension. Let i, j, and k index sites, seasons, and periods, respectively. Throughout this paper, we keep track of these indices in variable subscripts. For our motivating WNV data analysis, sites were PHUs, seasons ranged from 2002 to 2017, and periods concerned the epiweeks 18 to 44.

3.1 Standard occupancy models

Mackenzie et al. developed likelihood-based approaches in the early 2000s to accommodate imperfect detection when surveying animals (MacKenzie et al. 2002). We refer to this model and its many extensions (MacKenzie et al. 2017) as occupancy models. These models separate observation probabilities into products of occupancy and detection probabilities. Before generalizing to multiple seasons, we introduce the single-season occupancy model.

For an observation to be made, a species must have been both present in an area and detected during sampling. Let y denote a binary observation and z be a latent occupancy status. For our case study, y is 1 if there was a positive WNV test and 0 otherwise, and z is 1 if WNV was circulating in a mosquito population and 0 otherwise. There are three cases to consider: (a) observed $((y,z)=(1,1))$, (b) not observed but occupied $((y,z)=(0,1))$, and (c) not occupied $((y,z)=(0,0))$. (Detection can only occur if there is occupancy, hence $((y,z)=(1,0))$ is impossible.) If the species was observed at least once in the season, the third case did not apply. If the species was never observed, this was either because it did not occupy the site or because of repeated failures to detect it. The data-generating process is formally stated as:

$$Y_{ik} \sim \text{Bernouilli}(Z_i \cdot p_{ik})$$
$$Z_i \sim \text{Bernouilli}(\psi_i),$$

where vectors $\boldsymbol{\psi}$ and $\mathbf{p}$ are occupancy and detection probabilities. Detection probabilities can be different over time and among sites, whereas occupancy probabilities may only change by site. If a site did not have a survey for a given period, no data contributes to the model likelihood. Given presence-absence data $\mathbf{Y}$, the likelihood is as follows:

$$\begin{aligned}\mathcal{L}(\boldsymbol{\psi},\mathbf{p}|\mathbf{y}) = &\left(\prod_i \left(\psi_i \prod_k p_{ik}^{y_{ik}}(1-p_{ik})^{1-y_{ik}}\right)\cdot \underbrace{I\left(\sum_k y_{ik} \geq 1\right)}_{\text{Observed at least once}}\right)\\ &\times \left(\prod_i \left(\psi_i \prod_k (1-p_{ik}) + (1-\psi_i)\right)\cdot \underbrace{I\left(\sum_k y_{ik} = 0\right)}_{\text{Never observed}}\right)\end{aligned} \qquad (1)$$

A colonization-extinction framework applies to multiple sampling seasons in which site occupancy probabilities may have changed. This model involves additional parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\varepsilon}$ for colonization and extinction probabilities (MacKenzie et al. 2003). Occupancy probabilities are calculated recursively:

$$\psi_{ij} = \psi_{i(j-1)}(1-\varepsilon_{i(j-1)}) + (1-\psi_{i(j-1)})\gamma_{i(j-1)}$$

With too many parameters, maximum likelihood methods are unlikely to converge, so the standard practice is to model these component probabilities with probit or logistic regression (MacKenzie et al. 2002). An appealing aspect of the colonization-extinction model is that these components may depend on different covariates than the occupancy and detection components.

3.2 Time-varying occupancy

Single-season occupancy and colonization-extinction models are restrictive in their assumption that site occupancy was constant throughout a sampling season. We weaken this assumption, letting site occupancy vary by period. Thus, the data-generating process and likelihood are:

$$Y_{ijk} \sim \text{Bernouilli}(Z_{ijk} \cdot p_{ijk})$$
$$Z_{ijk} \sim \text{Bernouilli}(\psi_{ijk})$$
$$\mathcal{L}(\psi,\mathbf{p}|\mathbf{y}) = \prod_{ijk} (\underbrace{\psi_{ijk}p_{ijk}}_{\text{(a)}})^{y_{ijk}} (\underbrace{\psi_{ijk}(1-p_{ijk})}_{\text{(b)}} + \underbrace{(1-\psi_{ijk})}_{\text{(c)}})^{1-y_{ijk}} \qquad (2)$$

In our methods, we achieve time-varying occupancy ψ_{ijk} by including covariates $\mathbf{x}_{ijk}$ that varied between sampling periods, since ψ_{ijk} is modeled with sigmoid regression. Excluding period-dependent occupancy covariates means ψ_{ijk} is the same for all periods k. However, when a species was never observed at a site, the likelihoods (1) and (2) differ slightly. Namely,

$$\prod_k (\psi_{ij}(1-p_{ijk}) + (1-\psi_{ij})) \neq (1-\psi_{ij}) + \psi_{ij}\prod_k (1-p_{ijk})$$

Our likelihood includes non-occupancy in the product. While this distinction may be important for likelihood optimization methods, we did not explore it further as it did not affect our MCMC sampling routines. Another difference is that seasonal effects are handled jointly with other occupancy effects. We interpret seasonal effects to have impacted occupancy directly, rather than impacting occupancy indirectly through colonization and extinction. For instance, we might conclude that harsh winters decreased occupancy whereas MacKenzie et al. (2003) would say that harsh winters increased extinction and decreased colonization. In this respect, the time-varying occupancy model appears suitable for a species that is widespread and mobile, responding uniformly to macro environmental changes. This scenario is the case for WNV, as it is maintained in migratory avian hosts and transmitted by flying insect vectors.

3.3 Gibbs samplers for occupancy models

Classical Bayesian methods for modeling binary data take the perspective that the outcomes depend on a latent regression structure; namely,

$$u_l = \begin{cases} 1, & v_l \geq 0 \\ 0, & v_l < 0 \end{cases}$$

$$\tilde{u}_l = \eta_l + \epsilon_l$$

where η_l is a linear predictor and ϵ_l is an error distribution. The errors follow normal distributions for probit regression and logistic distributions for logistic regression. Albert and Chib (1993) demonstrated for probit regression that the latent variables could be sampled from truncated-normal distributions. This strategy of sampling latent variables in a hierarchical model is referred to as data-augmentation. Conditional on the augmented variables, regression effects can be shown to be *a posteriori* normally-distributed. Polson et al. (2013) showed that such a scheme is possible for Bayesian logistic regression as well. Their method is nearly the same except their latent variables are sampled from Pólya-Gamma (PG) distributions, where PG random variables can be represented as infinite sums of scaled Gamma random variables. Since the regression effects remain *a posteriori* multivariate normal (MVN) conditional on the latent PG variables, we can sample them in two steps in a Gibbs way.

Combining Bayesian sigmoid regressions for occupancy and detection together results in a blocked Gibbs sampler. Dorazio and Rodriguez (2012) and Clark and Altwegg (2019) have shown as much for probit and logistic regression, respectively. For time-varying occupancy models, we have to draw z_{ijk} at each period. We implemented blocked Gibbs samplers for time-varying occupancy with probit or logit link functions. We added further hierarchy to our samplers with inverse-Wishart (IW) conjugacy for the covariances of the MVN effects. There are many alternative ways to sample these posteriors distributions from exact (Metropolis et al. 1953; Hoffman et al. 2014) to approximate (Rue et al. 2009; Blei et al. 2017) MCMC methods. Polson et al. (2013) argue in simulated and real data studies that the PG method does not depend on careful hyperparameter tuning for proposal densities and is more efficient than Metropolis-Hastings methods, especially for complex model frameworks like ours. This conclusion has been verified in the case of large spatial occupancy models (Clark and Altwegg 2019, Table 3).

Before presenting our samplers, we lay out some notation. Let $\mathbf{X}$ and $\mathbf{W}$ be four-dimensional occupancy and detection covariate arrays with column vectors $\mathbf{x}_{ijk}$ and $\mathbf{w}_{ijk}$. $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ denote occupancy and detection effects, and multivariate normal means and covariances are $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Using this notation and link function f, the data-generating process is as follows:

$$\begin{aligned} Y_{ijk} &\sim \text{Bernoulli}(Z_{ijk} \cdot p_{ijk}) \\ Z_{ijk} &\sim \text{Bernoulli}(\psi_{ijk}) \\ f(\psi_{ijk}) &= \mathbf{x}'_{ijk}\boldsymbol{\beta} \\ f(p_{ijk}) &= \mathbf{w}'_{ijk}\boldsymbol{\alpha} \\ \boldsymbol{\beta} &\sim \text{Normal}(\boldsymbol{\mu}_{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}) \\ \boldsymbol{\alpha} &\sim \text{Normal}(\boldsymbol{\mu}_{\boldsymbol{\alpha}}, \boldsymbol{\Sigma}_{\boldsymbol{\alpha}}) \\ \boldsymbol{\Sigma}_{\boldsymbol{\beta}} &\sim \text{Inverse-Wishart}(v_{\boldsymbol{\beta}}, \boldsymbol{\Lambda}_{\boldsymbol{\beta}}) \\ \boldsymbol{\Sigma}_{\boldsymbol{\alpha}} &\sim \text{Inverse-Wishart}(v_{\boldsymbol{\alpha}}, \boldsymbol{\Lambda}_{\boldsymbol{\alpha}}) \end{aligned}$$

Link functions $\text{logit}(\eta) = \log(\eta/(1-\eta))$ and $f(\eta) = \Phi^{-1}(\eta)$, where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable, characterize Bayesian logistic and probit regression for our hierarchical model. We decided to model with the logit link function because interpretation with respect to log-odds is more intuitive and preferred in statistical ecology (Northrup and Gerber 2018).

