Supplementary Material - A Poisson-multinomial spatial model for simultaneous outbreaks with application to arboviral diseases

Alexandra M. Schmidt∗1, Laís P. Freitas2, Oswaldo G. Cruz3, and Marilia S. Carvalho3

1Department of Epidemiology, Biostatistics and Occupational Health, McGill University
2Programa de Pós-Graduação em Epidemiologia em Saúde Pública, Escola Nacional de Saúde Pública Sergio Arouca (ENSP), Fundação Oswaldo Cruz, Rio de Janeiro, Brazil
3Programa de Computação Científica, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

This supplemental material is organized as follows. Section A describes the posterior full conditionals of the parameters described in model M4 of the main text. Section B provides a brief overview of the model comparison criteria used in the main paper. Section C provides further details about the convergence of the MCMC algorithm and additional results of some of the components of the model. Finally, Section D shows the results of different artificial data analyses used to investigate if the components of the proposed model in Section 2 of the paper are identifiable.

∗Corresponding author: Alexandra M. Schmidt, Department of Epidemiology, Biostatistics and Occupational Health, McGill University, 2001 McGill College, Suite 1200, H3A 1G1, Montreal, QC, Canada. E-mail: alexandra.schmidt@mcgill.ca. Homepage: http://alex-schmidt.research.mcgill.ca/ .

A Posterior full conditional distributions under the proposed model

This section describes the resultant posterior full conditional distributions based on model M4 as described in Section 2 of the main text. Let $\Theta = (\beta, \alpha, \theta, \phi, \Sigma)'$ be the parameter vector of the model. Assuming that there are observations available of the number of cases of each of the diseases in neighborhood $i$, $y_i = (y_1, y_2, y_3)'_i$; assuming $total_i = \sum_{k=1}^3 y_{ik}$, and defining $\text{tot} = (total_1, \ldots, total_n)'$ and $y = (y_1, \ldots, y_n)'$, the likelihood function is given by

$$p(\text{tot}, y \mid \Theta) \propto \prod_{i=1}^n \exp\{-\exp[(x_i' \beta + \theta_i_1)]\} \exp[total_i(x_i' \beta + \theta_i_1)]$$

$$\times \left[ \frac{1}{1 + \sum_{k=2}^3 \exp(x_i' \alpha_k + \theta_i_k)} \right]^{y_{i1}} \left[ \frac{\exp(x_i' \alpha_2 + \theta_i_2)}{1 + \sum_{k=2}^3 \exp(x_i' \alpha_k + \theta_i_k)} \right]^{y_{i2}}$$

$$\times \left[ \frac{\exp(x_i' \alpha_3 + \theta_i_3)}{1 + \sum_{k=2}^3 \exp(x_i' \alpha_k + \theta_i_k)} \right]^{y_{i3}}$$

$$= \prod_{i=1}^n \exp\{-\exp[(x_i' \beta + \theta_i_1)]\} \left[ \frac{1}{1 + \sum_{k=2}^3 \exp(x_i' \alpha_k + \theta_i_k)} \right]^{total_i}$$

$$\times \prod_{k=2}^3 \exp[y_{ik}(x_i' (\alpha_k + \beta) + \theta_i_k + \theta_i_1)].$$

The prior specification is as follows, for $\beta$ and $\alpha$ we assign independent zero mean normal prior distributions with some reasonably large variance. The components of the vector $\phi$ follow a conditional autoregressive prior distribution. For the random effects $\theta_i$, $i = 1, 2, \cdots, n$ we assign conditionally independent multivariate normal prior distributions with mean $\phi_i \mathbf{1}_3$ and covariance matrix $\Sigma$. For the covariance matrix $\Sigma$ we assign an inverse-Wishart prior distribution with $n_0$ degrees of freedom and an identity scale matrix. Therefore, the kernel of the prior distribution is given by

$$p(\Theta) = \left[ \prod_{i=1}^3 N(\beta_i; 0, C_0) \right] \left[ \prod_{i=1}^3 N(\alpha_i; 0, C_0) \right] \prod_{i=1}^n N \left( \phi_i; \sum_{j \in E_i} \phi_j, \frac{\sigma^2}{n_i} \right) N_3(\theta_i; \phi_i \mathbf{1}_3, \Sigma)$$

$$\times \text{IW}(\Sigma; n_0, \mathbf{1}_3) HC(\sigma^2; 0, 1),$$

where $N(x; b, B)$ stands for the pdf of the normal distribution with mean $b$ and variance $B$ evaluated at $x$, $\text{IW}(C; n_0, S)$ stands for the pdf of the inverse Wishart distribution
with \( n_0 \) degrees of freedom, scale covariance matrix \( S \) evaluated at covariance matrix \( C \), and \( HC(\sigma^2; 0, 1) \) stands for the half-Cauchy pdf with location 0, scale 1 evaluated at \( \sigma^2 \).

