M1 Writeup

March 11, 2015

Contents

1 Notation **2**

2 System dynamics model **3**

2.1 Gravity model 3

2.2 Model fit diagnostics 4

3 Structure generation **4**

3.1 Structures 4

3.1.1 Random Network 4

3.1.2 Starred Alley 5

3.1.3 Ring 6

3.1.4 Grid 7

3.1.5 Bowtie 8

3.1.6 Scalefree 10

3.2 Covariates 11

3.3 Network distance 12

4 Simulations **12**

5 Results **13**

5.1 Simultaneous perturbation experiment 13

5.2 Agent jitter and num chunks experiment 13

5.3 Toy Structures 13

5.4 WNS 13

1 Notation

Each location $i \in \mathcal{L}$ lies in a R^2 at coordinate pair C_i. The matrix Ω is a binary matrix indicating adjacent locations. For any two locations i and j, $\Omega_{i,j} = 1$ if location j is adjacent to location i. We define the network such that locations are adjacenty to themselves, $\Omega_{i,i} = 1$ for all i. All networks defined will be undirected thus $\Omega_{i,j} = \Omega_{j,i}$ for all i, j. The adjacency matrix is the defining component of the network. For some of the networks, the adjacency matrix can be defined without knowing the coordinate pairs, but for other networks the coordinate pairs are required to generate the adjacency matrix. Regardless of the network, we outline every step to generate the network in its entirety.

When defining networks, some additonal notation may be added to keep the process clear and

concise.

To model the spread of the disease, we assume that once a location becomes infected, it remains infected thereafter. The infection status for location i at time t is $X_{i,t} = 1$ if infected and $X_{i,t} = 0$ otherwise. Thus $X_{i,t} \geq X_{i,t'}$ for all $t > t'$.

Let the subscript operator $[\cdot]$ denote the usual order statistics; let $\lfloor \cdot \rfloor$ denote the floor operator; let $\lceil \cdot \rceil$ denote the ceiling operator.

2 System dynamics model

2.1 Gravity model

To model the evolution of the disease, we model the probability that an infected location i infects an uninfected location j at time point t. We denote this probability as $p_{i,j,t}$. If we assume independence across events i infecting j and i' infecting j, the probability of j becoming infected at time t is

$$P(X_{j,t} = 1 | X_{\cdot,t-1}) = X_{j,t-1} \bigvee \left[1 - \prod_{i:X_{i,t-1}=1} (1 - p_{i,j,t}) \right].$$

To define $p_{i,j,t}$ we use a form of the gravity model.

$$\text{logit } p_{i,j,t} = \beta_0 + \beta_1 U_j - \alpha \frac{d_{i,j}}{(c_i c_j)^\nu} - \eta g_j(\tau_{i,t}) - \rho A_{j,t-1} - \gamma A_{i,t-1}$$

where $\tau_{i,t} = \sum_{t'=0}^{t} X_{i,t'}$ is the number of time points i has been infected. We compare different forms of $g_j(\cdot)$ discussed later. The name of the model comes from $d_{i,j}/(c_i, c_j)^\nu$ which is known as the gravity term. The effect of distance between two locations is diminished if the number of caves in the two locations is large. We assume as the number of caves increases, so does the number of bats effectively decreasing the effect of distance. The vector U_j is a vector of covariates as part of the state information. These covariates include the number of caves, average winter length, area, species richness, and number of hibernating bat species.

For model estimation, we use a Bayesian framework. The priors on the parameters are as follows $\beta_0 \sim N(0, 100)$, $\beta_1 \sim N(0, 10)$, $\log(\alpha) \sim N(0, 1)$, $\log(\nu) \sim N(0, 1)$. The priors for ρ and γ will be specified later. For WNS, the observed data does not include information about treatments. This requires an ad-hoc selection of the true treatment effects. The prior distributions during the simulations for these parameters will over estimate the true effect.

2.2 Model fit diagnostics

When estimating the posterior distributions, we compared multiple forms for $g(\cdot)$.

- "zero": $g_j(x) = 0$
- "linear": $g_j(x) = x - 1$
- "exp": $g_j(x) = exp(x - 1) - 1$
- "exp_caves": $g_j(x) = exp(\frac{\max_k c_k + 1}{c_j + 1}(x - 1)) - 1$

The table below shows the Bayesian P-value for each of the four transformations above. Using an exponential transformation greatly improves the estimated p-values. Going forward, we choose the "exp_caves" transformation as providing the best fit.