3.3.1 Logistic regression for time-varying occupancy

We outline an algorithm (Algorithm 1) to perform Gibbs sampling for a Bayesian time-varying occupancy model with logit link functions. (A similar algorithm (Algorithm 2) for probit link functions is available in the Appendix.) The data augmentation strategy of Polson et al. (2013) is to draw continuous augmented variables from PG distributions and then use them to sample regression effects. For occupancy modeling, we employ data augmentation to sample detection effects as well if the latent status is occupied for the current iteration. The supplement of Clark and Altwegg (2019) includes algebraic derivations for this sampler.

For the sampling algorithms, we use superscripts $^{(t)}$ to keep track of iterations, subscripts $_{\boldsymbol{\alpha}}$ and $_{\boldsymbol{\beta}}$ to distinguish between detection and occupancy, and accents to denote augmented variables $\tilde{y}_{ijk}$ and $\tilde{z}_{ijk}$ associated with observations y_{ijk} and latent occupancy z_{ijk}. The arrays $\mathbf{X}$ and $\mathbf{W}$ are collapsed to matrix form by stacking one on top of another the row vectors $\mathbf{x}'_{ijk}$ and $\mathbf{w}'_{ijk}$ for non-missing presence-absence observations. Intermediate mean $\mathbf{m}$ and covariance $\mathbf{V}$ terms are computed each time for MVN updates. Lastly, hyperpriors $\boldsymbol{\mu} = \mathbf{0}$, $\boldsymbol{\Lambda} = \mathbf{I}$ (identity matrix), and $v =$ dim(design matrix) are chosen to be weakly informative. A review of prior specifications for occupancy models can be found in Northrup and Gerber (2018).

Algorithm 1 Sample model effects for time-varying occupancy model (logit)

```
Require: data $\mathbf{y}$, $\mathbf{W}$, $\mathbf{X}$; priors $\boldsymbol{\mu}_{\boldsymbol{\alpha}}$, $\boldsymbol{\mu}_{\boldsymbol{\alpha}}$, $\nu_{\boldsymbol{\alpha}}$, $\nu_{\boldsymbol{\beta}}$, $\boldsymbol{\Lambda}_{\boldsymbol{\alpha}}$, $\boldsymbol{\Lambda}_{\boldsymbol{\beta}}$; iterations $T$
Initialize $\boldsymbol{\alpha}^{(0)} = \boldsymbol{\mu}_{\boldsymbol{\alpha}}$, $\boldsymbol{\beta}^{(0)} = \boldsymbol{\mu}_{\boldsymbol{\beta}}$, $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{(0)} = \boldsymbol{\Lambda}_{\boldsymbol{\alpha}}$, $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(0)} = \boldsymbol{\Lambda}_{\boldsymbol{\beta}}$
Compute $p_{ijk}^{(0)} = \text{logit}^{-1}(\mathbf{w}'_{ijk}\boldsymbol{\alpha}^{(0)})$, $\psi_{ijk}^{(0)} = \text{logit}^{-1}(\mathbf{x}'_{ijk}\boldsymbol{\beta}^{(0)})$ for all $i, j, k$
for t in 1:T do
    // occupancy component
    for all indices $i, j, k$ where $y_{ijk}$ non-missing do
        $z_{ijk}^{(t)} \sim \begin{cases} \text{Bernoulli}(1), & y_{ijk} = 1 \\ \text{Bernoulli}\left(\frac{\psi_{ijk}^{(t-1)}(1-p_{ijk}^{(t-1)})}{\psi_{ijk}^{(t-1)}(1-p_{ijk}^{(t-1)})+(1-\psi_{ijk}^{(t-1)})}\right), & y_{ijk} = 0 \end{cases}$
        $\tilde{z}_{ijk}^{(t)} \sim \text{Pólya-Gamma}(1, \mathbf{x}'_{ijk}\boldsymbol{\beta}^{(t-1)})$
        $\mathbf{V}_{\boldsymbol{\beta}}^{(t)} = ((\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(t-1)})^{-1} + \mathbf{X}'\text{diag}(\tilde{\mathbf{z}}^{(t)})\mathbf{X})^{-1}$
        $\mathbf{m}_{\boldsymbol{\beta}}^{(t)} = \mathbf{V}_{\boldsymbol{\beta}}^{(t)}((\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(t-1)})^{-1}\boldsymbol{\mu}_{\boldsymbol{\beta}} + \mathbf{X}'(\mathbf{z}^{(t)} - 1/2 \cdot \mathbf{1}))$
        $\boldsymbol{\beta}^{(t)} \sim \text{Normal}(\mathbf{m}_{\boldsymbol{\beta}}^{(t)}, \mathbf{V}_{\boldsymbol{\beta}}^{(t)})$
        $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(t)} \sim \text{Inverse-Wishart}(1 + \nu_{\boldsymbol{\beta}}, \boldsymbol{\Lambda}_{\boldsymbol{\beta}} + (\boldsymbol{\beta}^{(t)})'\boldsymbol{\beta}^{(t)})$
        $\psi_{ijk}^{(t)} = \text{logit}^{-1}(\mathbf{x}'_{ijk}\boldsymbol{\beta}^{(t)})$
    end for
    // detection component
    for all indices $i, j, k$ where $z_{ijk}^{(t)} = 1$ do
        $\tilde{y}_{ijk}^{(t)} \sim \text{Pólya-Gamma}(1, \mathbf{w}'_{ijk}\boldsymbol{\alpha}^{(t-1)})$
        $\mathbf{V}_{\boldsymbol{\alpha}}^{(t)} = ((\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{(t-1)})^{-1} + \mathbf{W}'\text{diag}(\tilde{\mathbf{y}}^{(t)})\mathbf{W})^{-1}$
        $\mathbf{m}_{\boldsymbol{\alpha}}^{(t)} = \mathbf{V}_{\boldsymbol{\alpha}}^{(t)}((\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{(t-1)})^{-1}\boldsymbol{\mu}_{\boldsymbol{\alpha}} + \mathbf{W}'(\mathbf{y}^{(t)} - 1/2 \cdot \mathbf{1}))$
        $\boldsymbol{\alpha}^{(t)} \sim \text{Normal}(\mathbf{m}_{\boldsymbol{\alpha}}^{(t)}, \mathbf{V}_{\boldsymbol{\alpha}}^{(t)})$
        $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{(t)} \sim \text{Inverse-Wishart}(1 + \nu_{\boldsymbol{\alpha}}, \boldsymbol{\Lambda}_{\boldsymbol{\alpha}} + (\boldsymbol{\alpha}^{(t)})'\boldsymbol{\alpha}^{(t)})$
        $p_{ijk}^{(t)} = \text{logit}^{-1}(\mathbf{w}'_{ijk}\boldsymbol{\alpha}^{(t)})$
    end for
end for
return $\boldsymbol{\alpha}^{(1:T)}, \boldsymbol{\beta}^{(1:T)}, \boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{(1:T)}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(1:T)}$
```

3.4 Spatial occupancy models

Dependence among areal units may confound inference in spatial models. Hughes and Haran (2013) recommend incorporating synthetic covariates that are orthogonal to fixed effects covariates to model positive dependence via spatial random effects (SREs) $\boldsymbol{\theta} \sim \text{Normal}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$. This approach has been taken before by Johnson et al. (2013) and Clark and Altwegg (2019) in fitting large spatial occupancy models.

Let $\mathbf{X}_s$ be occupancy covariates that are fixed for each site and $\mathbf{A}$ be an adjacency matrix, i.e., $A_{il} = 1$ if i and l touch and $A_{il} = 0$ otherwise. The Moran operator for $\mathbf{X}_s$ is $\mathbf{P}_{\mathbf{X}_s}^{\perp}\mathbf{A}\mathbf{P}_{\mathbf{X}_s}^{\perp}$, where $\mathbf{P}_{\mathbf{X}_s}^{\perp} = \mathbf{I} - \mathbf{X}_s(\mathbf{X}_s'\mathbf{X}_s)^{-1}\mathbf{X}_s'$. Eigenvectors from the spectral decomposition of the Moran operator become the synthetic covariates attached to the SREs; their corresponding eigenvalues indicate positive and/or negative spatial dependence. Let matrix $\mathbf{M}$ contain the orthogonal spatial covariates. The occupancy component is now modeled by the equation $f(\psi_{ijk}) = \mathbf{x}'_{ijk}\boldsymbol{\beta} + \mathbf{m}'_i\boldsymbol{\theta}$, and the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ is defined as $\sigma^2_{\boldsymbol{\theta}}(\mathbf{M}'\mathbf{Q}\mathbf{M})^{-1}$, where $\mathbf{Q}$ is the intrinsic conditionally autoregressive precision matrix (Besag and Kooperberg 1995). Inverse-Gamma (IG) conjugacy is assumed for the variance scalar $\sigma^2_{\boldsymbol{\theta}}$. In the Appendix, we present MCMC samplers for time-varying occupancy models with SREs.

