Hierarchical Spatial Structure in Compartmental Epidemiological models

James Bender

Sections

Copyright notice

Abstract

Epidemiological modelling is essential to understand and predict the spread of infectious diseases and plan effective interventions. Disease spread is an inherently spatial process, particularly in diseases that spread through contact with body fluids, and spatially structured epidemiological models provide an opportunity to incorporate a wealth of geographical information into the modelling process. Here, we present a compartmental meta-population model encoding the hierarchical spatial structure of the Australian Statistical Geography Standard (ASGS). Relationships between spatial groupings are represented in a mixing matrix, which influences the spread of infection between patches (i.e. Patches in the same spatial grouping are more likely to spread disease to each other compared to patches outside this grouping). As an example, we model the Greater Melbourne Statistical Area (SA) with patches representing SA2, SA4, and SA4 level groupings, and stochastically simulate the course of an epidemic as a Continuous Time Markov Chain (CTMC). We compare the results of simulations with varying influence of the spatial relationships between patches (through the construction of the mixing matrix), different levels of spatial resolution (which level of the spatial hierarchy is represented by compartmental patches), and different infectious disease parameters (e.g. R_0), as well as with the results of a spatial compartmental model incorporating an alternative source of geographic information (i.e. empirically derived mobility data). We conclude by speculating on the potential use of such geographically informed epidemiological models in the management of infectious diseases.

Declaration

Use only one of the following declarations.

Standard thesis

This thesis is an original work of my research and contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Student name:

Student signature:

Date:

Publications during enrolment

Remove this section if you do not have publications.

The material in sec-intro has been submitted to the journal *Journal of Impossible Results* for possible publication.

The contribution in sec-litreview of this thesis was presented in the International Symposium on Nonsense held in Dublin, Ireland, in July 2022.

Reproducibility statement

This thesis is written using Quarto with renv (Ushey, 2022) to create a reproducible environment. All materials (including the data sets and source files) required to reproduce this document can be found at the Github repository `github.com/SusanSu/thesis`.

Thesis including published works declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes ?? original papers published in peer reviewed journals and ?? submitted publications. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the Department of Econometrics & Business Statistics under the supervision of ??

(The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.)

In the case of (??insert chapter numbers) my contribution to the work involved the following:

I have / have not renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student name:

Student signature:

Date:

I hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor name:

Main Supervisor signature:

Date:

Acknowledgements

I would like to thank my pet goldfish for …

In accordance with Chapter 7.1.4 of the research degrees handbook, if you have engaged the services of a professional editor, you must provide their name and a brief description of the service rendered. If the professional editor's current or former area of academic specialisation is similar your own, this too should be stated as it may suggest to examiners that the editor's advice to the student has extended beyond guidance on English expression to affect the substance and structure of the thesis.

Free text section for you to record your acknowledgment and gratitude for the more general academic input and support such as financial support from grants and scholarships and the non-academic support you have received during the course of your enrolment. If you are a recipient of the "Australian Government Research Training Program Scholarship", you are required to include the following statement:

> "This research was supported by an Australian Government Research Training Program (RTP) Scholarship."

You may also wish to acknowledge significant and substantial contribution made by others to the research, work and writing represented and/or reported in the thesis. These could include significant contributions to: the conception and design of the project; non-routine technical work; analysis and interpretation of research data; drafting significant parts of the work or critically revising it so as to contribute to the interpretation.

Part I

Background

1 Introduction & Background

1.1 Modelling Infectious Epidemics

"Epidemic models are used to inform decisions on disease prevention, surveillance, control and treatment and can be applied to new epidemics" (**bjørnstad2020?**)

1.2 Compartment Models

"Compartmental models are the most frequently used type of epidemic model. In this class of models, individuals can be in a finite number of discrete states. Some of these states are simply labels that specify the various traits of individuals. Of these, some will be changing with time, such as age class, and others will be fixed, such as sex or species. Other states indicate the progress of an infection: for example, an individual can upon becoming infected, typically first enter a state of latency, then progress to a state of infectiousness, and then lose infected status to progress to a recovered/immune state. With each state one can associate the subpopulation of individuals who are in that particular state at the given time (e.g. a female in a latent state of infection). Often the same symbol is used as a label for a state and to denote the corresponding subpopulation size, either as a fraction or as a number (e.g. I or Y for individuals in an infectious state)" - (**diekmann2010?**)

