A spatio-temporal model for the relative risk of COVID-19 in Spain

Abdollah Jalilian[1] Jorge Mateu[2]

[1]Razi University, Kermanshah, Iran
[2]Universitat Jaume I, Castellón, Spain

April 16, 2020

Suppose that the spatial domain of interest S is partitioned into regions $A_1, \ldots, A_m$, defined by states, communities, provinces, counties, municipalities or towns. Let $N_t(A_i)$ denote the number of new cases in region A_i, $i = 1, \ldots, m$, at time $t = 1, \ldots, T$. We assume that $N_t(A_i)$ is a Poisson random variable with mean $E_{it}\theta_{it}$, where E_{it} is the expected number of cases in region A_i at time t. The nonnegative random quantities θ_{it}, $i = 1, \ldots, m$, $t = 1, \ldots, T$, are called the area-specific relative risks in the disease mapping literature (Lawson, 2018, Section 5.1.4). These relative risks account for the extra-Poisson variability (overdispersion) resulting from unmeasured confounders and model misspecification (Wakefield, 2007).

We consider the log linear model

$$\theta_{it} = \exp\left(\mu + \beta_1 t + \beta_2 t^2 + \beta_3 \texttt{popden}_i + \beta_4 \texttt{emp}_i + \zeta_t + \eta_i + \xi_{it}\right),$$

for the relative risks where

- $\beta_1 t + \beta_2 t^2$ is a quadratic temporal trend term motivated by the general agreement that the outbreak of many infectious diseases have a parabolic trend; i.e., rising, reaching a peak and then declining,
- $\texttt{popden}_i$ is the population density of region A_i,
- $\texttt{emp}_i$ is the number of employed people in region A_i,
- ζ_t is a temporal random effect,
- η_i is a spatial random effect representing the spatial correlation due to commuting between regions, and
- ξ_{it} is a spatio-temporal random effect representing dependencies due to spatial and temporal proximity.

We assume that $\mathbb{E}\boldsymbol{\zeta} = \mathbb{E}\boldsymbol{\eta} = 0$ and

$$\mathbb{V}\text{ar}\boldsymbol{\zeta} = \frac{1}{\tau_{\text{temp}}}\mathbf{I}_T, \quad \mathbb{V}\text{ar}\boldsymbol{\eta} = \frac{1}{\tau_c}\left(\mathbf{I}_m - \frac{\omega}{e_{\max}}\tilde{\mathbf{C}}\right)^{-1}$$

where $\tilde{\mathbf{C}} = \mathbf{C} + \mathbf{C}^{\mathsf{T}}$, $\mathbf{C} = [C_{i,j}]$, $C_{i,j}$ denotes the number of individuals commuting from region A_i to region A_j on a daily basis, $e_{\max}$ is the largest eigenvalue of $\tilde{\mathbf{C}}$, $\tau_{\text{temp}}, \tau_c > 0$ and $0 \leq \omega < 1$.

We further assume that $\mathbb{E}\boldsymbol{\xi} = 0$ and

$$\mathbb{V}\text{ar}\boldsymbol{\xi} = \frac{1}{\tau}\boldsymbol{\Gamma} \otimes \left((1-\varphi)\mathbf{I}_m + \varphi\mathbf{Q}^{-}\right),$$

where $\tau > 0$, $0 \leq \varphi \leq 1$, $\boldsymbol{\Gamma} = [\gamma(i - i')]$ is the covariance matrix of an AR(p) temporal process with

$$\gamma(h) = \begin{cases} 1 & h = 0 \\ \sum_{l=1}^{p} \phi_l \gamma(|h| - l) & |h| > 0 \end{cases}$$

and $\mathbf{Q}^{-}$ denotes the generalized inverse of the precision matrix $\mathbf{Q}$ with entries

$$Q_{i,i'} = \begin{cases} n_i & i = i' \\ -1 & i \sim i' \\ 0 & \text{otherwise} \end{cases}$$

Here, $i \sim i'$ indicates that regions A_i and $A_{i'}$ are neighbors (share a border) and n_i is the number of neighbors of region A_i.

The model for $\boldsymbol{\xi}$ is a variant of the Besag, York and Mollié (BYM) model (Besag et al., 1991), with a parameterization suggested by (Riebler et al., 2016) that accounts for scaling.

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. 2

Lawson, A. B. (2018). *Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology*. Chapman and Hall/CRC, 3 edition. 1

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. 2

Wakefield, J. (2007). Disease mapping and spatial regression with count data. *Biostatistics*, 8(2):158–183. 1