



Behavioural Change Piecewise Constant Spatial Epidemic Models



Chinmoy Roy Rahul ^a, Rob Deardon ^{a, b, *}

^a Department of Mathematics and Statistics, Mathematical Sciences Building, University of Calgary, Calgary, T2N 1N4, AB, Canada

^b Faculty of Veterinary Medicine, University of Calgary, 3280 Hospital Dr NW, Calgary, T2N 4Z6, AB, Canada

ARTICLE INFO

Article history:

Received 31 May 2024

Received in revised form 8 October 2024

Accepted 24 October 2024

Available online 12 November 2024

Handling Editor: Dr Yiming Shao

Keywords:

epidemic models

piecewise spatial risk

behavioural change

Bayesian Markov chain Monte Carlo

foot-and-mouth disease

ABSTRACT

Human behaviour significantly affects the dynamics of infectious disease transmission as people adjust their behavior in response to outbreak intensity, thereby impacting disease spread and control efforts. In recent years, there have been efforts to incorporate behavioural change into spatio-temporal individual-level models within a Bayesian MCMC framework. In this past work, parametric spatial risk functions were employed, depending on strong underlying assumptions regarding disease transmission mechanisms within the population. However, selecting appropriate parametric functions can be challenging in real-world scenarios, and incorrect assumptions may lead to erroneous conclusions. As an alternative, non-parametric approaches offer greater flexibility. The goal of this study is to investigate the utilization of semi-parametric spatial models for infectious disease transmission, integrating an “alarm function” to account for behavioural change based on infection prevalence over time within a Bayesian MCMC framework. In this paper, we discuss findings from both simulated and real-life epidemics, focusing on constant piecewise distance functions with fixed change points. We also demonstrate the selection of the change points using the Deviance Information Criteria (DIC).

© 2024 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Infectious disease transmission models can be used to advise policymakers on preparing for and addressing (re)emerging infectious diseases, especially in the absence of ample data from controlled experiments. Given the influence of host behavior on disease transmission and policy effectiveness, there is a growing interest in integrating behavioural change responses to perceived threat from disease into disease transmission models (Verelst et al. [2016]).

A behavioural change epidemic model (BCEM) is characterized by individuals' responsiveness to external information regarding the disease, prompting them to adopt one or more preventive measures to mitigate the risk of contracting the disease (Funk et al. [2010]). Individuals' responses to external information can be categorized as either global, universally accessible and pertinent to all individuals, such as news disseminated by media channels and information provided by public health authorities, or local, contingent on physical or social proximity to the information source. Moreover, this information may pertain to actual risks (prevalence-based), perceptions of these risks (belief-based), or a combination of both (Funk et al. [2010]). Most research on BCEMs has concentrated on population-averaged frameworks, valued for their computational

* Corresponding Author

efficiency and applicability to large populations. However, these models assume uniform susceptibility and mixing patterns across individuals, potentially overlooking the need for more nuanced models in certain diseases and smaller populations to accurately represent transmission dynamics (Ward et al. [2023b]). There is growing interest in individual-level models within the behavioural change (BC) modeling literature due to their ability to incorporate population and behavioural heterogeneity, address clustering of vaccine sentiment, and analyze stochastic and localized outbreaks of infectious diseases, particularly in highly vaccinated populations, as in the case of diseases such as measles. Additionally, individual-level models are well-equipped to simulate social distancing behaviors, such as reduced contacts, as a preventive measure, particularly in the context of an underlying contact structure (Verelst et al. [2016]).

A compartmental framework is commonly employed in infectious disease models to depict how individuals transition between different disease states. For instance, at any given time point in an SIR framework, the susceptible (S) class comprises individuals vulnerable to infection, the infectious (I) class includes potential transmitters of the disease, and the removed (R) class represents those who have recovered with immunity, been isolated, or died.

Typically, inference for these models is conducted within a Bayesian Markov chain Monte Carlo (MCMC) framework, which combines the robust mathematical principles of Bayesian analysis with the computational power of MCMC methods, known for effectively handling intricate models (e.g., Robert and Casella [2002]).

In modeling infectious disease spread while considering behavioural change, parametric models are frequently employed, depending on strong underlying assumptions regarding disease transmission mechanisms within the population (Ward et al. [2023a,b]). However, making appropriate parametric assumptions can be challenging in real-world scenarios, and incorrect assumptions may lead to erroneous conclusions (Gardner et al. [2011]). As an alternative, non-parametric approaches offer greater flexibility by not assuming a simple parametric model (Kypraios and O'Neill, 2018, Rahul and Deardon [2024]).

In this paper, we establish a framework for basic non-parametric spatial BC infectious disease transmission models using piecewise constant spatial kernels within a Bayesian Markov Chain Monte Carlo (MCMC) framework, building upon the ILM framework outlined in Rahul and Deardon [2024]. These BC-ILMs enable dynamic changes in susceptibility or transmission rates driven by prevalence-based behavioural change effects.

The layout of the rest of the paper is as follows. In Section 2, we introduce our proposed semi-parametric spatial BC model and inference method. In Section 3, we demonstrate the simulation process and the findings in detail. In Section 4, we apply our model to 2001 UK foot-and-mouth disease (FMD) data. Then lastly, in Section 5, we recap the study and offer plans for further research.

2. Methodology

2.1. The General ILM

To understand the spatiotemporal dynamics of infectious disease transmission in heterogeneous populations, individual level models (ILMs) of transmission can be used (Deardon et al. [2010]). In spatial ILMs, individuals are considered separate entities across both time and space, with events occurring at discrete time t representing their occurrence within the continuous time interval $[t, t + 1)$.

This paper examines SIR (susceptible-infectious-removed) and SEIR (susceptible-exposed-infectious-removed) compartmental frameworks in simulated and real-world data scenarios, respectively. According to the SEIR framework, a person can be susceptible (S), exposed (E), infectious (I), or removed (R) at any discrete time point. Susceptible individuals have not yet been infected, while exposed individuals are infected but not yet contagious, experiencing a latent period (the length of which is denoted by μ_E). Infectious individuals can transmit the disease and are actively contagious for a duration known as the infectious period (denoted by μ_I). Those in the removed class may have recovered with immunity, died, or have been separated from the population, for example, through quarantine. Individuals in any of these states at a specific time are represented as members of the respective sets $S(t)$, $E(t)$, $I(t)$, or $R(t)$. In the SIR framework, the E state is omitted with individuals becoming infectious when they are infected.

Deardon et al. [2010] proposed the following general form of the epidemic ILM:

$$P(i, t) = 1 - \exp \left[\left\{ - \Omega_S(i) \sum_{j \in I(t)} \Omega_T(j) k(i, j) \right\} - \epsilon(i, t) \right]; \Omega_S(i), \Omega_T(j), k(i, j), \epsilon(i, t) \geq 0, \quad (1)$$

where $P(i, t)$ is the probability of a susceptible individual i being infected at discrete time point t . $I(t)$ is the set of individuals who are infectious at time t . $\Omega_S(i)$ and $\Omega_T(j)$ denote susceptibility and transmissibility functions, respectively. These functions encapsulate various risk factors, such as environmental or genetic factors, associated with susceptible individual i contracting the disease and infectious individual j transmitting it. The infection kernel $k(i, j)$ accounts for shared risk factors between pairs of infectious and susceptible individuals, which could be based on spatial distance or connection through a network. A common term used for an infectious kernel derived from spatial distance is a spatial kernel. Additionally, infections not explained by $\Omega_S(i)$, $\Omega_T(j)$, or $k(i, j)$ can be attributed to the sparks term $\epsilon(i, t)$, which often represents potential infections originating from outside the observed population.

2.2. The General Behavioural Change ILM (BC-ILM)

Based on the ILMs of Deardon et al. [2010], Ward et al. [2023b] proposed a new class of BC-ILM. These models incorporate an alarm function, denoted as g_t , that is a function of disease prevalence or incidence. It serves as a proxy for the perceived risk within the population over time. It is assumed that a higher alarm function is accompanied by a higher level of disease-mitigating behaviour. Thus, as the alarm function increases, infection rates should decrease. This alarm function overcomes the problem that arises from not having BC-related data, as it's often hard to get information about the factors that capture behavioural change in reality (e.g., information about wearing a mask, physical distancing, or reducing movement).

According to the framework proposed by Ward et al. [2023b], the alarm function g_t can be integrated into any component of equation (1) in a multiplicative manner that ensures that an increase in g_t leads to a decrease in $P(i, t)$. The general BC-ILM proposed by Ward et al. [2023b] is as follows:

$$P(i, t) = 1 - \exp \left[\left\{ -\Omega_S(i; \mathbf{g}_t) \sum_{j \in I(t)} \Omega_T(j; \mathbf{g}_t) k(i, j; \mathbf{g}_t) \right\} - \epsilon(i, t; \mathbf{g}_t) \right]. \tag{2}$$

2.3. Non-parametric Spatial BC-ILMs

Rahul and Deardon [2024] proposed a new non-parametric ILM framework by setting $\Omega_S(i) = \Omega_T(j) = 1$, $\epsilon(i, t) = 0$, and $k(i, j) = \tilde{k}(i, j)$ to equation (1) to give

$$P_{it}(\theta) = 1 - \exp \left\{ - \sum_{j \in I_t} \tilde{k}(i, j) \right\} \tag{3}$$

where,

$$\tilde{k}(i, j) = \begin{cases} \tilde{k}_l(d_{ij}) & \delta_{l-1} \leq d_{ij} < \delta_l; l = 1, \dots, n-1 \\ \tilde{k}_n(d_{ij}) & d_{ij} \geq \delta_{n-1}; l = n, \end{cases} \tag{4}$$

and where, $\delta_0, \delta_1, \dots, \delta_{n-1}$ are change points and $\delta_0 = 0$.

They considered $\tilde{k}_l(d_{ij}) = \alpha_l$ where $\alpha_l \in \mathbb{R}^+$, to define an **n-step piecewise constant model**, with infection kernel:

$$\tilde{k}(i, j) = \begin{cases} \alpha_1 & \delta_0 \leq d_{ij} < \delta_1, \\ \alpha_2 & \delta_1 \leq d_{ij} < \delta_2, \\ \vdots & \vdots \\ \alpha_{n-1} & \delta_{n-2} \leq d_{ij} < \delta_{n-1}, \\ \alpha_n & d_{ij} \geq \delta_{n-1}. \end{cases} \tag{5}$$

Here, similar to Ward et al. [2023b], we incorporate a smooth exponential alarm function in the equations (3) and (5), where the impact of behaviour change increases as prevalence grows. Specifically, the alarm function is given by,

$$g_t = 1 - \exp(-\lambda |I_{t-1}|), \tag{6}$$

where $\lambda > 0$ determines the rate at which g_t will approach an asymptote of one, and $|I_{t-1}|$ is the prevalence at time $t - 1$. Examples of this alarm function are shown in Figure A.1 (see Appendix).

Then, our semi-parametric BC-ILM framework is given by:

$$P_{it}(\theta) = 1 - \exp \left\{ - \sum_{j \in I_t} \tilde{k}(i, j; g_t) \right\} \tag{7}$$

where,

$$\tilde{k}(i, j; g_t) = \begin{cases} \tilde{k}_l(d_{ij}, g_t) & \delta_{l-1} \leq d_{ij} < \delta_l; l = 1, \dots, n-1 \\ \tilde{k}_n(d_{ij}, g_t) & d_{ij} \geq \delta_{n-1}; l = n, \end{cases} \tag{8}$$

and where, $\delta_0, \delta_1, \dots, \delta_{n-1}$ are change points and $\delta_0 = 0$.

Thus, we can derive an **n-step BC piecewise constant model**, with infection kernel given by:

$$\tilde{k}(i, j; g_t) = \begin{cases} \alpha_1 \times (1 - g_t) & \delta_0 \leq d_{ij} < \delta_1, \\ \alpha_2 \times (1 - g_t) & \delta_1 \leq d_{ij} < \delta_2, \\ \vdots & \vdots \\ \alpha_{n-1} \times (1 - g_t) & \delta_{n-2} \leq d_{ij} < \delta_{n-1}, \\ \alpha_n \times (1 - g_t) & d_{ij} \geq \delta_{n-1}, \end{cases} \tag{9}$$

where, $\alpha_1, \dots, \alpha_n \in \mathbb{R}^+$. We can see that in the BC kernel, the α parameters that define the infection rate over space are multiplied by $(1 - g_t)$. This ensures that as the alarm increases, suggesting protective behavioural change, the infection rates reduce. Note that, our n-step BC piecewise kernel has n-1 change points. Here, predefined change points are established prior to model fitting. However, it is also possible to estimate the location of the change points.