Following Bayes’ theorem, the posterior distribution is proportional to

\[
p(\Theta \mid y) \propto \prod_{i=1}^{n} \exp\{-\exp[(x_i'\beta + \theta_{i1})]\} \left[ \frac{1}{1 + \sum_{k=2}^{3} \exp(x_i'\alpha_k + \theta_{ik})} \right]^{\text{total}_i} \\
\times \prod_{k=2}^{3} \exp[y_{ik}(x_i'\alpha_k + \beta) + \theta_{ik} + \theta_{i1}] \\
\times \left[ \prod_{i=1}^{3} N(\beta_i; 0, C_0) \right] \left[ \prod_{i=1}^{3} N(\alpha_i; 0, C_0) \right] \\
\times \prod_{i=1}^{n} \left[ N \left( \phi_i; \frac{\sum_{j \in d_i} \phi_j - \sigma^2 n_i}{n_i} \right) \right] N_3 \left( \theta_i; \phi_i 1_3, \Sigma \right) \\
\times IW(\Sigma; n_0, I_3) HC(\sigma^2; 0, 1).
\]

The distribution above does not have a closed form. Samples from this distribution can be obtained through Markov chain Monte Carlo methods. See e.g. Gamerman and Lopes (2006) for a review. In particular, one possible algorithm is to use the Gibbs sampling with some steps of the Metropolis-Hastings algorithm. This is performed by sampling from the posterior full conditional distributions that result from the kernel of the posterior distribution. Next we describe the kernel of the posterior full conditionals of each of the parameters.

Let \( p(a \mid .) \) denote the posterior full conditional of the parameter \( a \) given all the other parameters.

**Posterior full conditional of \( \beta \)** The posterior full conditional of the fixed effects \( \beta \) in the log-relative risk of the total cases of vector borne diseases is proportional to

\[
p(\beta \mid .) \propto \prod_{i=1}^{n} \exp\{-\exp[(x_i'\beta + \theta_{i1})]\} \prod_{k=2}^{3} \exp[y_{ik}(x_i'\alpha_k + \beta) + \theta_{ik} + \theta_{i1}] \\
\times \left[ \prod_{i=1}^{3} N(\beta_i; 0, C_0) \right],
\]

which does not have a closed form. A Metropolis-Hastings step with a multivariate normal proposal distribution whose mean is based on the current value of \( \beta \) and some
tunned covariance matrix can be used to sample from this distribution.

**Posterior full conditional of** $\alpha$  
Let $\alpha = (\alpha_2, \alpha_3)'$, then

$$p(\alpha | .) \propto \frac{1}{1 + \sum_{k=2}^{3} \exp(x_k'\alpha_k + \theta_{ik})} \prod_{k=2}^{3} \exp[y_{ik}(x_k'(\alpha_k + \beta) + \theta_{ik} + \theta_{i1})]$$

$$\times \prod_{k=2}^{3} N(\alpha_k; 0, C_0),$$

which does not have a closed form. A Metropolis-Hastings step with a multivariate normal proposal distribution with mean based on the current value of $\alpha$ and some tunning covariance matrix can be used to sample from this distribution.

**Posterior full conditional of** $\theta_i$  
For each neighborhood $i$, the posterior full conditional of $\theta_i = (\theta_{i1}, \theta_{i2}, \theta_{i3})'$ is proportional to

$$p(\theta_i | .) \propto \exp\{-\exp[(x_i'\beta + \theta_{i1})]\} \left[\frac{1}{1 + \sum_{k=2}^{3} \exp(x_k'\alpha_k + \theta_{ik})}\right]^{total_i} \prod_{k=2}^{3} \exp[y_{ik}(x_k'(\alpha_k + \beta) + \theta_{ik} + \theta_{i1})] N_3(\theta_i; \phi_i1_3, \Sigma),$$

which does not have a closed form. A Metropolis-Hastings step with a multivariate normal proposal distribution with mean equal to the current value of $\theta_i$ and covariance matrix tunned to provide reasonable acceptance rates can be used to sample from this distribution.

