Integrating Compartment and Point Process Models for Spatio-Temporal Modeling of Infectious Diseases

André Victor Ribeiro Amaral [1,*], Jonatan A. González [1], Paula Moraga [1]

[1] Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST).

* E-mail: andre.ribeiroamaral@kaust.edu.sa

1. Objective

We are interested in modeling the infected individuals in space and time. To do this, we will

1. Fit a temporal (compartment) model, and
2. Use the previous step acquired information as the mean of a LGCP for the point pattern representing the infected individuals in the studied region and time interval.

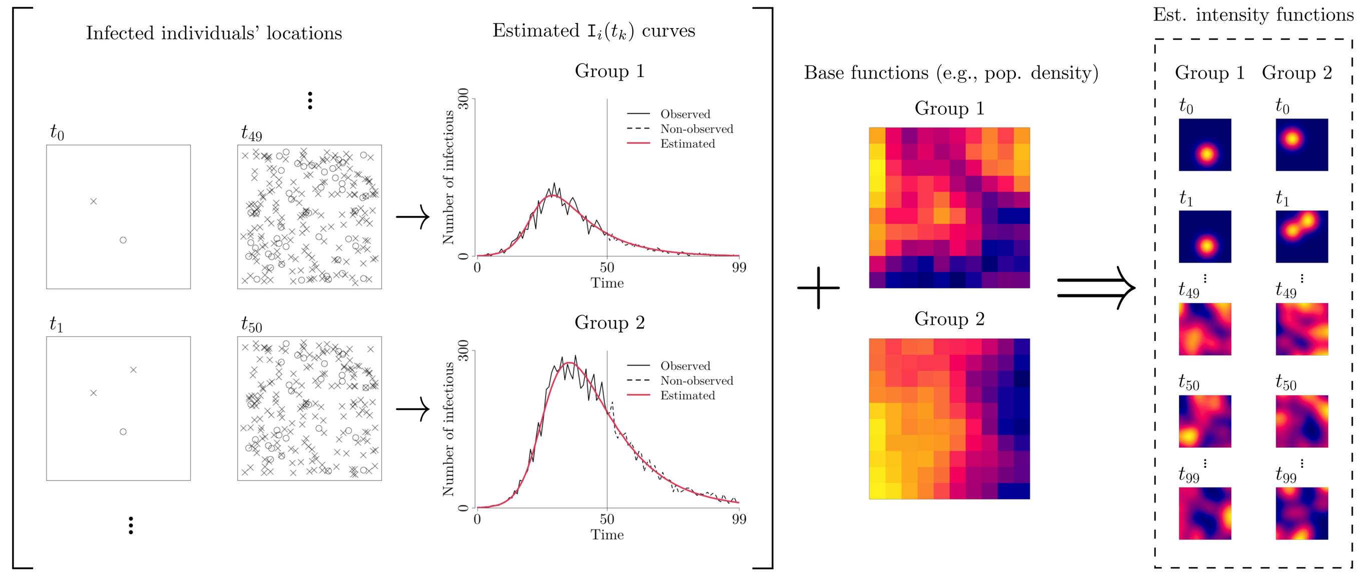

Figure 1: Two-step spatio-temporal modeling approach.

2. SIR modeling

The base-SIR model [2] is described as follows

Susceptible $\xrightarrow{\text{Transmission}}$ Infected $\xrightarrow{\text{Recovery}}$ Recovered

and it has the following assumptions

1. Homogeneous population with uniform mixing.
2. Constant infectious and recovery rates.
3. Preserved population mass.

Such a model can be extended in different ways.

Let $\mathtt{S}_i(t)$, $\mathtt{I}_i(t)$, and $\mathtt{R}_i(t)$ denote the number of susceptible, infected, and recovered individuals, respectively, at time t for age-group i. Then,

$$\begin{aligned}\frac{d\mathtt{S}_i(t)}{dt} &= -\beta \mathtt{S}_i(t) \textstyle\sum_{\text{all } j} C_{ij} \cdot \frac{\mathtt{I}_j(t)}{\mathtt{N}_j} \\ \frac{d\mathtt{I}_i(t)}{dt} &= +\beta \mathtt{S}_i(t) \textstyle\sum_{\text{all } j} C_{ij} \cdot \frac{\mathtt{I}_j(t)}{\mathtt{N}_j} - \gamma \mathtt{I}_i(t) \\ \frac{d\mathtt{R}_i(t)}{dt} &= +\gamma \mathtt{I}_i(t),\end{aligned} \tag{1}$$

such that C_{ij} is a contact matrix, $\mathtt{N}_i(t) = \mathtt{N}_i$, $\forall t$, and $\beta, \gamma > 0$. We will define a solution at $\{t_k\}_k$.

