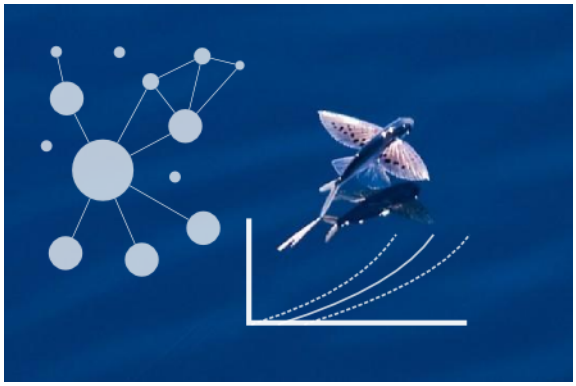
A spatial Poisson transmission model for Ebola (and other diseases)

Thibaut, Pierre, Anne, Nick G. and the Ebola team

October 17, 2014



Abstract

The model is a meta-population model using a known (spatial) connectivity matrix between patches and a simple kernel to model dispersal. The model is based on incidence data, with only infected individuals being known. Optionally, it can include infection from an unsampled reservoir. Unlike *outbreaker*, we are not trying to model individual ancestries, but merely dynamics within and between patches. In its simplest form, the model has only two parameters and its likelihood can be computed very fast. More complex extensions can allow for time-varying reproduction number, and/or time/spatially varying infection from the reservoir.

1 Notations

- I_t^i: observed incidence in patch i at time t (data)
- N_t^i: true, unobserved incidence in patch i at time t (augmented data)
- I_t, N_t: vectors of (observed, true) incidence of all patches
- I, T: matrix of observed / true incidence of all patches at all time steps $(I_t^1, \ldots, I_t^N)$
- T: the last date of the data
- D_{ij}: distance between i and j
- $w(.)$: the known probability mass distribution of the generation time / serial interval
- P: number of patches in the model
- n_t^i: the number of infected individuals in patch i at time t
- $d_{j \to i}$: intensity of dispersion from j to i
- δ: general dispersal parameter
- ϕ: background force of infection
- R: the effective reproduction number
- t_k: the date of infection of individual k
- $f_{\mathcal{P}}(a, b)$: the Poisson density for a observations and a rate b
- $k(c, d)$: a spatial kernel for a distance c and a parameter d

2 Model

We want to sample from the posterior distribution proportional to:

$$p(I, N, R, \phi, \delta, \omega) = p(I, N | R, \phi, \delta, \omega) p(R, \phi, \delta, \omega) \tag{1}$$

which can be rewritten

$$\underbrace{p(I|N, \omega) p(N|R, \phi, \delta)}_{likelihood} \underbrace{p(R) p(\phi) p(\delta) p(\omega)}_{priors} \tag{2}$$

The term $p(I|N, \omega)$ is the probability of the observed incidence given the true incidence N and the reporting probability ω. It is computed as:

$$p(I|N,\omega) = \prod_t \prod_i f_{\mathcal{B}}(I_t^i, N_t^i, \omega) \tag{3}$$

where $f_{\mathcal{B}}(x, a, b)$ is the Binomial p.m.f. for x successes, a draws and a probability b.

The term $p(N|R, \phi, \delta)$ is the probability of the true incidence given the infectivity in the system and the spatial processes at play. It is computed as:

$$p(N|R,\phi,\delta) = \prod_t \prod_i f_{\mathcal{P}}(N_t^i, \lambda_t^i) \tag{4}$$

where $f_{\mathcal{P}}$ if the p.m.f of a Poisson distribution and λ_t^i the force of infection towards patch i and time t. λ_t^i is a sum over of the forces of infection of all patches towards i (including $i \to i$). We note β_t^i the force of infection coming from infected individuals in patch i at time t, defined by the renewal equation:

$$\beta_t^i = R \sum_{k=1}^{n_t^i} w(t - t_k) \tag{5}$$

Where R is the reproduction number. Note that we can easily turn it into a time-dependent term R_t, in which case we will have to assume i) a functional form or ii) R_t constant between break-points. Similarly, it can be turned into a patch-specific R_i, in which case heterogeneity between patches can be modeled using a given distribution. This can get more complicated quickly.

The force of infection experienced by patch i at time t is then:

$$\lambda_t^i = \phi + \sum_{j=1}^{P} d_{j \to i} \beta_t^i \tag{6}$$

where $d_{j \to i}$ is the diffusion from j to i defined by a kernel k:

$$d_{j \to i} = k(D_{ij}, \delta) \tag{7}$$

where D_{ij} is the known distance between i and j (which needs not be Euclidean or symmetric).