In this context, a random-walk Metropolis-Hastings (Metropolis et al. [1953], Hastings [1970], Chib and Greenberg [1995]) approach is employed for parameter estimation within a Bayesian framework. Let θ be the vector of unknown parameters, and let X be the observed epidemic data recorded at discrete time points $t = 1, \dots, (T - 1)$. For the n-step BC piecewise constant model X consists of information on the location of each individual, and individual event times (e.g., infection and removal times for an SIR compartmental framework). Where relevant X might also contain individual-level covariates.

Then for the compartmental framework SEIR, the likelihood is

$$f(X|\theta) = \prod_{t=1}^{T-1} \left[\prod_{i \in E(t+1) \setminus E(t)} P(i, t) \right] \left[\prod_{i \in S(t+1)} [1 - P(i, t)] \right], \tag{10}$$

where, $E(t + 1) \setminus E(t)$ represents the group of individuals newly exposed at time $t + 1$; and $S(t + 1) \setminus S(t)$ is the set of individuals who are in susceptible state at time $t + 1$.

For an SIR compartmental framework, our likelihood function will be

$$f(X|\theta) = \prod_{t=1}^{T-1} \left[\prod_{i \in I(t+1) \setminus I(t)} P(i, t) \right] \left[\prod_{i \in S(t+1)} [1 - P(i, t)] \right], \tag{11}$$

where $I(t + 1) \setminus I(t)$ is the group of newly infectious or infected individuals at time $t + 1$.

We assume independent marginal priors for our parameters $\theta = (\alpha_1, \alpha_2, \dots, \alpha_n, \lambda)$, and combine with the likelihood as given in equation (10) or (11), depending on the respective compartmental framework (SEIR or SIR), to form the posterior distribution $\pi(\theta|X)$ up to a constant of proportionality. As stated previously, we assume the change-points, $\delta_0, \delta_1, \dots, \delta_{n-1}$, are known.

3. Simulation

Here, we detail a simulation study carried out to evaluate how well the suggested semi-parametric BC constant piecewise kernel spatial ILM characterized the dynamics of infectious diseases under three different scenarios of BC effects in an SIR compartmental framework.

Epidemics were simulated from a BC piecewise constant kernel ILM for fixed change points (δ 's) for three different values of BC parameter λ : $\lambda = 0$ (“no” BC effect or basic constant piecewise kernel), $\lambda = 0.02$ (a “medium” BC effect) and $\lambda = 0.04$ (a “strong” BC effect). Here, we consider three-step ($n = 3$) and seven step ($n = 7$) cases for the constant piecewise spatial kernel. For three-step cases, we consider a population of size $N = 250$ located with spatial coordinates $\underline{x} = (x_1, \dots, x_{250})$ and $\underline{y} = (y_1, \dots, y_{250})$, generated from independent uniform distributions $x \sim U[0, 10]$ and $y \sim U[0, 10]$. For seven-step cases, we consider a population of $N = 400$ individuals and the spatial coordinates sampled randomly from uniform distribution over a 25×25 unit square; i.e., $x \sim U[0, 25]$, $y \sim U[0, 25]$. The parameters utilized in the three-step and seven-step cases under various BC effect scenarios are shown in Figures 2 and A.5 to A.9 (see Appendix), respectively. Three randomly chosen individuals were selected to initiate the epidemics, with their infectious periods commencing at $t = 1$ and the infectious period $\mu_I = 3$ time units for all individuals. The simulations were carried out for a duration of $T = 20$ time units.

We used “vague” positive half-normal priors for the α 's ($\alpha_l \sim N^+(0, 10^5)$; $l = 1, 2, \dots, n$) and a “weakly informative” gamma distribution prior for λ ($\lambda \sim \text{Gamma}(1, 5)$). Note, we also explored alternative priors for λ to assess their influence on the posterior estimates. Specifically, we tested “weakly informative” and “vague” half-normal priors (e.g., $\lambda \sim N^+(0, 8)$ and $\lambda \sim N^+(0, 10^5)$) but did not obtain noticeably different results from those obtained with the $\text{Gamma}(1, 5)$ prior. Here, we present the results based on the $\text{Gamma}(1, 5)$ prior.

We assessed MCMC convergence by visually inspecting traceplots and utilizing Geweke's diagnostic (Geweke [1992]) for each model parameter. If any chains failed to converge according to Geweke's diagnostic test, then we used the Gelman-Rubin diagnostic (Gelman and Rubin [1992]) using three MCMC chains. We then calculate the posterior mean, median, and 95% percentile intervals (PIs) for each model parameter. Figures A.2 - A.4 (see Appendix) show the spatiotemporal dynamics of

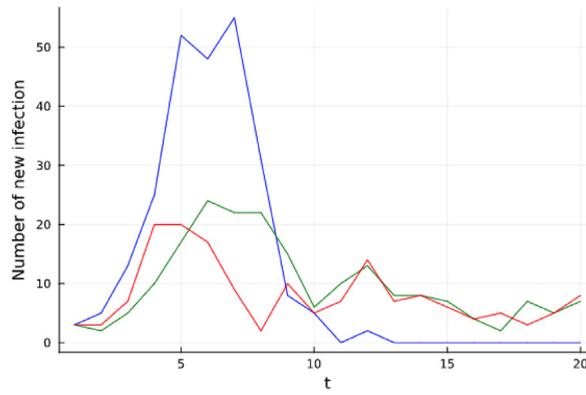


Figure 1. Epidemic curve under three different BC scenarios, where the blue, green, and red curves represent no ($\lambda = 0$), medium ($\lambda = 0.02$) and strong ($\lambda = 0.04$) BC effects, respectively.

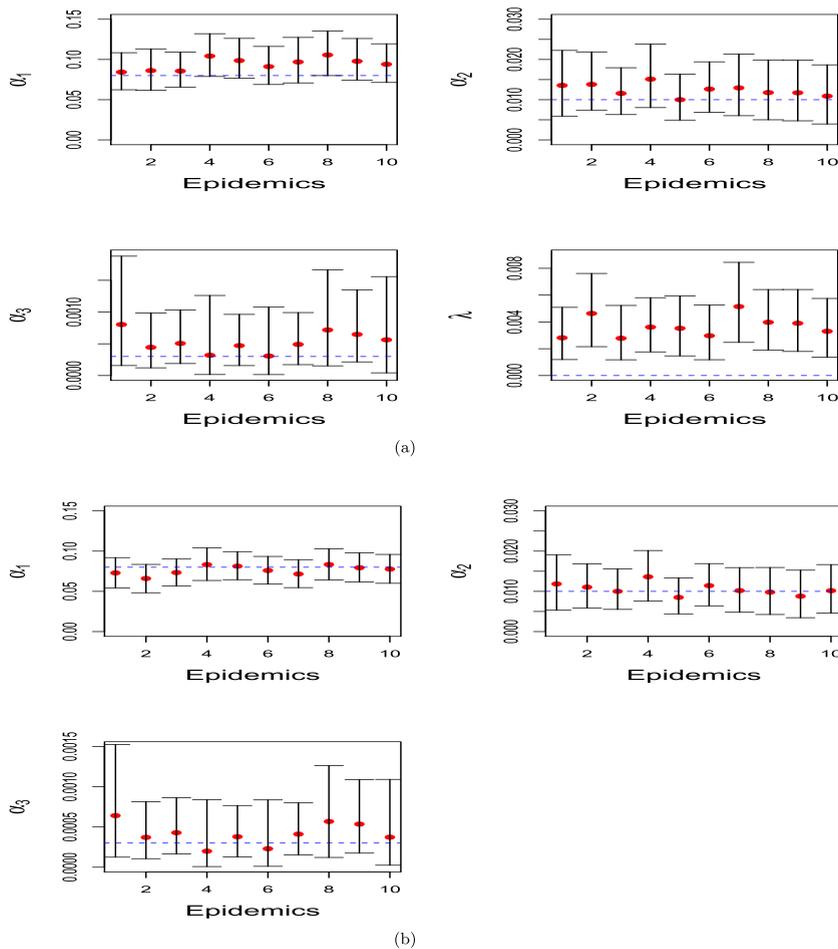


Figure 2. Posterior medians (red points) and 95% PIs for (a) α_1 , α_2 , α_3 and λ and (b) α_1 , α_2 , α_3 for 10 different simulated epidemics when the change points (δ 's) are considered fixed and λ are considered to be included or not in the model. The true parameter values $\alpha_1 = 0.08$, $\alpha_2 = 0.01$, $\alpha_3 = 0.0003$ and $\lambda = 0$ for three-step cases with fixed $\delta_1 = 1.5$, and $\delta_2 = 3$ are represented by the blue dashed line.

typical epidemics under the model and Figure 1 the epidemic curves of those epidemics, generated from the three-step piecewise constant kernel model with BC effects $\lambda = 0, 0.02$ and 0.04 , respectively.

Figure 2 shows the posterior medians and 95% PIs for the parameters of both the standard piecewise constant ILM and the BC piecewise constant ILM across ten distinct epidemics and populations. When $\lambda = 0$ (no BC) under three-step cases, all the parameters (i.e., α_1, α_2 , and α_3) are well estimated in the non-BC ILM, and all the 95% PIs capture the true parameter values that were used to generate the epidemics. In three-step cases the posterior medians of all parameters (apart from λ itself) closely approximate their true values, and all 95% PIs encompass the true parameter values employed for generating the epidemics. Thus, the incorporation of the BC effect into the model when it is not required still results in good estimates of the α parameters, as well as low estimates of λ itself.

Figure A.5 (see Appendix) shows the posterior medians and 95% PIs for the parameters of the BC and non-BC piecewise constant ILMs for ten distinct epidemics and populations, when $\lambda = 0.02$ (medium BC effect) under three-step cases. In three-step cases where BC is incorporated, the posterior medians of all the parameters are close to their true values, and most of the 95% PIs contain the true parameter values ($\alpha_1, \alpha_2, \alpha_3$, and λ) employed for generating the epidemics. However, in the non-BC models, α_1 and α_2 posterior estimates underestimate the true parameter value, and the 95% PIs do not encompass the true parameter values.

Figure A.6 (see Appendix) shows the posterior medians and 95% PIs for the parameters of the BC and non-BC piecewise constant model across ten distinct epidemics and populations, when $\lambda = 0.04$ (strong BC effect) in three-step scenarios. Once again, the true parameter values (i.e., $\alpha_1, \alpha_2, \alpha_3$, and λ) used to generate data for the three-step cases with fixed change points mostly fall within their posterior 95% PIs, with posterior medians generally close to the true parameter values. However, most of the α parameters are not well estimated when the BC effect is not incorporated into the model. Overall, this seems to show that we can generally estimate all the parameters from epidemic data generated from BC piecewise constant spatial ILMs well when fitting the true model to data. However, the spatial kernel estimates are biased if we fail to incorporate the BC effect into our model.

Figure 3 shows the DIC values of the models for the three-step cases discussed above for three different values of BC parameter λ for 10 different epidemics. We can see from Figure 3(a) that when there was no BC effect ($\lambda = 0$), the DIC correctly identifies the non-BC model as that with the best fit. However, Figures 3(b)-(c) suggest that when $\lambda = 0.02$ or 0.04 the DIC can successfully identify that the BC mechanism is required in each case.

Figures A.7 to A.9 (see Appendix) show the posterior medians and 95% PIs for the parameters of the BC and non-BC constant piecewise ILMs across ten epidemics, with $\lambda = 0, 0.02$, and 0.04 under seven-step scenarios. These cases demonstrate the performance of our model for a higher number of steps with varying degrees of BC. Once again, in the seven-step scenarios, the true parameter values employed for generating the data are contained within their respective posterior 95% PIs, with posterior medians generally reflecting proximity to the true parameter values.

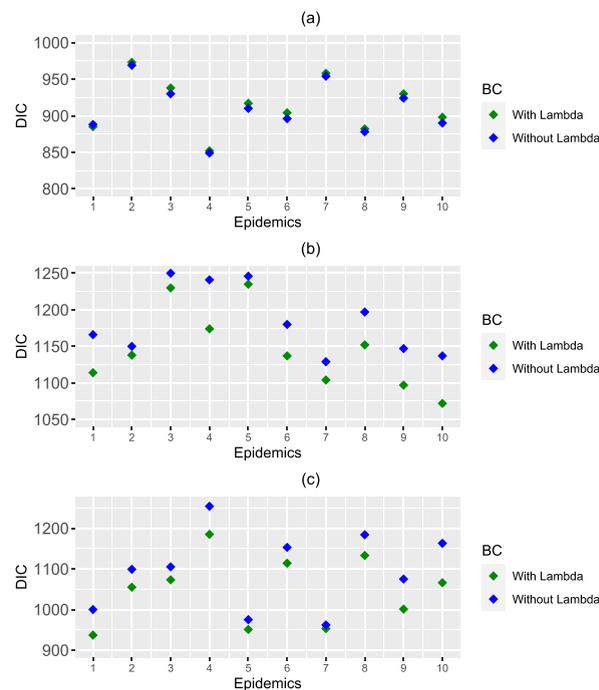


Figure 3. DIC values for three-step cases for 10 different simulated epidemics when the change points (δ 's) are considered fixed and BC effects parameter (a) $\lambda = 0$, (b) $\lambda = 0.02$ and (c) $\lambda = 0.04$ are considered to be included or not in the model.