3 Structure generation

3.1 Structures

3.1.1 Random Network

For each location $i = 1, \ldots, L$, $C_i \sim \{U(0, 1)\}^2$. To determine which locations are connected, we must first define some notation. Let $E_{i,j} = \|C_i - C_j\|_2$. The set of k nearby locations to location

	zero	linear	exp	exp_caves
n_inf	0.0044	0.6837	0.6604	0.6139
n_inf_2007	0.0507	0.2654	0.2942	0.3209
n_inf_2008	0.0000	0.4197	0.1263	0.1175
n_inf_2009	0.0000	0.6249	0.3884	0.3037
n_inf_2010	0.0001	0.6709	0.5682	0.4780
n_inf_2011	0.0013	0.6819	0.6420	0.5804
n_inf_2012	0.0234	0.6813	0.6710	0.6408
n_inf_2013	0.1357	0.6098	0.6757	0.6557
mean_year	0.7744	0.3494	0.6573	0.6276
mean_long	0.9122	0.3176	0.3735	0.4526
mean_lat	0.9260	0.3189	0.4078	0.5153
mean_dist_from_start	0.0484	0.6816	0.6064	0.5132
min_long	0.8960	0.3169	0.3612	0.4193
min_lat	0.8946	0.3170	0.3432	0.3931
max_long	0.0017	0.5111	0.2298	0.2070
max_lat	0.0744	0.6827	0.6384	0.5913
max_dist_from_start	0.1256	0.6834	0.6495	0.5994
Column Mean	0.2864	0.5186	0.4878	0.4723
Column Median	0.0625	0.5642	0.5280	0.4956
Column SD	0.3794	0.1616	0.1705	0.1508

Table 1: Bayesian p-values for postulated models.

i is $N_i^{(k)} = \{j : E_{i,j} \leq E_{i,[k+1]}\}$. Note the detail $[k+1]$ which is a consequence of $E_{i,i} = 0$. The adjacency matrix is defined as $\Omega_{i,j} = 1$ if $i \in N_j^{(k)}$ or $j \in N_i^{(k)}$.

3.1.2 Starred Alley

For a starred alley of size L, define

$$m_L = \arg\max_{m>0} f(m)\mathbb{1}_{\{f(m)\leq L\}}$$

where

$$f(m) = m + \left\lceil \frac{m}{2} \right\rceil \left(\left\lceil \frac{m}{2} \right\rceil - (m \bmod 2) + 1 \right).$$

Let locations $1, \ldots, m_L$ index the locations on the main horizontal line. Next define disjoint sets

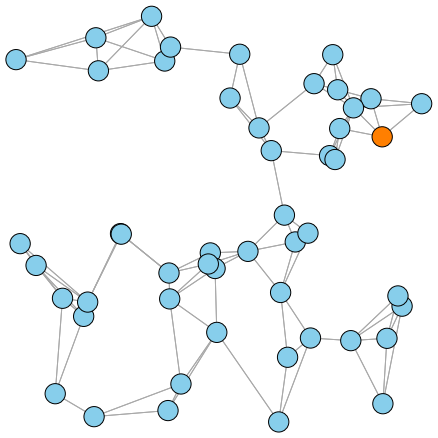

Figure 1: A random network with 50 locations

of locations N_i for \$i = 1,. . . ,m$_L$. The cardinality for N_i is defined as

$$|N_i| = \left\lfloor \frac{i}{2} \right\rfloor + \mathbb{1}_{\{(m_L - i) < (L - f(m_L))\}}.$$

For each location $j \in N_i$ set $\Omega_{i,j} = \Omega_{j,i} = 1$. For the first m_L locations, $\Omega_{i,i+1} = \Omega_{i+1,1} = 1$ for $i = 1, \ldots, m_L$.

To define the coordinates for each location, first set $C_i = \{i, 0\}$ for $i = 1, \ldots, m_L$. Then for each

3.1.3 Ring

Divide the number of locations L into two groups. Define $L_1 = L - L_2$ where $L_2 = \lceil L * 0.05 \rceil$. Let $\theta = 2\pi/(L_1 + 1)$ and $r = 1/(2\sin(\theta/2))$. Place location 1 at $C_1 = \{0, r\}$. Given location $i - 1$ at C_{i-1}, place location i at $C_i = rot(\theta)C_{i-1}$ for $i = 2, \ldots, L_1$ and

$$rot(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

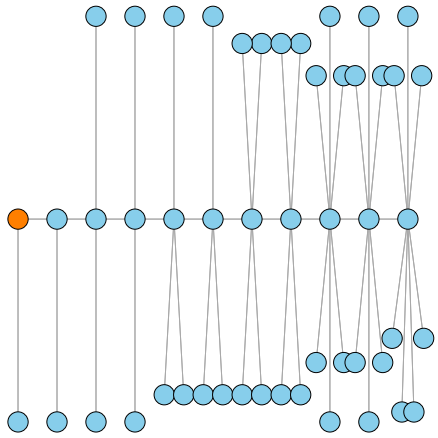

Figure 2: The starred alleyway network with 50 locations

Note that $C_1 \cdot C_{L_1} = \|C_1\|\|C_{L_1}\| \cos(2\theta)$.

