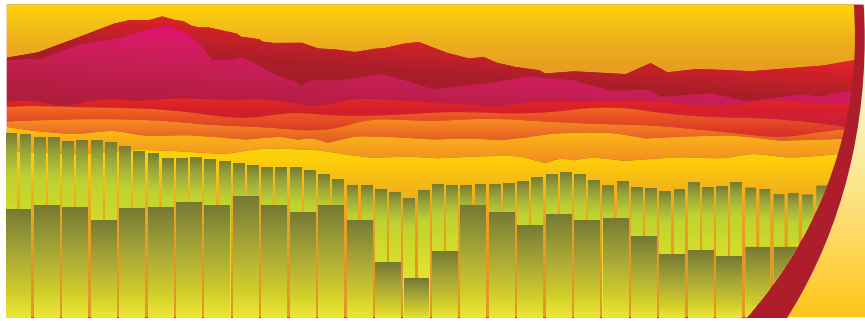


The University of Calgary
Biostatistics Centre (UCBC)

Bayesian Modelling of Epidemics
Part III: Individual-Level Models

Dr. Rob Deardon[1,2,3] **& Dr. Caitlin Ward**[1]

Department of Mathematics & Statistics[1]
Faculty of Veterinary Medicine[2]
Department of Community Health Sciences[3]

robert.deardon@ucalgary.ca
caitlin.ward@ucalgary.ca

ISBA, Montréal, June 2022

UNIVERSITY OF CALGARY

Outline

1. Individual-Level Models
1.1 General Framework
1.2 Susceptibility
1.3 Transmissibility
1.4 The Infection Kernel
1.5 Sparks Function

2. Some Examples
2.1 Simple Spatial Model
2.2 Spatial Model with Covariates
2.3 Network-based SI Model (Undirected and Unweighted)

3. Extras
3.1 More on Sparks Function

Outline

1. Individual-Level Models
1.1 General Framework
1.2 Susceptibility
1.3 Transmissibility
1.4 The Infection Kernel
1.5 Sparks Function

2. Some Examples
2.1 Simple Spatial Model
2.2 Spatial Model with Covariates
2.3 Network-based SI Model (Undirected and Unweighted)

3. Extras
3.1 More on Sparks Function

Outline

1. Individual-Level Models
1.1 General Framework
1.2 Susceptibility
1.3 Transmissibility
1.4 The Infection Kernel
1.5 Sparks Function

2. Some Examples
2.1 Simple Spatial Model
2.2 Spatial Model with Covariates
2.3 Network-based SI Model (Undirected and Unweighted)

3. Extras
3.1 More on Sparks Function

Individual-Level Models

So far, our infectious disease models have assumed homogeneity in the population, or a very simple population structure described by strata (e.g., gender, age)

This homogeneity refers to:

homogeneity of individual-level characteristics
homogeneous mixing (rates of contact) among the population

However, real populations have complex, heterogeneous structure in mixing and individuals with differing characteristics relevant to infection risk.

Therefore, we consider the idea of **individual level models**.

What individual-level characteristics might determine risk of infection?

Susceptibility: Risk factors associated with individuals who are potentially susceptible to the disease, contracting the disease

- e.g. age, genetics, environmental conditions (e.g. housing), vaccination history

Transmissibility: Risk factors associated with infectious individuals passing on the disease

- e.g. age, genetics, environmental conditions (e.g. housing), vaccination history

Connectivity of infectious and susceptible individuals

- spatial separation distance
- network-based separation distance

UNIVERSITY OF CALGARY

Compartmental (Multi-state) Framework

Here, we assume a **closed population** of n individuals: $i = 1, \ldots, n$

We assume that time is **discrete**, each time point representing an interval in real (continuous) time (e.g., a day):

We assume that at a given time point t, $t = 1, \ldots t_{max}$,
each individual can be in one of a number of states.

Possible Timelines

A number of **compartmental frameworks** are possible:

- SI
- SIR
- SEIR
- SIS
- SIRS
- etc.

Which we choose will depend on the disease we are studying
(and maybe the quality of our data...)

Today we will mainly concern ourselves with the **SI** and **SIR** compartmental frameworks.

