A Bayesian hierarchical model for disease mapping that accounts for scaling and heavy-tailed latent effects

Victoire Michal

McGill | Department of Epidemiology, Biostatistics and Occupational Health

Joint work with:

Alexandra M. Schmidt
Epidemiology, Biostatistics and Occ. Health
McGill University

Laís Picinini Freitas
School of Public Health
University of Montreal

StanConnect - October 31st, 2022

Set-up: Disease Mapping

Let a region of interest be partitioned into n non-overlapping areas with:

- Y_i, the number of cases in area i, $i = 1, \ldots, n$;
- E_i, the expected number at risk in that area (offset);
- $\boldsymbol{x}_i$, the row vector of covariates for area i.

Disease Mapping model

$$Y_i \sim \mathcal{P}(E_i \mu_i), \quad \text{with} \quad \ln(\mu_i) = \beta_0 + \boldsymbol{x}_i \boldsymbol{\beta} + b_i,$$

with β_0, the overall log risk, $\boldsymbol{\beta}$, the covariates coefficients and b_i, a random effect for area i.

$\Rightarrow$ **Goal:** propose a model for the b_i's that accounts for **spatial dependence** and identifies **outlying areas**.

Extreme risks (tail) | Different from their neighbours

Usual spatial effects models and their reparametrisation

- **BYM prior** (Besag et al., 1991)

$$b_i = \theta_i + u_i,$$

$$\begin{bmatrix} \theta_i \mid \sigma_\theta^2 \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\theta^2), \\ u_i \mid \boldsymbol{u}_{-i}, \sigma_u^2 \sim \mathcal{N}\left(\frac{1}{d_i}\sum_{j\sim i} u_j, \frac{\sigma_u^2}{d_i}\right) \end{bmatrix}$$

Issue: Unidentifiability of σ_θ^2 and σ_u^2 (MacNab, 2011)

Usual spatial effects models and their reparametrisation

- **BYM prior** (Besag et al., 1991)

$$b_i = \theta_i + u_i,$$
$$\begin{bmatrix} \theta_i \mid \sigma_\theta^2 \overset{i.i.d.}{\sim} \mathcal{N}(0,\sigma_\theta^2), \\ u_i \mid \boldsymbol{u}_{-i}, \sigma_u^2 \sim \mathcal{N}\left(\frac{1}{d_i}\sum_{j\sim i} u_j, \frac{\sigma_u^2}{d_i}\right) \end{bmatrix}$$

Issue: Unidentifiability of σ_θ^2 and σ_u^2 (MacNab, 2011)

- **Leroux prior** Leroux et al. (1999)

$$b_i \mid \boldsymbol{b}_{-i}, \lambda, \sigma_L^2 \sim \mathcal{N}\left(\frac{\lambda}{1-\lambda+\lambda d_i}\sum_{j\sim i} b_j, \frac{\sigma_L^2}{1-\lambda+\lambda d_i}\right)$$

Issue: Parameters are often incorrectly interpreted (Riebler et al., 2016).

Usual spatial effects models and their reparametrisation

- **BYM prior** (Besag et al., 1991)

$$b_i = \theta_i + u_i,$$
$$\begin{bmatrix} \theta_i \mid \sigma_\theta^2 \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\theta^2), \\ u_i \mid \boldsymbol{u}_{-i}, \sigma_u^2 \sim \mathcal{N}\left(\frac{1}{d_i}\sum_{j\sim i} u_j, \frac{\sigma_u^2}{d_i}\right) \end{bmatrix}$$

Issue: Unidentifiability of σ_θ^2 and σ_u^2 (MacNab, 2011)

$\Rightarrow$ **BYM2 prior** (Riebler et al., 2016)
Eases identifiability and interpretation

$$b_i = \sigma\left(\sqrt{1-\lambda}\theta_i + \sqrt{\lambda}u_i^\star\right),$$
$$\begin{bmatrix} \theta_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1), \\ u_i^\star = \frac{u_i}{h}, \quad \mathbb{V}(u_i^\star) \simeq 1 \end{bmatrix}$$

$\rightarrow$ σ^2 is the marginal overall variance: $\mathbb{V}(b_i) \simeq \sigma^2\left([1-\lambda]\times 1 + \lambda \times 1\right) = \sigma^2$.

- **Leroux prior** Leroux et al. (1999)

$$b_i \mid \boldsymbol{b}_{-i}, \lambda, \sigma_L^2 \sim \mathcal{N}\left(\frac{\lambda}{1-\lambda+\lambda d_i}\sum_{j\sim i} b_j, \frac{\sigma_L^2}{1-\lambda+\lambda d_i}\right)$$

Issue: Parameters are often incorrectly interpreted (Riebler et al., 2016).