With the first L_1 locations defined, we need to define the last L_2 locations. Place location $L_1 + i$ at $C_{L_1+i} = rot(\theta/L_2)C_{L_1+i-1}$ for $i = 1, \ldots, L_2$. Note that $C_L \cdot C_{L_1} = \|C_L\|\|C_{L_1}\| \cos(\theta)$ and $C_1 \cdot C_L = \|C_1\|\|C_L\| \cos(\theta)$.

To define the adjacency matrix, $A_{i,i+1} = A_{i+1,i} = 1$ for $i = 1, \ldots, L-1$ and $A_{1,L} = A_{L,1} = 1$.

3.1.4 Grid

First define L_1 to be the number of columns and L_2 to be the number of rows. These values are defined by

$$\{L_1, L_2\} = \underset{\substack{\ell_1,\ell_2>0 \\ \ell_1\ell_2=L \\ \ell_1\leq\ell_2}}{\arg\min} \quad |\ell_1 - \ell_2|$$

Location i is positioned at $C_i = \{(i \bmod L_2), \lfloor i/L_2 \rfloor\}$. The adjacency matrix is defined as $\Omega_{i,j} = 1$ if $\|C_i - C_j\| = 1$.

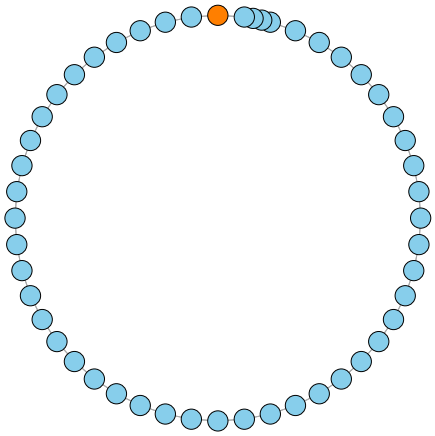

Figure 3: The ring network with 50 locations

3.1.5 Bowtie

The bowtie combines two grid networks with a random network in between. Spread of the disease begins in one grid and travels through the random network and into the other grid.

For a bowtie network of size L, create one random network of size $L_R = \lceil 0.1L \rceil + ((L - \lceil 0.1L \rceil) \bmod 2)$ and two grid networks of size $L_G = (L - L_R)/2$. Let i_R be the i^{th} location in the random network and i_{G_j} be the i^{th} location in the j^{th} grid.

To connect the three networks, we need to first identify which nodes in each network will connect to one of the other two networks. We will begin with the first grid. First, define L_1 and L_2 to be defined in the same way as in the grid description. Let $M_G = \lceil L_2/3 \rceil$ be the number of points to connect from the grid networks and $M_R = \lceil L_R/5 \rceil$. Define the following four sets of locations

$$\begin{aligned}
\Lambda_{G_1} &= \{L_G - 2L_1M_G + kL_1 : k = 1, \ldots, M_G\} \\
\Lambda_{G_2} &= \{L_G - 2L_1M_G + 1 + (k-1)L_1 : k = 1, \ldots, M_G\} \\
\Lambda_{R_1} &= \{i : C_{i_R,1} \leq C_{[M_R]_R,1}\} \\
\Lambda_{R_2} &= \{i : C_{i_R,1} \geq C_{[L_R - M_R + 1]_R,1}\}.
\end{aligned}$$

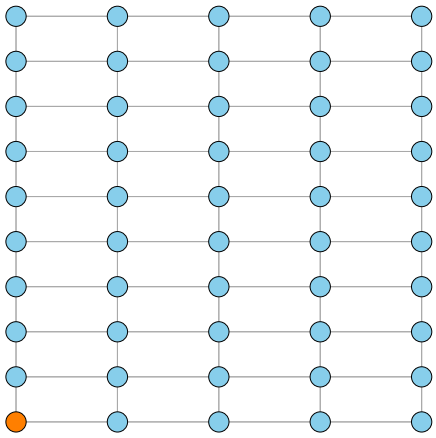

Figure 4: The grid network with 50 locations

To define the adjacency matrix, combine the adjacency matrices from the three separate networks as a block diagonal matrix. To define the additional edges, set $\Omega_{i_{G_1},j_R} = \Omega_{j_R,i_{G_1}} = 1$ for all $i \in \Lambda_{G_1}$ and $j \in \Lambda_{R_1}$ to connect the first grid to the random network. Then set $\Omega_{i_{G_2},j_R} = A_{j_R,i_{G_2}} = 1$ for all $i \in \Lambda_{G_2}$ and $j \in \Lambda_{R_2}$ to connect the second grid to the random network.