4 West Nile virus in Ontario

We analyzed the West Nile virus data in Ontario, Canada using our methodology (Algorithm 1). We used R to sample from the posterior distribution in a Gibbs way the parameters of Bayesian time-varying occupancy models. For the logistic model, we call the R package BayesLogit which implements an exact and efficient accept/reject sampler for PG random variables (Polson et al. 2019). In addition, we programmed utility tools for data preparation, model diagnostics, model evaluation, and model visualization.

4.1 Model selection procedure

We split our data into training, validation, and testing datasets. We held out seasons 2008, 2012, and 2016 as examples of low, high, and medium case counts. With the remaining 13 years, we trained and validated models on seventy-five and twenty-five percent of the observations.

One diagnostic we checked was the Watanabe-Akaike information criterion (WAIC) as a measurement of out-of-sample predictive accuracy for Bayesian models (Watanabe 2013; Gelman et al. 2014). We implemented WAIC as $-2 \cdot (\text{log pointwise predictive density} - p_{\text{WAIC}_1})$ (Gelman et al. 2013, pp. 169,173). We measured WAIC on the training dataset. Additionally, we made note of model effects in which ninety-five percent credible intervals did not overlap with zero. We interpreted such findings to mean that the effect has a strong statistical signal. Lastly, we formulated our own posterior predictive check for Bayesian occupancy models (Algorithm 2). First, we calculate presence probabilities for a given array by sampling posterior effects and applying their link functions. Second, we simulate presence-absence data 1000 times and construct summary statistics. Our main summary statistic is the relative presence, which is the average positive count divided by the true positive count. Relative presence evaluates how many binary 1 observations a model generates normalized by the true observations. It can be computed over subsets of the sites (PHUs), seasons (years), and periods (epiweeks) to diagnose strengths and weaknesses of a model. We looked at this posterior predictive check for select sites and seasons in the validation dataset to assess out-of-sample accuracy.

Algorithm 2 Posterior predictive check (relative presence)

```
Require: sites $\mathcal{I}$, seasons $\mathcal{J}$, periods $\mathcal{K}$; data $\mathbf{y}$, $\mathbf{W}$, $\mathbf{X}$; model $\mathcal{M}$, draws $L$
Subset $\mathbf{y}$, $\mathbf{W}$, $\mathbf{X}$ based on $\mathcal{I}$, $\mathcal{J}$, $\mathcal{K}$
Compute $Y = \sum_{\mathcal{I},\mathcal{J},\mathcal{K}} y_{ijk}$
for $l$ in 1:$L$ do
    Initialize $Y^{(l)} = 0$
    Sample $\boldsymbol{\alpha}^{(l)}, \boldsymbol{\beta}^{(l)}, \boldsymbol{\theta}^{(l)}$ from model $\mathcal{M}$
    for $i, j, k$ in $\mathcal{I}, \mathcal{J}, \mathcal{K}$ do
        Apply $\psi_{ijk}^{(l)} = f^{-1}(\mathbf{x}'_{ijk}\boldsymbol{\beta}^{(l)} + \mathbf{m}'_i\boldsymbol{\theta}^{(l)})$, $p_{ijk}^{(l)} = f^{-1}(\mathbf{w}'_{ijk}\boldsymbol{\alpha}^{(l)})$
        Sample $y_{ijk}^{(l)} \sim \text{Bernoulli}(\psi_{ijk}^{(l)} \cdot p_{ijk}^{(l)})$
        Update $Y^{(l)} = Y^{(l)} + y_{ijk}^{(l)}$
    end for
end for
Compute $\bar{Y} = \frac{1}{L}\sum_{l=1}^{L} Y^{(l)}$
return $\bar{Y} \div Y$
```

4.2 Model building

Given that our model space was large, we chose to construct models by iteratively adding covariates with plausible scientific meaning. All model covariates were derived from the datasets described in Section 2. We applied root and log transformations to covariates with skewed distributions, including land cover proportions (Agri_i and Urban_i for agricultural and urban land cover), mosquito genus proportions (Culex_{ijk}), survey count (Surveys_{ijk}), catch rate (CatchRate_{ijk}), the yearly count of weeks 1 to 17 with mean temperature below zero Celsius (Freezing_{ij}), and water level (Water_{ijk}). To keep the models parsimonious, we did not explore variable interactions.

For occupancy modeling, standard practice is to put a covariate in either the occupancy or the detection component but not both; otherwise, the effect may oscillate between the two components and fail to converge. We assigned covariates from the trap data to the detection component and those from the environmental literature review to the occupancy component. While mosquito genus proportions and vector abundances could explain occupancy, these were influenced by surveillance decisions at the PHU level.

Initially, we fit some exploratory models to establish intuition for which covariates contribute the most to model improvements. Bioclimatic effects tended to have credible intervals covering zero and did not improve WAIC scores. At closer inspection, MERRAclim variables displayed little variation over the province. We focused on agricultural and urban land types with a downstream application in mind of relating WNV in mosquito populations to human WNV cases. These land types carry great significance as well because they are habitats frequented by *Culex* species (Gorris et al. 2021). Population density from the 2016 census was strongly correlated with urban land type, so we did not use it in our models. Survey count, *Culex* proportion, and catch rate all appeared useful in calibrating detection probabilities.

Based on exploratory model building, we arrived at a good null model without time-varying occupancy. This model improved on an intercept only model in both WAIC and relative presence.

$$\text{Detection}_{ijk} = \alpha_0 + \alpha_1\text{Surveys}_{ijk} + \alpha_2\text{Culex}_{ijk} + \alpha_3\text{CatchRate}_{ijk}$$
$$\text{Occupancy}_i = \beta_0 + \beta_1\text{Agri}_i + \beta_2\text{Urban}_i$$

In an iterative fashion, we added a seasonal covariate, freezing weeks (Freezing_{ij}), and period covariates for temperature (Temp_{ijk}), water level (Water_{ijk}), and Gini-Simpson diversity among birds (Bird_{ijk}) to the occupancy component. Since environmental conditions in winter and spring can impact mosquito development, we explored lags of 2, 4, 6, and 8 weeks for the period covariates, finding that six week lags substantially improved training and validation WAIC scores. Next, we built in epidemic behavior through a custom covariate, the two week lagged count of known infected neighboring sites, including oneself. That is, a site only contributes to the count if WNV was detected there in the previous epiweek.

$$\begin{aligned}\text{Occupancy}_{ijk} = {} & \beta_0 + \beta_1\text{Agri}_i + \beta_2\text{Urban}_i \\ & + \beta_3\text{Freezing}_{ij} + \beta_4\text{Temp}_{ijk} \\ & + \beta_5\text{Water}_{ijk} + \beta_6\text{Bird}_{ijk} + \beta_7\text{Neighbors}_{ijk}\end{aligned}$$

Finally, we included five spatial random effects in the occupancy component. Our spatial design $\mathbf{X}_s$ included agricultural and urban land types and site averages for freezing weeks, temperature, water level, and bird diversity. Smaller spatial designs resulted in strong posterior correlations between land type fixed effects and SREs. For each model, we checked trace plots and the Gelman-Rubin diagnostic (Gelman et al. 2013) to ensure convergence of all regression effects.

We report model comparisons in Table 1. In addition to WAIC scores, for the validation data we computed our custom posterior predictive check, relative presence, for all sites, for the GTA region, and for seven sites above the 46th latitude (northern Ontario). Temperature and infected neighboring sites made the biggest difference in the iterative model build, which supports WNV occupancy varying by period. The first SRE captured a residual effect of longitude whereas the other SREs were less interpretable. Otherwise, including SREs did not impact predictive performance or change the posterior inference of occupancy and detection effects (Appendix, Figure 5). All models had poor predictive performance for northern Ontario. We followed up this finding by fitting these models on northern Ontario data only. We found that all posterior effects overlapped with zero, indicating that these covariates did not explain the observed WNV cases. Based on these

Table 1 Iteratively built models with WAIC applied to the training dataset and relative presence diagnostics applied to the validation dataset

Model	WAIC	Relative Presence		
	Training	All Sites	GTA	Northern
Intercept only	3562.037	1.050	0.425	5.189
Base model	2794.778	1.026	0.935	0.796
+ freezing weeks	2776.312	1.010	0.935	0.921
+ temperature	2582.637	0.981	0.789	0.549
+ water level, bird diversity	2584.243	0.972	0.786	0.644
+ infected neighbors	2293.325	0.951	0.823	0.779
+ 5 SREs	2293.407	0.949	0.806	0.747

diagnostics and our scientific understanding of WNV transmission dynamics, we selected the final model with SREs as our best model. We refit this model with training and validation data for 3 chains of 12,000 iterations with 2,000 iterations of burn-in.

4.3 Model results

Our final model identified associations that have plausible scientific interpretations and have been previously reported in the literature (Table 2). Urban zones were more strongly associated with WNV occupancy than agricultural land, though both land types supported the pathogen. Warmer temperatures appeared to provide favorable environmental conditions for mosquitoes and WNV to thrive. Likewise, colder winters affected WNV occupancy, possibly by decreasing survival during the mosquitoes' overwintering period. Decreases in eBird species diversity indicated more occupancy. This negative correlation was also observed in a separate analysis by Allan et al. (2009) who surmised that more bird species with low reservoir competence may exist in diverse communities. More traps and more mosquitoes, especially *Culex* mosquitoes, increased virus detection.