Usual spatial effects models and their reparametrisation

- **BYM prior** (Besag et al., 1991)

$$b_i = \theta_i + u_i,$$
$$\begin{bmatrix} \theta_i \mid \sigma_\theta^2 \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\theta^2), \\ u_i \mid \boldsymbol{u}_{-i}, \sigma_u^2 \sim \mathcal{N}\left(\frac{1}{d_i}\sum_{j\sim i} u_j, \frac{\sigma_u^2}{d_i}\right) \end{bmatrix}$$

Issue: Unidentifiability of σ_θ^2 and σ_u^2 (MacNab, 2011)

$\Rightarrow$ **BYM2 prior** (Riebler et al., 2016)
Eases identifiability and interpretation

$$b_i = \sigma\left(\sqrt{1-\lambda}\theta_i + \sqrt{\lambda}u_i^\star\right),$$
$$\begin{bmatrix} \theta_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1), \\ u_i^\star = \frac{u_i}{h}, \quad \mathbb{V}(u_i^\star) \simeq 1 \end{bmatrix}$$

$\rightarrow$ σ^2 is the marginal overall variance: $\mathbb{V}(b_i) \simeq \sigma^2\left([1-\lambda]\times 1 + \lambda \times 1\right) = \sigma^2$.

- **Leroux prior** Leroux et al. (1999)

$$b_i \mid \boldsymbol{b}_{-i}, \lambda, \sigma_L^2 \sim \mathcal{N}\left(\frac{\lambda}{1-\lambda+\lambda d_i}\sum_{j\sim i} b_j, \frac{\sigma_L^2}{1-\lambda+\lambda d_i}\right)$$

Issue: Parameters are often incorrectly interpreted (Riebler et al., 2016).

$\Rightarrow$ **Congdon prior** Congdon (2017)
Allows for spatial heteroscedasticity

$$b_i \mid \boldsymbol{b}_{-i}, \lambda, \sigma_C^2, \boldsymbol{\kappa} \sim \mathcal{N}\left(\frac{\lambda}{1-\lambda+\lambda d_i}\sum_{j\sim i}\kappa_j b_j, \frac{\sigma_C^2/\kappa_i}{1-\lambda+\lambda d_i}\right),$$
$$\kappa_i \overset{i.i.d.}{\sim} \text{Gamma}\left(\tfrac{\nu}{2}, \tfrac{\nu}{2}\right), \ \nu \sim \text{Exp}(1/\mu_\nu)$$

$\rightarrow$ $\kappa_i < 1$ indicates that area i is an outlier.

Proposed reparametrisation

$\Rightarrow$ We propose a model to identify outlying areas with interpretable parameters.

Proposed model: BYM2-Gamma

$$b_i = \frac{\sigma}{\sqrt{\kappa_i}}\left(\sqrt{1-\lambda}\theta_i + \sqrt{\lambda}u_i^\star\right), \quad \text{with}$$

$$\theta_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1) \quad \text{and scaled} \quad u_i^\star = \frac{u_i}{h}, \quad \text{for} \quad u_i \mid \boldsymbol{u}_{-i} \sim \mathcal{N}\left(\frac{1}{d_i}\sum_{j\sim i} u_j, \frac{1}{d_i}\right).$$

Proposed reparametrisation

$\Rightarrow$ We propose a model to identify outlying areas with interpretable parameters.

Proposed model: BYM2-Gamma

$$b_i = \frac{\sigma}{\sqrt{\kappa_i}}\left(\sqrt{1-\lambda}\theta_i + \sqrt{\lambda}u_i^\star\right), \quad \text{with}$$

$$\theta_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1) \quad \text{and scaled} \quad u_i^\star = \frac{u_i}{h}, \quad \text{for} \quad u_i \mid \boldsymbol{u}_{-i} \sim \mathcal{N}\left(\frac{1}{d_i}\sum_{j\sim i} u_j, \frac{1}{d_i}\right).$$

- Marginal variance: $\mathbb{V}(b_i) \simeq \frac{\sigma^2}{\kappa_i}$.

Proposed reparametrisation

$\Rightarrow$ We propose a model to identify outlying areas with interpretable parameters.

Proposed model: BYM2-Gamma

$$b_i = \frac{\sigma}{\sqrt{\kappa_i}}\left(\sqrt{1-\lambda}\theta_i + \sqrt{\lambda}u_i^\star\right), \quad \text{with}$$

$$\theta_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1) \quad \text{and scaled} \quad u_i^\star = \frac{u_i}{h}, \quad \text{for} \quad u_i \mid \boldsymbol{u}_{-i} \sim \mathcal{N}\left(\frac{1}{d_i}\sum_{j\sim i} u_j, \frac{1}{d_i}\right).$$

- Marginal variance: $\mathbb{V}(b_i) \simeq \dfrac{\sigma^2}{\kappa_i}$.
- Natural choice for the scale mixture parameters' prior:

$$\kappa_i \overset{i.i.d.}{\sim} \text{Gamma}(\nu/2, \nu/2), \quad \text{with} \quad \nu \sim \text{Exp}(1/\mu_\nu).$$

$\rightarrow$ $b_i \sim t_{\mu_\nu}$ for $\lambda = 0$,
$\rightarrow$ Sensible choice for the prior mean: $\mu_\nu = 4$ (Gelman et al., 2004),
$\rightarrow$ $\kappa_i < 1$ indicates that area i is an outlier.