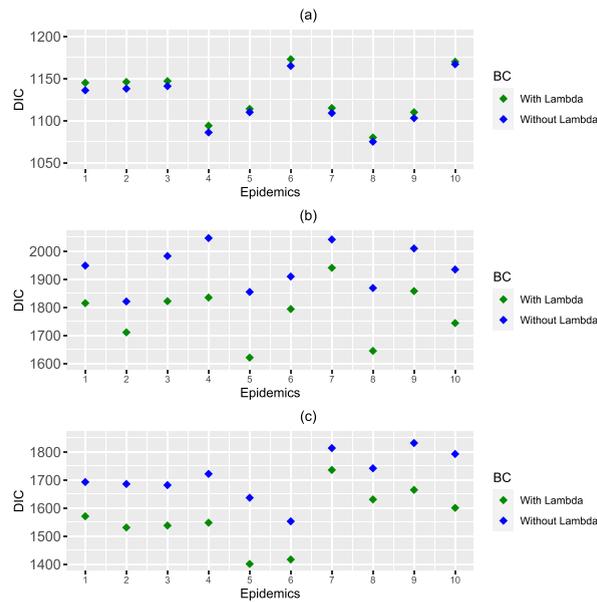


Figure 4. DIC values for seven-step cases for 10 different simulated epidemics when the change points (δ 's) are considered fixed and BC effects parameter (λ) = 0, (b) $\lambda = 0.02$ and (c) $\lambda = 0.04$ are considered to be included or not in the model.

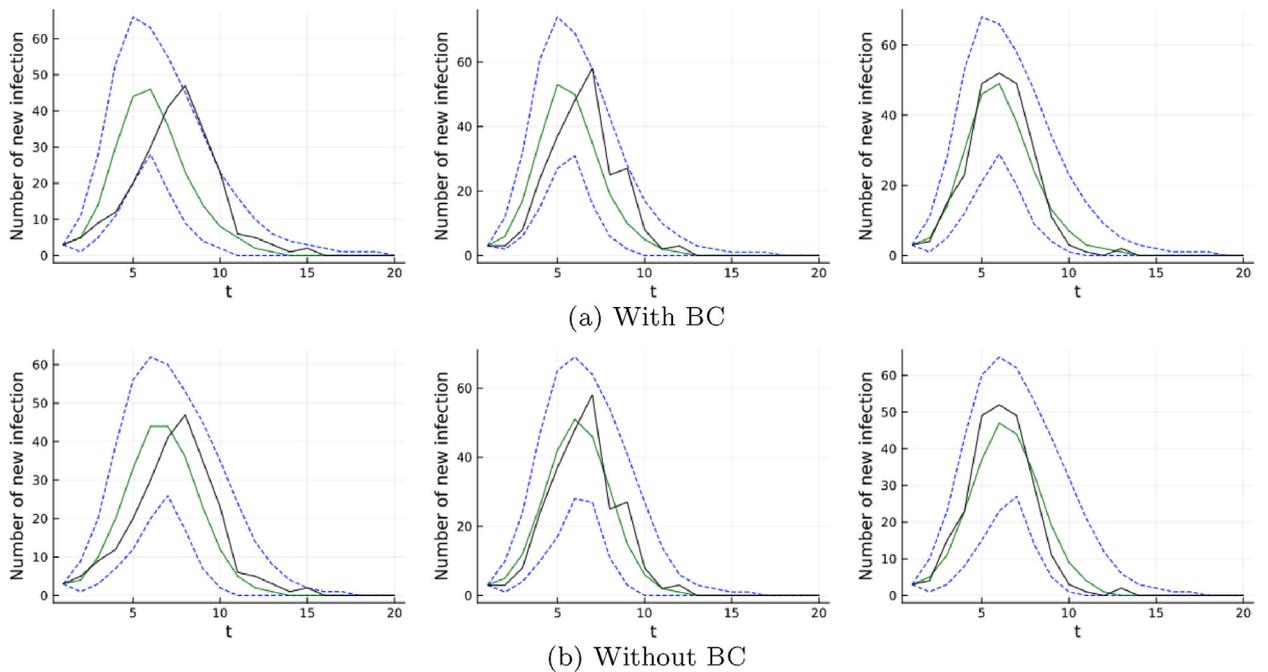


Figure 5. Posterior predictive distribution of the epidemic curve for three typical simulated epidemics for three-step cases with incorporating BC effects (a) $\lambda = 0$ and (b) without BC effect. The black solid line represents the true epidemic curve, the green solid line represents the estimated median curve and the blue dotted lines represent the 95% PI, respectively.

Figure 4 displays the DIC values of the models for seven-step cases across ten different epidemics, considering three different values of the BC parameter λ . From Figure 4(a), it is evident that when there was no BC effect ($\lambda = 0$), the DIC correctly identifies the non-BC model as having the best fit. However, Figures 4(b)–(c) indicate that when $\lambda = 0.02$ or 0.04 , the DIC effectively demonstrates the necessity of the BC mechanism in each case.

3.1. Posterior Predictive Ability

In this section, we present results regarding the posterior predictive distribution (PPD) of the epidemic curve for three randomly chosen epidemics, typical of those generated from the model with and without BC effects for three-step and seven-step cases, in order to evaluate the proposed model's predictive accuracy across various scenarios of BC effects, namely, $\lambda = 0$, 0.02 , and 0.04 .

Figure 5 shows the PPD of the epidemic curve for three simulated epidemics for both the BC and non-BC piecewise constant ILMs, when $\lambda = 0$ (no BC effect) under three-step cases. In three-step cases, under the non-BC ILMs, the true epidemic curves are very well enveloped by the 95% PIs of the epidemic curves under the true model. In addition, the true epidemic curves are also mostly enveloped by the 95% PIs of the epidemic curves under the true model when BC is incorporated in the same scenario. This suggests that the epidemic simulation remains reasonably robust even when the incorporation of the BC effect into the model is not necessary.

Figure A.10 (see Appendix) shows the PPD of the epidemic curve for three simulated epidemics for both the BC and non-BC piecewise constant ILMs, when $\lambda = 0.02$ (medium BC effect) under three-step cases. In three-step cases where BC is incorporated in the model, the true epidemic curves are very well enveloped by the 95% PIs of the epidemic curves under the true BC piecewise constant spatial model. However, the true epidemic curves are not fully captured by the 95% PIs of the epidemic curves under the piecewise constant spatial model when BC is not considered in the model.

Similarly, Figure A.11 (see Appendix) shows the PPD of the epidemic curve for three simulated epidemics for both the BC and non-BC piecewise constant ILMs, when $\lambda = 0.04$ (strong BC effect) under three-step cases. The true epidemic curves are very well enveloped by the 95% PIs of the epidemic curves under the true BC piecewise constant spatial model for three-step cases. However, the true epidemic curves are not entirely captured by the 95% PIs of the epidemic curves under the piecewise constant spatial model ignoring the BC effect. Further, the oscillatory nature of the epidemic curves are completely missed. Overall, this indicates that our approach is adaptable enough to capture epidemic dynamics when the true kernel is based on the spatial BC piecewise constant kernel. However, it fails to capture the epidemic dynamics if we neglect to integrate the BC effect into our model.

Figures A.12 – A.14 (see Appendix) show the PPD of the epidemic curve for three simulated epidemics for both the BC and non-BC piecewise constant ILMs, when $\lambda = 0$, 0.02 and 0.04 under seven-step cases. Once again, the results justify that our approach is flexible enough to capture epidemic dynamics with correct model (BC or non-BC) which generated the data.

4. UK 2001 Foot-and-Mouth Disease (FMD) Data

Here, we utilize the BC constant piecewise kernel models outlined in Section 2.3 within an SEIR framework to analyze the foot-and-mouth disease (FMD) data from the 2001 epidemic in the UK. We contrast the outcomes of these models with those obtained using the standard piecewise constant ILM as given in equations (3) and (5).

We analyze a portion of the UK FMD dataset from the county of Cumbria, located in the northwest of England, as used in Rahul and Deardon [2024]. This subset comprises 1,079 individual farms housing cattle and/or sheep, along with their corresponding x/y Cartesian coordinates. Each farm is considered to be an individual-level unit, with no consideration given to farm size or type here. Within the SEIR framework, when a farm initially susceptible to infection becomes infected, it enters an exposed or latent period before being capable of transmitting the infection to others. We assume a fixed and known latent period of 5 days ($\mu_E = 5$) before transitioning from the exposed (E) state to the infectious (I) state. The infectious period is determined by the cull date for each farm; if no cull date is available, we assign an infectious period of 4 days ($\mu_I = 4$). Our analysis focuses on a temporal subset of the complete dataset, encompassing infections where the recorded exposure times fall between 28 and 66 days after the initial infection on February 19, 2001. The national ban on animal movements was implemented on February 23, 2001 (Chis Ster and Ferguson, 2007). So here, analysis is conducted using data collected well after this date, and any observed behavioural change effect detected would be in addition to the ban on animal movements.

The model described in equation (7) is adjusted to incorporate a constant sparks term, denoted as $\epsilon(i, t) = \epsilon$. Specifically, our model takes the form:

$$P_{it}(\theta) = 1 - \exp \left\{ - \left(\sum_{j \in I_t} (\tilde{k}(i, j; g_t) + \epsilon) \right) \right\}. \quad (12)$$

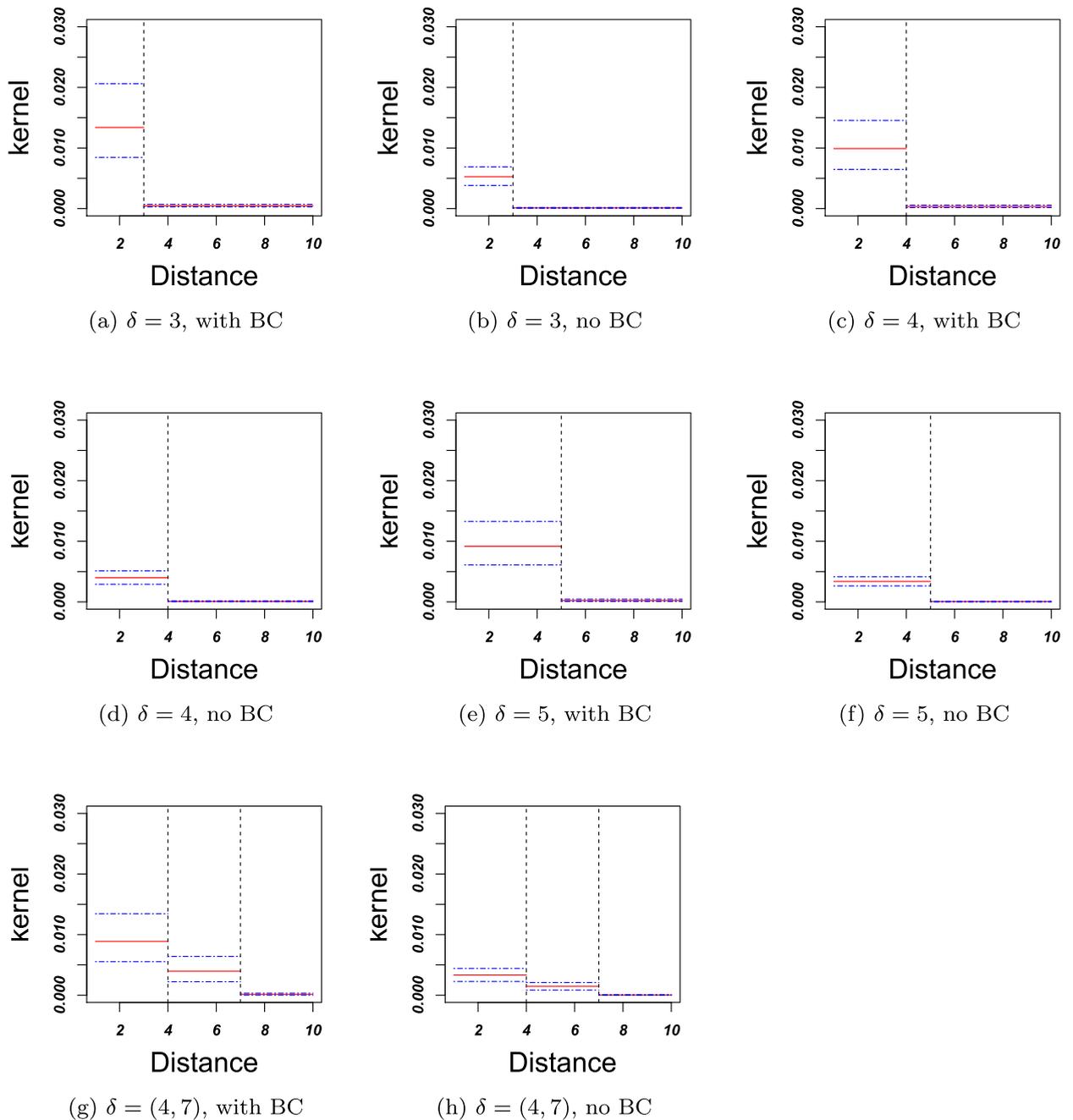


Figure 6. Posterior medians and 95% Pls for different change point (δ) models (a-h) with BC and Non-BC piecewise constant kernel which fitted to FMD data are represented by red and blue dashed lines, respectively.