Temperature and surveillance changes over the sampling season modified occupancy and detection. Figure 2 illustrates this pattern in the model predictions for a subset of PHUs in 2016. Heat waves in the early summer drove up occupancy rates for late summer. These seasonal peaks in occupancy aligned well with the pattern of human WNV cases from the Centers for Disease Control and Prevention (McDonald et al. 2021, Figure 1). More *Culex* mosquitoes were captured and tested in the late summer as well, increasing detection rates. Figure 2 displays temporal trends and uncertainty quantification, but it hides the inherently spatial nature of this model. Maps for modeled occupancy and detection in the 34th epiweek of 2016 capture this spatial behavior (Figures 3 and 4). Southeastern Ontario, especially GTA, had high occupancy rates. During the peak season, epiweeks 30 to 40, detection generally exceeded 0.50 in GTA and sometimes surpassed this mark in Ottawa (OTT), Windsor-Essex (WEC), and other PHUs. Animated maps for the entire 2016 sampling season are available online as Supplementary Materials.

We also assessed our model's posterior predictive performance for the test dataset (Table 3). For each PHU, we simulated 200 presences and averaged the results. PHUs

Table 2 Posterior effects for final model. The first eight rows are occupancy effects $\boldsymbol{\beta}$ and the last four rows are detection effects $\boldsymbol{\alpha}$

Covariate		Effect Quantiles			Gelman-Rubin
Variable	Transform	0.025	0.500	0.975	
Intercept		− 5.046	− 3.115	− 1.357	1.001
Agricultural	sqrt	− 0.141	0.613	1.419	1.002
Urban	sqrt	1.178	1.983	2.814	1.003
Freezing Weeks	sqrt	− 1.038	− 0.706	− 0.374	1.001
Temperature	lag 6	0.115	0.151	0.188	1.003
Water level	lag 6, sqrt	− 0.031	0.108	0.233	1.000
Bird Diversity	lag 6	− 1.360	− 0.276	0.847	1.001
Infected neighbors	lag 2	1.208	1.414	1.654	1.005
Intercept		− 10.100	− 9.004	− 7.982	1.001
Surveys	sqrt	0.440	0.557	0.682	1.001
Catch rate	cbrt	1.012	1.193	1.398	1.001
Culex proportion	sqrt	5.358	6.339	7.398	1.000

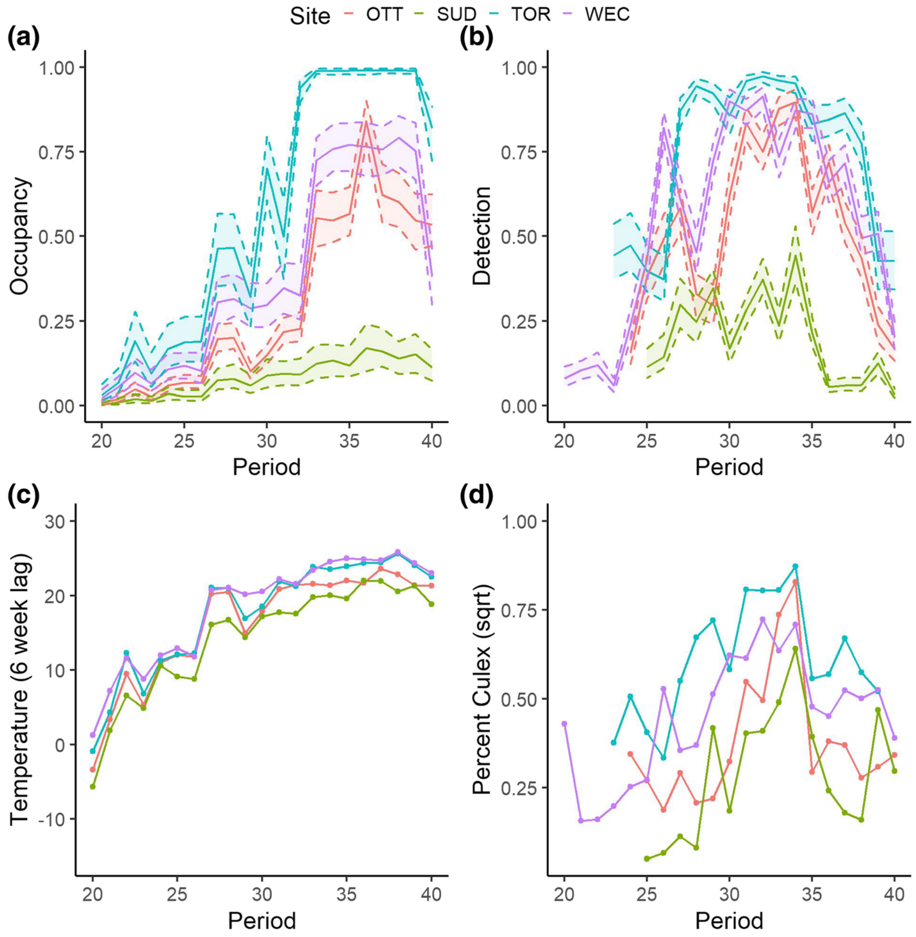

Fig. 2 (a) Occupancy and (b) detection probabilities for PHUs Ottawa (OTT), Sudbury (SUD), Toronto (TOR), and Windsor-Essex (WEC) in 2016. Solid lines are posterior medians, and bands are ninety-five percent credible intervals. (c) Lagged mean temperature and (d) *Culex* proportion are influential occupancy and detection covariates

with no presences were typically modeled to have fractional average presences. Meanwhile, PHUs with many presences were modeled to have many average presences. Metropolitan Toronto, one of the most surveyed PHUs, was remarkably well estimated by our method. In sum, our model simulated too few presences, especially for 2012; however, regression emphasizes mean behavior, and 2012 had the highest case counts of any season.

5 Discussion

We developed a new approach to solve binary regression problems in which the observed process is the product of a detection process and a latent binary process. Our method is in some sense a generalization of the single-season occupancy model (MacKenzie et al. 2002) but without the assumption of fixed occupancy. Our implementation relies on the state-of-the-art data augmentation strategy of Polson et al. (2013) to generate posterior samples in a Gibbs way. We suspect that these methods could be thoughtfully applied to analyze presence-absence data for other vector-borne pathogens in which survey detection is imperfect, e.g., the bacteria causing Lyme disease in ticks.

In an extended case study of West Nile virus in Ontario, Canada, we demonstrated that Bayesian time-varying occupancy models can attain good predictive performance and identify scientifically meaningful relationships between responses and covariates. Notably, we found strong evidence to suggest that warmer temperatures in general foster WNV spread in our study region. Careful data fusion with citizen science and other Internet data resources could provide value to future ecological and environmental studies. Our analysis also suggests that using traps to specifically capture *Culex* mosquitoes may assist in WNV surveillance efforts. Finally, our modeling is the first among studies on this dataset to examine mosquitoes and WNV across all of Ontario. Our efforts have laid groundwork in the pursuit of an omnibus model for the province.

Models for WNV occupancy in Ontario could be improved and extended in various ways. Compiling and integrating more covariate data of ecological importance may elucidate new associations. Normalized difference vegetation index (NDVI), a greenness measurement from

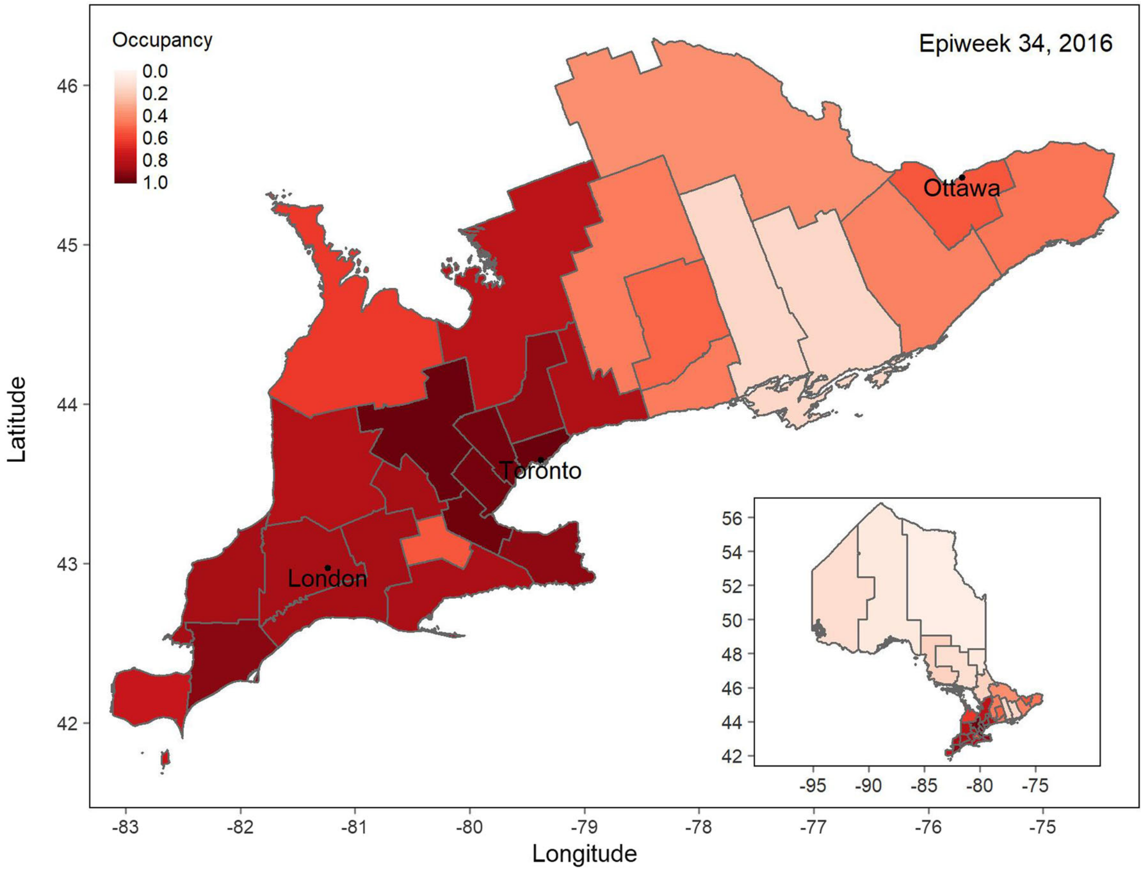

Fig. 3 Posterior medians for occupancy probabilities in Ontario on the 34th epiweek of 2016. The inset map juxtaposes southern Ontario relative to northern Ontario