Interpretation: Proposed model vs Congdon's

Proposed model:

$$b_i = \frac{\sigma}{\sqrt{\kappa_i}} \left(\sqrt{1-\lambda}\theta_i + \sqrt{\frac{\lambda}{h}} u_i \right)$$

Congdon's model:

$$b_i \mid \boldsymbol{b}_{-i} \sim \mathcal{N} \left(\frac{\lambda}{1-\lambda+\lambda d_i} \sum_{j \sim i} \kappa_j b_j, \frac{\sigma_C^2}{\kappa_i [1-\lambda+\lambda d_i]} \right)$$

Interpretation: Proposed model vs Congdon's

Proposed model:

$$b_i = \frac{\sigma}{\sqrt{\kappa_i}}\left(\sqrt{1-\lambda}\theta_i + \sqrt{\frac{\lambda}{h}}u_i\right)$$

- $\underline{\lambda = 0:}$
  $$b_i = \frac{\sigma}{\sqrt{\kappa_i}}\theta_i \overset{ind.}{\sim} \mathcal{N}\left(0, \frac{\sigma^2}{\kappa_i}\right)$$

Congdon's model:

$$b_i \mid \boldsymbol{b}_{-i} \sim \mathcal{N}\left(\frac{\lambda}{1-\lambda+\lambda d_i}\sum_{j\sim i}\kappa_j b_j, \frac{\sigma_C^2}{\kappa_i[1-\lambda+\lambda d_i]}\right)$$

- $\underline{\lambda = 0:}$
  $$b_i \overset{ind.}{\sim} \mathcal{N}\left(0, \frac{\sigma_C^2}{\kappa_i}\right)$$

Interpretation: Proposed model vs Congdon's

Proposed model:

$$b_i = \frac{\sigma}{\sqrt{\kappa_i}}\left(\sqrt{1-\lambda}\theta_i + \sqrt{\frac{\lambda}{h}}u_i\right)$$

- $\underline{\lambda = 0:}$
$$b_i = \frac{\sigma}{\sqrt{\kappa_i}}\theta_i \overset{ind.}{\sim} \mathcal{N}\left(0, \frac{\sigma^2}{\kappa_i}\right)$$

- $\underline{\lambda = 1:}$
$$b_i \mid \boldsymbol{b}_{-i} \sim \mathcal{N}\left(\frac{1}{d_i}\sum_{j\sim i}\sqrt{\frac{\kappa_j}{\kappa_i}}b_j, \frac{\sigma^2/h}{\kappa_i d_i}\right)$$

Congdon's model:

$$b_i \mid \boldsymbol{b}_{-i} \sim \mathcal{N}\left(\frac{\lambda}{1-\lambda+\lambda d_i}\sum_{j\sim i}\kappa_j b_j, \frac{\sigma_C^2}{\kappa_i[1-\lambda+\lambda d_i]}\right)$$

- $\underline{\lambda = 0:}$
$$b_i \overset{ind.}{\sim} \mathcal{N}\left(0, \frac{\sigma_C^2}{\kappa_i}\right)$$

- $\underline{\lambda = 1:}$
$$b_i \mid \boldsymbol{b}_{-i} \sim \mathcal{N}\left(\frac{1}{d_i}\sum_{j\sim i}\kappa_j b_j, \frac{\sigma_C^2}{\kappa_i d_i}\right)$$

Interpretation: Proposed model vs Congdon's

Proposed model:

$$b_i = \frac{\sigma}{\sqrt{\kappa_i}}\left(\sqrt{1-\lambda}\theta_i + \sqrt{\frac{\lambda}{h}}u_i\right)$$

- $\underline{\lambda = 0:}$
$$b_i = \frac{\sigma}{\sqrt{\kappa_i}}\theta_i \overset{ind.}{\sim} \mathcal{N}\left(0, \frac{\sigma^2}{\kappa_i}\right)$$

- $\underline{\lambda = 1:}$
$$b_i \mid \boldsymbol{b}_{-i} \sim \mathcal{N}\left(\frac{1}{d_i}\sum_{j\sim i}\sqrt{\frac{\kappa_j}{\kappa_i}}b_j, \frac{\sigma^2/h}{\kappa_i d_i}\right)$$

Congdon's model:

$$b_i \mid \boldsymbol{b}_{-i} \sim \mathcal{N}\left(\frac{\lambda}{1-\lambda+\lambda d_i}\sum_{j\sim i}\kappa_j b_j, \frac{\sigma_C^2}{\kappa_i[1-\lambda+\lambda d_i]}\right)$$

- $\underline{\lambda = 0:}$
$$b_i \overset{ind.}{\sim} \mathcal{N}\left(0, \frac{\sigma_C^2}{\kappa_i}\right)$$

- $\underline{\lambda = 1:}$
$$b_i \mid \boldsymbol{b}_{-i} \sim \mathcal{N}\left(\frac{1}{d_i}\sum_{j\sim i}\kappa_j b_j, \frac{\sigma_C^2}{\kappa_i d_i}\right)$$