**Posterior full conditional of** $\phi_i$  
For each neighborhood $i$, the posterior full conditional of the common spatial effect $\phi_i$ is proportional to

$$p(\phi_i | .) \propto N_3(\theta_i; \phi_i1_3, \Sigma) N\left(\phi_i; \frac{\sum_{j \in \delta_i} \phi_j}{n_i}, \frac{\sigma^2}{n_i}\right),$$

which, for each $\phi_i$, is the kernel of a univariate normal distribution with variance $\sigma^2_{\phi_i} = \left(1_3'\Sigma^{-1}1_3 + \frac{n_i}{\sigma^2}\right)^{-1}$ and mean $\mu_{\phi_i} = \sigma^2_{\phi_i} \left(1_3'\Sigma^{-1}\theta_i + \frac{1}{\sigma^2} \sum_{j \in \delta_i} \phi_j\right)$. As the conditional
autoregressive prior distribution is improper, to make sure the posterior distribution is proper, after sampling each of the $\phi_i$, $i = 1, 2, \cdots, n$ from the normal distribution above, a sum-to-zero constrain is imposed by centering them at each iteration of the MCMC.

**Posterior full conditional of $\Sigma$** The posterior full conditional of $\Sigma$ is proportional to

$$p(\Sigma \mid \cdot) \propto \prod_{i=1}^{n} N_3(\theta_i; \phi_i 1_3, \Sigma), IW(\Sigma; n_0, I_3),$$

which follows an Inverse Wishart distribution with $n_0 + n$ degrees of freedom, and scale covariance matrix equals $I_3 + \sum_{i=1}^{n} (\theta_i - 1_3 \phi_i)(\theta_i - 1_3 \phi_i)'$.

**Posterior full conditional of $\sigma^2$**

$$p(\sigma^2 \mid \cdot) \propto \prod_{i=1}^{n} N\left(\phi_i; \frac{\sum_{j \in \delta_i} \phi_j}{n_i}, \frac{\sigma^2}{n_i}\right) HC(\sigma^2; 0, 1),$$

which does not have a closed form. One can use a Metropolis-Hastings step with a lognormal proposal whose mean is based on the logarithm of the current value and some tuned variance.

To make the implementation of the MCMC easier, instead of coding ourselves the algorithm to sample from the posterior full conditionals above, we decided to use the software **Stan** (Carpenter et al., 2017) within the package **RStan** in **R** (R Core Team, 2020). The codes for the different models, together with an artificial dataset, are available from [https://github.com/laispfreitas/joint_DZC_model](https://github.com/laispfreitas/joint_DZC_model).

## B Model comparison criteria

We explore three different model comparison criteria, the Widely Available Information criterion (WAIC) (Watanabe, 2010), the logarithmic score (Czado et al., 2009) and the energy score (Gneiting et al., 2008). The following subsections describe the three criteria.
B.1 Widely Available (Watanabi-Akaike) information criterion

Watanabe (2010) proposed the widely available information criterion (WAIC). Different from DIC, WAIC averages over the posterior distribution rather than conditioning on a point estimate. As described in Gelman et al. (2014), the criterion is computed as

$$\text{WAIC} = -2(l_{ppd} - p_{WAIC}),$$ \hspace{1cm} (1)

where $l_{ppd}$ is the log pointwise predictive density, which measures the quality of the model fitting, and is computed as

$$\sum_{i=1}^{n} \log \left( \frac{1}{L} \sum_{s=1}^{L} l(y_i; \Theta^s) \right),$$ \hspace{1cm} (2)

with $\Theta^s$, denoting the $s$-th sampled value from the posterior distribution, $s = 1, \cdots, L$. The effective number of parameters is computed as (Gelman et al., 2014)

$$p_{WAIC} = \sum_{i=1}^{n} V_{s=1}^{L} (\log l(y_i; \Theta^s)),$$ \hspace{1cm} (3)

with $V_{s=1}^{L}(\cdot)$ corresponding to the sample variance. Smaller values of WAIC indicate better fitting models.

B.2 Logarithmic score (LogS)

The logarithmic score, which is a proper scoring rule (Gneiting et al., 2008; Czado et al., 2009), is defined as $-\log p(total_i, y_i)$, where $p(total_i, y_i)$ is the probability mass at neighbourhood $i$.