Transitions between compartments

We are concerned with modelling when the **transition** between compartments or states occurs

For the SIR framework this involves two transitions:

$S \rightarrow I$
$I \rightarrow R$

For the SI framework this involves only one transition:

$S \rightarrow I$

Let's keep things simple to begin with and start with an SI compartmental framework

General form of the individual level model (ILM): Deardon et al (2010)

In a homogeneous SI or SIR model, the
probability of susceptible individual i **being exposed/infected at time** t is given by:

$$P(i,t) = P(t) = 1 - \exp\left(-\lambda I_t\right) \tag{1}$$

where:

$\lambda > 0$ is the transmission rate; and
I_t is the **number** of infectious individuals at time point t

Here, we individualize the infection rate so that:

$$P(i,t) = 1 - \exp\left(-\lambda_{it}\right) \tag{2}$$

where:

$\lambda_{it} > 0$ is a function of I_t

General form of the individual level model (ILM): Deardon et al (2010)

The **probability of susceptible individual** i **being exposed/infected at time** t is given by:

$$P(i,t) = 1 - \exp\left(-\left[\mathcal{S}(i)\left(\sum_{j \in I(t)} \mathcal{T}(j)\kappa(i,j)\right) + \epsilon(i,t)\right]\right) \tag{3}$$

where

$\mathcal{S}(i)$ is a susceptibility function for individual i
$I(t)$ is the **set** of infectious individuals at time t
$\mathcal{T}(j)$ is a transmissibility function for individual j
$\kappa(i,j)$ is the infection kernel
$\epsilon(i,t)$ is the random sparks function

Additionally, it is assumed that $\mathcal{S}(i), \mathcal{T}(j), \kappa(i,j)$ and $\epsilon(i,t) \geq 0$ always

Outline

1. Individual-Level Models
1.1 General Framework
1.2 Susceptibility
1.3 Transmissibility
1.4 The Infection Kernel
1.5 Sparks Function

2. Some Examples
2.1 Simple Spatial Model
2.2 Spatial Model with Covariates
2.3 Network-based SI Model (Undirected and Unweighted)

3. Extras
3.1 More on Sparks Function

Susceptibility function: $\mathcal{S}(i)$

This function describes mathematically, risk factors associated with the susceptible individual contracting the disease

Can take any form and contain any covariates/risk factors we have recorded

Typically use a linear function:

$$\mathcal{S}(i) = \alpha_0 + \alpha_1 X_1(i) + \alpha_2 X_2(i) + \ldots + \alpha_{n_S} X_{n_S}(i)$$

where $\alpha_0, \ldots, \alpha_{n_S} \geq 0$ are parameters/coefficients to be estimated
and $X_1, \ldots, X_{n_S}$ are covariates (data).

Note:

We generally include a baseline α_0 in the susceptibility function
If we have no susceptibility covariates, we typically set

$$\mathcal{S}(i) = \alpha_0 = \alpha$$

Outline

1. Individual-Level Models
1.1 General Framework
1.2 Susceptibility
1.3 Transmissibility
1.4 The Infection Kernel
1.5 Sparks Function

2. Some Examples
2.1 Simple Spatial Model
2.2 Spatial Model with Covariates
2.3 Network-based SI Model (Undirected and Unweighted)

3. Extras
3.1 More on Sparks Function

Transmissibility function: $\mathcal{T}(j)$

This function describes mathematically, risk factors associated with the infectious individual passing on the disease

Can take any form and contain any covariates/risk factors we have recorded

Typically use a linear function:

$$\mathcal{T}(j) = \eta_1 Y_1(j) + \eta_2 Y_2(j) + \ldots + \eta_{n_T} Y_{n_T}(j)$$

where $\eta_1, \ldots, \eta_{n_T}$ are coefficients to be estimated
and $Y_1, \ldots, Y_{n_T}$ are covariates (data).

Note:

We generally do not include a baseline η_0 in the transmissibility function
Baseline normally included in the susceptibility function
If we have no transmissibility covariates, typically set

$$\mathcal{T}(j) = 1$$

Outline

1. Individual-Level Models
1.1 General Framework
1.2 Susceptibility
1.3 Transmissibility
1.4 The Infection Kernel
1.5 Sparks Function

2. Some Examples
2.1 Simple Spatial Model
2.2 Spatial Model with Covariates
2.3 Network-based SI Model (Undirected and Unweighted)

3. Extras
3.1 More on Sparks Function

Infection kernel: $\kappa(i, j)$

$\kappa(i, j)$ characterizes the risk of infection due to **risk factors shared by susceptible individual** i **and infectious individual** j