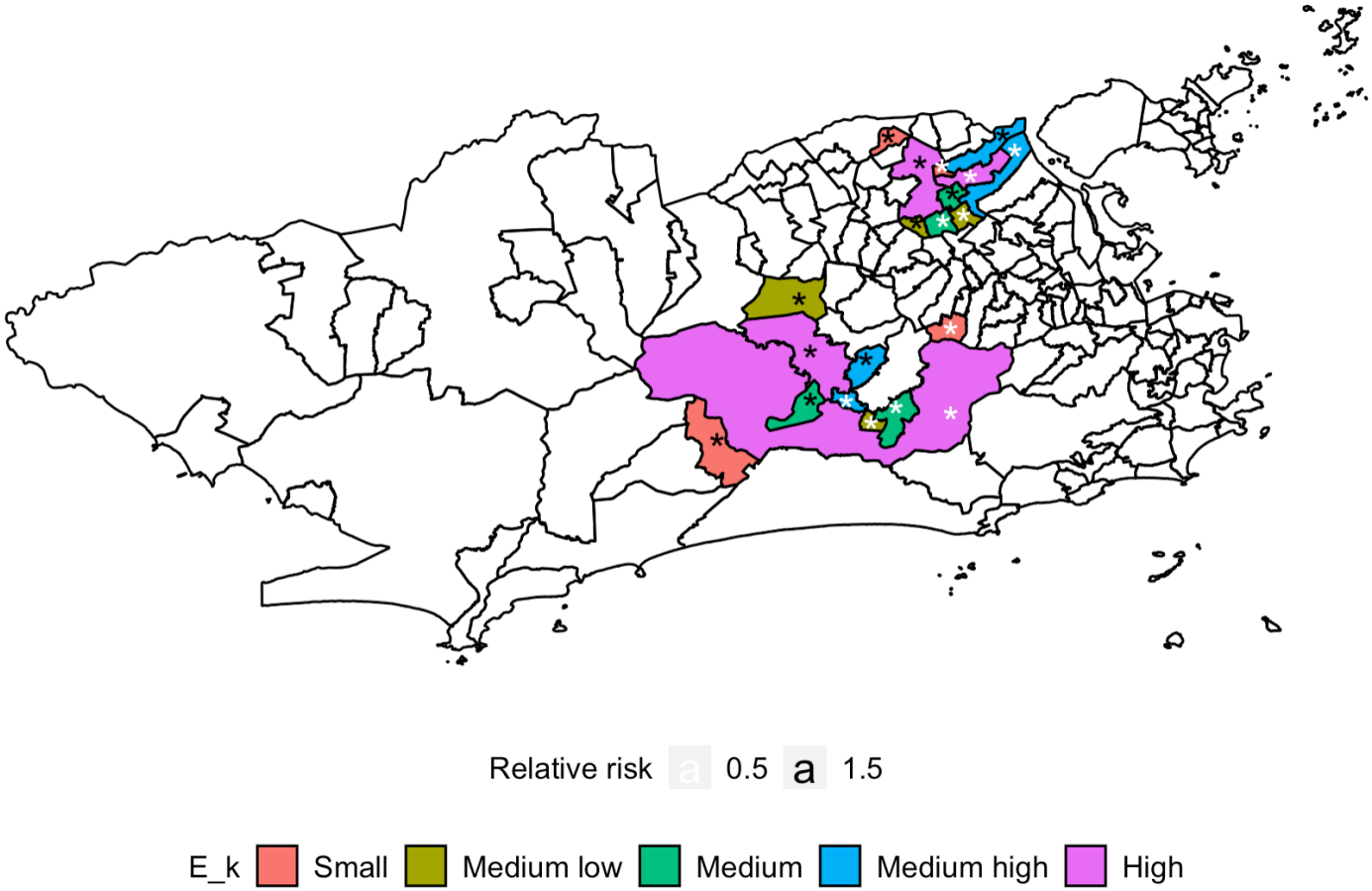
→ When the outliers are **far from each other**: BYM2-Gamma $\simeq$ Congdon;

→ **Clusters of outliers**: BYM2-Gamma >> Congdon as it borrows strength from the neighbouring κ's.

Simulation study: set-up (Richardson et al., 2004)

Replicate 200 datasets: $Y_i \sim \mathcal{P}(E_i \mu_i)$ with

- E_i's from the 2015-2016 Zika epidemic in Rio de Janeiro;
- $\mu_i = 1 \ \forall i$ except **2 clusters of 10 outliers** such that:



→ **Comparison between the proposed model and Congdon's (`rstan`)**
with priors: $\beta_0 \sim \mathcal{N}(0, 10^2)$, $\lambda \sim \mathcal{U}(0,1)$, and $\sigma, \sigma_C \sim \mathcal{N}_+(0,1)$.

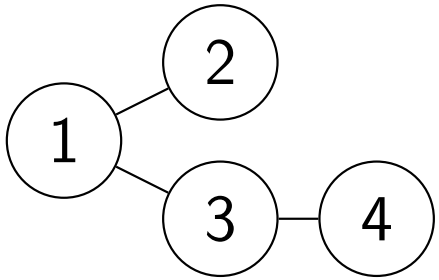
Stan code (Morris et al., 2019)

```
model{
  for(i in 1:N)
    y[i] ~ poisson_log(log_E[i] + beta0 +  sigma/sqrt(kappa[i]) *
                       (sqrt(1 - lambda) * theta[i] +
                        sqrt(lambda) * s[i]/sqrt(scaling_factor)) );

  target += -0.5 * dot_self(s[node1] - s[[node2]);
  sum(s) ~ normal(0, 0.001 * N); // soft sum-to-zero constraint on s

  theta ~ normal(0.0, 1.0);
  kappa ~ gamma(nu/2,nu/2);
  nu ~ exponential(1.0/4.0);
  ...
}
```

Where node1 and node2 store the neighbourhood structure as in the following example:




```
node1 = [1, 1, 3]
node2 = [2, 3, 4]
```

Simulation study: results

	Sensitivity		**Specificity**	
E_k	BYM2-Gamma	Congdon	BYM2-Gamma	Congdon
Small	39.9	34.9	100.0	99.9
Med. low	81.2	67.6	99.9	99.9
Medium	98.7	83.4	99.9	99.9
Med. high	99.9	91.7	99.9	99.9
High	100.0	94.4	100.0	99.8
Overall	83.9	74.4	100.0	99.9

Table: Percents of the true outliers identified (sensitivity) and percents of the true non-outliers identified (specificity) across the 200 replicates depending on the offset size.