Assuming there is a sample from the posterior distribution of the parameters of size $L$ available, the predictive distribution $p(total_i, y_i)$ is approximated using Monte Carlo integration, that is,

$$p(total_i, y(s_i)) \approx \frac{1}{L} \sum_{i=1}^{L} p(total_i, y_i) \mid \Theta^{(t)}),$$

where $p(total_i, y_i \mid \Theta^{(t)})$ is the joint probability function of the Poisson and multinomial distributions conditioned on the $t^{th}$ sampled value of the parameter vector $\theta$, evaluated
at \( total_i \) and \( y_i \), respectively. Then the logarithmic score is computed as

\[
\text{Log} S = \frac{1}{n} \sum_{i=1}^{n} - \log p(\text{total}_i, y_i)
\]

Smaller values indicate better fitting models.

### B.3 Energy score

The energy score (es) (Gneiting et al., 2008) is a multivariate generalization of the continuous ranked probability score (crps) (Gneiting and Raftery, 2007). The energy score is computed as

\[
es(P, y) = E_P ||Y - y|| - \frac{1}{2} E_P ||Y - Y'||,
\]

where \( ||.|| \) denotes the Euclidean norm and \( Y \) and \( Y' \) are independent random vectors with distribution \( P \) and \( y \) is the observed value. When a sample from the posterior distribution of the parameter vector is available, Gneiting et al. (2008) suggest to approximate (4) via

\[
es(P, y) = \frac{1}{L} \sum_{l=1}^{L} ||y_l - y|| - \frac{1}{2(L-1)} ||y_l - y_{l+1}||,
\]

where \( y_1, \ldots, y_L \) is a sample from the posterior predictive density. As we have \( n = 160 \) neighborhoods, the values shown in Table 1 of the main text are based on an average of \( \hat{es}(P, y) \). Smaller values indicate better fitting models. Note that as the Poisson-Multinomial model involves the total cases of vector borne diseases, to make the Poisson-Multinomial and multivariate Poisson models comparable, we computed the energy score based on the vector of cases of dengue, Zika and chikungunya; that is, the total cases of vector borne diseases was not involved in the computation of the energy score when computed for the Poisson-Multinomial model.
C  Further results of the analysis in Section 3 of the main paper

This Section provides some further results associated with the analysis described in Section 3 of the main text.

C.1  Convergence diagnostics

This Section provides some of the diagnostic tools used to check convergence of the chains. As mentioned in the main text, we run three independent chains for each of the models, considered a total of 10,000 iterations, discarded the first 3,000 and stored every seventh sampled value. Panels of Figure 1 show the traceplots of the fixed effects included in the Poisson-Multinomial model under parametrization M4. We checked traceplots for all the other parameters and they all look similar.

Figure 1: Traceplots, based on three independent chains, of the sampled values of the fixed effects (see equations (4) and (5) of the main text) of the Poisson-Multinomial model under model M4.

Following Vehtari et al. (2021), convergence was also checked through the effective sample size (n_eff) and Rhat. The literature suggests that values of Rhat smaller than 1.01 provide indication of convergence of the chains. Table 1 provides a summary of
the magnitude of \( n_{\text{eff}} \) and \( Rhat \) based on the resultant samples from all parameters involved in the model. Clearly, the smallest \( n_{\text{eff}} \) is greater than 1,000, and \( Rhat \) has maximum value equals 1.0073, suggesting that the sampled values can be considered an approximation of the target distribution of interest. Although we are showing these values only for one model, all the others resulted in similar values.

| Diagnostic | Min. | 1st Qu. | Median | Mean  | 3rd Qu. | Max. |
|------------|------|---------|--------|-------|---------|------|
| \( n_{\text{eff}} \) | 1114 | 1385 | 1448 | 1447 | 1515 | 1827 |
| \( Rhat \) | 0.9981 | 0.9990 | 0.9996 | 0.9998 | 1.0003 | 1.0073 |

Table 1: Summary of the effective sample size \( (n_{\text{eff}}) \) and \( Rhat \) criteria to check convergence of the chains. These summaries are based on all parameters estimated in the Poisson-Multinomial model under the M4 parametrization. Other fitted models provided similar results to these ones.