1.2.1 The SIR compartmental model

The simplest comparmental model of infectious disease spread is the SIR Compartmental Model, with three compartments: S - for individuals *susceptible* to the disease; I - for *infected* individuals; and R - for previously infected individuals who have *recovered* (or been otherwise *removed* from that compartment). With the simplifying assumption of a constant population size, N, i.e.

$$N = S + I + R \tag{1.1}$$

Individuals move between compartments in a fixed set of ways, they may either become infected (moving from S -> I), or recover from infection (I -> R).

At each time unit an infected individual can come into contact with, on average, $k\frac{S}{N}$ Susceptible individuals. π is the probability of infecting somebody on coming in contact, so $\beta = k\pi$ is the average rate at which an infected indivudual will infect a susceptible. Infected individuals recover at the constant rate, γ, with $1/\gamma$ the mean recovery time.

Parameter	Interpretation
β	Transmission rate
γ	Recovery rate

Thus, an SIR model can be represented by the schematic

1.3 SIR Dynamics

1.3.1 ODE representation

Given the rate parameters defined above, the change in compartment composition over time can be described by the system of differential equations

$$\begin{aligned} \frac{dS}{dt} &= -\beta I \frac{S}{N} \\ \frac{dI}{dt} &= \beta I \frac{S}{N} - \gamma I \\ \frac{dR}{dt} &= \gamma I \end{aligned} \tag{1.2}$$

Importantly, if the total population size is known, given Equation **??**, R = N - S + I and the entire system can be described by two of the equations in Equation **??**.

EXAMPLE

Limitations of Deterministic ODE

While Equation **??** provides a neat solution for the expected behaviour of a epidemic, it fails to caputure the variability inherent in a complex process like disease spread.

1.3.2 Stochastic Process (Markov Chain) Representation

A Stochastic Process is a collection of random variables, X_t.

Markov chains

A Markov chain is a sequence of random variables $X_0, X_1, ...$ taking values in S with the property that

$$\begin{aligned} P\left(X_{n+1} = j \mid X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = i\right) \\ = P\left(X_{n+1} = j \mid X_n = i\right), \end{aligned} \tag{1.3}$$

for all $x_0, \ldots, x_{n-1}, i, j \in S$, and $n \geq 0$. That is, the state at the next time step is determined only by the state at the current time step.

Time Homogeneity

While not a general property of Markov chains, all of those considered in this work will have the additional property of *time homogeneity*:

$$P\left(X_{n+1} = j \mid X_n = i\right) = P\left(X_1 = j \mid X_0 = i\right),$$

i.e. the probabilities in Equation **??** do not vary with t

SIR DTMC

An SIR Compartment model an be described by a Markov chain of two independent random variables S(t) and I(t) representing the number of susceptible or individuals infected at time t respectively (as in the ODE case, a third RV R(t) denoting the number of recovered individuals is fully determined when the population size is known, and can be left out from this characterisation).

There are two potential events (tbl-SIR_Events) resulting in a change of state i.e. from X_t to $X_{t+\delta t}$

$$p_{(s+k,i+j),(s,i)}(\Delta t) = \begin{cases} \frac{\beta i s}{N}\Delta t, & (k,j) = (-1,1) \\ \gamma i \Delta t, & (k,j) = (0,-1) \\ 1 - \left[\frac{\beta i s}{N} + \gamma i\right]\Delta t, & (k,j) = (0,0) \\ 0 & otherwise \end{cases} \tag{1.4}$$

Transition Matrix

1.3.3 Stochastic simulation

Simoni et al. (2019)

2 Spatial Metapopulation Model

2.1 Motivation

2.2 Formulation

Jacquez et al. (1988)

We employ a basic SIR compartmental model, assuming no births and deaths, with a total of N individuals.

$$S + I + R = N \tag{2.1}$$

We divide the population into n connected sub-populations,hereafter 'patches', with N_i denoting the population of patch i (note that $N = \sum_{i=1}^{n} N_i$).