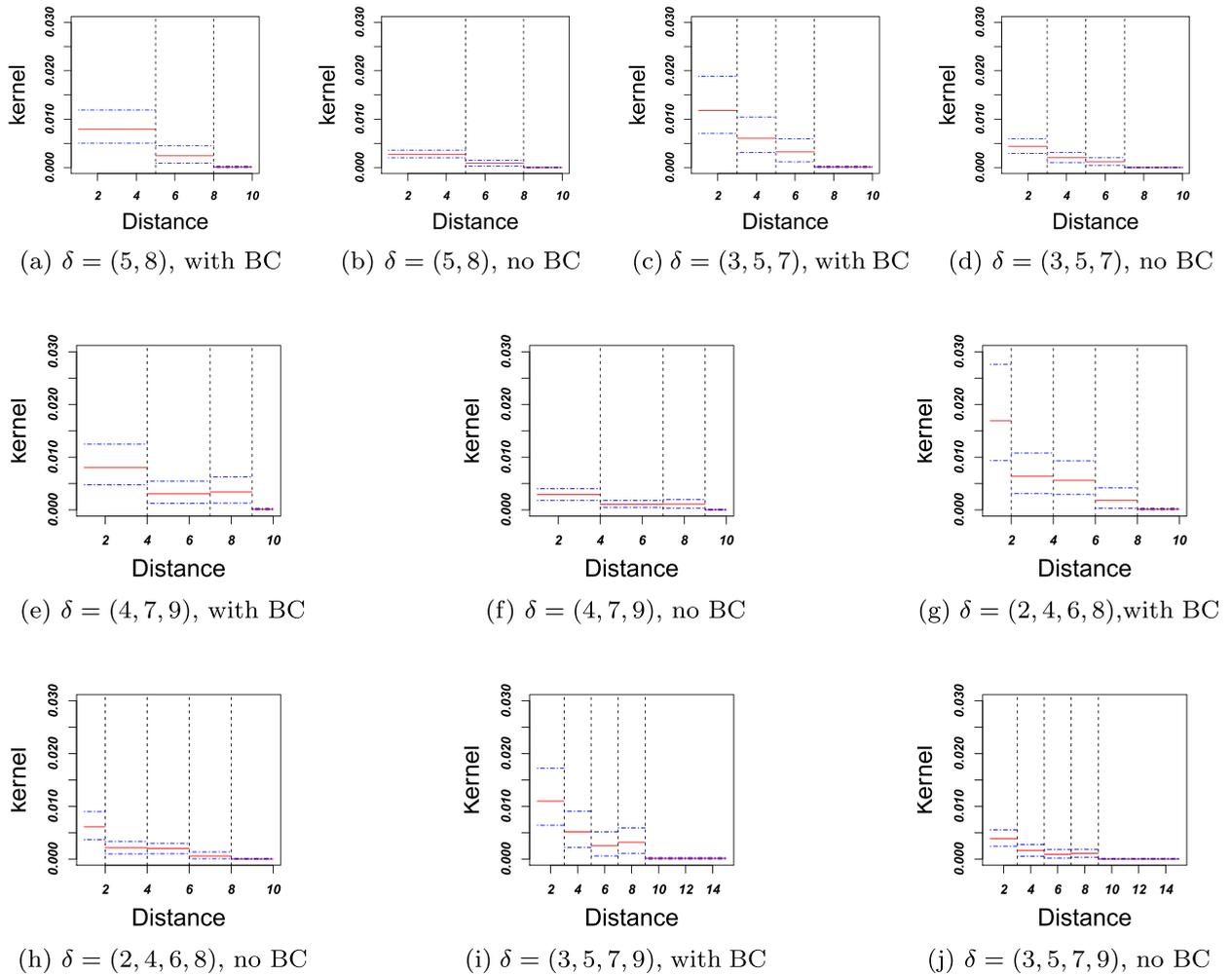


Figure 7. Posterior medians and 95% PIs for different change point (δ) models (a-j) with BC and Non-BC piecewise constant kernel which fitted to FMD data are represented by red and blue dashed lines, respectively.

The inclusion of the sparks term serves to introduce a small, random chance of a farm becoming infected at any particular moment. This accounts for the possibility of infections occurring randomly within the observed subset of the UK foot-and-mouth disease data, potentially originating from sources external to the observed spatial region.

Under our SEIR framework, the form of the likelihood is given by equation (10). For BC piecewise constant kernel models, we once again use “vague” positive half-normal priors for the α 's ($\alpha_l \sim N^+(0, 10^5)$; $l = 1, 2, \dots, n$) and a “weakly informative” gamma distribution prior for λ ($\lambda \sim \text{Gamma}(1, 5)$). We evaluate MCMC convergence using traceplot visual inspection and Geweke’s diagnostic (Geweke [1992]) for each model parameter.

4.1. Results

Figures 6 and 7 show the fitted BC and non-BC piecewise constant kernel models for different change points under the posterior when the models have been fitted to the 2001 UK FMD data. This represents the spatial kernel when $g_t = 0$ (baseline behaviour) under the BC-ILMs. We see that the 95% PI of non-BC piecewise constant kernel models are narrower than those of BC piecewise constant kernel models with the same change points. However, similar spatial dynamics are suggested for all models under baseline behaviour.

Table 1 presents the DIC values for both BC and non-BC piecewise constant kernel models across various change points. It is evident that the BC constant piecewise kernel ILMs exhibit lower DIC values compared to their non-BC counterparts for corresponding change points. However, the discrepancy in DIC values between BC and non-BC models with the same change

Table 1

DIC values and λ estimates (median) with 95% PIs for different semi-parametric BC and DIC values for different semi-parametric non-BC piecewise constant kernel ILMs when fitted to FMD data.

Model	Change point δ	DIC for with BC	DIC for without BC	λ (95% CI)
Two-step	3	2769	2775	0.0385 (0.0256 - 0.0523)
Two-step	4	2758	2765	0.0369 (0.0223 - 0.0496)
Two-step	5	2741	2749	0.0396 (0.0264 - 0.0527)
Three-step	(4,7)	2735	2742	0.0387 (0.0249 - 0.0517)
Three-step	(5,8)	2734	2740	0.0399 (0.0268 - 0.0526)
Four-step	(3,5,7)	2732	2738	0.0401 (0.0262 - 0.0545)
Four-step	(4,7,9)	2727	2735	0.0409 (0.0274 - 0.0539)
Five-step	(2,4,6,8)	2728	2733	0.0414 (0.0285 - 0.0542)
Five-step	(3,5,7,9)	2725	2732	0.0422 (0.0294 - 0.0543)

points is relatively small, ranging from 5 to 8, though this would be considered a “significant” reduction in DIC. Additionally, we note that the DIC values decrease for both BC and non-BC piecewise constant kernel ILMs as the number of change points increases, indicating a preference for greater model flexibility. Specifically, the smallest DIC value of 2725 is attained by the five-step BC model with change points occurring at distances 3, 5, 7, and 9. Table 1 also shows the posterior median estimates of the BC effects (λ) along with their 95% CIs for the same BC piecewise constant kernel models across various change points examined in this analysis. It’s noticeable that the estimates are highly consistent, and the smallest 95% CI is also achieved by the five-step BC model with change points occurring at distances 3, 5, 7, and 9. The fact that $\lambda \approx 0.04$ in each case suggest that the BC mechanism is similarly estimated under each kernel, and that λ estimation is fairly robust to the choice of change points.

Overall, our results suggest evidence of a moderate BC effect. This might well be expected. First we would expect to see alarm (and maybe biosecurity) to increase as the number of cases increases within any region. However, since the major change in behavioural occurred before the window of data observed (i.e., government restrictions on animal movements) that behavioural change might not be expected to be very severe.

5. Discussion

This study introduces a framework for semi-parametric spatial infectious disease individual-level models that incorporate BC piecewise constant kernels, and it contrasts the outcomes with those obtained using standard piecewise constant kernels within a Bayesian Markov Chain Monte Carlo (MCMC) framework. Through simulated epidemic scenarios, we demonstrate that spatial individual-level models (ILMs) utilizing BC piecewise constant kernels can accurately estimate the true parameters when the correct model is applied to the data. Additionally, even when the BC effect is unnecessary, incorporating it into the model still yields accurate parameter estimates, typically with low estimates of the BC effect parameter. Furthermore, the accurate models demonstrate strong posterior predictive ability and effectively capture the true epidemic curve. We also show that there is evidence of BC effects on the spatial dynamics of spread of FMD in the UK 2001 epidemic.

There are various avenues for future research that we could pursue. We have utilized a prevalence-based exponential alarm function in the constant piecewise kernel to model any BC effects. Alternatives exist, however. One possible approach could involve incorporating different types of prevalence-based alarm functions, such as the Hill-type or threshold alarm function (Ward et al. [2023b]). We can also consider alarm functions based on incidence or smoothed averages of prevalence or incidence over time. Here, we may also be interested in optimizing the exact form of the metric (e.g., window size) used in the alarm function.

Up to this point, we have focused on utilizing a BC piece-wise constant kernel. However, there is potential value in exploring the development of semi-parametric models that incorporate BC polynomial piece-wise spatial kernels (e.g., linear or cubic; see, for example, Silverman [1985], Friedman et al. [2001], Kwong et al. [2013]). This approach would offer a different perspective from our current use of BC piece-wise constant kernels.

In this paper, we employed the widely used Deviance Information Criterion (DIC) to carry out model selection. Our simulation study results do seem to suggest that the DIC works well within this context. However, other criteria, such as the Watanabe-Akaike Information Criterion (WAIC), or other variants of the DIC (Deeth et al. [2015]) could also be considered.

Our study exclusively examined models without covariates, yet in real-world scenarios, individual or group-level factors are likely to influence infection transmission or susceptibility. Expanding our model to include covariates would allow for the consideration of more complex dynamics. For instance, while our foot-and-mouth disease models omitted the number of cattle and sheep on the farm to emphasize the spatial aspect of disease spread, these variables are frequently employed as covariates in similar models (Deardon et al. [2010], Ward et al. [2023b]).

In our research, we initially assumed a fixed rate parameter in our alarm function. However, in real-world scenarios, human behavior evolves over time and in response to infection prevention measures. To address this, we could incorporate a

continuous function for the rate parameter, enabling a time-varying alarm function that better captures the effects of changing behavior over time. This might capture effects such as “lockdown fatigue”.

Finally, we explored models with fixed and known latent and/or infectious periods to manage computational complexity and maintain simplicity in our analysis. However, this assumption may not hold true in many scenarios, as these periods can significantly vary between individuals, and the timing of exposure, infection onset, and removal may often be censored. Addressing this uncertainty can involve techniques like data-augmented MCMC, albeit at the expense of significantly increased computational workload.

Chinmoy Roy Rahul: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Conceptualization. Rob Deardon: Writing – review & editing, Validation, Supervision, Methodology, Investigation, Conceptualization

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant program (RGPIN/03292 – 2022) and the Alberta Innovates Advance – NSERC Alliance program (222302037).

Appendix A

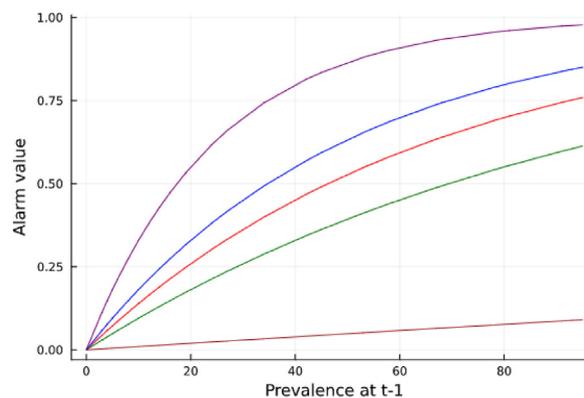


Figure A.1. Exponential Alarm function with different BC effects (λ) where brown, green, red, blue, and purple lines represents alarm values for $\lambda = 0.001, 0.01, 0.015, 0.02,$ and $0.04,$ respectively.