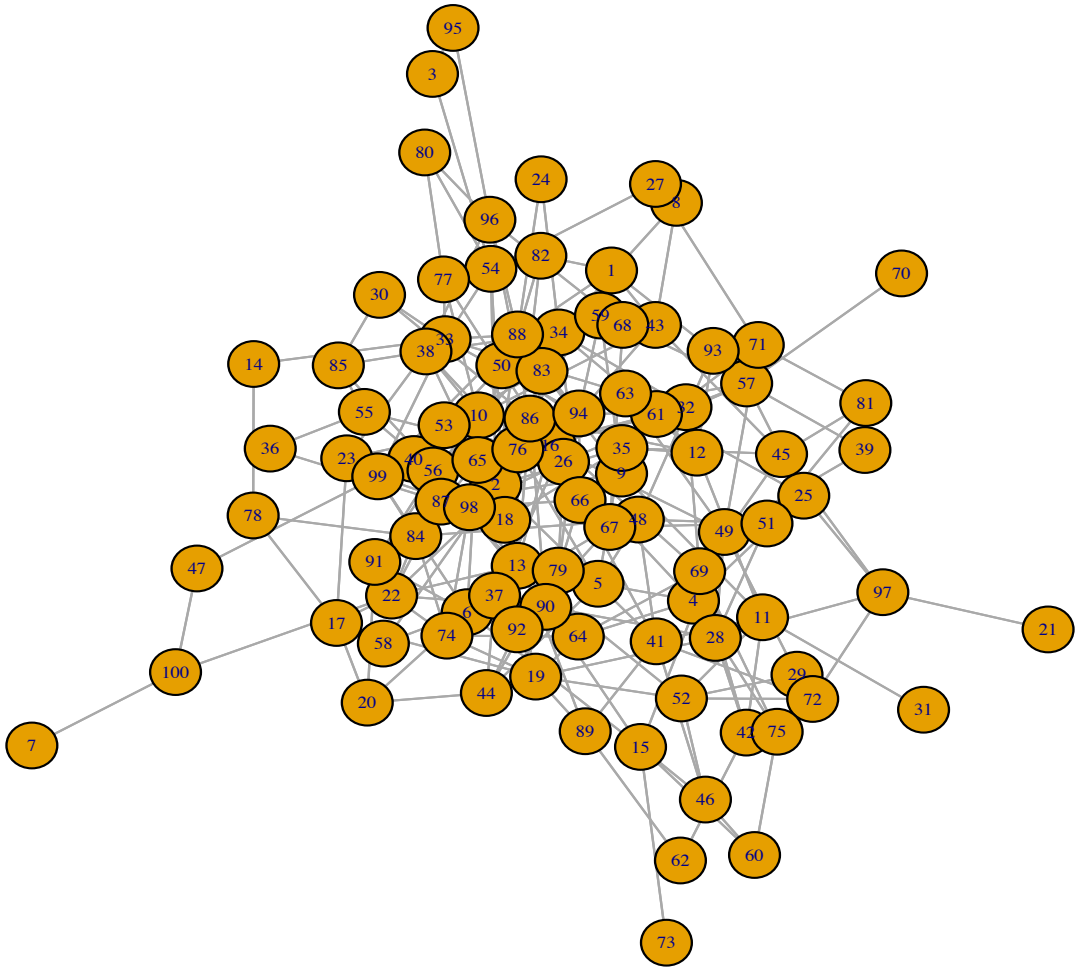
Generally fall into two types (or can be a combination):

- a **network-based** contact measure between two individuals
- a **distance (e.g. spatial)** measure between two individuals

UNIVERSITY OF CALGARY

Simple Network Infection kernel: $\kappa(i, j)$

A network can be represented by a graph. For example:



Simple Network Infection kernel: $\kappa(i, j)$

Mathematically this is typically represented by an **adjacency** or **contact matrix**:

$$C = \begin{bmatrix} 0 & c_{12} & c_{13} & \dots & c_{1n} \\ c_{21} & 0 & c_{23} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & & \vdots \\ c_{n1} & c_{n2} & c_{n3} & \dots & 0 \end{bmatrix}$$

Note that the diagonal entries $c_{11}, c_{22}, c_{33}, \dots, c_{nn}$ are set equal to zero

Simple Network Infection kernel: $\kappa(i, j)$

Consider a simple contact measure, c_{ij} as follows.

$$c_{ij} = c_{ji} = \begin{cases} 1 & \text{if individuals } i \text{ and } j \text{ know each other} \\ 0 & \text{otherwise} \end{cases}$$

Then let

$$\kappa(i, j) = c_{ij}$$

Disease transmission can only occur through individuals connected within this network (unless $\epsilon(i, t) > 0$)

This is an example of a **undirected, binary** contact network

Simple Network Infection kernel: $\kappa(i, j)$

We can also allow for directed networks:

undirected: $c_{ij} = c_{ji}$
directed: $c_{ij} \neq c_{ji}$

We can allow for weighted networks:

instead of c_{ij} being either 0 or 1 (i.e., binary)
we could have c_{ij} being any real number (i.e., continuous), often between 0 and 1

E.G., a weighted, directed network to model supply or trade between farms

Multiple Networks Infection kernel: $\kappa(i, j)$

We can also consider multiple networks representing different facets of our disease system as follows:

Then let

$$\kappa(i, j) = c_{ij}^{(0)} + \beta_1 c_{ij}^{(1)} + \beta_2 c_{ij}^{(2)} + \dots$$

Each $c^{(\cdot)}$ represents a different network

Each of these networks can any combination of directed or undirected, and binary or weighted

UNIVERSITY OF CALGARY

Multiple Networks Infection kernel: $\kappa(i, j)$

Note that the first network, $c_{ij}^{(0)}$ has no parameter associated with it, but the later ones do

This is because the "strength" of the first network is described by the baseline α_0 in the susceptibility function

Simple Distance Infection kernel: $\kappa(i, j)$

Various measures of distance d_{ij} between susceptible individual i and infectious individual j could be used:

Spatial measures:

Euclidean (straight line distance on map)
geographical distance (allowing for curvature of the earth)
road-distance

other measures:

genetic/antibody similarity
travel time-based distance
number of nodes in the network that need to be visited to connect i and j

Euclidean distance is most common of these in use by far

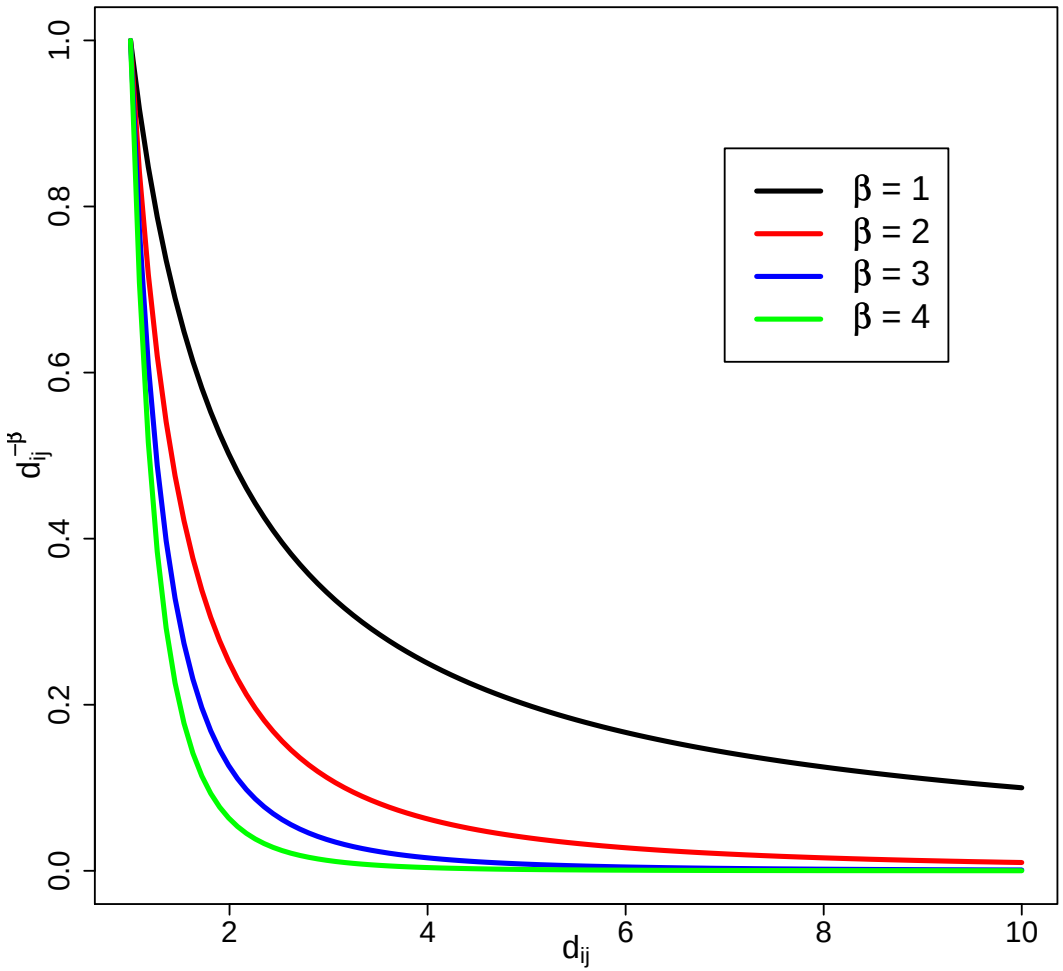
Simple Distance-based Infection kernel: $\kappa(d(i,j))$