remote sensing, could indicate microchanges to land type at the temporal scale of days/weeks (DeMets et al. 2020a, b). Another useful predictor would be based on standing water and other basins in which mosquitoes breed, given that water level demonstrated a suggestive direction of effect (Shutt et al. 2021). Similar to freezing weeks, a seasonal covariates for precipitation could investigate the effects of droughts and flooding. On the other hand, we could study WNV occupancy at a finer spatial mesh. From the onset, we aggregated to 34 areal sites, some as large as states. While there are existing spatial methods to analyze coordinate locations (Yue and Speckman 2010), working with site locations as points would be challenging, both to collect covariate data and address sparsity concerns. SREs may also be important to account for spatial confounding. Such a model could provide locally informative updates on epidemics.

From a policy standpoint, we require statistical analysis like that presented in this paper to inform public health efforts and to start to quantify the relative impacts of time-varying occupancy versus time-varying sampling. Under a changing climate, vector-borne pathogens have been colonizing new regions and increasing in frequency, again highlighting the utility of being able to predict viral occupancy in systems that are seasonal. Our methods can be implemented on a personal laptop and employed in real-time to accommodate epidemic responses. With adequate vector surveillance, public health officials can now quantify the probability of occupancy and detection of an environmentally-driven pathogen and mitigate its spread in

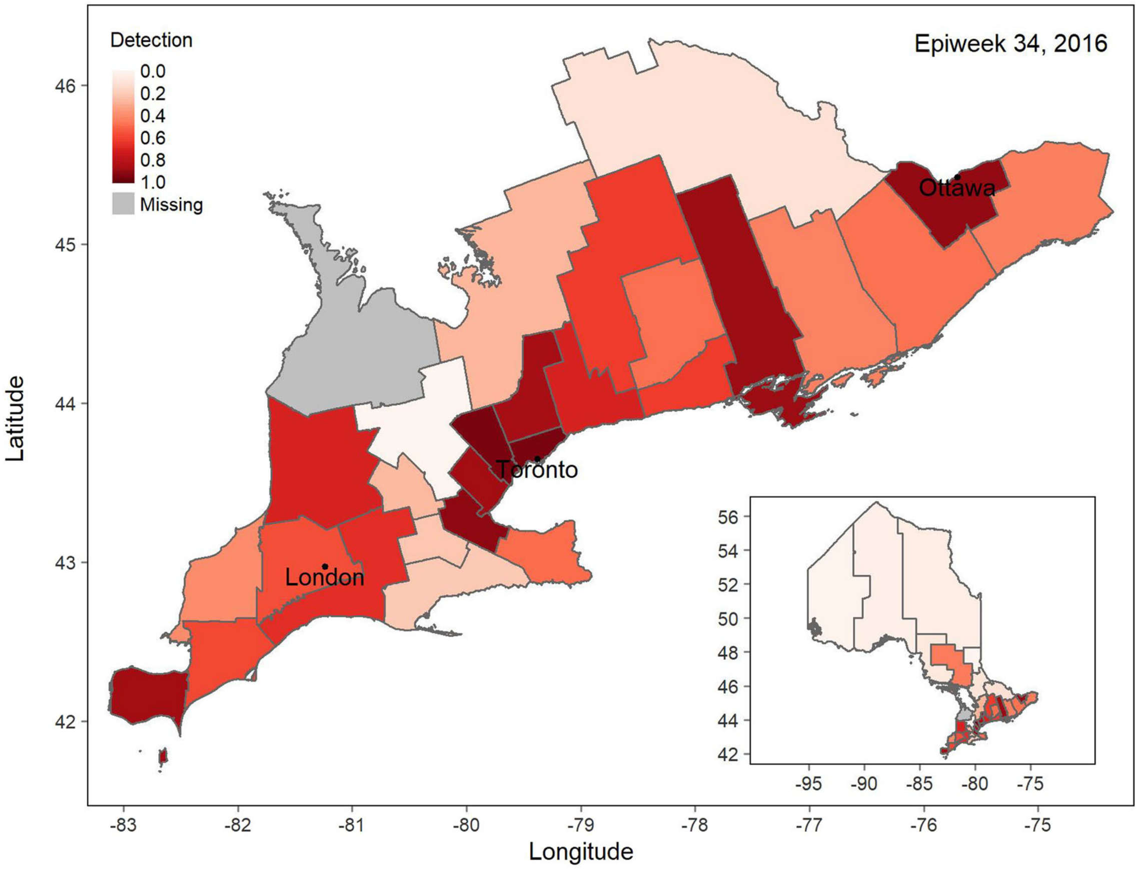

Fig. 4 Posterior medians for detection probabilities in Ontario on the 34th epiweek of 2016. Detection was missing when no surveys were done. The inset map juxtaposes southern Ontario relative to northern Ontario

a timely manner. Our method can also be used to improve parameterization of other model types, such as differential equation and agent-based disease transmission models. By providing estimates of the probability through time that a virus is present and that it is detected, we can better calibrate the data-fitting process for spatio-temporal mechanistic models fit to mosquito virus data.

Table 3 Model average versus actual presences for test data

PHU	2008		2012		2016		Total	
	Model	Actual	Model	Actual	Model	Actual	Model	Actual
ALG	0.09	0	0.21	0	0.06	0	0.36	0
BRN	0.93	1	1.55	3	0.61	0	3.09	4
CHK	2.18	0	3.13	3	2.90	0	8.21	3
DUR	2.28	0	5.11	6	3.11	6	10.50	12
ELG	1.20	0	3.86	6	2.10	2	7.16	8
EOH	0.72	0	1.87	2	1.28	1	3.87	3
GBO	0.00	0	0.03	0	0.00	0	0.03	0
HAL	4.53	4	6.15	10	6.64	9	17.32	23
HAM	4.39	3	8.82	9	7.14	9	20.35	21
HDN	1.82	1	0.98	1	0.83	0	3.63	2
HKP	0.28	0	1.21	0	0.66	1	2.15	1
HPE	0.33	0	1.88	6	1.10	1	3.31	7
HUR	0.49	0	3.38	1	2.27	3	6.14	4
KFL	0.23	0	1.12	3	0.15	0	1.50	3
LAM	0.71	0	4.77	6	1.46	1	6.94	7
LGL	0.55	0	1.42	0	0.60	0	2.57	0
MSL	1.30	0	4.67	8	2.14	4	8.11	12
NIA	1.56	0	5.37	9	3.25	5	10.18	14
NPS	0.33	0	0.49	0	0.32	0	1.14	0
NWR	0.12	0	0.06	0	0.10	0	0.28	0
OTT	1.45	0	3.81	9	3.46	8	8.72	17
PEE	5.95	7	9.25	12	7.62	9	22.82	28
PQP	0.04	0	0.04	0	0.04	0	0.12	0
PTC	0.31	0	4.32	7	1.34	0	5.97	7
REN	0.21	0	0.61	0	0.26	0	1.08	0
SMD	1.35	0	3.22	2	1.42	2	5.99	4
SUD	0.57	0	0.29	1	0.26	0	1.12	1
THB	0.05	0	0.05	0	0.03	0	0.13	0
TOR	7.26	6	11.26	12	8.79	10	27.31	28
TSK	0.10	0	0.12	0	0.09	0	0.31	0
WAT	1.60	0	2.31	5	2.00	1	5.91	6
WDG	1.16	0	4.77	2	0.10	0	6.03	2
WEC	3.89	7	6.86	9	5.70	9	16.45	25
YRK	2.25	2	7.15	10	3.80	1	13.2	13
Total	50.23	31	110.14	142	71.63	82	232	255