Now that the adjacency marix is defined, we need to finish the network. The last step is scaling and adjusting the coordinates from the three networks. We define this as an ordered list of operations.

1. For $i = 1, \ldots, L_R$: $C_{i_R} = C_{i_R}/2$.
2. Define $K = (\max_i C_{i_R,1} - \min_i C_{i_R,1})/2$.
3. For $i = 1, \ldots, L_R$: $C_{i_R,1} = C_{i_R,1} - \min_j C_{j_R,1} + \max_j C_{j_{G_1},1} + K$
4. For $i = 1, \ldots, L_G$: $C_{i_{G_2},1} = C_{i_{G_2},1} - \min_j C_{j_{G_2},1} + \max_j C_{j_R,1} + K$.

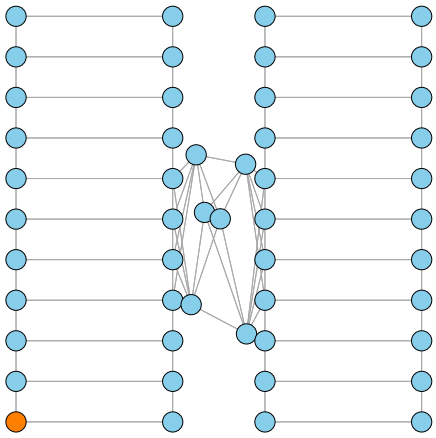

Figure 5: The bowtie network with 50 locations

3.1.6 Scalefree

The scalefree network is designed so that some locations are highly connected while others are only connected by a couple edges. This type of network was originally desigined to mimic the network of webpages. The internet houses many websites that are highly connected (e.g. Google, Yahoo, etc.) while the majority are not.

To generate the network of size L, we will define the process via induction. Define $\Omega^{(i)}$ to be the adjacency network on iteration $i = 1, \ldots, L$. A network of size 1 is the trivial network. Given a network of size $N-1$, place the additonal N^{th} location adjacent to an existing location $i \in \{1, \ldots, N-1\}$ with probability

$$P(\Omega^{(N)}_{N,i} = 1 | \Omega^{(N-1)}) = \frac{\sum_{j!=i} \Omega^{(N-1)}_{j,i}}{\sum_{j,k\,:\,j>k} \Omega^{(N-1)}_{j,k}}.$$

Generating a scale-free network in this fashion results in a Barabasi-Network. The defining feature of a scale-free network is the proportion of nodes with k edges is proportional to $k^{-\gamma}$ for

some γ. In a Barabasi-Network $\gamma = 3$.

Given the adjacency matrix, we use the Fruchterman-Reingold algorithm to position the nodes in a two dimensional plane.

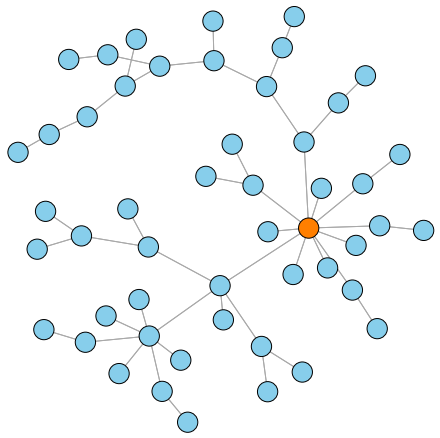

Figure 6: The scalefree network with 50 locations

3.2 Covariates

Define X to be a $n \times p$ centered and scaled covariate matrix. The covariates are multivariate normal $MVN_p(\mu, \Sigma)$ where $\mu_i = 2\widetilde{C}_{i,1}^2 + 1\widetilde{C}_{i,2}$ and $\widetilde{C}$ are the centered and scaled coordinates. The covariance matrix has the structure $\Sigma_{ip+s,jp+t} = \rho e^{(-\tau E_{i,j} - \eta|s-t|)}$. Thus, the variance of the individual covariates are 1 and the correlation decays as $E_{i,j}$ increases and as $|s-t|$ increases.