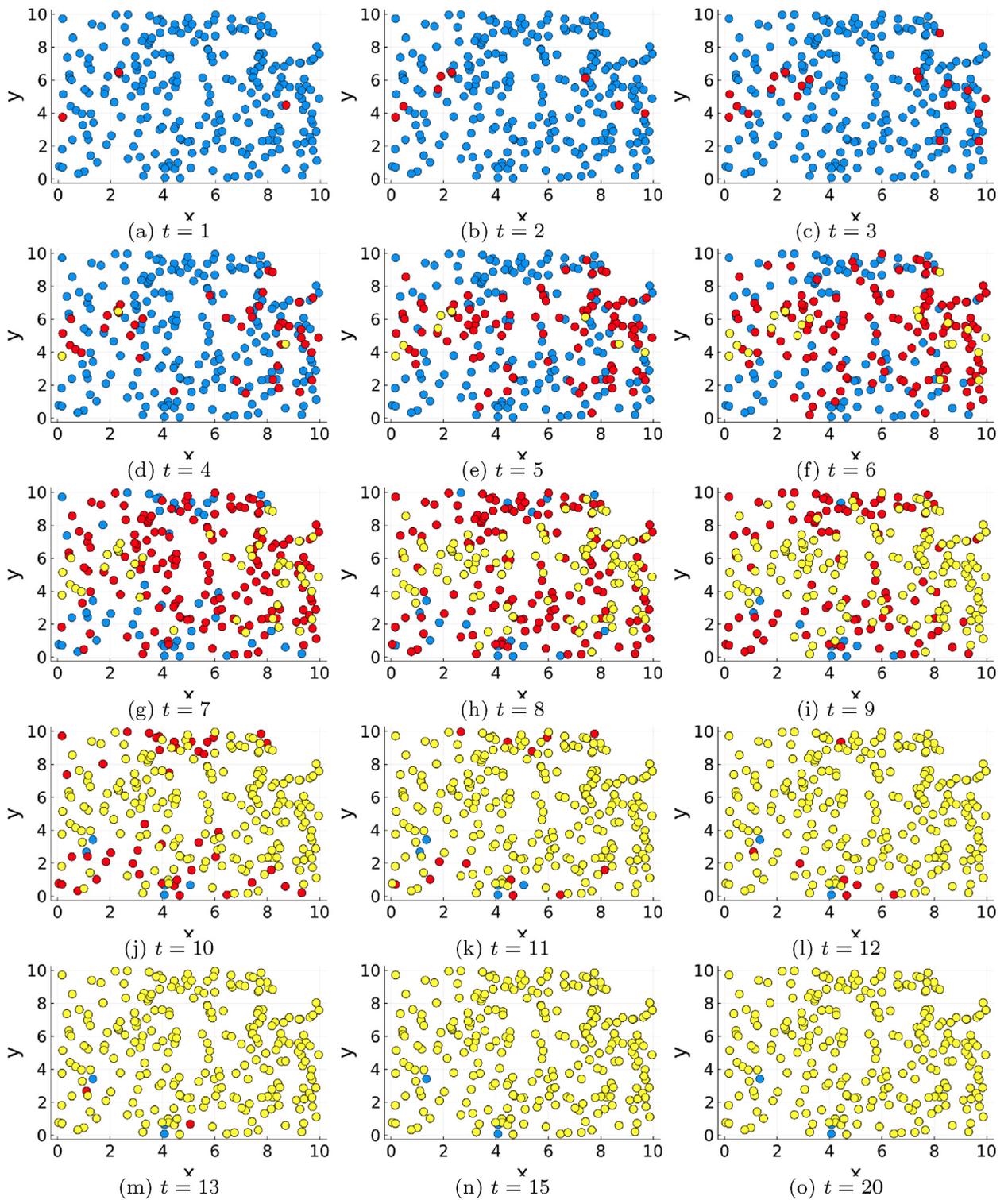


Figure A.2. Typical simulated epidemic realization across a grid of individuals for basic piecewise constant kernel ILM, where susceptible individuals are denoted by blue circles, infected individuals are denoted by red circles and removed individuals are denoted by yellow circles.

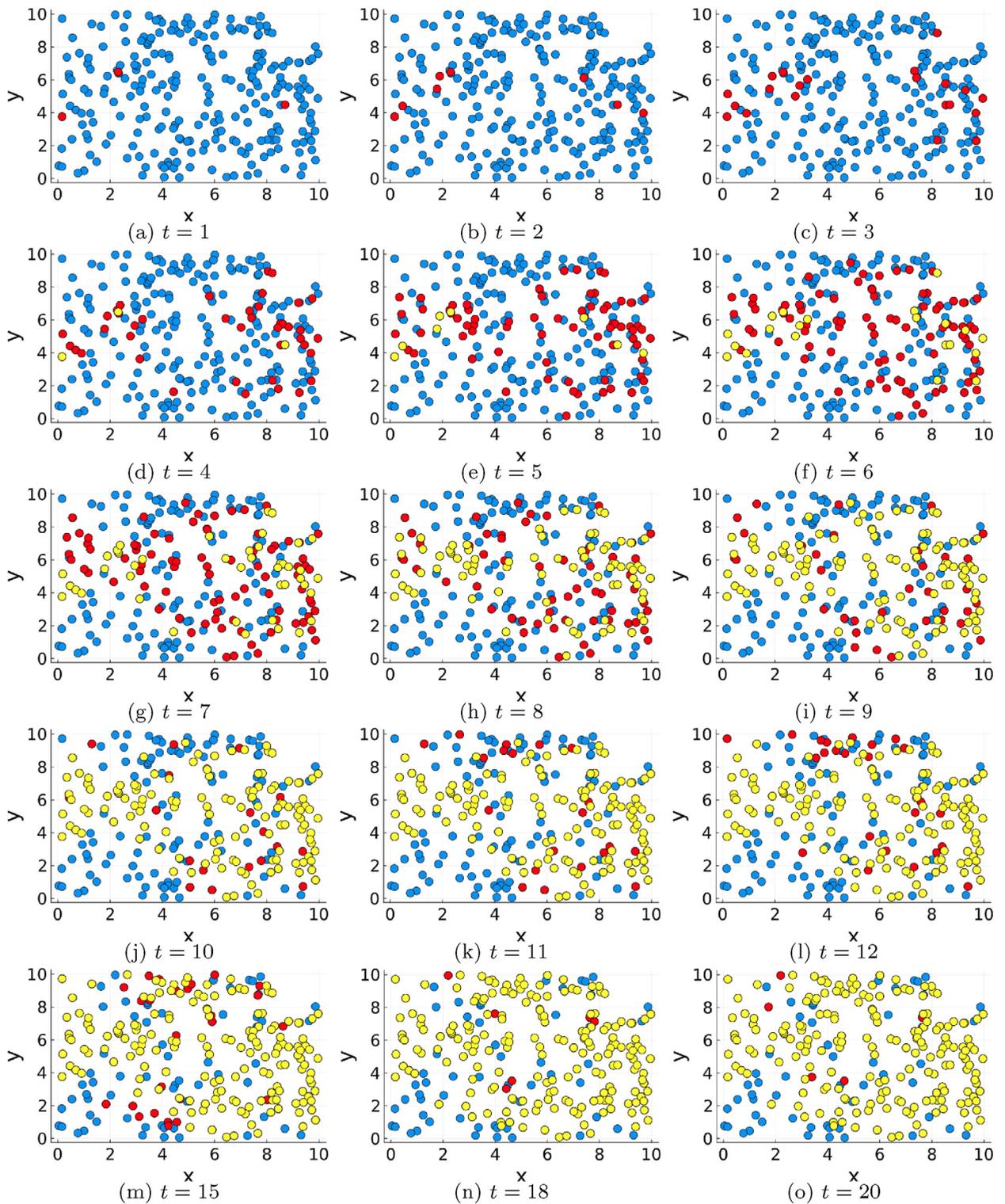


Figure A.3. Typical simulated epidemic realization across a grid of individuals for piecewise constant kernel ILM with medium BC effect, where susceptible individuals are denoted by blue circles, infected individuals are denoted by red circles and removed individuals are denoted by yellow circles.

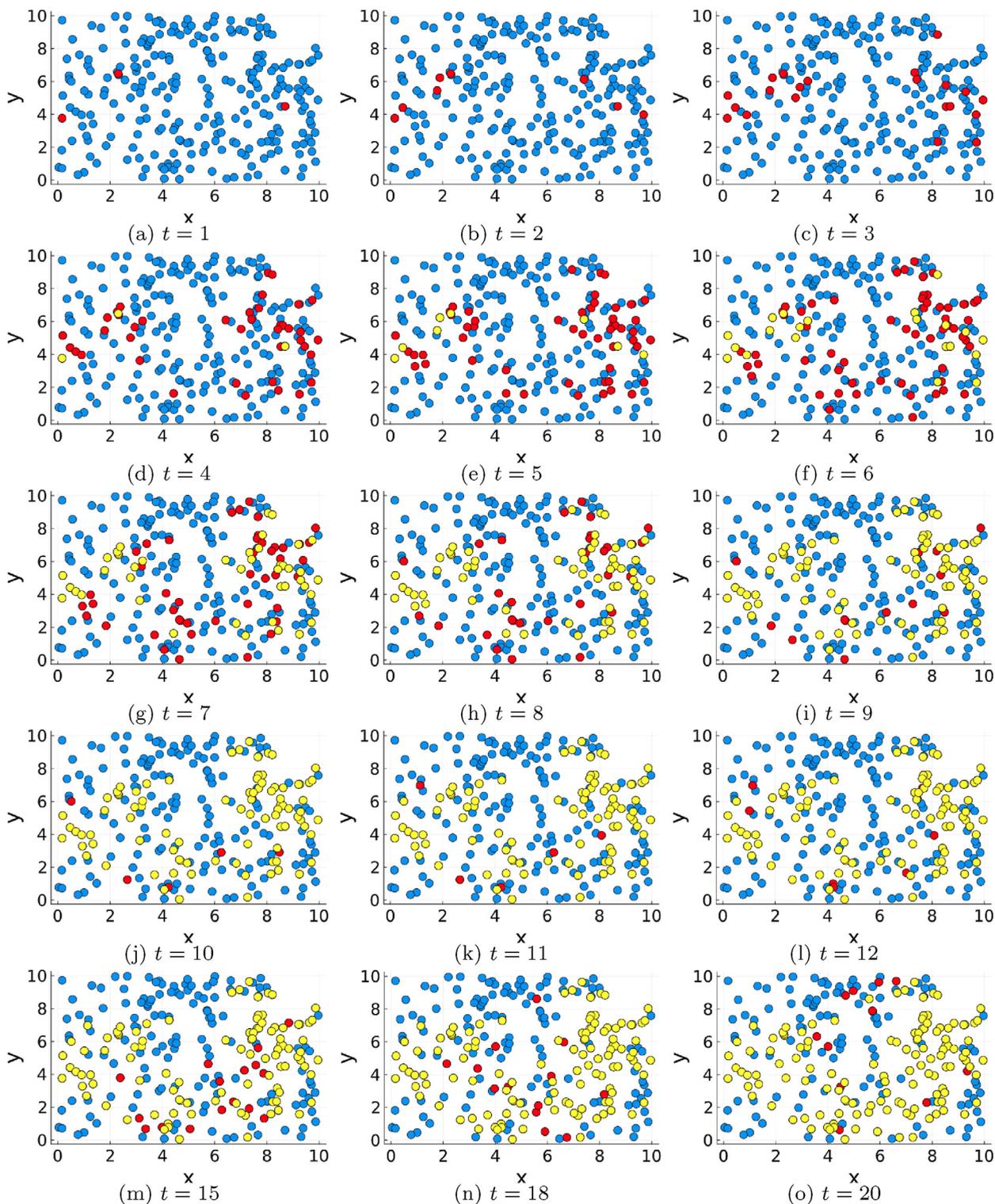


Figure A.4. Typical simulated epidemic realization across a grid of individuals for picewise constant kernel ILM with strong BC effect, where susceptible individuals are denoted by blue circles, infected individuals are denoted by red circles and removed individuals are denoted by yellow circles.

