

Detecting changes in the transmission rate of a stochastic epidemic model

Jenny Huang | Raphaël Morsomme | David Dunson | Jason Xu 

Department of Statistical Science, Duke University, Durham, North Carolina, USA

Correspondence

Jason Xu, Department of Statistical Science, Duke University, Durham, NC, USA.

Email: jason.q.xu@duke.edu

Funding information

National Institutes of Health, Grant/Award Numbers: R01ES027498, R01ES028804; National Science Foundation, Grant/Award Numbers: DMS-2230074, PIPP-2200047

Throughout the course of an epidemic, the rate at which disease spreads varies with behavioral changes, the emergence of new disease variants, and the introduction of mitigation policies. Estimating such changes in transmission rates can help us better model and predict the dynamics of an epidemic, and provide insight into the efficacy of control and intervention strategies. We present a method for likelihood-based estimation of parameters in the stochastic susceptible-infected-removed model under a time-inhomogeneous transmission rate comprised of piecewise constant components. In doing so, our method simultaneously learns change points in the transmission rate via a Markov chain Monte Carlo algorithm. The method targets the exact model posterior in a difficult missing data setting given only partially observed case counts over time. We validate performance on simulated data before applying our approach to data from an Ebola outbreak in Western Africa and COVID-19 outbreak on a university campus.

KEYWORDS

Bayesian change point detection, COVID-19, data-augmented MCMC, incidence data, likelihood-based inference, stochastic epidemic models

1 | INTRODUCTION

This article considers the statistical task of estimating epidemic parameters from observational data when the rate of disease transmission varies over time. In particular, we seek to model the transmission rate under an susceptible-infected-removed (SIR) model so that it can flexibly capture changes over time, yet remain parsimonious enough to retain tractable inference and to avoid overfitting the available data.

Recent outbreaks have been characterized by changes in social distancing behaviors and economic policies, control or mitigation measures, the emergence of new disease variants, and many other factors which may lead to changes in the disease transmission rate.^{1,2} The common assumption that the transmission rate is fixed over time may no longer be appropriate in such cases, and classical SIR models with a constant transmission parameter, denoted β , often struggle to fit observed trends in data from recent epidemics such as COVID-19 and Ebola. As a result, several studies have posited time-varying rates, relaxing the assumption of a constant rate of transmission made in classical SIR models. Many of these approaches operate within deterministic models^{3,4} and their variants. However, deterministic models do not lend themselves to natural probabilistic interpretations so that resulting estimates do not come equipped with measures of uncertainty. When a sampling model, such as Gaussian or binomial emissions, is employed on top of a deterministic mean model, the uncertainty quantification should only be interpreted as a reflection of the measurement error rather than the

underlying stochasticity of the model. In order to account for uncertainty from the process itself more interpretably, we focus on the stochastic version of the SIR model.

In the stochastic epidemic model, the spread of disease is described according to a probability law rather than a system of differential equations.⁵ The canonical formulation specifies that the disease moves through the population as a Markov jump process (MJP), which agrees with the deterministic formulation in the large-population limit.⁶ Unlike diffusion approximations for large populations,^{7,8} the Markov jump process formulation is defined on a discrete state space and is appropriate in early as well as later stages of an outbreak. These *stochastic* epidemic models (SEM) form a class of compartmental models of disease that replace deterministic transition rates by instantaneous jump probabilities, as a result evolving as continuous-time stochastic processes. When it comes to learning how the rate of transmission in such models changes over time, a much smaller literature has explored inference under the general stochastic epidemic model.

One reason this space is relatively underexplored stems from computational difficulties that arise when the data are only partially informative. Epidemic data are commonly reported as incomplete summaries of a process that evolves continuously through time. For instance, a common setting in observational studies provides *incidence data*, comprised of new infection counts given on some fixed schedule (eg, weekly, monthly). With incidence data, however, quantities such as the infection and removal times defining the likelihood are unknown. In this article, we will focus on this missing data setting, where only the number of new infections at a discrete set of times is observed. Here, evaluating the *marginal* likelihood of the partially observed data directly requires access to the finite-time transition probabilities of the process.⁹ These quantities are difficult to compute in our model, requiring integration over the high-dimensional space of all configurations of the trajectories between observed times consistent with our observed data.¹⁰ Although integrating over this space is possible in principle,^{11,12} available methods are computationally intensive and often impractical beyond the most simple cases.¹³ Our data-augmented methodology performs straightforward, efficient MCMC directly on the model parameters, providing a complementary approach to the flexible but quite computationally intensive class of simulation-based methods.^{8,14-17}

To further introduce change points in the transmission rate under these models, some studies have taken a fairly rigid approach by fixing change point locations based on knowledge of the exact dates on which policy changes occurred.^{18,19} Unfortunately, this rigid approach can lead to biased results, failing to account for lag times and ruling out the possibility of change points that are unaccounted for. Here, we stress the importance of learning change points flexibly from data. While we place this task in the framework of estimating change points, classical change point detection algorithms are poorly suited when the goal is to detect changes in the *transmission rate* under a model-based approach. The transmission rate is related to the observed time series data through the mechanistic model. In particular, the overall infection rate is already varying over time, given by $\beta S(t)I(t)$. In replacing the constant transmission rate β by a function $\beta(t)$, one should seek formulations that balance flexibility with enough parsimony so that it does not overwhelm the model-based dynamics defined via interactions between $S(t)$ and $I(t)$.

To this end, our article seeks to capture changes in the transmission rate by replacing the constant force with a piecewise constant function $\beta(t)$, learning where to best allow change points in its segments from the data. There is a mature body of literature focused on multiple change point detection.²⁰⁻²³ Frequentist methods commonly rely on a test statistic, such as the likelihood ratio or CUSUM statistic, to test for changes in the distribution of the data, followed by a model selection step, typically using a Lasso-type penalty, to determine the number of parameters defining the signal.²⁴⁻²⁶ In contrast, Bayesian approaches formulate the multiple change point problem in terms of a sequence of hidden discrete state variables that evolve as a Markov process, and use methods such as MCMC or dynamic programming to estimate the posterior distribution of change points.^{21,23} The majority of these existing methods concern independent and identically distributed random variables, and early works have explored Bayesian approaches to the change point problem within the context of sequential data.²⁷ This is not the case under the SIR model, however, where the number of new infections depends on both the transmission rate and the sizes of the disease compartments: event times follow a Poisson process with intensity function $\beta S(t)I(t)$. For these reasons, general-purpose change point detection algorithms are ill-equipped for our task within this mechanistic model. Again, we seek to model heterogeneity in the *transmission rate* through change points, rather than to detect change points in the sequence of observed counts directly.