Simulation study: results

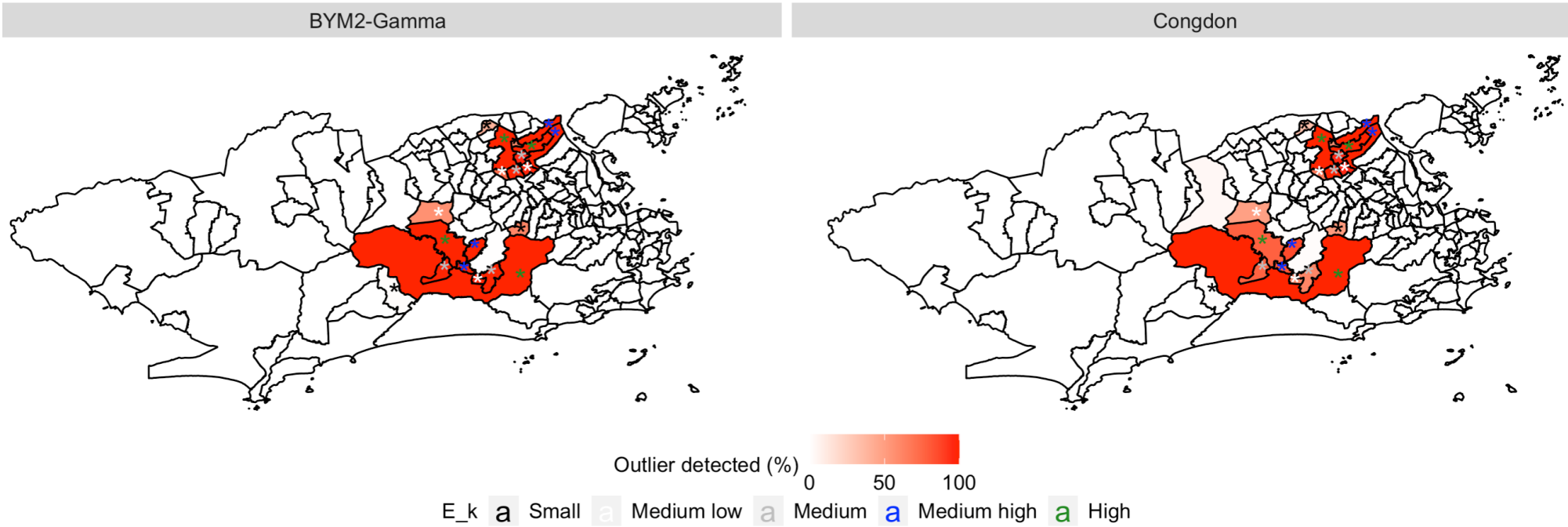


Figure: Maps of the percentages of outliers detected across the 200 replicates depending on the offset size. *: Contaminated districts.