We need the mass of each individual node to compute the gravity term in the gravity model. In the case of White-Nose Syndrome, the mass is the number of caves. Arbitrarily set the first covariate to be the mass of the node. Define $M_i = \lfloor X_{i,1} - X_{[1],1} \rfloor + 1$ to be the mass of location i.

3.3 Network distance

In order for the structure of the network to have a strong impact on the disease spread, we define distance as the geodesic distance raised to a power. The distance between adjacent nodes (i and j such that $A_{i,j} = 1$) is defined to be 1. For non-adjacent nodes, the distance is the number of edges along the shortest path raised to a power. To determine the appropriate power, let $p_{i,j}^{(1)}$ be the hypothetical probability node i infects node j where i and j have geodesic distance 1. Let $p_{i,j}^{(2)}$ be the hypothetical probability if infection if i and j have geodesic distance 2. For their mass, use the mean mass of the network. The log odds-ratio is equal to

$$-\frac{\alpha}{m^\rho} + \frac{\alpha 2^z}{m^\rho}.$$

Set this equation equal to $\log(0.5)$ and solve for z.

4 Simulations

The simulations begin at time point 0 and end at $T = 15$. Observed data from WNS has data from 8 years. When simulating the spread of the disease under intervention, treatments are not given until time point 8 to mimic the observed data.

At time point 8, we estimate the posterior distributions of the indexing parameters in the system dynamics model. Note that there are no treatments yet in the data, so the posterior for the treatment effects are simply the prior. Using the estimated posterior distributions, the simultaneous perturbation algorithm is run to maximize the posterior mean reward at time T. The reward funtion at time t is the current proportion of locations infected, $Y^t(\pi) \triangleq \frac{1}{L}\sum_i X_{i,t}$.

At $t = 9$, the process is the same with one addition. Since the data now include a year of treatments, we run the tuning algorithm for simultaneous perturbation to adaptively improve the performance. Time points $t = 10, \ldots, T$ are the same as 8.

The prior for the treatment effects are set to be optimistic. They are normal with mean 4 times

as large as the true effect and variance 1.

5 Results

5.1 Simultaneous perturbation experiment

To calibrate the simultaneous perturbation algorithm, we ran a full factorial experiment over the following levels.

- $A \in \{30, 50\}$
- $B \in \{1, 10\}$
- $C \in \{2, 5\}$
- $L \in \{1, 1.25\}$
- $T \in \{1, 2\}$

For each of the 32 combinations, the table below shows the mean and standard deviation of the value across all replications of that factor combination. The results are sorted by the mean value.

5.2 Agent jitter and num chunks experiment

5.3 Toy Structures

5.4 WNS

valueMean	valueSd	timeMean	Afac	Bfac	Cfac	Tfac	Lfac	combo
0.554	0.016	2.757	+	-	-	+	+	12
0.560	0.021	2.741	-	+	-	+	+	20
0.565	0.030	2.700	-	-	-	+	+	28
0.565	0.017	2.789	-	+	+	+	+	16
0.568	0.019	2.817	+	+	-	+	+	4
0.569	0.009	2.746	-	-	+	+	+	24
0.574	0.014	0.909	-	+	-	+	-	21
0.574	0.013	0.920	-	+	+	+	-	17
0.576	0.017	0.915	+	+	-	+	-	5
0.578	0.029	2.848	+	+	+	+	+	0
0.578	0.014	0.931	+	-	+	+	-	9
0.581	0.032	0.898	-	-	-	+	-	29
0.581	0.017	2.825	+	-	+	+	+	8
0.584	0.014	0.917	-	-	+	+	-	25
0.584	0.019	0.920	+	-	-	+	-	13
0.587	0.022	0.942	+	+	+	+	-	1
0.631	0.023	0.791	-	+	-	-	-	23
0.649	0.026	0.789	+	-	-	-	-	15
0.650	0.012	0.821	+	+	-	-	-	7
0.658	0.017	0.811	-	-	-	-	-	31
0.662	0.011	0.729	-	+	+	-	-	19
0.664	0.029	0.773	+	+	+	-	-	3
0.667	0.016	0.721	-	-	+	-	-	27
0.667	0.015	0.740	+	-	+	-	-	11
0.668	0.030	2.072	-	-	-	-	+	30
0.670	0.040	1.964	-	-	+	-	+	26
0.671	0.028	2.124	-	+	+	-	+	18
0.675	0.018	2.040	+	-	-	-	+	14
0.686	0.024	2.094	+	+	+	-	+	2
0.688	0.025	1.973	+	-	+	-	+	10
0.700	0.020	2.108	+	+	-	-	+	6
0.710	0.024	2.078	-	+	-	-	+	22

Table 2: Results from the simultaneous perturbation experiment.