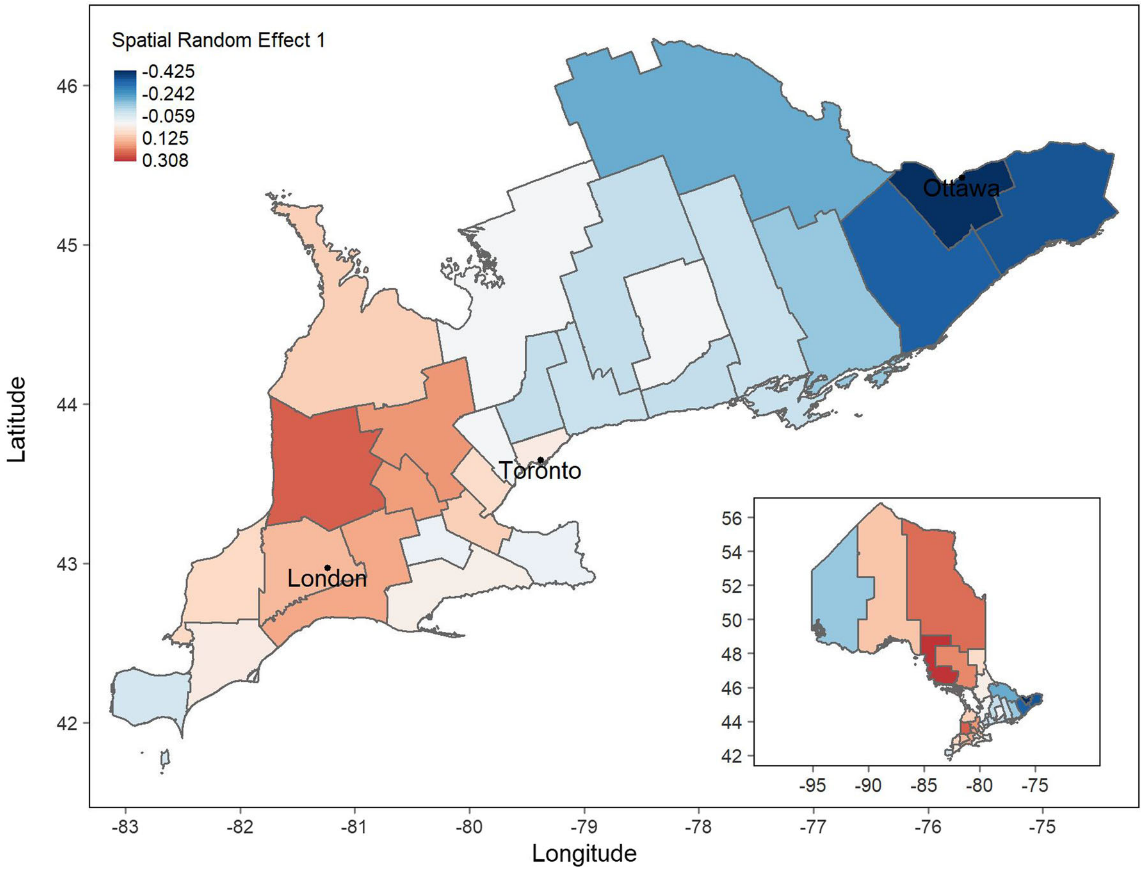

Fig. 5 Areal covariate corresponding to the first SRE. Negative (positive) values are in blue (red). The inset map juxtaposes southern Ontario relative to northern Ontario

Table 4 Posterior SREs $\boldsymbol{\theta}$ for final model

Variable	Transform	0.025	0.500	0.975	Gelman-Rubin
SRE 1		-0.114	-0.003	0.094	1.001
SRE 2		-0.061	0.004	0.095	1.001
SRE 3		-0.062	0.000	0.069	1.000
SRE 4		-0.062	0.002	0.077	1.000
SRE 5		-0.044	0.006	0.095	1.000

Appendix

Probit regression for time-varying occupancy

The data augmentation strategy of Albert and Chib (1993) is to draw continuous augmented variables from truncated-normal distributions and then use them to sample the regression effects. Intermediate mean $\mathbf{m}$ and covariance $\boldsymbol{\Sigma}$ terms are adjusted as well. Otherwise, the sampler, Algorithm 3, is similar to Algorithm 1.

Algorithm 3 Sample model effects for time-varying occupancy model (probit)

```
Require: data $\mathbf{y}$, $\mathbf{W}$, $\mathbf{X}$; priors $\boldsymbol{\mu_\alpha}$, $\boldsymbol{\mu_\alpha}$, $\nu_{\boldsymbol{\alpha}}$, $\nu_{\boldsymbol{\beta}}$, $\boldsymbol{\Lambda_\alpha}$, $\boldsymbol{\Lambda_\beta}$; iterations $T$
Initialize $\boldsymbol{\alpha}^{(0)} = \boldsymbol{\mu_\alpha}$, $\boldsymbol{\beta}^{(0)} = \boldsymbol{\mu_\beta}$, $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{(0)} = \boldsymbol{\Lambda_\alpha}$, $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(0)} = \boldsymbol{\Lambda_\beta}$
Compute $p_{ijk}^{(0)} = \Phi(\mathbf{w}'_{ijk}\boldsymbol{\alpha}^{(0)})$, $\psi_{ijk}^{(0)} = \Phi(\mathbf{x}'_{ijk}\boldsymbol{\beta}^{(0)})$ for all indices $i, j, k$
for t in 1:T do
    // occupancy component
    for all indices $i, j, k$ where $y_{ijk}$ non-missing do
        $z_{ijk}^{(t)} \sim \begin{cases} \text{Bernoulli}(1), & y_{ijk} = 1 \\ \text{Bernoulli}\left(\frac{\psi_{ijk}^{(t-1)}(1-p_{ijk}^{(t-1)})}{\psi_{ijk}^{(t-1)}(1-p_{ijk}^{(t-1)})+(1-\psi_{ijk}^{(t-1)})}\right), & y_{ijk} = 0 \end{cases}$
        $\tilde{z}_{ijk}^{(t)} \sim \begin{cases} \text{Truncated-Normal}(\mathbf{x}'_{ijk}\boldsymbol{\beta}^{(t-1)}, -\infty, 0), & z_{ijk}^{(t)} = 0 \\ \text{Truncated-Normal}(\mathbf{x}'_{ijk}\boldsymbol{\beta}^{(t-1)}, 0, \infty), & z_{ijk}^{(t)} = 1 \end{cases}$
        $\mathbf{V}_{\boldsymbol{\beta}}^{(t)} = ((\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(t-1)})^{-1} + \mathbf{X}'\mathbf{X})^{-1}$
        $\mathbf{m}_{\boldsymbol{\beta}}^{(t)} = \mathbf{V}_{\boldsymbol{\beta}}^{(t)}((\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(t-1)})^{-1}\boldsymbol{\mu_\beta} + \mathbf{X}'\tilde{\mathbf{z}}^{(t)})$
        $\boldsymbol{\beta}^{(t)} \sim \text{Normal}(\mathbf{m}_{\boldsymbol{\beta}}^{(t)}, \mathbf{V}_{\boldsymbol{\beta}}^{(t)})$
        $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(t)} \sim \text{Inverse-Wishart}(1 + \nu_{\boldsymbol{\beta}}, \boldsymbol{\Lambda_\beta} + (\boldsymbol{\beta}^{(t)})'\boldsymbol{\beta}^{(t)})$
        $\psi_{ijk}^{(t)} = \Phi(\mathbf{x}'_{ijk}\boldsymbol{\beta}^{(t)})$
    end for
    // detection component
    for all indices $i, j, k$ where $z_{ijk}^{(t)} = 1$ do
        $\tilde{y}_{ijk}^{(t)} \sim \begin{cases} \text{Truncated-Normal}(\mathbf{w}'_{ijk}\boldsymbol{\alpha}^{(t-1)}, -\infty, 0), & y_{ijk} = 0 \\ \text{Truncated-Normal}(\mathbf{w}'_{ijk}\boldsymbol{\alpha}^{(t-1)}, 0, \infty), & y_{ijk} = 1 \end{cases}$
        $\mathbf{V}_{\boldsymbol{\alpha}}^{(t)} = ((\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{(t-1)})^{-1} + \mathbf{W}'\mathbf{W})^{-1}$
        $\mathbf{m}_{\boldsymbol{\alpha}}^{(t)} = \mathbf{V}_{\boldsymbol{\alpha}}^{(t)}((\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{(t-1)})^{-1}\boldsymbol{\mu_\alpha} + \mathbf{W}'\tilde{\mathbf{y}}^{(t)})$
        $\boldsymbol{\alpha}^{(t)} \sim \text{Normal}(\mathbf{m}_{\boldsymbol{\alpha}}^{(t)}, \mathbf{V}_{\boldsymbol{\alpha}}^{(t)})$
        $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{(t)} \sim \text{Inverse-Wishart}(1 + \nu_{\boldsymbol{\alpha}}, \boldsymbol{\Lambda_\alpha} + (\boldsymbol{\alpha}^{(t)})'\boldsymbol{\alpha}^{(t)})$
        $p_{ijk}^{(t)} = \Phi(\mathbf{w}'_{ijk}\boldsymbol{\alpha}^{(t)})$
    end for
end for
return $\boldsymbol{\alpha}^{(1:T)}, \boldsymbol{\beta}^{(1:T)}, \boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{(1:T)}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(1:T)}$
```