3. Point Process modeling

Let $\xi(t_k)$ be a log-Gaussian Cox process driven by $\Lambda(\mathbf{u}; t_k)$. In particular,

$$\Lambda(\mathbf{u}; t_k) = \mu(\mathbf{u}; t_k) \cdot \exp\{\zeta(\mathbf{u}; t_k)\}, \tag{2}$$

where $\zeta(\mathbf{u}; t_k)$ is a stationary Gaussian process with constant mean function given by $-\sigma^2/2$, and $\text{Cov}(\zeta(\mathbf{u}_1; t_k), \zeta(\mathbf{u}_2; t_k)) = \sigma^2 \rho(h; t_k)$.

Specific choices of $\mu(\mathbf{u}; t_k)$ were discussed by [1].

4. Temporal modeling (practical)

For a set of initial values $(\mathtt{S}_i(0), \mathtt{I}_i(0), \mathtt{R}_i(0))$, $\forall i$, and initial guesses for β and γ, we can solve the system of ODEs for $\mathtt{S}_i(t_k)$, $\mathtt{I}_i(t_k)$, and $\mathtt{R}_i(t_k)$ with a numerical method. We will name solutions $\mathtt{S}_i^{\mathtt{ODE}}(t_k)$, $\mathtt{I}_i^{\mathtt{ODE}}(t_k)$, and $\mathtt{R}_i^{\mathtt{ODE}}(t_k)$.

Now, suppose that we have obtained $\mathtt{i}_i(t_k)$, $\forall i, k$. One way to model such data is assuming that they come from a certain probability distribution with mean given by the ODE solution $\mathtt{I}_i^{\mathtt{ODE}}(t_k)$. In particular,

$$\mathtt{I}_i(t_k) \sim \text{Negative Binomial}(\mathtt{I}_i^{\mathtt{ODE}}(t_k), \varphi), \tag{3}$$

such that φ is the overdispersion parameter.

In that way, we will have to

1. Set initial values for β, γ, and φ.
2. Solve Model (1) for $\mathtt{S}_i(t_k)$, $\mathtt{I}_i(t_k)$, and $\mathtt{R}_i(t_k)$.
3. Plug the $\mathtt{I}_i^{\mathtt{ODE}}(t_k)$ curve into the mean component of Model (3) and evaluate it.
4. Update β, γ, and φ, and get back to (2.) until reach convergence.

Here, we used `RStan` [4] to estimate the posterior distribution of $\boldsymbol{\theta} = (\beta, \gamma, \varphi)^\top$.

5. ST modeling (practical)

The final model is specified as follows

$$\begin{aligned}\mathtt{N}_i(t_k) | \Lambda_i(\mathbf{u}; t_k) &= \lambda_i(\mathbf{u}; t_k) \sim \text{Poisson}\left(\textstyle\int_{\mathcal{U}} \lambda_i(\mathbf{u}; t_k) d\mathbf{u}\right), \ \forall i, k \\ \Lambda_i(\mathbf{u}; t_k) &= \mu_i(\mathbf{u}; t_k) \cdot \exp\{\zeta_i(\mathbf{u}; t_k)\} \\ \mu_i(\mathbf{u}; t_k) &= \lambda_{0,i}(\mathbf{u}; t_k) \cdot \mathtt{I}_i(t_k) \\ \zeta_i(\mathbf{u}; t_k | \boldsymbol{\eta}_i) &\sim \text{Gaussian Process}(\beta_{0,i}, \phi_i(h; t_k | \boldsymbol{\eta}_i)) \\ \boldsymbol{\eta}_i &\sim \text{priors},\end{aligned} \tag{4}$$

such that $\phi_i(h; t_k | \boldsymbol{\eta}_i)$ is a covariance function, and $\boldsymbol{\eta}_i$ is the vector of parameters.

Finally, Model (4) was fitted using `R-INLA` [3].

6. Case Study

We consider as a study region an area of approx. 3 km^2 in São Paulo, Brazil. Then, we divided people into three age groups: 0–19, 20–59, 60+.