Simulation study: results

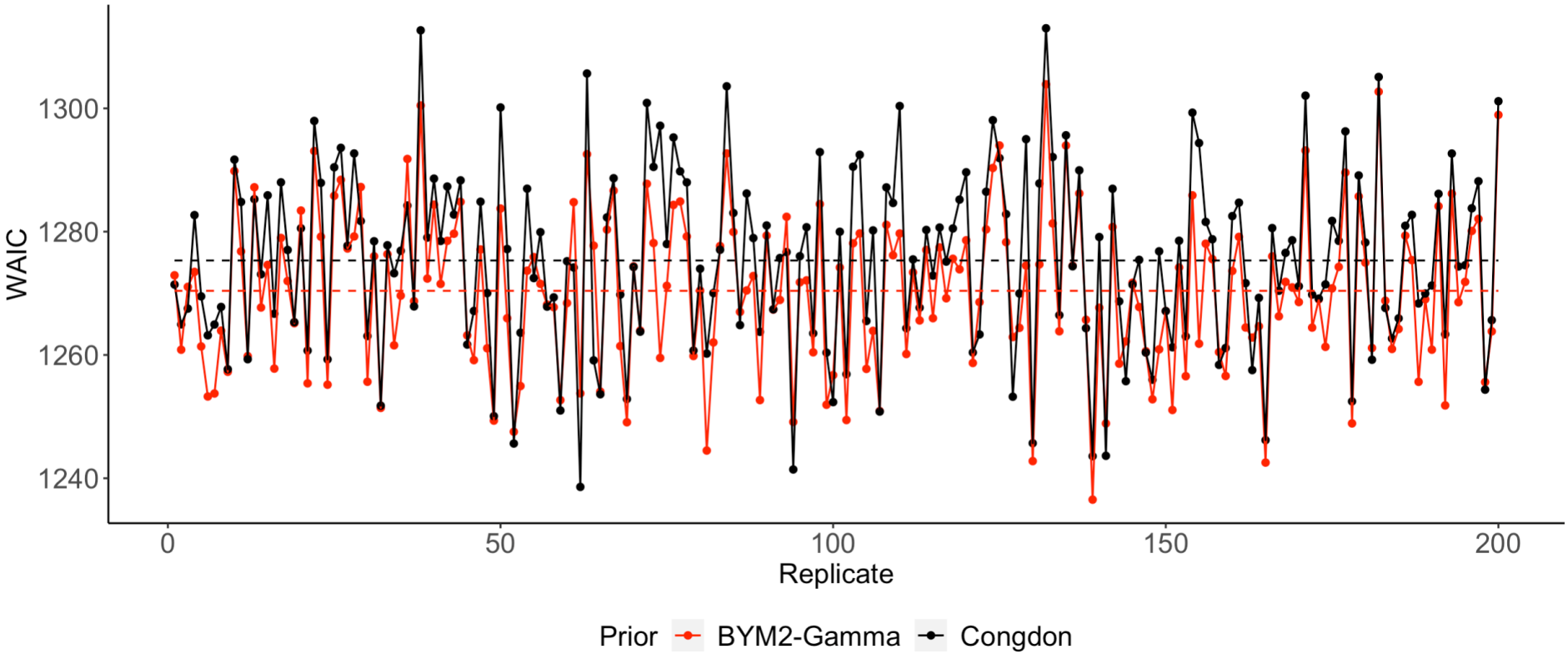


Figure: WAIC across the 200 replicates for the proposed model and Congdon's. Dashed lines: mean WAIC for each model (1270.4 vs 1275.3)

Data analysis: set-up

Data available for the $n = 160$ districts of Rio de Janeiro:

- Y_i's: Number of Zika cases recorded during the 2015-2016 epidemic;
- $E_i = P_i \frac{\sum_j Y_j}{\sum_j P_j}$: Expected count, using the population size, P;
- x_i's: Socio-development index.

$$\Rightarrow Y_i \sim \mathcal{P}\left(E_i \exp\left[\beta_0 + \beta x_i + b_i\right]\right)$$

Models fitted to the data:

BYM2-Gamma : $b_i = \frac{\sigma}{\sqrt{\kappa_i}}\left(\sqrt{1-\lambda}\theta_i + \sqrt{\lambda}u_i^\star\right)$

Congdon: $b_i \mid \boldsymbol{b}_{-i} \sim \mathcal{N}\left(\frac{\lambda \sum_{j\sim i} \kappa_j b_j}{1-\lambda+\lambda d_i}, \frac{\sigma_C^2/\kappa_i}{1-\lambda+\lambda d_i}\right)$

with priors:

- $\kappa_i \overset{i.i.d.}{\sim} \mathrm{Ga}(\nu/2, \nu/2)$,
- $\nu \sim \mathrm{Exp}(1/4)$,

Data analysis: set-up

Data available for the $n = 160$ districts of Rio de Janeiro:

- Y_i's: Number of Zika cases recorded during the 2015-2016 epidemic;
- $E_i = P_i \frac{\sum_j Y_j}{\sum_j P_j}$: Expected count, using the population size, P;
- x_i's: Socio-development index.

$$\Rightarrow Y_i \sim \mathcal{P}\left(E_i \exp\left[\beta_0 + \beta x_i + b_i\right]\right)$$

Models fitted to the data:

BYM2-Gamma : $b_i = \frac{\sigma}{\sqrt{\kappa_i}}\left(\sqrt{1-\lambda}\theta_i + \sqrt{\lambda}u_i^\star\right)$

Congdon: $b_i \mid \boldsymbol{b}_{-i} \sim \mathcal{N}\left(\frac{\lambda \sum_{j\sim i} \kappa_j b_j}{1-\lambda+\lambda d_i}, \frac{\sigma_C^2/\kappa_i}{1-\lambda+\lambda d_i}\right)$

BYM2: $b_i = \sigma\left(\sqrt{1-\lambda}\theta_i + \sqrt{\lambda}u_i^\star\right)$

Leroux: $b_i \mid \boldsymbol{b}_{-i} \sim \mathcal{N}\left(\frac{\lambda \sum_{j\sim i} b_j}{1-\lambda+\lambda d_i}, \frac{\sigma_L^2}{1-\lambda+\lambda d_i}\right)$

with priors:

- $\kappa_i \overset{i.i.d.}{\sim} \text{Ga}(\nu/2, \nu/2)$,
- $\nu \sim \text{Exp}(1/4)$,
- $\beta_0, \beta \sim \mathcal{N}(0, 10^2)$,
- $\lambda \sim \mathcal{U}(0,1)$,
- $\sigma, \sigma_C, \sigma_L \sim \mathcal{N}_+(0,1)$.