To jointly address these considerations, we propose a fully Bayesian inferential framework for learning changes in the transmission rate of the general stochastic epidemic model. In related work on time-varying parameters in epidemic models, Reference 13 fit the force of infection with second-order B -splines using data with exact recovery times and unknown infection times. However, they do so in a non-parametric model that ignores the mass action assumption of the SIR, effectively replacing the intensity $\beta S(t)I(t)$ entirely by a quite flexible function $\beta(t)$. In contrast, we preserve the interaction term between susceptible and infectious individuals that undergirds the SIR dynamics, learning a rate

function $\beta(t)S(t)I(t)$ where $\beta(t)$ is a step function that only features change points when there is sufficient evidence from the data. This leads to a more parsimonious model, while facilitating samplers targeting the posterior distribution of the transmission rate parameters. Interpretable outputs can then be used to assess the efficacy of proposed public health interventions such as travel restrictions, school closures, and lockdowns.

In terms of Bayesian computation, References 10 and 13 employ reversible jump (RJ) MCMC for posterior computation of time-varying epidemic parameters. The dimension varies under this approach when proposing to increase or decrease the number of steps, necessitating a RJ-MCMC sampler. Efficient dimension changing moves can be notoriously difficult to design, however, and RJ-MCMC algorithms often exhibit slow mixing.¹³ Here we present an alternate approach more similar in spirit to the hidden Markov multiple change point model,²¹ bypassing the need for the reversible jump step. That is, rather than estimating the sequence of change points directly, we introduce a latent sequence of indicators, $\{\Delta_t\}$, that take the value 1 at change point locations and 0 otherwise. When the length of the observed time series is T , the method we propose simply samples a fixed-dimensional sequence of indicators of length $T - 1$. To account for the temporal dependence between change point locations, we infer the hidden states of the Markov process $\{\Delta_t\}$. Further leveraging on recent developments in data augmentation (DA) MCMC for stochastic SIR models,²⁸ we propose a simple Metropolis-within-Gibbs sampling scheme.

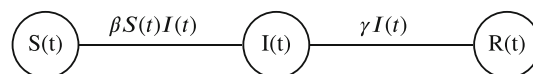
The remainder of the article is structured as follows: Section 2 provides a description of the data, an introduction to the stochastic SIR model, and a detailed description of our change point model and computational algorithm. In Section 3, the model is applied to simulated data followed by applications to Ebola and COVID-19 outbreaks with suspected change points. Section 4 closes with a discussion and highlights avenues for future work.

2 | METHODS

2.1 | Data setting and the stochastic SIR model

In this article, we focus on the *incidence data* setting in which new infection counts are collected at discrete reporting times. This is typical of observational studies of epidemic data, but presents a challenge for exact model-based inference due to the missing information between observation times.⁸ In particular, given a set of observation times $t_{0:K}$, the observed incidence consists of a K -dimensional vector $\mathbf{I}_{1:K} = (I_1, \dots, I_K)$, where I_k denotes the number of new cases reported during the interval $(t_{k-1}, t_k]$.

We are interested in learning the parameters of a stochastic SIR model to these data. This widely used compartmental model describes the spread of an epidemic through a population divided into three compartments based on disease status: susceptible (S), infectious (I), and removed (R) individuals. Individuals transition from S to I when they become infected and from I to R when they recover or are removed (eg, death). While the deterministic version is defined through a system of differential equations, the stochastic SIR model evolves as a continuous-time Markov chain (CTMC) $X = \{X(t); t > 0\}$ where the components $X(t) = \{S(t), I(t), R(t)\}$ track the respective compartments $S(t)$, $I(t)$, and $R(t)$ denoting the number of susceptible, infected and removed individuals in the population at time t ⁵ (figure shown below).



The population is assumed to be closed and homogeneously mixing, ignoring demographic dynamics such as births and deaths.²⁹ This implies that contacts between individuals follow independent Poisson processes with rate β , which can be interpreted as the infection rate of the contagious disease. We further assume that the infectious periods follow independent exponential distributions with rate γ for exposition. As a result, the stochastic SIR model is a bivariate Markov process that is characterized by the transition probabilities below:

$$P(S(t+h), I(t+h)) = (k, j) \mid (S(t), I(t)) = (s, i) = \begin{cases} \beta S(t)I(t)h + o(h) & (k, j) = (s-1, i+1) \\ \gamma I(t)h + o(h) & (k, j) = (s, i-1) \\ 1 - (\beta S(t)I(t)h + \gamma I(t)h) + o(h) & (k, j) = (s, i) \\ o(h) & \text{otherwise.} \end{cases}$$

The transition probabilities above describe how the process whose current state at time t is denoted (s, i) evolves. Over a small time interval h , the expressions give the probabilities of infection, recovery, and no event occurring, the only

possibilities up to $o(h)$. In particular, we see that the *instantaneous rates* of the process can be understood as limits of these probabilities divided by h as the time interval decreases to zero: for instance, doing so for the probability of infection formalizes what it means to call $\beta S(t)I(t) = \lim_{h \downarrow 0} \frac{\beta S(t)I(t)h}{h}$ as the infection rate.

2.2 | Time-varying infection rate

The model as specified above posits that the transmission rate β per contact between susceptible and infectious person is constant over time. Here, we relax this standard assumption to allow this rate $\beta(t)$ to vary over time, and choose to model it as a piecewise constant function, without specifying in advance the points at which its constant values change. This choice keeps the original mass action dynamics of SIR, while allowing for jumps in the per-contact transmission rate. That is, given some J change point locations $(c_1, \dots, c_J) \subset (t_1, \dots, t_K)$,

$$\beta(t) = \beta_k \forall t \in [c_k \leq t \leq c_{k+1}]. \quad (1)$$

Under this piecewise constant formulation, the points $\{c_k\}$ are naturally interpreted as change point locations. To learn their locations jointly with parameter inference, we introduce a sequence of latent binary variables which denote the presence of such change points. At each observation time t_k , the latent indicator variable Δ_t takes on the value of 1 to indicate the existence of a change in transmission rate and 0 if there is no change.

$$\Delta_t = \begin{cases} 1 & \text{if } t = c_k \text{ for some } k \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

In particular, the locations of the 1's in the sequence $\{\Delta_t\}$ encode the locations of change points $\{c_k\}$ where $\beta(t)$ takes on a new value. Hence, our formulation will reduce the task of inferring a time-varying infection rate $\beta(t)$ to estimating the change point sequence $\{\Delta_t\}$, and performing efficient parameter updates jointly on the locally constant segments.