Let S_i, I_i, R_i respectively denote the number of susceptible, infected and recovered individuals in Patch i (noting also that $N_i = S_i + I_i + R_i$). Also, since $R_i = N_i - S_i - I_i$, we will not explicitly show R_i in the equations hereafter (thought this could be calculated at any time)

Individuals coming into contact with an infective, whether within or without patch, become infected at a rate β.

While individuals do not migrate between patches, infectives from patch i infect - individuals from j with a proportion ϕ_{ij}.

The force of infection in patch i given by

$$\Lambda_i = \beta \sum_{j}^{n} \phi_{ij} I_j$$

And thus the force of infection vector

$$\Lambda = \beta I \Phi \tag{2.2}$$

Where I is the vector of patch infection counts, and Φ is the mixing matrix.

2.2.1 Construction of mixing matrix

Instead of constructing the mixing matrix from empirical OD data, we specified the matrix analytically to encode the hierarchical structure of the SA classification.

The main constraint employed was that patches within the same level L SA region would mix together more strongly than those not. Nevertheless, because of a higher level enclosing group (in the case of Melbourne, the state level grouping), we only need to specify a series of coefficients which correspond to the lowest level grouping in which two patches co-occur. Furthermore, to ensure the matrix rows sum to one, mixing is also proportional to the fraction of the population of the grouping region comprised by the patch.

for any two patches i and j,

i, j are patches

S_i^L is the set of patches in level L N_i is the population of i N_i^L is the population of S_i^L $\xi^L(\delta)$ is the proportion of mixing that occurs within S_i^L but notS_i^{L-1}

The intra level mixing coefficients ξ_L comprise the vector ξ. Constraining the elements of ξ s.t. $\sum_{L=1}^{L} \xi_i = 1$ (where L is the highest level grouping) ensures the rows of the mixing matrix sum to unity.

-More verbosity

-Be more explicit about SA Levels

-Mention Mixing with PPMM

2.3 Implementation

The model was implemented as continuous time markov chain, with discrete state variables

$$S_i(t), I_i(t) \in 0, 1, 2, ...N_i$$

where $t \in [0, \infty)$

at $t = 0$, a randomly selected patch, i, is seeded with $I = 10$ invectives, such that

$$S_i(0) = N_i - 10, I_i(0) = 10$$

and the remaining patches are disease free

$$S_j(0) = N_j, I_j(0) = 0, i \neq j$$

Algorithm 2.1 Quicksort

```
1: procedure QUICKSORT(A, p, r)
2:     if p < r then
3:         q = PARTITION(A, p, r)
4:         QUICKSORT(A, p, q − 1)
5:         QUICKSORT(A, q + 1, r)
6:     end if
7: end procedure
8: procedure PARTITION(A, p, r)
9:     x = A[r]
10:    i = p − 1
11:    for j = p to r − 1 do
12:        if A[j] < x then
13:            i = i + 1
14:            exchange A[i] with A[j]
15:        end if
16:        exchange A[i] with A[r]
17:    end for
18: end procedure
```

at $t > 0$ possible changes in state ('events') are infection at a patch i (if $S_i \geq 1$ & $I \geq 1$), or recovery at patch i (if $I_i \geq 1$). Rates for these events are given in Table **??**

Table 2.1: Event Rates

Event	State change	Rate
Infection in patch i	$\{S_i - 1, I_i + 1\}$	$S_i\Lambda_i$
Recovery in patch i	$\{I_i - 1\}$	γI_i

Due to the Markov property, events occur after a exponentially distributed time with parameter equal to the sum of all event rates

$$\Delta t \sim Exp(\lambda)$$

where

$$\lambda = \sum_{i}^{n} S_i\Lambda_i + \gamma I_i$$

at $t + \delta t$, a random event from Table **??** will occur with probability directly proportional to it's rate.

When the system enters a disease free equilibrium state (i.e. $I = 0$), the implementation is complete.

2.4 Example: Greater Melbourne Area Spatial Metapopulation model

Part II

Current Study

3 Code & Data

4 Package dependencies