Application: exploratory data analysis

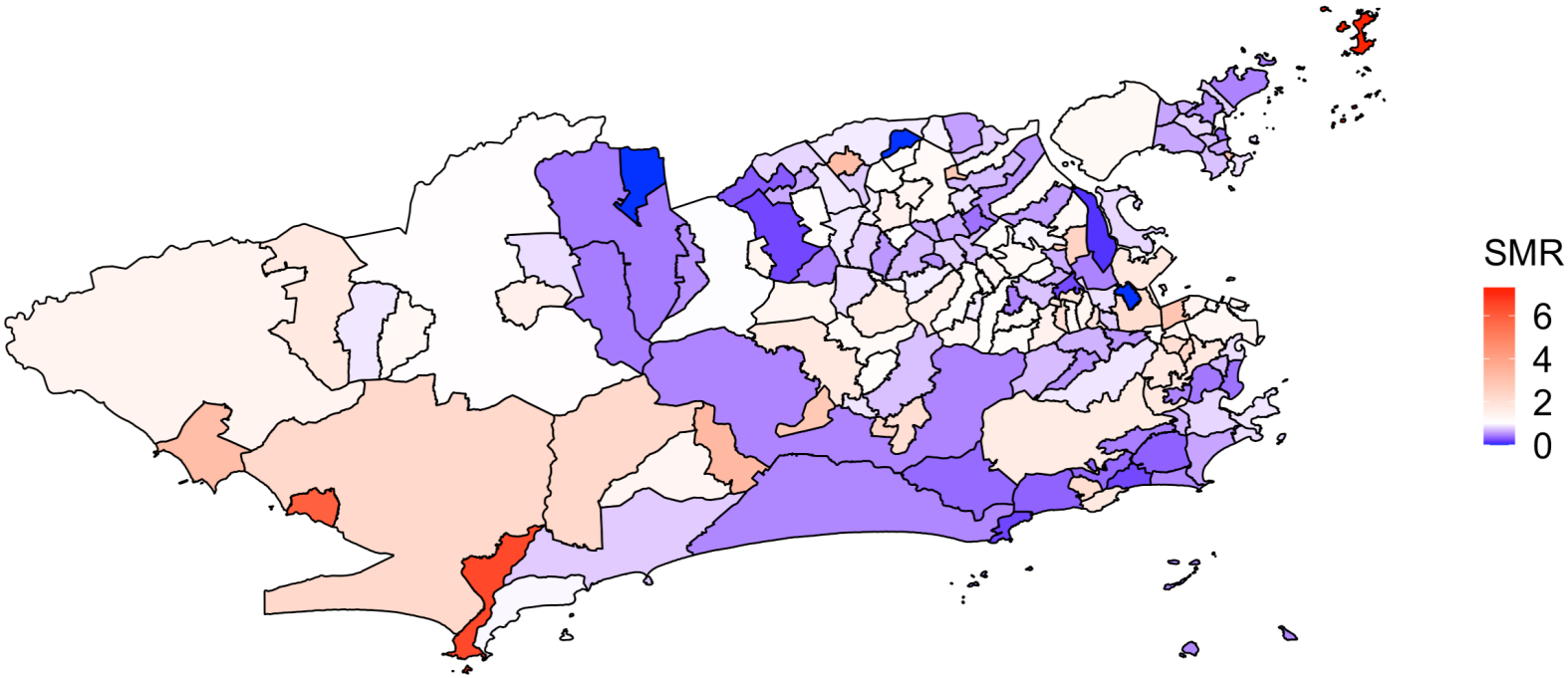


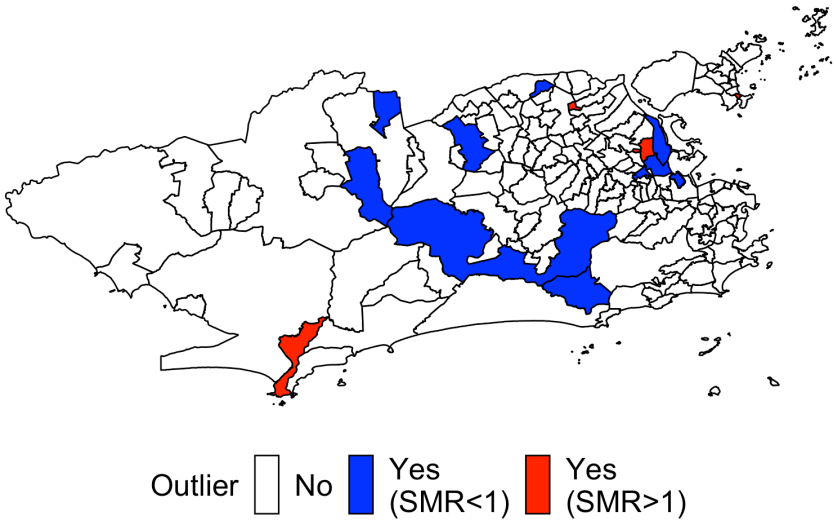
Figure: Standardised morbidity ratio (SMR, Y/E) for the Zika counts across the 160 neighbourhoods of Rio de Janeiro