2.3 | The complete data likelihood

As a CTMC, the time until the next infection event follows an exponential distribution with rate $\beta(t)S(t)I(t)$, and the time until next removal event follows an exponential distribution with rate $\gamma I(t)$. Considering the epidemic on one of the intervals $(c_k, c_{k+1}]$, we obtain the following likelihood of the SIR model with a piecewise infection rate

$$\begin{aligned} \mathcal{L}(\theta; \mathbf{X}) &= \prod_k \left(\prod_{j \in \mathcal{I}_k} \beta_k I(\tau_j^I) \prod_{l \in \mathcal{R}_k} \gamma \exp \left\{ - \int_{c_k}^{c_{k+1}} \beta_k I(t) S(t) + \gamma I(t) dt \right\} \right) \\ &= \gamma^{n_R} \prod_k \left(\beta_k^{n_{I_k}} \prod_{j \in \mathcal{I}_k} I(\tau_j^I) \exp \left\{ - \beta_k \int_{c_k}^{c_{k+1}} I(t) S(t) dt + \gamma \int_{c_k}^{c_{k+1}} I(t) dt \right\} \right). \end{aligned} \quad (3)$$

Here, $\mathbf{X}$ denotes an epidemic trajectory consisting of all infection and removal times, k is the index on the step of the piecewise transmission rate, and $\{c_k\}$ are the sequence of change point times. The index sets $\mathcal{I}_k = \{j : \tau_j^I \in (c_k, c_{k+1}]\}$ and $\mathcal{R}_k = \{j : \tau_j^R \in (c_k, c_{k+1}]\}$ correspond to the individuals that are infected and removed during the interval $(c_k, c_{k+1}]$. The quantity $n_{I_k} = |\mathcal{I}_k|$ denotes the number of infection and removal events over this same time interval, and $n_R = \sum_k |\mathcal{R}_k|$ denotes the total number of removal events. Note that due to the Markov property, the likelihood over an interval where the transmission rate is locally constant, $(c_k, c_{k+1}]$, takes on the same form as the usual stochastic SIR likelihood with constant parameter β_k for an outbreak restricted to that interval, and the likelihood of the process on the entire observation period is a product over these intervals.

2.4 | Model parsimony and promoting sparsity in change points

To control model complexity and avoid overfitting, we next seek to impose parsimony on the rate function given by Equation (1). It is undesirable to assign each k th observation interval its unique β_k value; the number of model parameters

increases with the number of times we observe incidence, and we may expect non-identifiability when $\beta(t)$ is too flexible since $I(t)$ is unknown and the overall infection rate to be learned takes the form $\beta(t)S(t)I(t)$.

An intuitive way to avoid this is by promoting *local constancy*: that is, we would like to encourage runs of constant values of β_k for multiple observation intervals, $(t_{k-1}, t_k]$, in a row a priori. Then, the rate function only changes when sufficiently supported by the data. We may naturally encode this through prior probability distributions on the vector of changes Δ . We employ a simple Markov process for how Δ changes, governed by the transition matrix

$$\Pi = \begin{bmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{bmatrix}. \quad (4)$$

Here π_{ij} denotes the probability of transitioning from state i to j from one time point to the next.

Note if there is a change point between the previous and current time step (ie, $\Delta_t = 1$) the probability of another change point occurring at the next time step is π_{11} , while the probability of the rate remaining constant is represented by the transition into a non-change point state π_{10} . Likewise, if there was no change point between the previous and current time step (ie, $\Delta_t = 0$), the probability of a change point at the next time is π_{01} while that of remaining in a constant state is π_{00} . We see that this prior promotes local constancy as long as $\pi_{\cdot 1}$ is small—that is, it encourages sequences Δ to have runs of zeros a priori. This effect is akin to first-order methods to promote local constancy such as the intuition behind the fused Lasso³⁰ when $\pi_{11} = \pi_{01}$, but allows the user to additionally penalize back-to-back change points by setting $\pi_{11} < \pi_{01}$. Doing so is desirable in settings when rapid changes in policy, behavior, or the emergence of new variants may be unrealistic. The prior for Δ induces a hidden Markov model that captures the temporal dependence of the infection rate between a two-state latent variable over time.^{31,32} An analysis of the impact of hyperparameter choices for π on change point recovery is included in the Supplemental Material (Figure S2).

While we design priors on Π to promote sparsity in change points, note that the Bayesian view of Occam's razor³³ automatically protects against over-fitting to an extent. In particular, as the number of change points increases, the size of the effective model increases in turn as it entails more segment-specific parameters β_k to estimate. Viewed as a Bayesian model selection problem, the posterior probability on the model corresponding to a particular number of change points is proportional to the prior probability multiplied by a marginal likelihood obtained by integrating out the β_k parameters. This marginal likelihood is formed as a product of likelihoods conditional on the β_k by their prior, followed by integrating out the β_k 's. In this sense, the marginal likelihood automatically tends to decrease as the number of effective parameters grow. Hence, the posterior will favor parsimonious models with relatively few change points, only adding change points if they are clearly supported by the data.

This observation is well supported empirically, detailed later in our simulation studies. On each constant segment of the transmission rate β_k , we place a gamma prior with weakly informative hyperparameters, $\beta_k \sim \text{Gamma}(1, 1)$. Even under these weakly informative priors, we will find that the posterior distribution is able to concentrate successfully around the true value (Figure S5), indicating that the data are quite informative of transmission rate within each segment.

2.5 | Posterior computation via Metropolis-within-Gibbs

The model development in the previous sections is designed to admit efficient posterior sampling schemes. Here, we derive a data-augmented Markov chain Monte Carlo (DA-MCMC) sampler targeting the posterior distribution of the parameters of interest. Recall the observed incidence data consist of counts of new infections at discrete time points, providing an incomplete view of the continuous-time epidemic process. In particular, the times at which an individual in the population either becomes infected or removed are unavailable. The complete-data likelihood in Equation (3) is therefore out of reach, and strictly speaking, is related to the likelihood of the observed incidence via marginalizing over all possible sets of missing event times compatible with the observations. This entails an unwieldy integration step over the space of latent complete-data trajectories, for which no closed form expression is available.