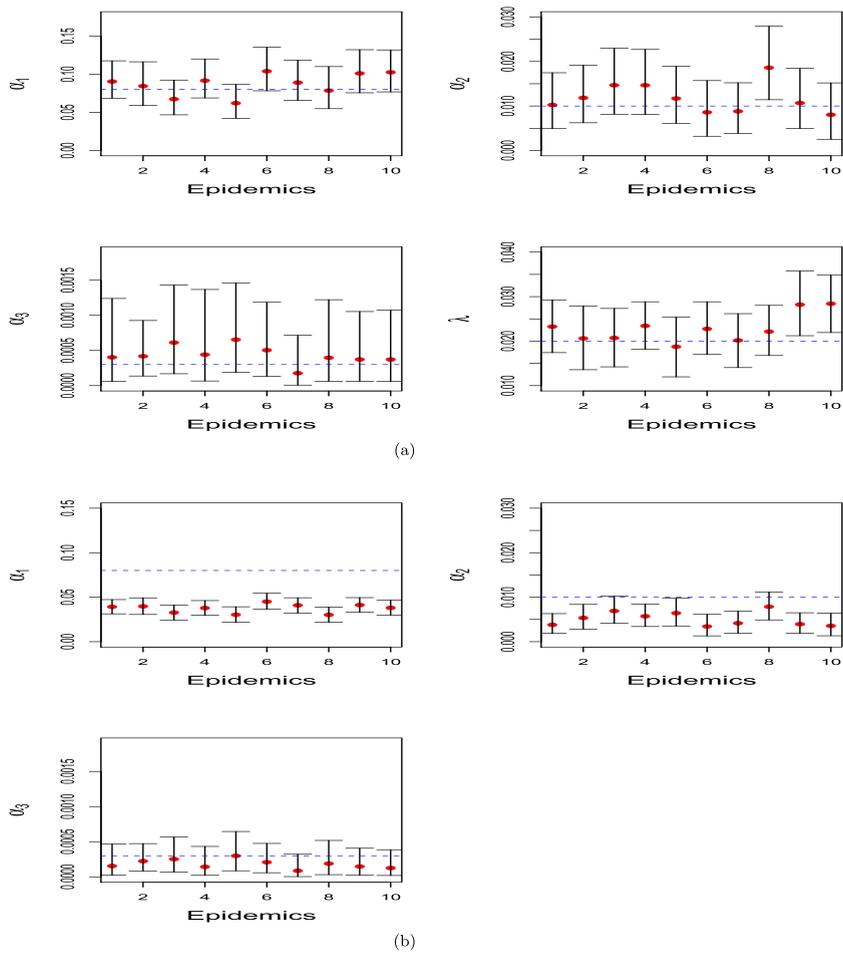


Figure A.5. Posterior medians (red points) and 95% PIs for (a) α_1 , α_2 , α_3 and λ and (b) α_1 , α_2 , α_3 for 10 different simulated epidemics when the change points (δ 's) are considered fixed and λ are considered to be included or not in the model. The true parameter values $\alpha_1 = 0.08$, $\alpha_2 = 0.01$, $\alpha_3 = 0.0003$ and $\lambda = 0.02$ for three-step cases with fixed $\delta_1 = 1.5$, and $\delta_2 = 3$ are represented by the blue dashed line.

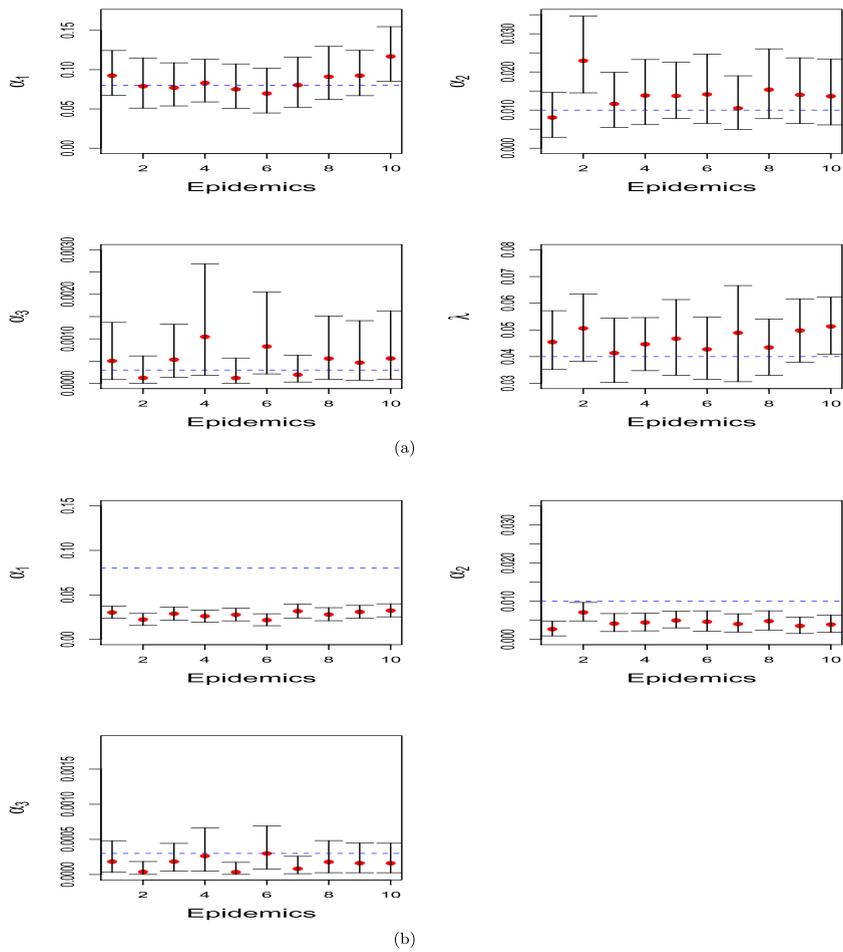


Figure A.6. Posterior medians (red points) and 95% PIs for (a) α_1 , α_2 , α_3 and λ and (b) α_1 , α_2 , α_3 for 10 different simulated epidemics when the change points (δ 's) are considered fixed and λ are considered to be included or not in the model. The true parameter values $\alpha_1 = 0.08$, $\alpha_2 = 0.01$, $\alpha_3 = 0.0003$ and $\lambda = 0.04$ for three-step cases with fixed $\delta_1 = 1.5$, and $\delta_2 = 3$ are represented by the blue dashed line.

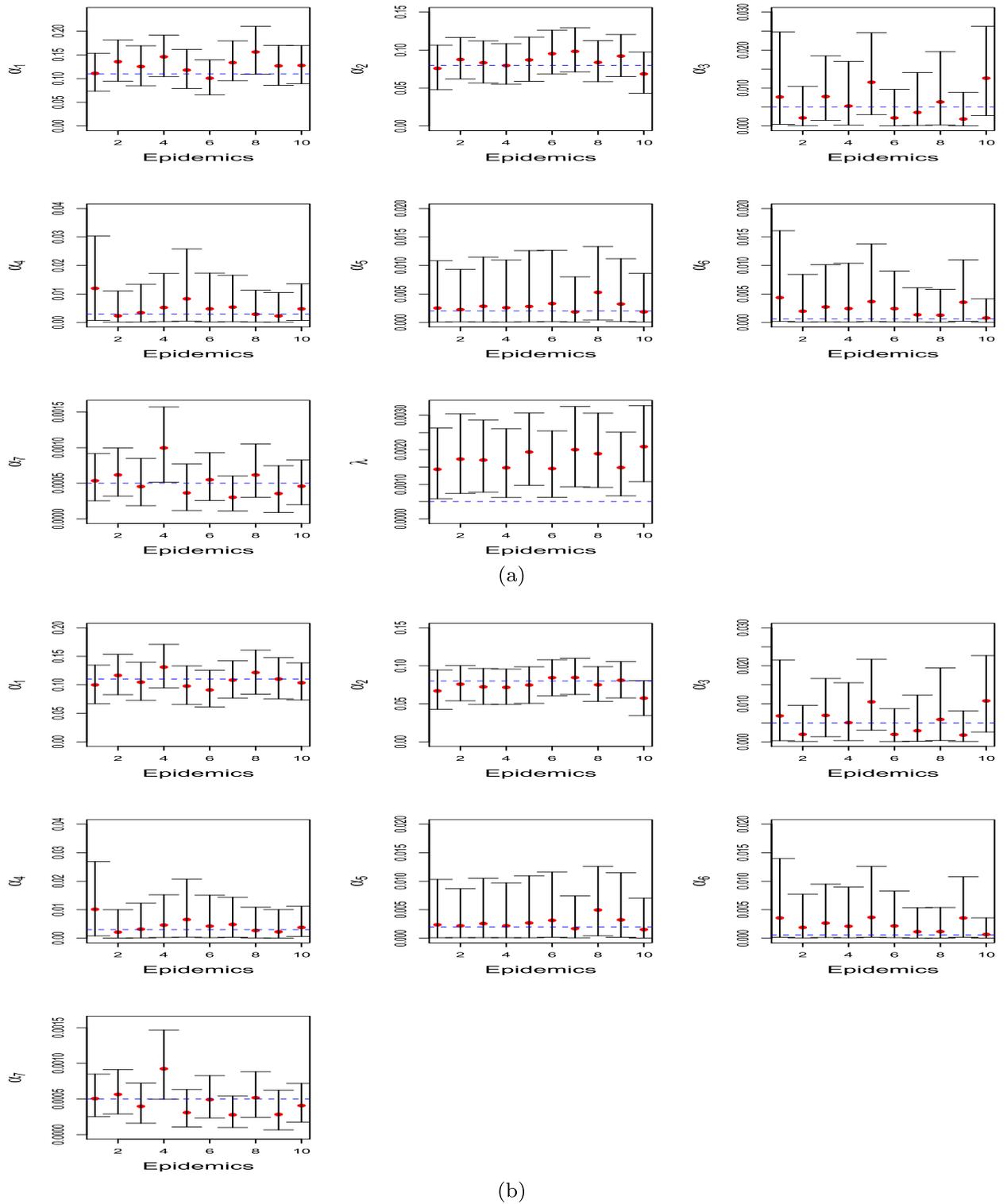


Figure A.7. Posterior medians (red points) and 95% PIs for (a) α_1 to α_7 and λ and (b) α_1 to α_7 for 10 different simulated epidemics when the change points (δ 's) are considered fixed and λ are considered to be included or not in the model. The true parameter values $\alpha_1 = 0.11$, $\alpha_2 = 0.08$, $\alpha_3 = 0.005$, $\alpha_4 = 0.003$, $\alpha_5 = 0.002$, $\alpha_6 = 0.0006$, $\alpha_7 = 0.0005$ and $\lambda = 0$ for seven-step cases with fixed $\delta_1 = 2.5$, $\delta_2 = 4$, $\delta_3 = 5.5$, $\delta_4 = 6.5$, $\delta_5 = 7.5$, and $\delta_6 = 8.5$ are represented by the blue dashed line.