Spatial logistic regression for time-varying occupancy

The Gibbs sampler for Bayesian logistic regression with SREs requires modifications to a PG parameter and $\mathbf{m}$ and $\mathbf{V}$ terms. We also have to sample from a MVN for the SREs $\boldsymbol{\theta}$ and from an IG for the variance parameter $\sigma_{\boldsymbol{\theta}}^2$. Otherwise, the sampler is as in the main article. Below we show these updates to the occupancy component step in Algorithm 1.

$$z_{ijk}^{(t)} \sim \begin{cases} \text{Bernoulli}(1), & y_{ijk} = 1 \\ \text{Bernoulli}\left(\dfrac{\psi_{ijk}^{(t-1)}(1-p_{ijk}^{(t-1)})}{\psi_{ijk}^{(t-1)}(1-p_{ijk}^{(t-1)})+(1-\psi_{ijk}^{(t-1)})}\right), & y_{ijk} = 0 \end{cases}$$

$$\tilde{z}_{ijk}^{(t)} \sim \text{Pólya-Gamma}(1, \mathbf{x}'_{ijk}\boldsymbol{\beta}^{(t-1)} + \mathbf{m}'_i\boldsymbol{\theta}^{(t-1)})$$

$$\mathbf{V}_{\boldsymbol{\beta}}^{(t)} = ((\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(t-1)})^{-1} + \mathbf{X}'\text{diag}(\tilde{\mathbf{z}}^{(t)})\mathbf{X})^{-1}$$

$$\begin{aligned}\mathbf{m}_{\boldsymbol{\beta}}^{(t)} = {} & \mathbf{V}_{\boldsymbol{\beta}}^{(t)}((\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(t-1)})^{-1}\boldsymbol{\mu_\beta} \\ & + \mathbf{X}'(\mathbf{z}^{(t)} - 1/2 \cdot \mathbf{1} - \text{diag}(\tilde{\mathbf{z}}^{(t)})\mathbf{M}\boldsymbol{\theta}^{(t-1)}))\end{aligned}$$

$$\boldsymbol{\beta}^{(t)} \sim \text{Normal}(\mathbf{m}_{\boldsymbol{\beta}}^{(t)}, \mathbf{V}_{\boldsymbol{\beta}}^{(t)})$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(t)} \sim \text{Inverse-Wishart}(1 + v_{\boldsymbol{\beta}}, \boldsymbol{\Lambda_\beta} + (\boldsymbol{\beta}^{(t)})'\boldsymbol{\beta}^{(t)})$$

$$\mathbf{V}_{\boldsymbol{\theta}}^{(t)} = ((\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(t-1)})^{-1} + \mathbf{M}'\text{diag}(\tilde{\mathbf{z}}^{(t)})\mathbf{M})^{-1}$$

$$\begin{aligned}\mathbf{m}_{\boldsymbol{\theta}}^{(t)} = {} & \mathbf{V}_{\boldsymbol{\theta}}^{(t)}((\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(t-1)})^{-1}\boldsymbol{\mu_\theta} \\ & + \mathbf{M}'(\mathbf{z}^{(t)} - 1/2 \cdot \mathbf{1} - \text{diag}(\tilde{\mathbf{z}}^{(t)})\mathbf{X}\boldsymbol{\beta}^{(t)}))\end{aligned}$$

$$\boldsymbol{\theta}^{(t)} \sim \text{Normal}(\mathbf{m}_{\boldsymbol{\theta}}^{(t)}, \mathbf{V}_{\boldsymbol{\theta}}^{(t)})$$

$$\begin{aligned}(\sigma_{\boldsymbol{\theta}}^2)^{(t)} \sim {} & \text{Inverse-Gamma}(a_{\boldsymbol{\theta}} + \text{length}(\boldsymbol{\theta})/2, b_{\boldsymbol{\theta}} \\ & + (\boldsymbol{\theta}^{(t)})'(\mathbf{M}'\mathbf{Q}\mathbf{M})^{-1}\boldsymbol{\theta}^{(t)}/2)\end{aligned}$$

$$\psi_{ijk}^{(t)} = \text{logit}^{-1}(\mathbf{x}'_{ijk}\boldsymbol{\beta}^{(t)} + \mathbf{m}_i\boldsymbol{\theta}^{(t)})$$

Spatial probit regression for time-varying occupancy

The Gibbs sampler for Bayesian probit regression with SREs requires similar modifications. Otherwise, the sampler is as in Algorithm 3. Below we show these updates to the occupancy component step.

$$z_{ijk}^{(t)} \sim \begin{cases} \text{Bernoulli}(1), & y_{ijk} = 1 \\ \text{Bernoulli}\left(\dfrac{\psi_{ijk}^{(t-1)}(1-p_{ijk}^{(t-1)})}{\psi_{ijk}^{(t-1)}(1-p_{ijk}^{(t-1)})+(1-\psi_{ijk}^{(t-1)})}\right), & y_{ijk} = 0 \end{cases}$$

$$\tilde{z}_{ijk}^{(t)} \sim \begin{cases} \text{Truncated-Normal}(\mathbf{x}'_{ijk}\boldsymbol{\beta}^{(t-1)} + \mathbf{m}'_i\boldsymbol{\theta}^{(t-1)}, -\infty, 0), & z_{ijk}^{(t)} = 0 \\ \text{Truncated-Normal}(\mathbf{x}'_{ijk}\boldsymbol{\beta}^{(t-1)} + \mathbf{m}'_i\boldsymbol{\theta}^{(t-1)}, 0, \infty), & z_{ijk}^{(t)} = 1 \end{cases}$$

$$\mathbf{V}_{\boldsymbol{\beta}}^{(t)} = ((\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(t-1)})^{-1} + \mathbf{X}'\mathbf{X})^{-1}$$

$$\mathbf{m}_{\boldsymbol{\beta}}^{(t)} = \mathbf{V}_{\boldsymbol{\beta}}^{(t)}((\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(t-1)})^{-1}\boldsymbol{\mu_\beta} + \mathbf{X}'(\tilde{\mathbf{z}}^{(t)} - \mathbf{M}\boldsymbol{\theta}^{(t-1)}))$$

$$\boldsymbol{\beta}^{(t)} \sim \text{Normal}(\mathbf{m}_{\boldsymbol{\beta}}^{(t)}, \mathbf{V}_{\boldsymbol{\beta}}^{(t)})$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{(t)} \sim \text{Inverse-Wishart}(1 + v_{\boldsymbol{\beta}}, \boldsymbol{\Lambda_\beta} + (\boldsymbol{\beta}^{(t)})'\boldsymbol{\beta}^{(t)})$$

$$\mathbf{V}_{\boldsymbol{\theta}}^{(t)} = ((\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(t-1)})^{-1} + \mathbf{M}'\mathbf{M})^{-1}$$

$$\mathbf{m}_{\boldsymbol{\theta}}^{(t)} = \mathbf{V}_{\boldsymbol{\theta}}^{(t)}((\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(t-1)})^{-1}\boldsymbol{\mu_\theta} + \mathbf{M}'(\tilde{\mathbf{z}}^{(t)} - \mathbf{X}\boldsymbol{\beta}^{(t)}))$$

$$\boldsymbol{\theta}^{(t)} \sim \text{Normal}(\mathbf{m}_{\boldsymbol{\theta}}^{(t)}, \mathbf{V}_{\boldsymbol{\theta}}^{(t)})$$

$$\begin{aligned}(\sigma_{\boldsymbol{\theta}}^2)^{(t)} \sim {} & \text{Inverse-Gamma}(a_{\boldsymbol{\theta}} + \text{length}(\boldsymbol{\theta})/2, b_{\boldsymbol{\theta}} \\ & + (\boldsymbol{\theta}^{(t)})'(\mathbf{M}'\mathbf{Q}\mathbf{M})^{-1}\boldsymbol{\theta}^{(t)}/2)\end{aligned}$$

$$\psi_{ijk}^{(t)} = \Phi(\mathbf{x}'_{ijk}\boldsymbol{\beta}^{(t)} + \mathbf{m}'_i\boldsymbol{\theta}^{(t)})$$

Spatial random effects

Figure 5 shows that the areal covariate corresponding to the first SRE uncovered a residual effect of longitude. Table 4 reports that all SREs had posterior quantiles tightly centered about zero.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00477-022-02257-4.

Acknowledgements We appreciate and thank Public Health of Ontario for sharing their mosquito data. We are also grateful to Morgan E. Gorris, Andrew W. Bartlow, and Deborah P. Shutt for their guidance and help in working with the mosquito data. The findings and conclusions in this report are those of the authors and do not necessarily reflect the official position of their affiliations, including Los Alamos National Laboratory (LANL) and University of Washington (UW).

Author Contributions Conceptualization: SDT, CAM, KAK; Methodology: SDT, KAK; Data retrieval and management: SDT, KAK; Statistical analysis: SDT, KAK; Computer programming: SDT; Writing-original draft: SDT; Writing -review and editing: SDT, CAM, KAK; Funding: CAM, KAK; Supervision: CAM, KAK.

Funding SDT gratefully acknowledges support from a National Defense Science and Engineering Graduate Fellowship. Funding sources played no role in the design, analysis, and presentation of this research project. 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-21-31764. 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.