Instead, we model the unobserved event times as latent variables, tackling the challenging marginalization step by using a data augmentation approach.³⁴ This allows us to conduct inference targeting the *exact* model posterior distribution over the parameters and latent variables. Though the MCMC algorithm must explore an expanded space, parameter updates now take advantage of the complete-data likelihood given a configuration of the latent variables that “complete” the observed data.

Specifically, we alternate between three types of updates: (i) updates of the transition rate parameters Π , (ii) joint updates of the change point locations Δ and steps of the transmission rate β , and (iii) updates of the latent infection and removal times X .

First, we update π_{01} and π_{11} ; these parameters are assigned independent beta priors. Because the total number of change points follows a binomial distribution, it follows that the full conditional distribution remains beta distributed. This allows us to update Π using a Gibbs sampler (Equation (5)). Specifically, we use the prior $\pi_{ij} \sim \text{Beta}(a_{ij}, b_{ij})$, with expectation $a_{ij}/(a_{ij} + b_{ij})$ and variance $(a_{ij}b_{ij})/[(a_{ij} + b_{ij})^2(a_{ij} + b_{ij} + 1)]$. The conditional distribution for the transition probabilities is then

$$\pi_{01}|\Delta \sim \text{Beta}(a_{01} + n_{01}, b_{01} + n_{00}) \quad (5)$$

where n_{01} is the number of transitions in Δ from state 0 to 1 and n_{00} is the number of transitions from state 0 to 0. Similarly, $\pi_{11}|\Delta \sim \text{Beta}(a_{11} + n_{11}, b_{11} + n_{10})$.

Next, the parameters $\{\Delta_t\}$ and $\beta(t)$ are proposed jointly in a Metropolis-Hastings step. Specifically, we propose $\{\Delta_t\}$ as a realization of a Markov chain with transition matrix Π . The likelihood for the hidden state sequence is

$$P(\Delta|v_1, \Pi) = v_1 \prod_{i=2}^T \pi_{\Delta_{i-1}\Delta_i} = \pi_{00}^{n_{00}} \cdot \pi_{01}^{n_{01}} \cdot \pi_{10}^{n_{10}} \cdot \pi_{11}^{n_{11}} \quad (6)$$

where n_{ij} is the number of transitions in $\{\Delta_t\}$ from state i to j and $v_1 = \text{Pr}(\Delta_1 = 1)$ is the initial probability, which follows the same distribution as π_{01} . Within this joint Metropolis-Hastings update, we may now propose $\beta(t)$ segment-by-segment conditional on $\{\Delta_t\}$. Each constant segment of $\beta(t)$ is assigned a gamma prior; the full conditional distribution remains gamma-distributed by conjugacy. Hence, for each of these segments, we place the prior $\beta_k \sim \text{Ga}(a_0, b_0)$, denoting the parametrization with expectation a_0/b_0 and variance a_0/b_0^2 . The conditional distribution of each segment $(\beta_k|\Delta, X)$ then follows

$$\beta_k|\Delta, X \sim \text{Gamma}\left(a_0 + n_T, b_0 + \int_{c_k}^{c_{k+1}} I(t)S(t)dt\right). \quad (7)$$

One can choose how many components of Δ to propose per iteration of this joint update; in our applications, proposing one new component per iteration tends to admit a healthy acceptance probability between 0.3 and 0.6.

Finally, we update the latent data X using a Metropolis-Hastings step²⁸ conditional on β and γ . Since the number of latent infection and recovery times can be quite large, the success of DA-MCMC hinges on the construction of an efficient proposal for the latent data. The model specification in the previous section was carefully designed to remain tractable with these considerations in mind: in particular, it allows us to make use of efficient proposals from the *piecewise decoupled SIR* (PD-SIR) process. Briefly, the PD-SIR closely resembles the SIR, with the only difference being that the infection times follow a linear death process within each time interval $(c_{t-1}, c_t]$.²⁸ The proposal process is rate-linear, leading to closed form expressions for conditionally generating latent data that match the observed infection counts. Denoting the density of the PD-SIR by $q(\cdot)$, We accept X^* with probability

$$\alpha_2 = \min \left\{ 1, \frac{L(\Delta^j, \beta^j; X^*)q(X^{(j-1)}|\Delta^j, \beta^j)}{L(\Delta^j, \beta^j; X^{(j-1)})q(X^*|(\Delta, \beta)^j)} \right\}.$$

That is, though the latent data are proposed using a recent technique from an approximating process, the Metropolis-Hastings correction ensures that the accepted draws correspond to the exact model posterior under the time-inhomogeneous SIR model. Below, we provide a pseudocode summary of the overall sampler in Algorithm 1.

3 | RESULTS

3.1 | Simulation study

To validate the performance of our MCMC, we begin with simulation studies on synthetic datasets. First, we simulate an epidemic from a piecewise constant transmission rate. As a result, we produce a trajectory with clear deviations from a homogeneous SIR trajectory, which reflects the nature of the observed data in both the Ebola and COVID-19 case studies.