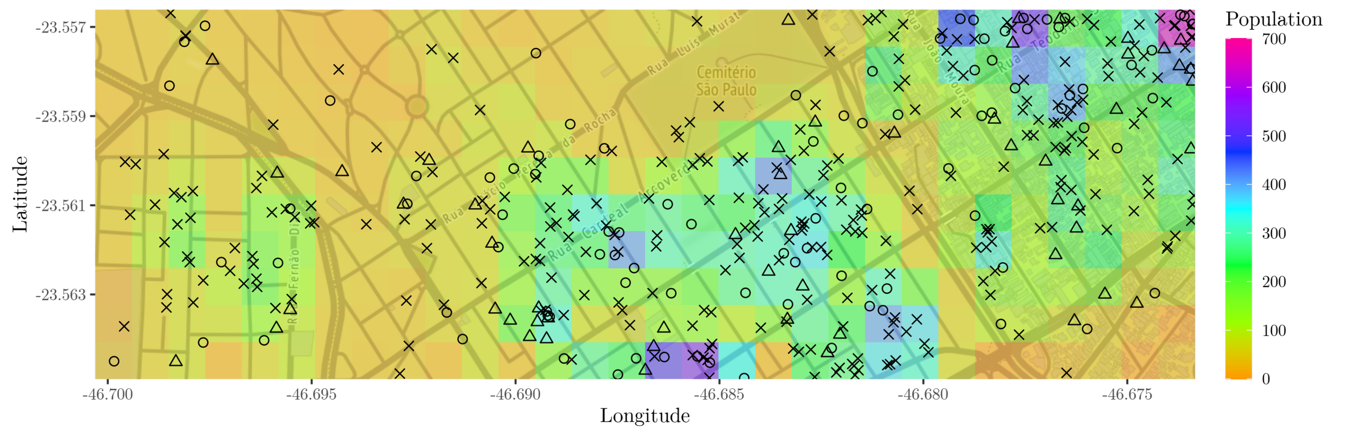

Figure 2: São Paulo (Brazil) with the overlapped grid for the estimated population and infec. individuals' locations.

For two scenarios (`FC` and `EP`), we simulated the temporal curves and the spatio-temporal intensity functions. After fitting different models,

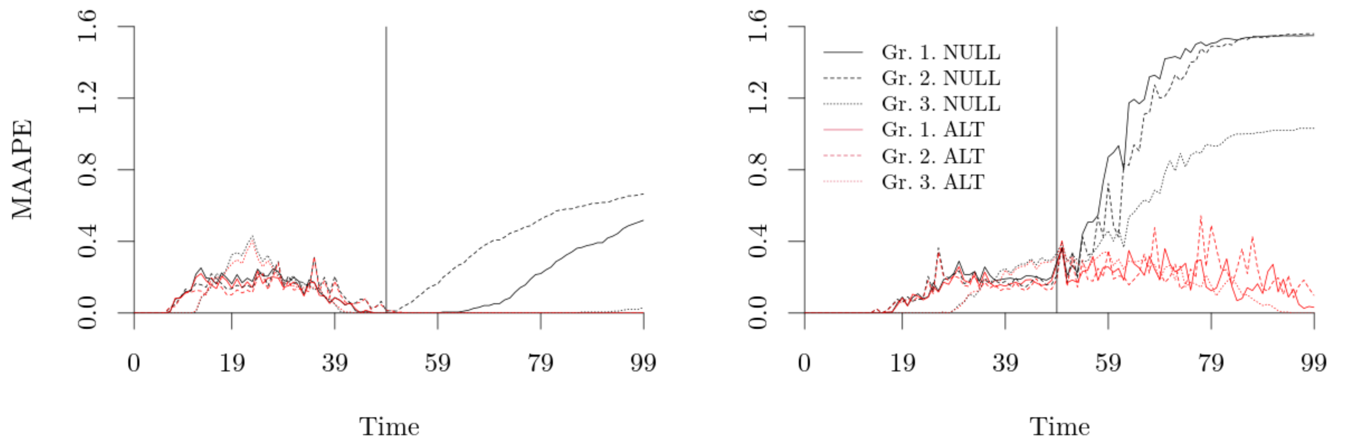

Figure 3: Computed errors for groups 0–19, 20–59, 60+. Models were fitted with data up to t_{49} (vertical solid line).

References

[1] Peter J Diggle. Spatio-temporal point processes: methods and applications. *Monographs on Statistics and Applied Probability*, 107:1, 2006.

[2] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

[3] Håvard Rue, Sara Martino, and Nicolas Chopin. Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)*, 71(2):319–392, 2009.

[4] Stan Development Team. RStan: the R interface to Stan, 2021. R package version 2.21.3.