Typically a simple relationship such as a **power law** is used to model the effect of distance:

$$\kappa(i,j) = d_{ij}^{-\beta}$$

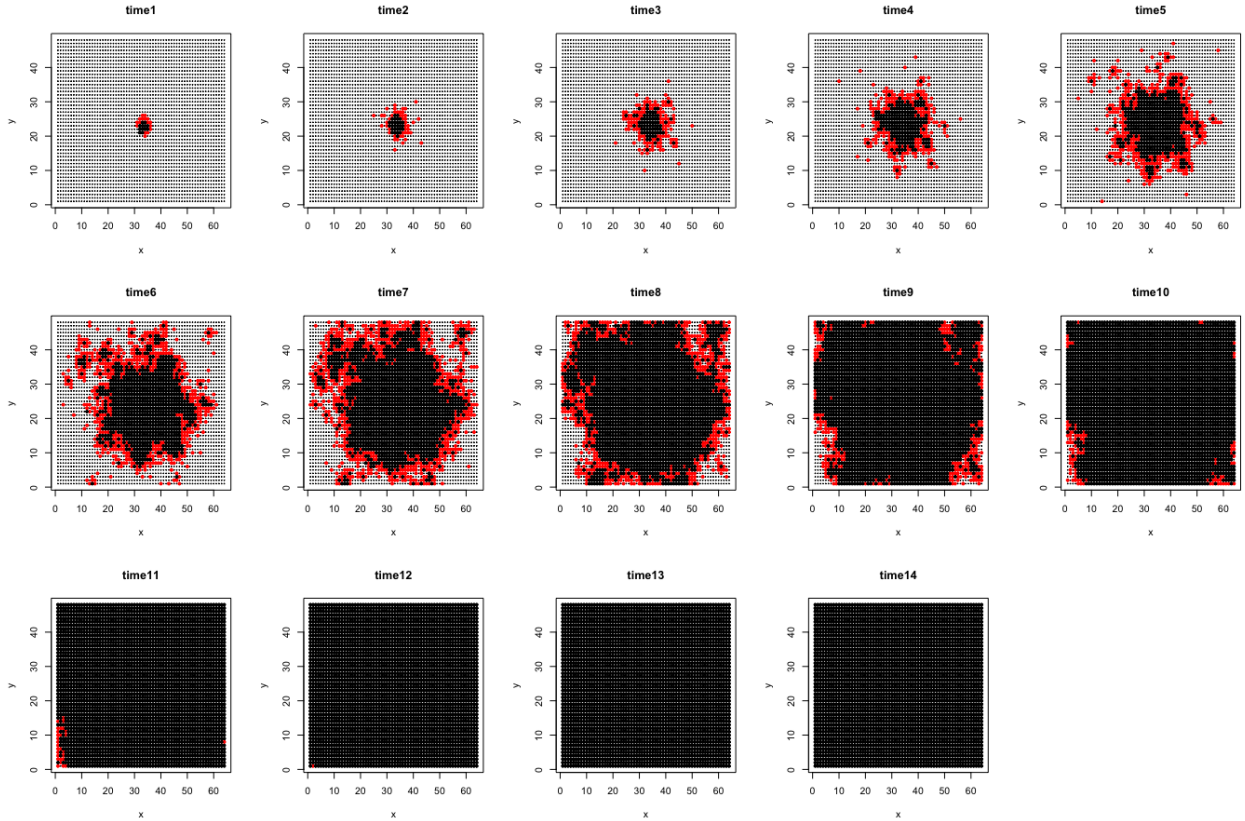
where $\beta \geq 0$ is a spatial parameter

Power-law spatial kernel: $\kappa(i,j) = d_{ij}^{-\beta}$



UNIVERSITY OF CALGARY

Power-law spatial-ILM simulation across grid



From Vrbik et al (2012), *Bayesian Analysis*, 7(3), 615 - 638..

UNIVERSITY OF CALGARY

Alternative spatial kernels

One alternative to the power law kernel is an exponential kernel:

$$\kappa(i,j) = \exp^{-\beta d_{ij}}$$

This is similar in nature to the power law kernel

Another alternative is the **neighbourhood kernel**:

$$\kappa(i,j) = \mathbb{I}[d_{ij} < r]$$

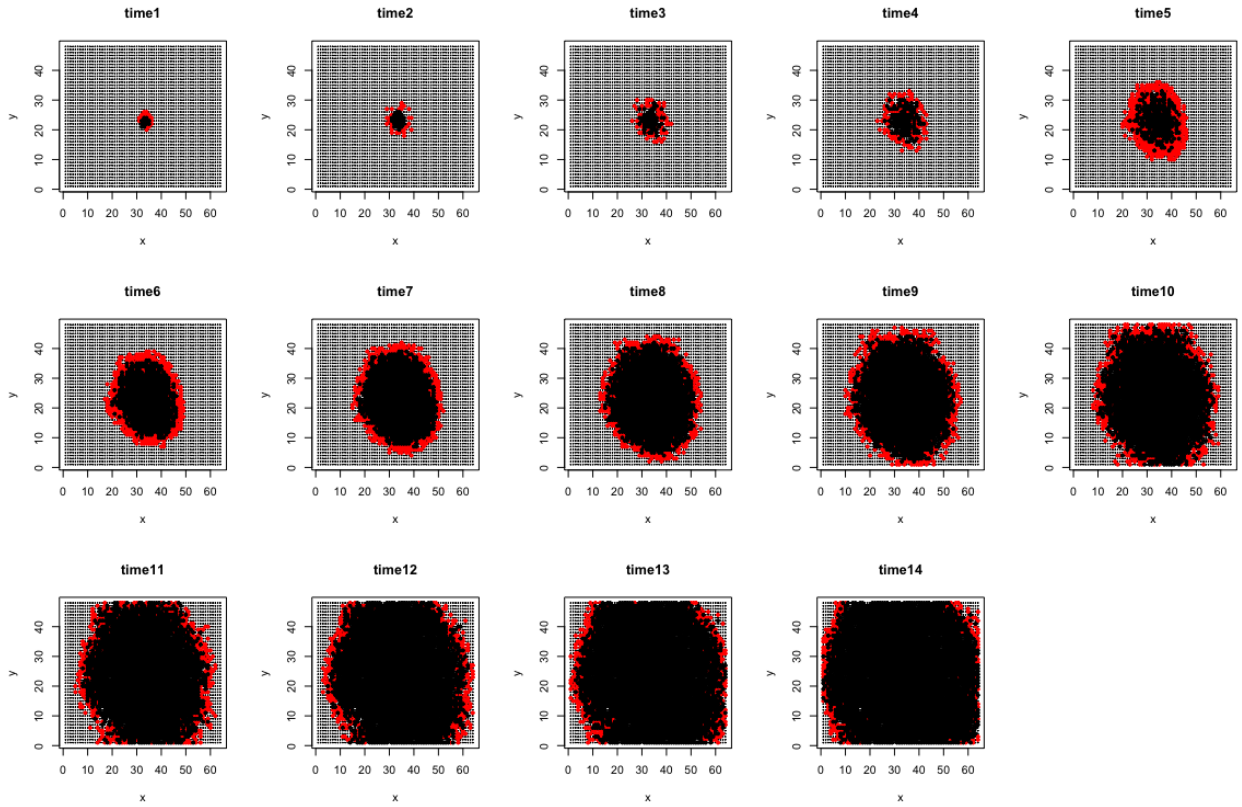
where:

$$\mathbb{I}[d_{ij} < r] = \begin{cases} 1 & \text{if } d_{ij} < r \\ 0 & \text{if } d_{ij} \geq r \end{cases}$$

and

$r > 0$ is a 'spatial' parameter (radius)

Neighbourhood ILM simulation across grid



From Vrbik et al (2012) to appear in Bayesian Analysis, 7(3).

UNIVERSITY OF CALGARY

Outline

1. Individual-Level Models
1.1 General Framework
1.2 Susceptibility
1.3 Transmissibility
1.4 The Infection Kernel
1.5 Sparks Function

2. Some Examples
2.1 Simple Spatial Model
2.2 Spatial Model with Covariates
2.3 Network-based SI Model (Undirected and Unweighted)

3. Extras
3.1 More on Sparks Function

Sparks Function

The sparks function is used to allow for infections **unexplained by other parts of the model**

Often used to represent infection coming in from **outside the observed population**.

It also allows for infections unexplained by an otherwise **poorly fitting model**.

For example:

missing covariates
using a spatial proxy for network
poor functional assumptions

Sparks Function

Typically, $\epsilon(i,t) = \varepsilon$ is used

Thus, every susceptible individual in the population has the same additional chance of being infected at every time point regardless of who they are and when we are in time:

$$P(i,t) = 1 - \exp\left(-\left[\mathcal{S}(i)\left(\sum_{j \in I(t)} \mathcal{T}(j)\kappa(i,j)\right) + \varepsilon\right]\right) \tag{4}$$

Outline

1. Individual-Level Models
1.1 General Framework
1.2 Susceptibility
1.3 Transmissibility
1.4 The Infection Kernel
1.5 Sparks Function

2. Some Examples
2.1 Simple Spatial Model
2.2 Spatial Model with Covariates
2.3 Network-based SI Model (Undirected and Unweighted)

3. Extras
3.1 More on Sparks Function

Software

We have two **R packages** for modelling individual-level disease transmission

First, in discrete time:

V. Warriyar, ***W. Almutiry*** & **R. Deardon**
EpiILM: Spatial and Network Based Individual Level Models for Epidemics.
https://CRAN.R-project.org/package=EpiILM

V. Warriyar, ***W. Almutiry*** & **R. Deardon** (2020) "Individual Level Modelling of Infectious Disease Data: EpiILM" in the *R Journal*, 12(1), 199-217.

https://arxiv.org/abs/2003.04963

Then, in continuous time:

W. Almutiry, ***V. Warriyar*** & **R. Deardon**
EpiILMCT: Continuous Time Individual-Level Models of Infectious Disease.
https://cran.r-project.org/web/packages/EpiILMCT/index.html

W. Almutiry, ***V. Warriyar*** & **R. Deardon** (2021) "Continuous Time Individual-Level Models of Infectious Disease: EpiILMCT" in the *Journal of Statistical Software* 98(10), 1-44. https://www.jstatsoft.org/article/view/v098i10.

https://github.com/waleedalmutiry/EpiILMCT

R: EpiILM

We will focus initially on discrete-time models.

Install the R package EpiILM

```
> install.packages("EpiILM")
```

Call the EpiILM library in R

```
> library(EpiILM)
```

What's inside ?

```
> help(package=EpiILM)
```

Or search on web

https://cran.r-project.org/web/packages/EpiILM/EpiILM.pdf

Outline

1. Individual-Level Models
1.1 General Framework
1.2 Susceptibility
1.3 Transmissibility
1.4 The Infection Kernel
1.5 Sparks Function

2. Some Examples
2.1 Simple Spatial Model
2.2 Spatial Model with Covariates
2.3 Network-based SI Model (Undirected and Unweighted)

3. Extras
3.1 More on Sparks Function

UNIVERSITY OF CALGARY

Simple Spatial Model

Example:

Imagine we are modelling the spread of a transmissible disease through a series of farms.

We have no individual-level covariates about the animals or farms, other than the spatial locations of the farms.

It seems reasonable to treat the farms themselves as individual level unit.

Further infection from outside the population of farms is expected to be negligible/non-existent.

We also assume that once the farm's animals have been infected, that farm will stay infectious over the foreseeable future.

Simple Spatial Model

For this model:

Susceptibility:	$\mathcal{S}(i)$	$=$	α
Transmissibility:	$\mathcal{T}(j)$	$=$	1
Infection Kernel:	$\kappa(i,j)$	$=$	$d_{ij}^{-\beta}$
Sparks:	$\epsilon(i,t)$	$=$	0

So our model is:

$$P(i,t) = 1 - \exp\left(-\alpha \sum_{j \in I(t)} d_{ij}^{-\beta}\right)$$

where

d_{ij} is the spatial distance between individuals i and j
α is an infectivity parameter
β is a spatial parameter

Simple Spatial Model

Parameter settings:

$\alpha = 0.3$ and $\beta = 5$

Assume an SI power-law spatial model

Example: Simulating from Model

File: “RMarkdown Part3a.Rmd”

Outline

1. Individual-Level Models
1.1 General Framework
1.2 Susceptibility
1.3 Transmissibility
1.4 The Infection Kernel
1.5 Sparks Function

2. Some Examples
2.1 Simple Spatial Model
2.2 Spatial Model with Covariates
2.3 Network-based SI Model (Undirected and Unweighted)

3. Extras
3.1 More on Sparks Function

Spatial Model with Binary Covariate

Example:

Imagine we have the same set up as in the previous example, except now we wish to take account of the effect of the biosecurity measures on the farm have upon the susceptibility of the farm.

Consider:

$$X_{bio}(i) = \begin{cases} 0 & \text{if farm } i \text{ is a HIGH biosecurity farm} \\ 1 & \text{if farm } i \text{ is a LOW biosecurity farm} \end{cases}$$

We also assume that farms stop being infectious after 3 time units

Spatial Model with Covariates

For this model:

$$\begin{array}{lrcl} \text{Susceptibility:} & \mathcal{S}(i) & = & \alpha_0 + \alpha_1 X_{bio}(i) \\ \text{Transmissibility:} & \mathcal{T}(j) & = & 1 \\ \text{Infection Kernel:} & \kappa(i,j) & = & d_{ij}^{-\beta} \\ \text{Sparks:} & \epsilon(i,t) & = & 0 \end{array}$$

So our model is:

$$P(i,t) = 1 - \exp\left(-(\alpha_0 + \alpha_1 X_{bio}(i)) \sum_{j \in I(t)} d_{ij}^{-\beta} \right)$$

where

d_{ij} is the spatial distance between individuals i and j
α_0 is a baseline infectivity parameter
α_1 is the biosecurity effect parameter
β is a spatial parameter