Algorithm 1. Data-Augmented MCMC for sampling from the posterior distribution

Require: observed data Y and initial parameters $(\Pi^{(0)}, \Delta^{(0)}, \beta(t)^{(0)})$
return $\{(\Pi^{(j)}, \Delta^{(j)}, \beta(t)^{(j)})\}_{j=0}^N$
 $X^{(0)} \leftarrow \text{draw}(\cdot | (\Delta, \beta)^{(0)}, Y)$ (initial latent data using the PD-SIR process)
for $j = 2, \dots, N$ **do**
 1. Gibbs Update on $\Pi_{2 \times 2}$:
 $\pi_{01}^{(j)} | \Delta^{(j-1)} \sim \text{Beta}(a_{01} + n_{01}, b_{01} + n_{00})$
 $\pi_{11}^{(j)} | \Delta^{(j-1)} \sim \text{Beta}(a_{11} + n_{11}, b_{11} + n_{10})$
 2. Metropolis-Hastings Update on $(\Delta, \beta(t))$:
 $\Delta^* | \Pi^j \sim \nu_1 \prod_{i=1}^n \pi_{\Delta_{i-1}, \Delta_i}$
 for $k = 1, \dots, K$ **do**
 $\beta_k^* | \Delta^* \sim \text{Gamma}\left(a_0 + n_{T_k}, b_0 + \int_{c_k}^{c_{k+1}} I(t)S(t)dt\right)$
 end for
 Accept (Δ^*, β^*) with probability $\min\left\{1, \frac{L(\Delta^*, \beta^*; X)P(\beta^* | \Delta^*)q(\beta^{j-1} | \Delta^{j-1})q(\Delta^{j-1} | \Pi^j)}{L(\Delta^{j-1}, \beta^{j-1}; X)P(\beta^{j-1} | \Delta^{j-1})q(\beta^* | \Delta^*)q(\Delta^* | \Pi^j)}\right\}$.
 3. Metropolis-Hastings Update on Latent Data X :
 $X^* \leftarrow \text{draw}(\cdot | (\Delta, \beta)^j, Y)$ (generated using the PD-SIR process)
 Accept X^* with probability $\min\left\{1, \frac{L(\Delta^j, \beta^j; X^*)q(X^{(j-1)} | \Delta^j, \beta^j)}{L(\Delta^j, \beta^j; X^{(j-1)})q(X^* | (\Delta, \beta)^j)}\right\}$.
end for

Next, we demonstrate that our model is still able to produce a close, yet interpretable, piecewise approximation of the truth under a misspecified setting in which the transmission rate value varies smoothly over time.

The shape of the trajectory and placement of change points in the first simulation setting are designed to resemble the Ebola outbreak data. In this setting, we take $S_0 = 10\,000$, $I_0 = 10$, with change points located at observation times times 3 and 10. The true parameter values are set at $\beta = (1.75e - 4, 1.25e - 4, 0.75e - 4)$, and $\gamma = 1$. We observe the outbreak at twelve observation intervals, and in the realization we considered here, the observed incidence counts are given by $\mathbf{I} = (32, 81, 170, 189, 201, 241, 243, 251, 257, 272, 120, 85)$. For reference, the maximum likelihood estimate for the transmission rate, obtained under knowledge of the complete data and true change point locations, is $\hat{\beta} = (1.92e - 4, 1.21e - 4, 0.75e - 4)$.

The observed incidence data are plotted in Figure 1. This trajectory displays clear deviations from the typical shape of the SIR, indicating that the homogeneous model may be a poor choice, suggesting the presence of potential change points. We chose hyperparameters to be uninformative; for each segment of β , we chose the weakly informative prior $\text{Gamma}(1,1)$. For the prior on Π , we used Jeffrey's prior on both π_{01} and π_{11} . We assume the recovery rate γ is known, fixing it to the true value.

The algorithm generates samples efficiently, taking less than 10 min to run 50 000 iterations of the chain on a single laptop machine, and achieves an effective sample size per second (ESS/s) of 2.39 in this setting. The results under this simulation setting are displayed in column one of Figure 1. Notably, even with uninformative priors on the number of change points, the algorithm is successful in assigning high probability only to the locations of true change points, empirically validating the Bayesian Occam's razor principle.³³ Moreover, we observe very little bias in estimates for each segment of the step-function, with the true transmission rate values well-contained within the 95% equal-tailed Bayesian credible intervals. This is noteworthy, since only 12 observed time points in the epidemic trajectory were supplied to the algorithm toward recovering such estimates. The posterior reflects less precision about the positioning of the second change point.

Although the credible intervals are generally narrow, we note that the variance of the estimates were larger near the front and end of the observed outbreak. This is evident in the posterior draws of $\beta(t)$ shown in (Figure 1). This can be understood both epidemiologically and statistically: it is more difficult to conduct inference on epidemic parameters observing only the tail or early stages of an epidemic. Indeed, note that the gamma distribution describing the conditional for β_k is $(a_0 + n_T)/(b_0 + \int_{c_k}^{c_{k+1}} I(t)S(t)dt)^2$; the term $I(t)$ appearing in the denominator tends to be small at the beginning of an outbreak or at its end, resulting in wider credible intervals. This perspective is