C.2 Posterior mean of the random effects \( \theta_{ik} \)

The columns of Figure 2 show the posterior mean \( \theta_{ik}, i = 1, 2, \ldots, n \) and \( k = 1, 2, 3 \) under the Poisson-Multinomial and multivariate Poisson parametrizations under models M4 and M3, respectively. Note that these maps are not comparable, as these latent effects have different interpretations under each of the parametrizations of \( \lambda_i \) and \( \pi_{ik} \) (see equations (4)-(5) and (11) of the main text). The first column shows the posterior mean of \( \theta_{ik} \) under the Poisson model. These maps show how the latent effects adjust after accounting for the covariates in equations (4)-(5) (maps on the first column) and (10) (maps on the second column). Focusing first on the maps on the left column, the random effect in the equation of the total was estimated at higher values on the northern-eastern portion of the city. There is a different spatial pattern between the estimated random effects in the equation for “Zika-dengue” and “chik.-dengue”. Clearly, for the equation “Zika-dengue” the random effect assumed higher values on the western portion of the city, suggesting that Zika was more spread across the city than chikungunya, when compared to dengue. Focusing on the maps on the left hand side we obtain the usual
interpretation from Poisson models, each map is showing how the random effect in the log-relative risk of each disease adjusted after accounting for the available covariates. While dengue and chikungunya resulted in high, positive values of the random effects mostly on the south-eastern region of the city, the random effect in the log-relative risk for Zika also assumed positive values in the west region of the city. Suggesting that during the observed period the disease had spread across the whole city.

![Maps showing posterior mean of random effects](image)

**Figure 2:** Posterior mean of the random effects $\theta_{ik}, \ i = 1, 2, \ldots, n$ and $k = 1, 2, 3$ under model M4 in the Multinomial-Poisson (equations (4)-(5), left column) and model M3 in the multivariate Poisson parametrization (equation (10), right column).
D Artificial Data Analysis

Data generating mechanism  Using the neighborhoods of Rio de Janeiro as our geographical region and the available covariates, we investigated if we were able to recover the values of the parameters used to generate data from different model specifications. The general structure of the model used to generate the artificial data is as follows:

\[
\begin{align*}
total_i \mid \lambda_i & \sim Poi(E_i \lambda_i) \\
y_i \mid \pi_i & \sim Multinomial(total_i, \pi_i), \quad \text{for } i = 1, 2, \cdots, n,
\end{align*}
\]

where \(E_i\) was computed as described in the beginning of Section 2. We consider three sets of artificial data. The first assumes \(\theta_{ik} = 0, \forall i = 1, 2, \cdots, n \) and \(k = 1, 2, 3\). This is to check that we are able to recover the coefficients used to generate the data. The second set of artificial data is generated from the Poisson-multinomial equations (4)-(5) with random effects \(\theta_{ik}\) fixed at the estimated values from the available data under model M4 of the main text. And the third set of artificial data is generated from the Poisson-Multinomial model assuming the estimates obtained under model M5 for the real data.

D.1 Data generated from model without random effects

Initially we fixed \(\beta\) and \(\alpha_k, k = 2, 3\) at the respective estimated values obtained from the real data analysis. Then we generated \(L = 500\) datasets, from the Poisson-multinomial model with parameters:

\[
\begin{align*}
\log \lambda_i &= X_i' \beta \quad \text{and} \\
\log \left( \frac{\pi_{ik}}{\pi_{i1}} \right) &= X_i' \alpha_k, \quad k = 2, 3.
\end{align*}
\]

Next, for each of the \(L = 500\) datasets, we fitted the same model used to generate the datasets to check if we were able to recover the values of the fixed effects. Table 2 shows the bias and root mean squared errors (rMSE) in the estimation of the fixed effects. Clearly we are able to recover the true values of the parameters well as the resultant bias for the different fixed effects are negligible.
Equations associated with each of the coefficients

| Covariate       | Total       | Zika-dengue | chik.-dengue |
|-----------------|-------------|-------------|--------------|
| Intercept       | $8.128 \times 10^{-7}$ | $2.311 \times 10^{-4}$ | $1.038 \times 10^{-4}$ |
|                 | (0.0025)    | (0.0145)    | (0.0078)     |
| SDI             | $7.066 \times 10^{-5}$ | $-1.083 \times 10^{-4}$ | $-9.188 \times 10^{-4}$ |
|                 | (0.0029)    | (0.0066)    | (0.0304)     |
| Pct green area  | $1.996 \times 10^{-4}$ | $-9.059 \times 10^{-5}$ | $2.482 \times 10^{-4}$ |
|                 | (0.0031)    | (0.0198)    | (0.0091)     |
| Pop. density    | $5.656 \times 10^{-5}$ | $8.7767 \times 10^{-5}$ | $2.478 \times 10^{-4}$ |
|                 | (0.0037)    | (0.0080)    | (0.0112)     |

Table 2: Bias, together with the root mean squared error (in brackets), in the estimation of the fixed effects in equation (4) for the relative risk of the total number of cases, and for the odds ratio of a neighborhood having Zika or chikungunya in comparison to dengue (see equation (5)).