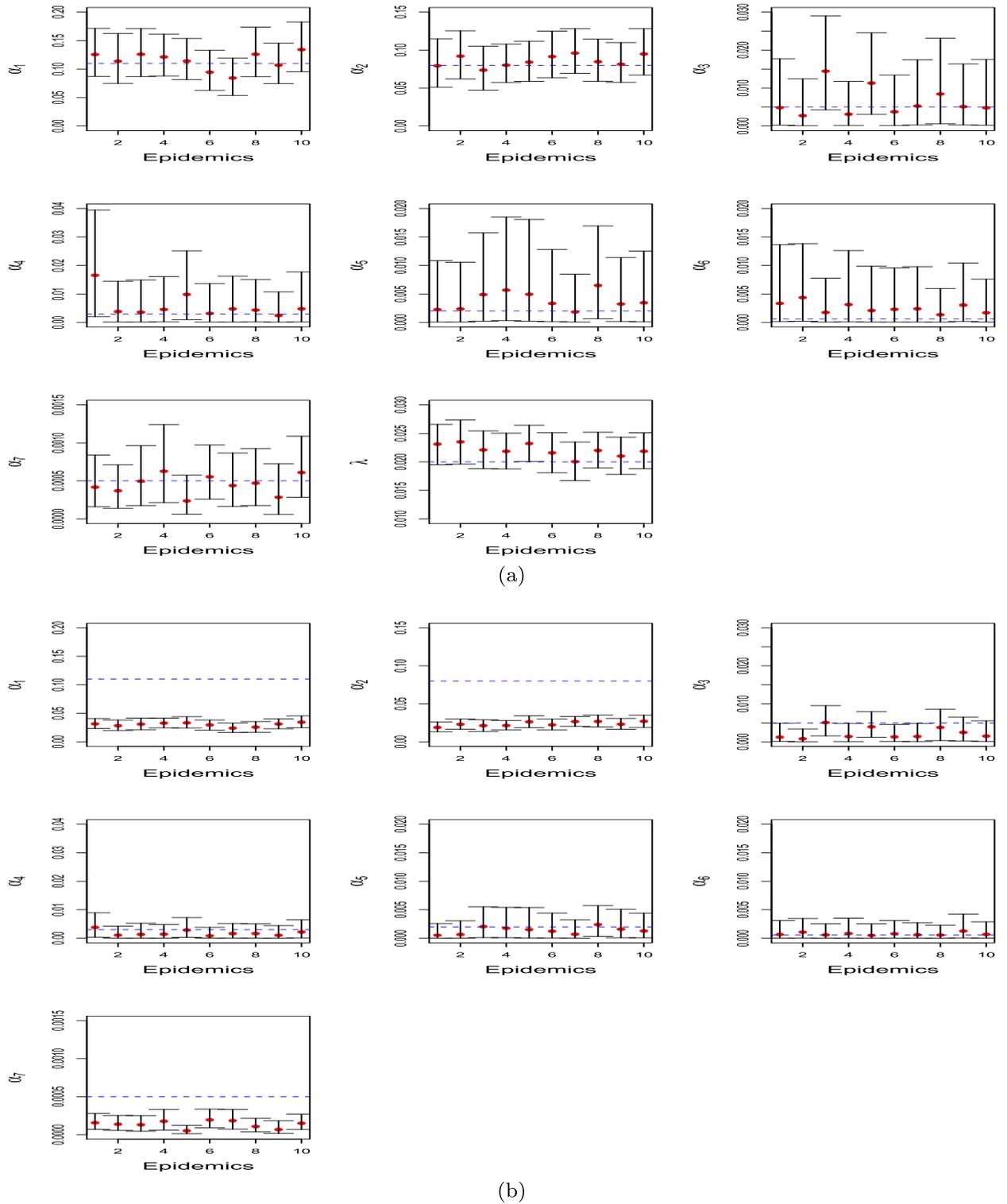


Figure A.8. Posterior medians (red points) and 95% PIs for (a) α_1 to α_7 and λ and (b) α_1 to α_7 for 10 different simulated epidemics when the change points (δ 's) are considered fixed and λ are considered to be included or not in the model. The true parameter values $\alpha_1 = 0.11, \alpha_2 = 0.08, \alpha_3 = 0.005, \alpha_4 = 0.003, \alpha_5 = 0.002, \alpha_6 = 0.0006, \alpha_7 = 0.0005$ and $\lambda = 0.02$ for seven-step cases with fixed $\delta_1 = 2.5, \delta_2 = 4, \delta_3 = 5.5, \delta_4 = 6.5, \delta_5 = 7.5,$ and $\delta_6 = 8.5$ are represented by the blue dashed line.

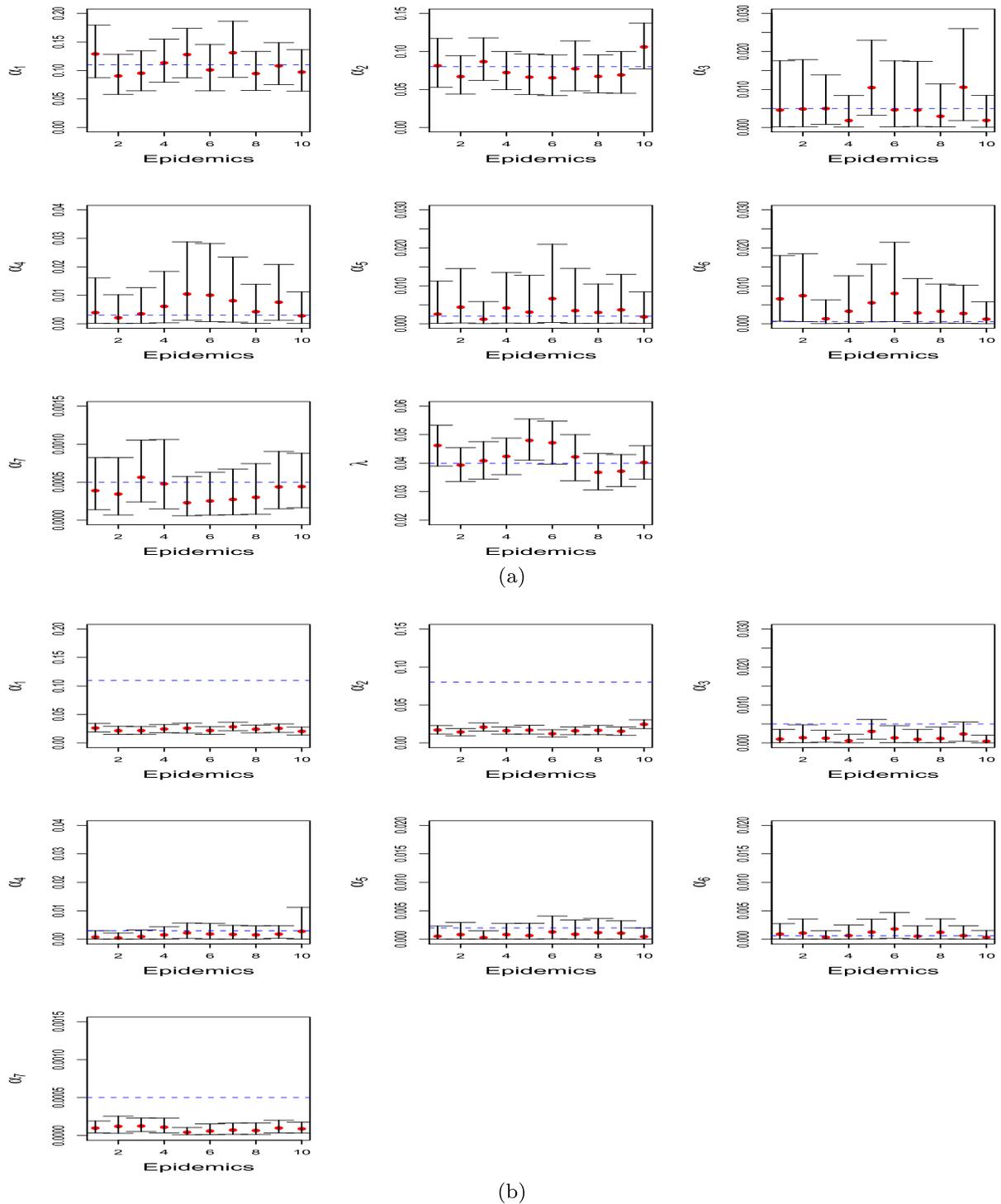


Figure A.9. Posterior medians (red points) and 95% PIs for (a) α_1 to α_7 and λ and (b) α_1 to α_7 for 10 different simulated epidemics when the change points (δ 's) are considered fixed and λ are considered to be included or not in the model. The true parameter values $\alpha_1 = 0.11$, $\alpha_2 = 0.08$, $\alpha_3 = 0.005$, $\alpha_4 = 0.003$, $\alpha_5 = 0.002$, $\alpha_6 = 0.0006$, $\alpha_7 = 0.0005$ and $\lambda = 0.04$ for seven-step cases with fixed $\delta_1 = 2.5$, $\delta_2 = 4$, $\delta_3 = 5.5$, $\delta_4 = 6.5$, $\delta_5 = 7.5$, and $\delta_6 = 8.5$ are represented by the blue dashed line.

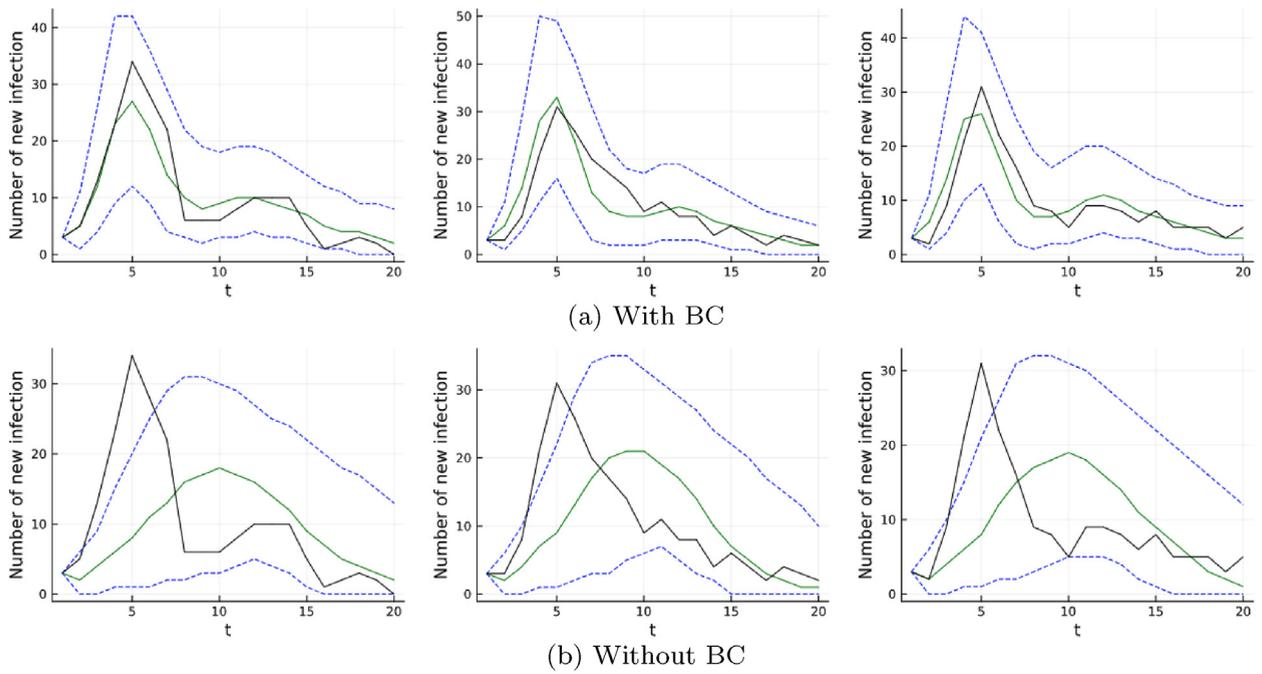


Figure A.10. Posterior predictive distribution of the epidemic curve for three typical simulated epidemics for three-step cases with incorporating BC effects (a) $\lambda = 0.02$ and (b) without BC effect. The black solid line represents the true epidemic curve, the green solid line represents the estimated median curve and the blue dotted lines represent the 95% PI, respectively.

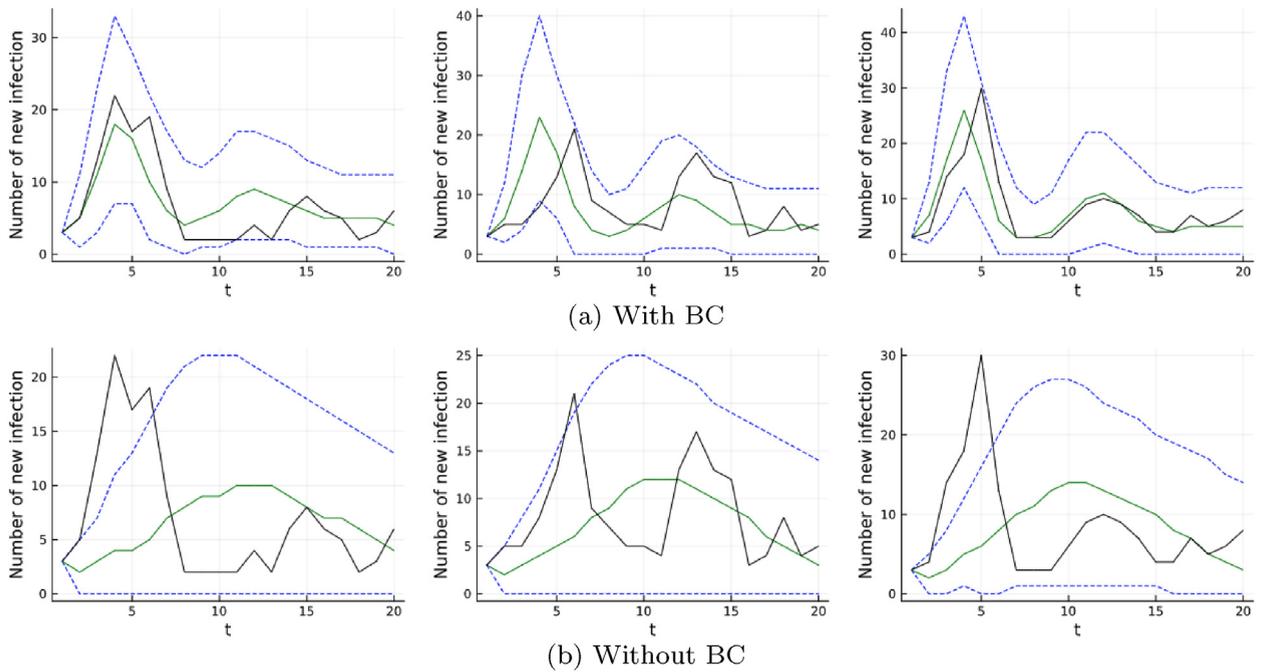


Figure A.11. Posterior predictive distribution of the epidemic curve for three typical simulated epidemics for three-step cases with incorporating BC effects (a) $\lambda = 0.04$ and (b) without BC effect. The black solid line represents the true epidemic curve, the green solid line represents the estimated median curve and the blue dotted lines represent the 95% PI, respectively.