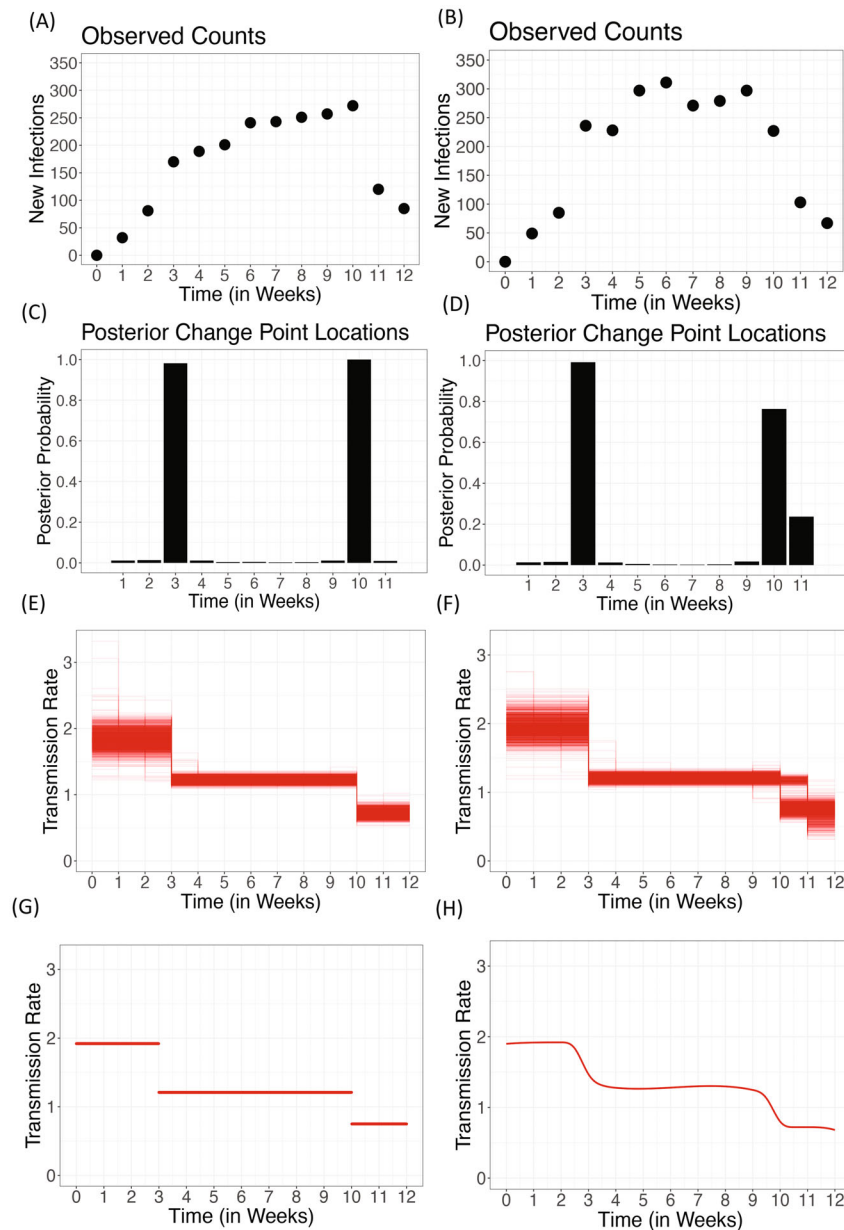


FIGURE 1 Simulation study results: Setting 1 is displayed in the left column and setting 2 in the right column. (A) Simulated incidence data under setting 1. (B) Simulated incidence data under setting 2. (C) The marginal posterior probabilities of a change point occurring at each week under setting 1. (D) The marginal posterior probabilities of a change point occurring at each week under setting 2. (E) Plot of the posterior samples for $\beta(t)$ under setting 1. (F) Plot of the posterior samples for $\beta(t)$ under setting 2. (G) True piecewise transmission rate value under setting 1. (H) True smoothly-varying transmission rate value under setting 2. (All transmission rate values are multiplied by a factor of 10 000 for readability.)

unavailable in the time-homogeneous case when inference is based around one constant value β that describes the entire outbreak.

To evaluate the convergence properties of the Markov chain, we run 3 separate chains with over-dispersed starting values $\beta_0 = (0.1 \cdot \beta, \beta, 10 \cdot \beta)$. Despite being initialized in various low-density regions, each chain appears to converge as evident from both the trace plots and a multivariate Gelman-Rubin convergence diagnostic value of less than 1.01 (Figure S5). The frequency of flips between ones and zeros for each $\{\Delta_k\}$ is a good indication that the MCMC mixes well and explores with high frequency the space of candidate change point configurations 11.

In reality, the drops in transmission rate following a policy intervention may not be abrupt. To test the performance of our model in this setting, we next run our model on an epidemic that was generated by a smoothly-varying, rather than piecewise, transmission rate. Specifically, this smoothly-varying rate was generated using a cubic B-spline with cut-points at times 2, 2.5, 3, 3.5, 4, 9, 9.5, 10, 10.5, and 11. Keeping the same initial conditions, $S_0 = 10\,000$ and $I_0 = 10$, this new transmission rate resulted in a trajectory with observed counts $\mathbf{I} = (49, 85, 236, 228, 297, 311, 271, 279, 297, 227, 103, 67)$ (Figure 1B). As in the first simulation setting, we keep hyperparameters uninformative, and assumed that the recovery rate γ is known, fixing it at the true value.

In contrast to the piecewise setting, in which both change points produced visibly notable deviations from the time-homogeneous SIR curve in the observed infection counts (Figure 1A), the continuous transmission rate presented a visibly more challenging problem. Specifically, the observed case count data of Figure 1B, show a less visible deviation from the typical time-homogeneous SIR, indicating that the homogeneous model may be a reasonable model under this observed trajectory. In particular, the more gradual decline in transmission rate around weeks nine through eleven (shown in Figure 1H) are less abrupt. However, the sampler still suggests with high posterior probability a drop in transmission rate around weeks ten through eleven, show that our model is able to give an interpretable, yet close, piecewise approximation of $\beta(t)$ (Figure 1F).

3.2 | Comparison to existing methods

We next compare the proposed model to the time-homogeneous stochastic SIR model described in Reference 28, the fixed change point models described in References 18 and 19, and popular general-purpose change point detection algorithms which ignore SIR dynamics.³⁵ We see that these previous approaches resulted in biased estimates for $\beta(t)$ in at least one segment (Figure 2). The time-homogeneous model's estimate for β fails to capture the presence of change points in the data, resulting in a biased estimate that is the average of three notably disparate regimes (Figure 2). A fixed change point model assumes the changes in transmission occur at known locations, frequently assumed to be the exact time of intervention. This is frequently done in practice in the epidemic modeling literature^{18,19}. Although these rigid change point

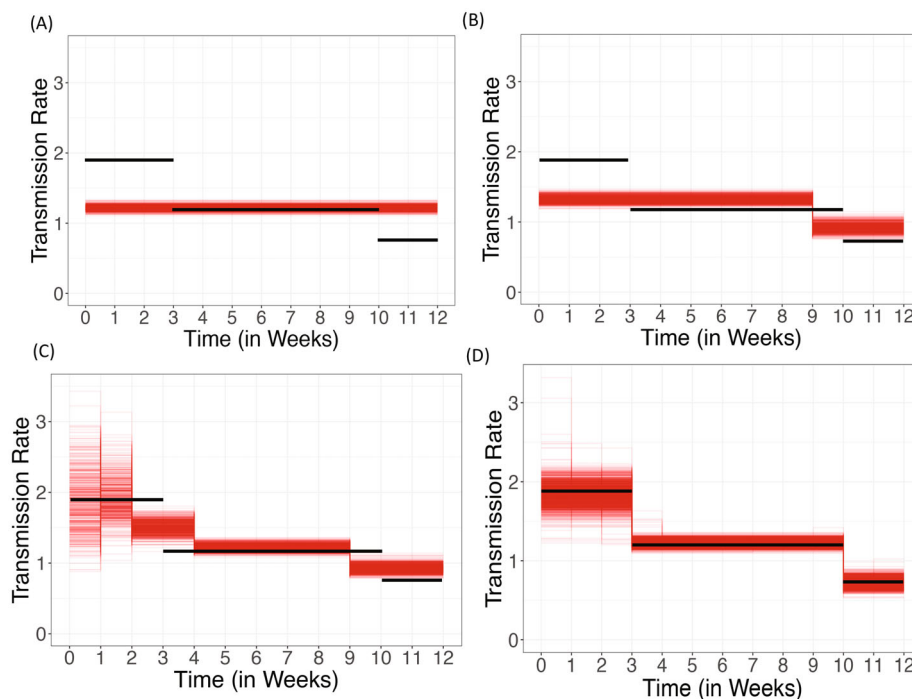


FIGURE 2 (A) Samples from the posterior transmission rate over time, recovered from the time-homogeneous model. (B) Samples from the posterior transmission rate over time, recovered from the fixed change point model. (C) Samples of the posterior transmission rate over time, recovered using binary segmentation and PELT. (D) Samples from the posterior transmission rate over time, recovered using our proposed model. Black lines indicate the true transmission rate values.