D.2 Data generated from Model M4 in the main text

In this Section we generate data from model M4 in the main text. More specifically, we assume

$$
total_i \mid \lambda_i \sim \text{Poi}(E_i \lambda_i) \tag{7}
$$

$$
y_i \mid total_i, \pi_i \sim \text{Multinomial}(total_i, \pi_i), \text{ for } i = 1, 2, \cdots, n,
$$

where $E_i$ was computed as described in the beginning of Section 2. The parameters $\lambda_i$ were fixed according to equation (4), and each $\pi_{ik}$ was fixed according to equation (5). The values of $\theta_{ik}$ were fixed at the posterior mean of the estimated values obtained from the data analyzed in the manuscript. The covariates were standardized and the fixed effects used to generate the data are based on the estimates obtained from fitting model M4 to the available data.

Table 3 shows the resultant bias and root mean squared error in the estimation of
the fixed effects. Clearly, both are quite small but slightly bigger when compared to the model that does not include the random effects.

| Covariate          | Total | Zika-dengue | chik.-dengue |
|--------------------|-------|-------------|--------------|
|                    | \((\lambda_i)\) | \((\pi_{i2}/\pi_{i1})\) | \((\pi_{i3}/\pi_{i1})\) |
| Intercept          | 0.0025 | -0.0023     | -0.0162      |
|                    | (0.0077) | (0.0145)   | (0.0256)     |
| SDI                | 0.0054 | 0.0141      | -0.0172      |
|                    | (0.014)  | (0.0245)   | (0.0304)     |
| Pct green area     | 0.0002 | 0.0040      | -0.0019      |
|                    | (0.010)  | (0.0198)   | (0.0228)     |
| Pop. density       | 0.0018 | -0.0201     | 0.0211       |
|                    | (0.016)  | (0.0339)   | (0.0406)     |

Table 3: Bias, together with the root mean squared error (in brackets), in the estimation of the fixed effects in equation (4) for the relative risk of the total number of cases, and for the odds ratio of a neighborhood having Zika or chikungunya in comparison to dengue (see equation (5)).

Panels of Figure 3 show the estimated quantiles of 2.5%, 50% and 97.5% across the \(L = 500\) datasets for \(\theta_{i1}\) (first row), \(\theta_{i2}\) (second row) and \(\theta_{i3}\) (third row). The true values of each \(\theta_{ik}\), \(i = 1, 2, \cdots, n\) and \(k = 1, 2, 3\) are represented by an open triangle. All the true values seem to be well recovered.

**Simulating data from a Multivariate Poisson model**

To investigate if this small bias observed in Table 3 was due to the modelling of the total cases of *Aedes* borne diseases together with the distribution of the number of cases of each of the diseases, or if it was due to the inclusion of the random effects in the equations of \(\log \lambda_i\) and \(\log \frac{\pi_{ik}}{\pi_{i1}}\) we performed another simulation study. We simulated data from a multivariate Poisson model that includes random effects in the modelling of the log relative risks.
Figure 3: Summaries (mean: open circle and quantiles of 2.5% and 97.5%: vertical segments) of the posterior mean of the random effects $\theta_{i1}$ (first row), $\theta_{i2}$ (second row) and $\theta_{i3}$ (third row) based on $L = 500$ datasets generated from the model in equation (7). The open triangles represent the true values used to generate the data.

More specifically, we generated data from

$$Y_{ik} \mid \gamma_{ik} \sim \text{Poisson}(E_{ik}\delta_{ik}), \quad i = 1, 2, \ldots, n \text{ and } k = 1, 2, 3,$$

$$\log \delta_{ik} = X_i\beta + \theta_{ik}.$$  

(8)

To generate the data, parameters were fixed at the estimates obtained after fitting model M4 under the multivariate Poisson specification to the available data. Table 4 shows the resultant bias and root mean squared error obtained under this scenario. It is clear that the magnitude of the bias and rMSE are similar to the ones obtained under the Poisson-Multinomial likelihood on Table 3. In other words, the slight increase we observe in the
bias when compared to the simple model that does not include random effects, seems to be due to the inclusion of the random effects. This suggests that we are able to estimate the parameters of the proposed model.