Data Availability An R package and code for the analysis will be made available on GitHub. Interested parties can contact Public Health Ontario to request access to the mosquito data. Animated choropleth maps for the 2016 sampling season are available online as Supplementary Materials.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Albers SJ (2017) tidyhydat: Extract and tidy Canadian hydrometric data. J Open Source Softw 2(20):511. https://doi.org/10.21105/joss.00511

Albert JH, Chib S (1993) Bayesian analysis of binary and polychotomous response data. J Amer Stat Assoc 88(422):669–679. https://doi.org/10.1080/01621459.1993.10476321

Allan BF, Langerhans RB, Ryberg WA et al (2009) Ecological correlates of risk and incidence of West Nile virus in the United States. Oecologia 158(4):699–708. https://doi.org/10.1007/s00442-008-1169-9

Bartlow AW, Manore C, Xu C et al (2019) Forecasting zoonotic infectious disease response to climate change: Mosquito vectors and a changing environment. Vet Sci 6(2):40. https://doi.org/10.3390/vetsci6020040

Besag J, Kooperberg C (1995) On conditional and intrinsic autoregressions. Biometrika 82(4):733–746. https://doi.org/10.1093/biomet/82.4.733

Blei DM, Kucukelbir A, McAuliffe JD (2017) Variational inference: A review for statisticians. J Am Stat Assoc 112(518):859–877. https://doi.org/10.1080/01621459.2017.1285773

Ciota AT, Kramer LD (2013) Vector-virus interactions and transmission dynamics of West Nile virus. Viruses 5(12):3021–3047. https://doi.org/10.3390/v5123021

Clark AE, Altwegg R (2019) Efficient Bayesian analysis of occupancy models with logit link functions. Ecol Evol 9(2):756–768. https://doi.org/10.1002/ece3.4850

Darsie RF Jr, Ward RA (1981) Identification and geographical distribution of the mosquitoes of North America, north of Mexico. Tech. rep, Walter Reed Army Inst of Res Wash DC

DeMets S, Ziemann A, Manore C, et al (2020a) Improving mosquito population predictions in the Greater Toronto Area using remote sensing imagery. In: 2020 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI). IEEE, pp 78–81, https://doi.org/10.1109/SSIAI49293.2020.9094591

DeMets SA, Ziemann A, Manore C, et al (2020b) Too big, too small, or just right? The influence of multispectral image size on mosquito population predictions in the greater Toronto area. In: Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imagery XXVI, vol 11392. SPIE, pp 224–231, https://doi.org/10.1117/12.2558128

Dorazio RM, Rodriguez DT (2012) A Gibbs sampler for Bayesian analysis of site-occupancy data. Methods Ecol Evol 3(6):1093–1098. https://doi.org/10.1111/j.2041-210X.2012.00237.x

Dunnington D (2017) rclimateca: fetch climate data from Environment Canada

Gelman A, Carlin JB, Stern HS, et al (2013) Bayesian data analysis, 3rd edn. Chapman and Hall/CRC, https://doi.org/10.1201/b16018

Gelman A, Hwang J, Vehtari A (2014) Understanding predictive information criteria for Bayesian models. Stat Comput 24(6):997–1016. https://doi.org/10.1007/s11222-013-9416-2

Giordano BV, Turner KW, Hunter FF (2018) Geospatial analysis and seasonal distribution of West Nile virus vectors (Diptera: Culicidae) in southern Ontario, Canada. Int J Environ Res Public Health 15(4):614. https://doi.org/10.3390/ijerph15040614

Gorris ME, Bartlow AW, Temple SD et al (2021) Updated distribution maps of predominant Culex mosquitoes across the

Americas. Parasites & Vectors 14(1):1–13. https://doi.org/10.1186/s13071-021-05051-3

Hadfield J, Brito AF, Swetnam DM, et al (2019) Twenty years of West Nile virus spread and evolution in the Americas visualized by Nextstrain. PLOS Pathog 15(10) e1008,042. https://doi.org/10.1371/journal.ppat.1008042

Hoffman MD, Gelman A, Others, (2014) The No-U-Turn sampler: adaptively setting path lengths in hamiltonian monte carlo. J Mach Learn Res 15(1):1593–1623

Hooten MB, Hobbs NT (2015) A guide to Bayesian model selection for ecologists. Ecol Monogr 85(1):3–28. https://doi.org/10.1890/14-0661.1

Hughes J, Haran M (2013) Dimension reduction and alleviation of confounding for spatial generalized linear mixed models. J R Stat Soc Series B Stat Methodol 75(1):139–159. https://doi.org/10.1111/j.1467-9868.2012.01041.x

Johnson DS, Conn PB, Hooten MB et al (2013) Spatial occupancy models for large data sets. Ecology 94(4):801–808. https://doi.org/10.1890/12-0564.1

Kesavaraju B, Farajollahi A, Lampman RL et al (2012) Evaluation of a rapid analyte measurement platform for West Nile virus detection based on United States mosquito control programs. Amer J Tropical Med Hyg 87(2):359. https://doi.org/10.4269/ajtmh.2012.11-0662

MacKenzie DI, Nichols JD, Lachman GB et al (2002) Estimating site occupancy rates when detection probabilities are less than one. Ecology 83(8):2248–2255. https://doi.org/10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2

MacKenzie DI, Nichols JD, Hines JE et al (2003) Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. Ecology 84(8):2200–2207. https://doi.org/10.1890/02-3090

MacKenzie DI, Nichols JD, Royle JA et al (2017) Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Elsevier. https://doi.org/10.1016/C2012-0-01164-7

McDonald E, Mathis S, Martin SW, et al (2021) Surveillance for West Nile virus disease-United States, 2009–2018

Metropolis N, Rosenbluth AW, Rosenbluth MN et al (1953) Equation of state calculations by fast computing machines. J Chem Phys 21(6):1087–1092. https://doi.org/10.1063/1.1699114

Niemi J (2020) Package “MMWRweek”

Northrup JM, Gerber BD (2018) A comment on priors for Bayesian occupancy models. PLOS One 13(2):e0192,819. doi10.1371/journal.pone.0192819

Polson NG, Scott JG, Windle J (2013) Bayesian inference for logistic models using Pólya-Gamma latent variables. J Amer Stat Assoc 108(504):1339–1349. https://doi.org/10.1080/01621459.2013.829001

Polson NG, Scott JG, Windle J, et al (2019) Package “BayesLogit”

Royle JA, Dorazio RM (2006) Hierarchical models of animal abundance and occurrence. J Agric Biol Environ Stat 11(3):249–263. https://doi.org/10.1198/108571106X129153

Royle JA, Nichols JD (2003) Estimating abundance from repeated presence-absence data or point counts. Ecology 84(3):777–790. https://doi.org/10.1890/0012-9658(2003)084[0777:EAFRPA]2.0.CO;2

Rue H, Martino S, Chopin N (2009) Approximate Bayesian inference for latent Gaussian models by using integrated nested laplace approximations. J R Stat Soc Series B Stat Methodol 71(2):319–392. https://doi.org/10.1111/j.1467-9868.2008.00700.x

Shutt DP, Goodsman DW, Hemez ZJL, et al (2021) A process-based model with temperature, water, and lab-derived data improves predictions of daily mosquito density, https://doi.org/10.1101/2021.09.08.458905

Simpson EH (1949) Measurement of diversity. Nature 163(4148):688–688. https://doi.org/10.1038/163688a0

Sullivan BL, Wood CL, Iliff MJ et al (2009) eBird: a citizen-based bird observation network in the biological sciences. Biol Conserv 142(10):2282–2292. https://doi.org/10.1016/j.biocon.2009.05.006

Tuanmu MN, Jetz W (2014) A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. Glob Ecol Biogeogr 23(9):1031–1045. https://doi.org/10.1111/geb.12182

Turell MJ, Dohm DJ, Sardelis MR et al (2005) An update on the potential of North American mosquitoes (Diptera: Culicidae) to transmit West Nile virus. J Med Entomol 42(1):57–62. https://doi.org/10.1093/jmedent/42.1.57

Vega GC, Pertierra LR, Olalla-Tárraga MÁ (2017) MERRAclim, a high-resolution global dataset of remotely sensed bioclimatic variables for ecological modelling. Sci Data 4(1):1–12. https://doi.org/10.1038/sdata.2017.78

Wang J, Ogden NH, Zhu H (2011) The impact of weather conditions on Culex pipiens and Culex restuans (Diptera: Culicidae) abundance: a case study in Peel region. J Med Entomol 48(2):468–475. https://doi.org/10.1603/ME10117

Watanabe S (2013) A widely applicable Bayesian information criterion. J Mach Learn Res 14(Mar):867–897

Willis AD, Martin BD (2020) Estimating diversity in networked ecological communities. Biostatistics. https://doi.org/10.1093/biostatistics/kxaa015

Yoo EH (2014) Site-specific prediction of West Nile virus mosquito abundance in Greater Toronto Area using generalized linear mixed models. Int J Geogr Inf Sci 28(2):296–313. https://doi.org/10.1080/13658816.2013.837909

Yoo EH, Chen D, Diao C et al (2016) The effects of weather and environmental factors on West Nile virus mosquito abundance in Greater Toronto Area. Earth Interactions 20(3):1–22. https://doi.org/10.1175/EI-D-15-0003.1

Yue Y, Speckman PL (2010) Nonstationary spatial Gaussian Markov random fields. J Comput Graph Stat 19(1):96–116. https://doi.org/10.1198/jcgs.2009.08124

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.