models allow for a piecewise transmission rate, they lead to biased estimates of the effects of control interventions by assuming, rather than learning, the change point locations. Of course these models fail to account for additional change points that are not known or specified by the user, resulting in biased estimates of the magnitudes of the drops. Specifically, the first piecewise segment of the estimate for $\beta(t)$ in Figure 2B has a bias of 0.103 while the second has a bias of -0.183 .

Finally, to show the inaccuracies that arise when we disregard the SIR dynamics altogether, we compare our method to two classical change point detection algorithms, binary segmentation and PELT.³⁵ Both algorithms were designed to detect change points in the rate of a Poisson model. However, these two methods do not benefit from the additional information given by SIR compartment sizes and dynamics toward detecting change points. As a consequence, it picks up spurious change points at weeks 1 and 2 and misplaces the locations of the change points at weeks 3 and 9. This in turn results in biased estimates of each piecewise segment of the transmission rate (Figure 2C).

This illustrates the advantage of learning change points of the transmission rate function, over casting it in a classical setting operating on the count data themselves in a way that is model agnostic. Within these comparisons, we see that our proposed method is advantageous for detecting change points in the context of epidemic count data.

3.3 | Applications to outbreaks of Ebola and COVID-19

We next consider case studies on data collected from the Ebola outbreak in Guinea in 2014, and a COVID-19 outbreak on a university campus in 2021.

The Ebola dataset we study consists of weekly infection counts from January 1 to December 31, 2014, collected in Guéckédou, Guinea, a prefecture with 292 000 inhabitants. These data were collected over a time period in which control measures were implemented at various times,³⁶ making it a natural candidate for modeling the force of infection to vary over time. Applying our model with piecewise components at a monthly resolution, we obtain results displayed in the left column of Figure 3. Figure 3A shows the observed cases of Ebola reported over time on a weekly basis, with the red lines marking places of the most likely change point occurrences, according to the model. The model suggests the most likely change point locations to be in March and August. Figure 3C shows the marginal posterior probabilities of a change point occurring at each month. Figure 3E shows a plot of the posterior samples of transmission rate.

The change points in the data may give insights into the efficacy of the control measures implemented in Guinea in 2014. Interpreting these results, we note that March marked the closing of public schools in Guéckédou and the confirmation of Ebola as the infectious agent.³⁶ August marked a closing of borders with Sierra Leone and Liberia and the implementation of exit screening and flight cancellations. Our findings in terms of changes in $\beta(t)$ coinciding with these measures suggest they are associated in slowing the spread of the outbreak, as reflected in a lower inferred force of transmission. An important quantity in the spread of an epidemic is the effective reproduction number $R(t) = \frac{\beta(t)}{\gamma} S(t)$ which corresponds to the expected number of secondary infections arising from one infectious individual at time t . The posterior trajectories of $R(t)$ drop in value from 2.25 to 1.08, and then then to 0.74 following August. These results indicate that the virus was initially spreading rapidly in the population, $R(t) > 1$, before being brought under control at the end of the year, $R(t) < 1$. Our model did not, however, pick up a change in October when the government banned country-wide celebrations of Eid,³⁶ suggesting that this measure may have had a less significant effects on the transmission rate as reflected in the data.

We next consider inference using data from Duke University's COVID-19 Testing Tracker in the spring of 2021. The data consist of weekly new positive case counts for students, faculty, and staff across campus from January 11 to April 25, 2021. During this time, Duke conducted sample testing without respect to reported symptoms. At the end of week 9, the university implemented a stay-in-place order (March 8-14), during which students were required to stay in their residences at all times except for essential activities related to food, health, or safety. Figure 3 shows the results of our model applied to these incidence data.

Figure 3B shows the observed cases of COVID-19 reported over time on a weekly basis, with the red lines marking all places of likely change point occurrences, according to the model. Figure 3D shows the marginal posterior probabilities of a change point occurring at each month. Figure 3F shows a plot of the posterior samples of transmission rate. Here, our method picks up a pronounced change, where we observe a 4.5 fold decrease in β from week 9 to week 10, suggesting that the university-wide stay-in-place order was highly effective in reducing the spread of the virus (Figure 3). Moreover, our model estimates that $R(t)$ dropped below 1 following the stay-in-place order, suggesting that the spread of the virus was at least temporarily under control following the mitigation policy. A reviewer remarks that changes in $R(t)$ can also