| Covariate       | Equations associated with each of the coefficients | Poisson model |
|-----------------|---------------------------------------------------|---------------|
|                 | Dengue (λ_i)                                        |               |
| Intercept       | 0.0008                                             | -0.0005       | -0.0149       |
|                 | (0.001)                                            | (0.008)       | (0.022)       |
| SDI             | 0.0088                                             | 0.0024        | -0.0212       |
|                 | (0.017)                                            | (0.010)       | (0.030)       |
| Pct green area  | 0.0086                                             | -0.0027       | 0.012         |
|                 | (0.017)                                            | (0.010)       | (0.021)       |
| Pop. density    | 0.012                                              | -0.0119       | 0.0311        |
|                 | (0.026)                                            | (0.018)       | (0.043)       |

Table 4: Bias, together with the root mean squared error (in brackets), in the estimation of the fixed effects in equation (8) for the relative risk of each of the diseases.

Panels of Figure 4 show the summaries of the posterior mean of the random effects θ_{ik}. Again, these results are similar to the ones obtained under the Poisson-Multinomial model.

D.3 Data generated from Model M5 in the main text

In this Section we generate data from model M5 in the main text, wherein the random effect θ_i follows a multivariate CAR prior distribution. Again, we assume

\[ total_i \mid \lambda_i \sim \text{Poi}(E_i \lambda_i) \]  
\[ y_i \mid total_i, \pi_i \sim \text{Multinomial}(total_i, \pi_i), \quad \text{for } i = 1, 2, \cdots, n, \]

where \( E_i \) was computed as described in the beginning of Section 2. The parameters \( \lambda_i \) were fixed according to equation (4), and each \( \pi_{ik} \) was fixed according to equation (5).
Figure 4: Summaries (mean: open circle and quantiles of 2.5% and 97.5%: vertical segments) of the posterior mean of the random effects $\theta_{i1}$ (first row), $\theta_{i2}$ (second row) and $\theta_{i3}$ (third row) based on $L = 500$ datasets generated from the model in equation (8). The open triangles represent the true values used to generate the data.

The values of $\theta_{ik}$ were fixed at the posterior mean of the estimated values obtained from the data analyzed in the manuscript under model M5.

We generated $L = 500$ datasets from this model and fitted the same model to check if we were able to recover the true values of the parameters. Table 5 shows the resultant bias and root mean squared error in the estimation of the fixed effects. Like the results obtained from data generated under model M4 the bias and RMSE are relatively small suggesting that we are able to recover the values of the parameters used to generate the data.

Panels of Figure 5 show the estimated quantiles of 2.5%, 50% and 97.5% across the $L = 500$ datasets for $\theta_{i1}$ (first row), $\theta_{i2}$ (second row) and $\theta_{i3}$ (third row). The true values of each $\theta_{ik}$, $i = 1, 2, \cdots, n$ and $k = 1, 2, 3$ are represented by an open triangle. Again, all the true values seem to be reasonably well recovered.

We ran other simulation studies that provided similar results to the ones shown here. The different simulation studies provide enough evidence that we are able to estimate
Equations associated with each of the coefficients

| Covariate | Total \((\lambda_i)\) | Zika-dengue \((\pi_{i2}/\pi_{i1})\) | chik.-dengue \((\pi_{i3}/\pi_{i1})\) |
|-----------|------------------|-------------------|------------------|
| Intercept | 0.0014           | 0.0081            | -0.0051          |
|           | (0.0071)         | (0.0166)          | (0.0204)         |
| SDI       | 0.0052           | 0.0065            | -0.0045          |
|           | (0.0223)         | (0.0645)          | (0.0295)         |
| Pct green area | 0.0004      | -0.0160          | -0.0066          |
|           | (0.0111)         | (0.0266)          | (0.0260)         |
| Pop. density | 0.0018       | -0.0228          | 0.0069           |
|           | (0.0141)         | (0.0342)          | (0.0349)         |

Table 5: Bias, together with the root mean squared error (in brackets), in the estimation of the fixed effects in equation (4) for the relative risk of the total number of cases, and for the odds ratio of a neighborhood having Zika or chikungunya in comparison to dengue (see equation (5)).
Figure 5: Summaries (mean: open circle and quantiles of 2.5% and 97.5%: vertical segments) of the posterior mean of the random effects $\theta_{i1}$ (first row), $\theta_{i2}$ (second row) and $\theta_{i3}$ (third row) based on $L = 500$ datasets generated from the model in equation (9). The open triangles represent the true values used to generate the data.
E A Negative-Binomial-Multinomial model