Spatial Model with Covariates

Epidemic simulation:

Parameters: $\alpha_0 = 0.3$, $\alpha_1 = 0.6$ and $\beta = 5$

Function argument for covariate information (X_{bio}): "Sformula"

Function argument for infectious periods: "infectious"

Type will now be “SIR”

File: “RMarkdown Part3b.Rmd”

Outline

1. Individual-Level Models
1.1 General Framework
1.2 Susceptibility
1.3 Transmissibility
1.4 The Infection Kernel
1.5 Sparks Function

2. Some Examples
2.1 Simple Spatial Model
2.2 Spatial Model with Covariates
2.3 Network-based SI Model (Undirected and Unweighted)

3. Extras
3.1 More on Sparks Function

UNIVERSITY OF CALGARY

Network-based model (Undirected and Unweighted)

Example:

Imagine we are modelling the spread of a disease through a series of farms in a region.

We have no individual-level covariates about the animals or farms, other than whether the farms share the same supply feed truck.

It seems reasonable to treat the farms themselves as individual level unit.

We also assume that once the farm's animals have been infected, that farm will stay infectious over the foreseeable future.

Whether infection from outside the population of farms is occurring is unknown...

Network-based model (Undirected and Unweighted)

For this model:

$$\begin{array}{lrcl} \text{Susceptibility:} & \mathcal{S}(i) & = & \alpha \\ \text{Transmissibility:} & \mathcal{T}(j) & = & 1 \\ \text{Infection Kernel:} & \kappa(i,j) & = & c_{ij} \\ \text{Sparks:} & \epsilon(i,t) & = & \varepsilon \end{array}$$

So our model is:

$$P(i,t) = 1 - \exp\left(-\left[\alpha\left(\sum_{j \in I(t)} c_{ij}\right) + \varepsilon\right]\right)$$

where

α is an infectivity parameter

$$c_{ij} = c_{ji} = \begin{cases} 1 & \text{if i and j share a supplier} \\ 0 & \text{otherwise} \end{cases}$$

Simple Spatial Model

Parameter settings:

$\alpha = 0.1$ and $\varepsilon = 0$ (to begin)

Assume an SI network-based model

Contact Network (Adjacency) Matrix

Data File: “ISBA_network.csv”

Example: Simulating from Model

File: “RMarkdown Part3c.Rmd”

Outline

1. Individual-Level Models
1.1 General Framework
1.2 Susceptibility
1.3 Transmissibility
1.4 The Infection Kernel
1.5 Sparks Function

2. Some Examples
2.1 Simple Spatial Model
2.2 Spatial Model with Covariates
2.3 Network-based SI Model (Undirected and Unweighted)

3. Extras
3.1 More on Sparks Function

Outline

1. Individual-Level Models
1.1 General Framework
1.2 Susceptibility
1.3 Transmissibility
1.4 The Infection Kernel
1.5 Sparks Function

2. Some Examples
2.1 Simple Spatial Model
2.2 Spatial Model with Covariates
2.3 Network-based SI Model (Undirected and Unweighted)

3. Extras
3.1 More on Sparks Function

Sparks Function

Typically, $\epsilon(i,t) = \varepsilon$ is used

Thus, every susceptible individual in the population has the same additional chance of being infected at every time point regardless of who they are and when we are in time:

$$P(i,t) = 1 - \exp\left(-\left[\mathcal{S}(i)\left(\sum_{j \in I(t)} \mathcal{T}(j)\kappa(i,j)\right) + \varepsilon\right]\right) \tag{5}$$

Sparks Function

However, it often makes sense to allow the sparks function to depend upon susceptible individual i and/or time

For example, consider the situation where the sparks term is representing infections coming in from outside the observed population

Individuals highly susceptible to getting the disease due to some genetic/immunity factor, will be highly susceptible from both the observed population and the population outside

Thus, we might use

$$\epsilon(i,t) = \epsilon(i) = \varepsilon\mathcal{S}(i)$$

Sparks Function

We might also want the chances of infection from outside the observed population to increase as the number of infectious individuals within the observed population increases

i.e., allow the chances of infection from outside to be guided by the observed epidemic curve

Thus, we might use

$$\epsilon(i,t) = \epsilon(t) = \varepsilon|\mathcal{I}(t)|$$

where

$|\mathcal{I}(t)|$ is the number of infectious individuals at time point t

Sparks Function

Or we could do both.

Thus, we might use

$$\epsilon(i,t) = \varepsilon \mathcal{S}(i)|\mathcal{I}(t)|$$

where

$|\mathcal{I}(t)|$ is the number of infectious individuals at time point t