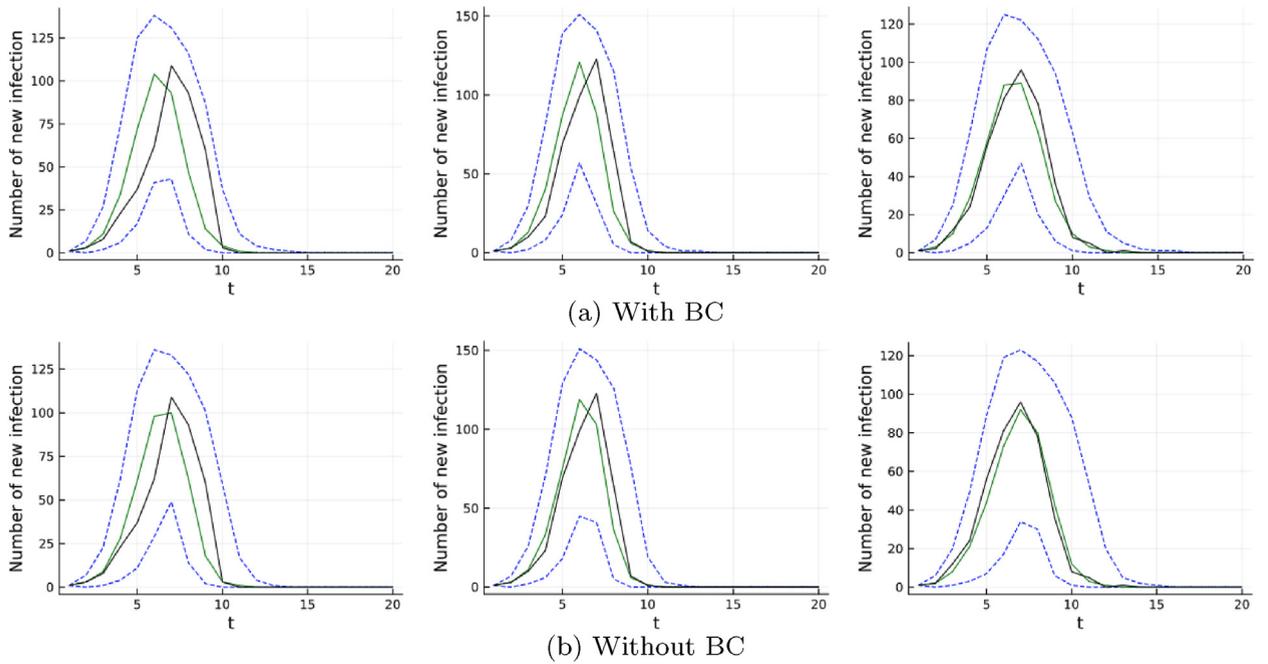


Figure A.12. Posterior predictive distribution of the epidemic curve for three typical simulated epidemics for seven-step cases with incorporating BC effects (a) $\lambda = 0$ and (b) without BC effect. The black solid line represents the true epidemic curve, the green solid line represents the estimated median curve and the blue dotted lines represent the 95% PI, respectively.

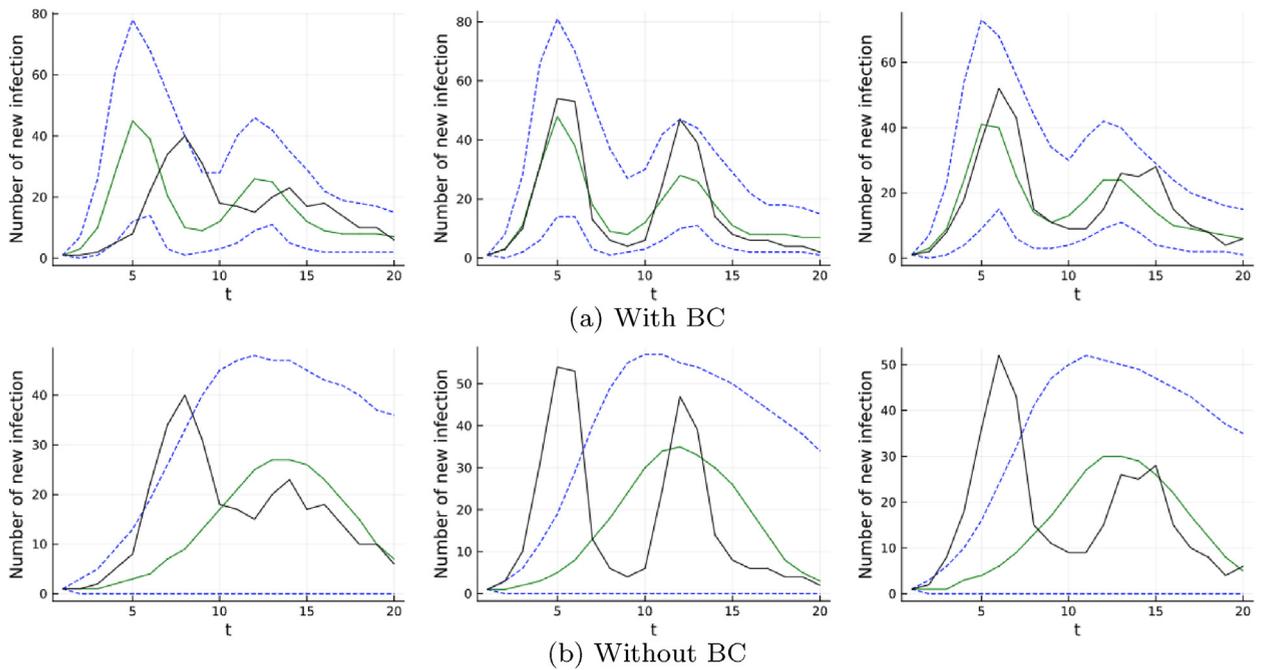


Figure A.13. Posterior predictive distribution of the epidemic curve for three typical simulated epidemics for seven-step cases with incorporating BC effects (a) $\lambda = 0.02$ and (b) without BC effect. The black solid line represents the true epidemic curve, the green solid line represents the estimated median curve and the blue dotted lines represent the 95% PI, respectively.

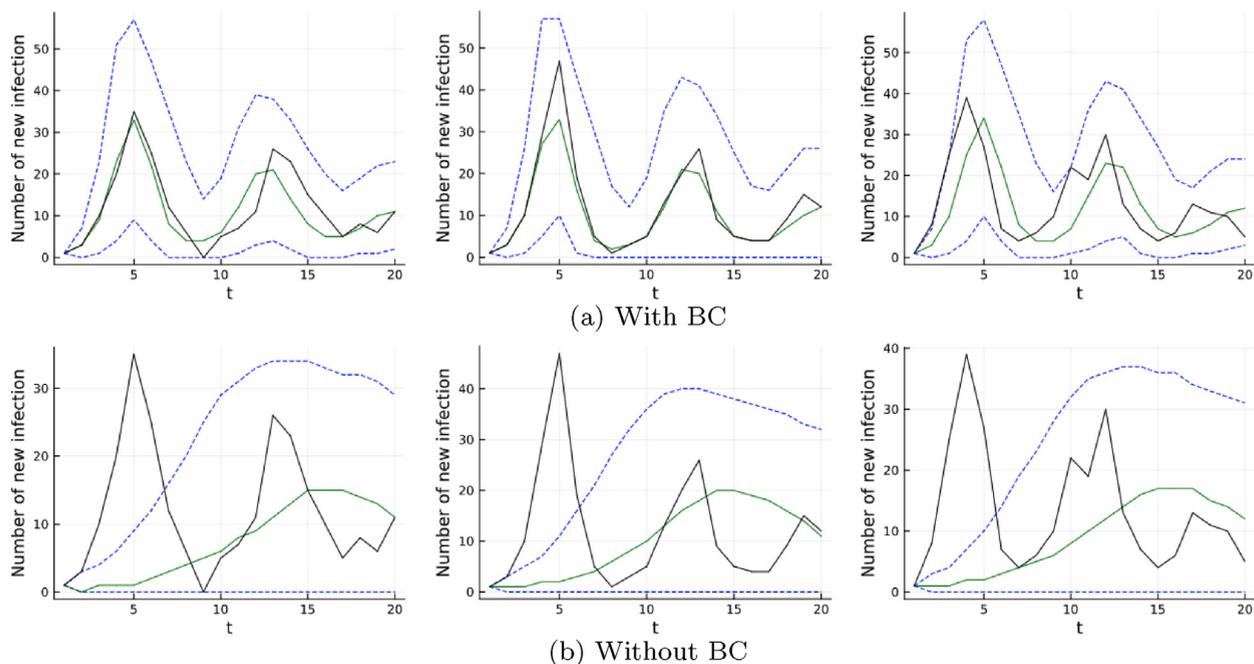


Figure A.14. Posterior predictive distribution of the epidemic curve for three typical simulated epidemics for seven-step cases with incorporating BC effects (a) $\lambda = 0.04$ and (b) without BC effect. The black solid line represents the true epidemic curve, the green solid line represents the estimated median curve and the blue dotted lines represent the 95% PI, respectively.

References

- Chib, S., & Greenberg, E. (1995). Understanding the Metropolis–Hastings algorithm. *The American Statistician*, 49(4), 327–335.
- Chis Ster, I., & Ferguson, N. M. (2007). Transmission parameters of the 2001 foot and mouth epidemic in Great Britain. *PLoS one*, 2(6), Article e502.
- Deardon, R., Brooks, S. P., Grenfell, B. T., Keeling, M. J., Tildesley, M. J., Savill, N. J., Shaw, D. J., & Woolhouse, M. E. (2010). Inference for individual-level models of infectious diseases in large populations. *Statistica Sinica*, 20(1), 239.
- Deeth, L. E., Deardon, R., & Gillis, D. J. (2015). Model choice using the deviance information criterion for latent conditional individual-level models of infectious disease spread. *Epidemiologic Methods*, 4(1), 47–68.
- Friedman, J., Hastie, T., Tibshirani, R., et al. (2001). *The Elements of Statistical Learning*. Springer Series in Statistics.
- Funk, S., Salathé, M., & Jansen, V. A. (2010). Modelling the influence of human behaviour on the spread of infectious diseases: a review. *Journal of the Royal Society Interface*, 7(50), 1247–1256.
- Gardner, A., Deardon, R., & Darlington, G. A. (2011). Goodness-of-fit measures for individual-level models of infectious disease in a bayesian framework. *Spatial and Spatio-temporal Epidemiology*, 2(4), 273–281.
- Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457–472.
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments. *Bayesian Statistics*, 4, 641–649.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57, 97–109.
- Kwong, G. P., Poljak, Z., Deardon, R., & Dewey, C. E. (2013). Bayesian analysis of risk factors for infection with a genotype of porcine reproductive and respiratory syndrome virus in Ontario swine herds using monitoring data. *Preventive Veterinary Medicine*, 110(3–4), 405–417.
- Kypraios, T., & O’Neill, P. D. (2018). Bayesian nonparametrics for stochastic epidemic models. *Statistical Science*, 33(1), 44–56.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6), 1087–1092.
- Rahul, C. R., & Deardon, R. (2024). Individual-level models of disease transmission incorporating piecewise spatial risk functions. *ArXiv Preprint arXiv:2405.00835*.
- Robert, C. P. and Casella, G. (2002). *Monte Carlo Statistical Methods*. Springer.
- Silverman, B. W. (1985). Some aspects of the spline smoothing approach to non-parametric regression curve fitting. *Journal of the Royal Statistical Society: Series B (Methodological)*, 47(1), 1–21.
- Verelst, F., Willem, L., & Beutels, P. (2016). Behavioural change models for infectious disease transmission: a systematic review (2010–2015). *Journal of The Royal Society Interface*, 13(125).
- Ward, C., Deardon, R., & Schmidt, A. M. (2023a). Bayesian modeling of dynamic behavioral change during an epidemic. *Infectious Disease Modelling*, 8(4), 947–963.
- Ward, M. A., Deardon, R., & Deeth, L. E. (2023b). A framework for incorporating behavioural change into individual-level spatial epidemic models. *Canadian Journal of Statistics*, Article e11828.