An alternative to the Poisson-Multinomial model is to assume that the total cases of vector borne diseases in each neighborhood follows a negative binomial instead of a Poisson distribution. In this case, one can follow the Negative-binomial 2 parametrization and assume that given the mean $\mu_i$ and a dispersion parameter $\tau$, the total of vector borne diseases at a neighborhood $i$ is a realization from the following probability function

$$p(\text{total}_i \mid \mu_i, \tau) = \left( \frac{\text{total}_i + \tau - 1}{\text{total}_i} \right) \left( \frac{\mu_i}{\mu_i + \tau} \right)^{\text{total}_i} \left( \frac{\tau}{\mu_i + \tau} \right)^\tau,$$

where $E(\text{total}_i \mid \mu_i, \tau) = \mu_i$ and $V(\text{total}_i \mid \mu_i, \tau) = \mu_i + \frac{\mu_i^2}{\tau}$. Following a similar structure of the model for the relative risk of the total number of cases described in equation (4) of the main text, we assume

$$\log(\mu_i) = X_i^T \beta + \theta_{i1}.$$

And the model for $y_i$ conditioned on the total cases in neighborhood $i$ follows a multinomial distribution with probability vector $\pi_i$. The resultant likelihood function for each neighborhood $i$ is given by

$$p(\text{total}_i, y_i \mid \Theta) = \frac{(\text{total}_i + \tau - 1)!}{\text{total}_i! (\tau - 1)!} \left( \frac{\mu_i}{\mu_i + \tau} \right)^{\text{total}_i} \left( \frac{\tau}{\mu_i + \tau} \right)^\tau \frac{\text{total}_i!}{y_{i1}! y_{i2}! y_{i3}!} \prod_{k=1}^{3} (\pi_{ik})^{y_{ik}}. (10)$$

Note that in this case, regardless of the way we parametrize the probability vector $\pi_i$, the total in each area will always bring information to the likelihood function.

We fitted the parametrization proposed under model M4 in the main text assuming the negative binomial distribution for the total cases and the multinomial distribution parametrized as in equation (5) of the main text. The resultant WAIC was 4304.8, the log-score (logS) was 2028.75 and the energy score (es) was 5.42. Although es resulted in a value close to the Poisson-Multinomial model, WAIC and logS were much greater than their respective values under the Poisson-Multinomial model. Moreover, the dispersion parameter $\tau$ showed some sensitivity to its prior specification. The result above was obtained assuming a half-normal prior distribution for $\tau$. If a half-Cauchy is used, the dispersion parameter is estimated at quite high values, suggesting that the component
\( \frac{\mu^2}{\tau} \) of the conditional variance of the total is close to zero. Therefore, this model might be overparametrized for this dataset. For this reason, we did not explore the negative-binomial model further for this dataset.

**References**

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P. and Riddell, A. (2017) Stan: A probabilistic programming language. *Journal of Statistical Software, Articles*, 76, 1–32.

Czado, C., Gneiting, T. and Held, L. (2009) Predictive model assessment for count data. *Biometrics*, 65, 1254–1261.

Gamerman, D. and Lopes, H. F. (2006) *Markov Chain Monte Carlo - Stochastic Simulation for Bayesian Inference*. 2nd Edition, Chapman & Hall/CRC.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. and Rubin, D. B. (2014) *Bayesian Data Analysis, 3rd edition*. New York: Chapman & Hall / CRC.

Gneiting, T. and Raftery, A. E. (2007) Strictly proper scoring rules, prediction and estimation. *Journal of the American Statistical Association*, 102, 359–378.

Gneiting, T., Stanberry, L. I., Grimit, E. P., Held, L. and Johnson, N. A. (2008) Assessing probabilistic forecasts of multivariate quantities, with applications to ensemble predictions of surface winds (with discussion). *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 17, 211–264.

R Core Team (2020) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B. and Bürkner, P.-C. (2021) Rank-normalization, folding, and localization: An improved \( \hat{R} \) for assessing convergence of MCMC. *Bayesian Analysis*, 16, 667–718.
Watanabe, S. (2010) Asymptotic equivalence of Bayes cross validation and widely appli-
cable information criterion in singular learning theory. *Journal of Machine Learning
Research, 11*, 3571–3594.