Application: results

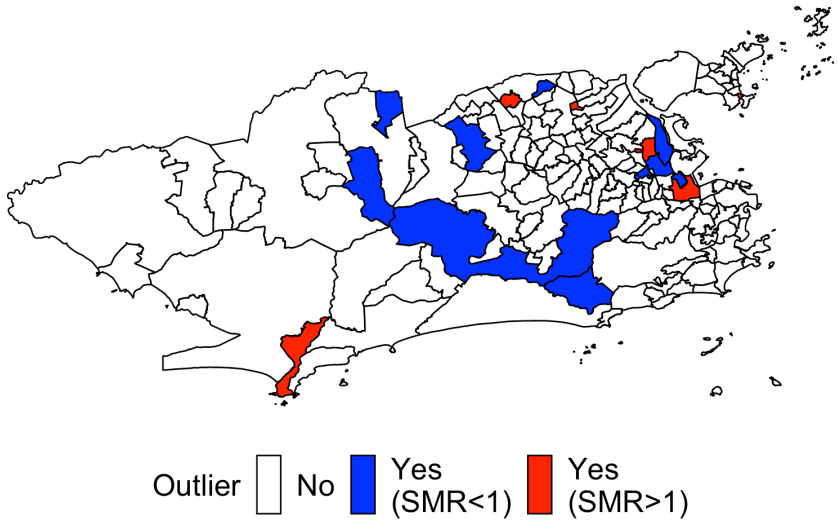
	BYM2	BYM2-Gamma	Congdon	Leroux
Model fit				
WAIC	1371.2	1335.6	1337.5	1375.1
p_W	88.6	80.0	81.0	89.2
Parameters' posterior summaries				
	Mean (CI)	Mean (CI)	Mean (CI)	Mean (CI)
β_0	1.6 (0.4,2.8)	2.5 (1.7,3.4)	2.4 (1.4,3.2)	1.6 (0.7,2.6)
β	-2.8 (-4.8,-0.8)	-4.3 (-5.6,-2.9)	-4.0 (-5.4,-2.6)	-2.9 (-4.4,-1.3)
λ	0.7 (0.4,0.9)	0.7 (0.3,0.9)	0.8 (0.5,0.9)	0.6 (0.2,0.9)
σ	0.8 (0.7,0.9)	**0.4 (0.3,0.5)**	0.6 (0.4,0.8)	0.7 (0.6,0.8)
ν	-	1.1 (0.6,1.9)	1.9 (1.3,2.8)	-

Table: Model assessment (WAIC) and parameters' posterior summaries: mean and 95% credible interval (CI) for BYM2, BYM2-Gamma, Congdon and Leroux.

Application: results



(a) BYM2-Gamma



(b) Congdon

Figure: Maps of the outliers as indicated by $\kappa_u < 1$, where κ_u is the upper bound of the posterior 95% credible interval of κ.

Discussion

We proposed a **Disease Mapping model**:

- Able to identify areas with potentially outlying disease risks;
  - $\rightarrow$ Extreme risks;
  - $\rightarrow$ Outliers with respect to the neighbours.

Discussion

We proposed a **Disease Mapping model**:

- Able to identify areas with potentially outlying disease risks;
  - → Extreme risks;
  - → Outliers with respect to the neighbours.
- Scaled latent spatial effects yielded interpretable parameters;
  - → Easy to assign priors for any spatial structure.

Discussion

We proposed a **Disease Mapping model**:

- Able to identify areas with potentially outlying disease risks;
  - → Extreme risks;
  - → Outliers with respect to the neighbours.
- Scaled latent spatial effects yielded interpretable parameters;
  - → Easy to assign priors for any spatial structure.
- Simulation studies helped assess our model's performance;
  - → Able to recover the true parameters (when generate from the model);
  - → More precise than Congdon's when detecting clusters of contaminated districts;
  - → Always performed best than Congdon's in terms of WAIC.

Discussion

We proposed a **Disease Mapping model**:

- Able to identify areas with potentially outlying disease risks;
  - → Extreme risks;
  - → Outliers with respect to the neighbours.
- Scaled latent spatial effects yielded interpretable parameters;
  - → Easy to assign priors for any spatial structure.
- Simulation studies helped assess our model's performance;
  - → Able to recover the true parameters (when generate from the model);
  - → More precise than Congdon's when detecting clusters of contaminated districts;
  - → Always performed best than Congdon's in terms of WAIC.
- Analysis of 2015-2016 Zika epidemic in Rio de Janeiro;
  - → Detected some areas with outlying risks (zero cases and higher risks);
  - → Our model performed similarly to Congdon's in terms of WAIC.

Thank you !

References

Besag, J., York, J., and Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the institute of statistical mathematics*, 43(1):1–20.

Congdon, P. (2017). Representing spatial dependence and spatial discontinuity in ecological epidemiology: a scale mixture approach. *Stochastic Environmental Research and Risk Assessment*, 31(2):291–304.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2004). *Bayesian data analysis*. CRC press.

Leroux, B. G., Lei, X., and Breslow, N. (1999). Estimation of disease rates in small areas: a new mixed model for spatial dependence. In *Statistical models in epidemiology, the environment, and clinical trials*, pages 179–191. Springer.

MacNab, Y. C. (2011). On gaussian markov random fields and bayesian disease mapping. *Statistical Methods in Medical Research*, 20(1):49–68.

Morris, M., Wheeler-Martin, K., Simpson, D., Mooney, S. J., Gelman, A., and DiMaggio, C. (2019). Bayesian hierarchical spatial models: Implementing the Besag York Mollié model in stan. *Spatial and spatio-temporal epidemiology*, 31:100301.

Richardson, S., Thomson, A., Best, N., and Elliott, P. (2004). Interpreting posterior relative risk estimated in disease-mapping studies. *Environmental Health Perspectives*, 112(9):1016–1025.

Riebler, A., Sørbye, S. H., Simpson, D., and Rue, H. (2016). An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical methods in medical research*, 25(4):1145–1165.