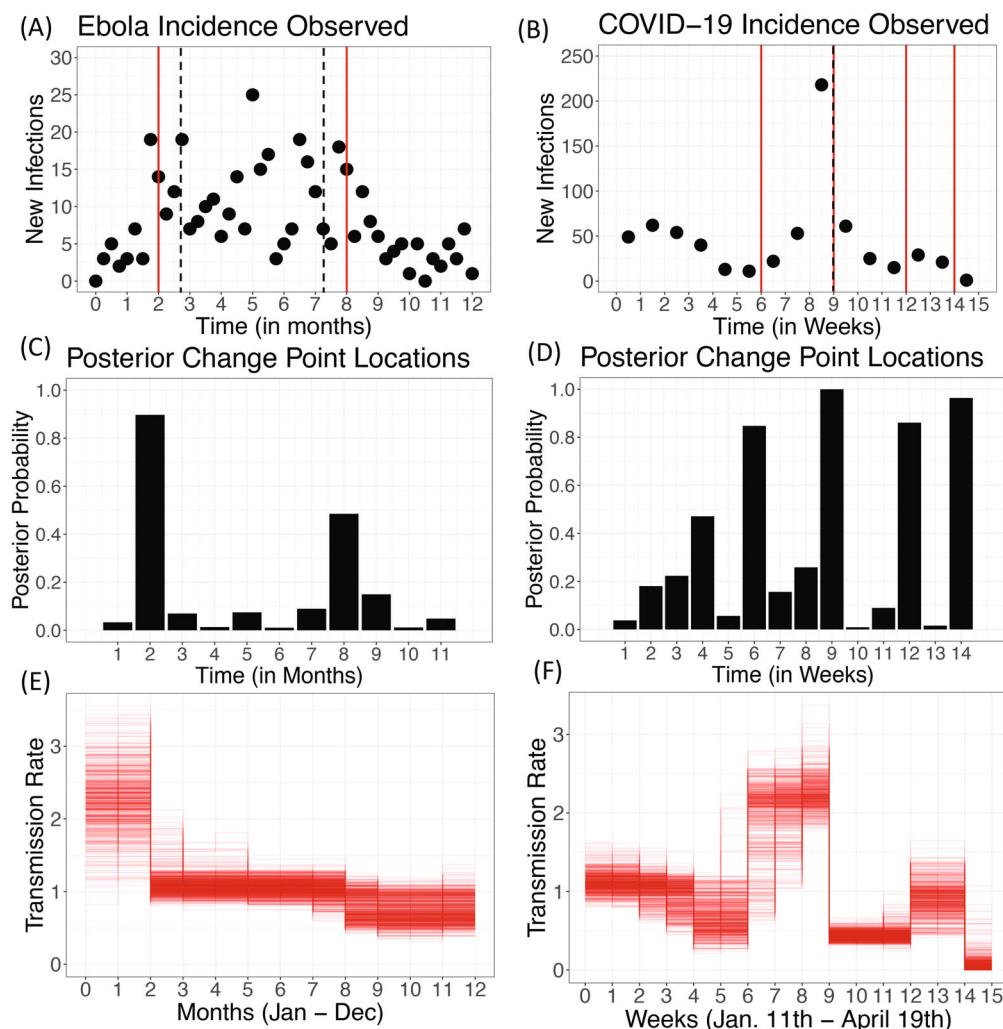


FIGURE 3 Application to two real outbreaks: Ebola and COVID-19. (A) Counts of weekly new infections of Ebola cases in Guéckédou, Guinea from January to December, 2014. Red lines mark the positions with high posterior probability of being a change point. Black dotted lines mark the locations of policy changes. (B) Counts of weekly new infections of COVID-19. Red lines indicate positions with high posterior probability of being a change point. Black dotted lines mark the location of a policy change. (C) The marginal posterior probabilities of a change point occurring at each month for the Ebola outbreak. (D) The marginal posterior probabilities of a change point occurring at each week for the COVID-19 outbreak. (E) Plot of the posterior samples for the transmission rate for Ebola in Guéckédou, Guinea. (F) Plot of the posterior samples for the transmission rate for COVID-19. (All transmission rate values are scaled by the population size for readability.)

be attributed to depletion of the S compartment, but because in both applications here the number of total cases is small relative to the population, it is reasonable to interpret change points largely in terms of changes in transmission rate.

Following a reviewer's suggestion, Figure 4 presents a side-by-side comparison of the recovered transmission rate among the four methods in Section 3.2. The top row of Figure 4 compares the four methods run on the Ebola data and the bottom row on the COVID-19 data.

Model diagnostics

To assess model fit for each of the four methods compared in Section 3.2, we perform posterior predictive checks based on the number of new cases over time. For each data set, we simulate epidemic trajectories from the posterior predictive distribution and then compute the trajectory of new cases over time in each simulation. This produces a posterior predictive distribution of cases over time, as plotted in Figures 5 and 6. For our proposed model, the

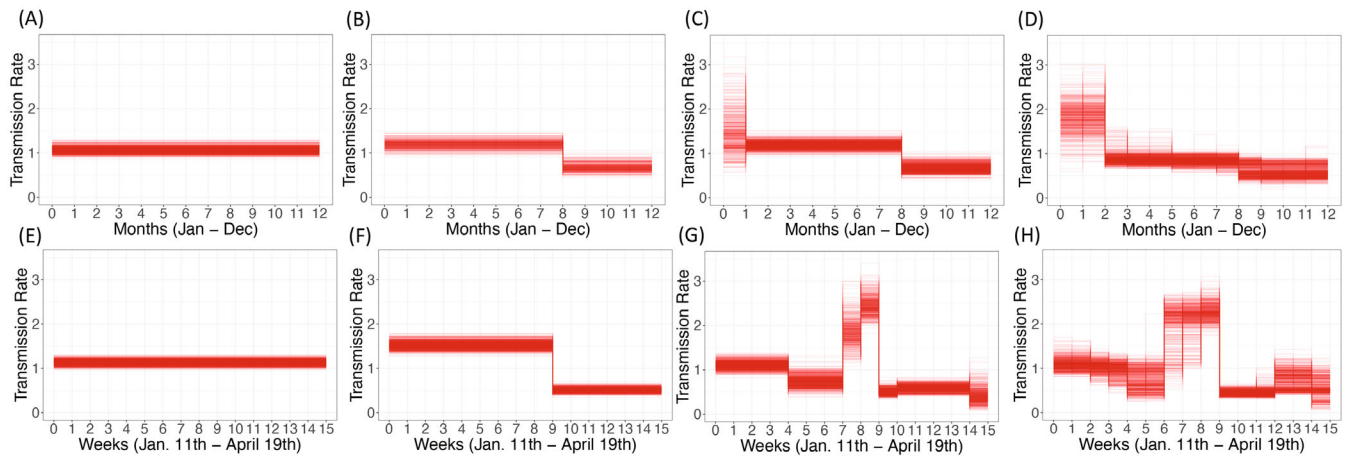


FIGURE 4 Ebola (top row): (A) Samples from the posterior transmission rate over time, recovered from the time-homogeneous model. (B) Samples from the posterior transmission rate over time, recovered from the fixed change point model. (C) Samples of the posterior transmission rate over time, recovered using binary segmentation and PELT. (D) Samples from the posterior transmission rate over time, recovered using our proposed model. COVID-19 (bottom row): (E) Samples from the posterior transmission rate over time, recovered from the time-homogeneous model. (F) Samples from the posterior transmission rate over time, recovered from the fixed change point model. (G) Samples of the posterior transmission rate over time, recovered using binary segmentation and PELT. (H) Samples from the posterior transmission rate over time, recovered using our proposed model.

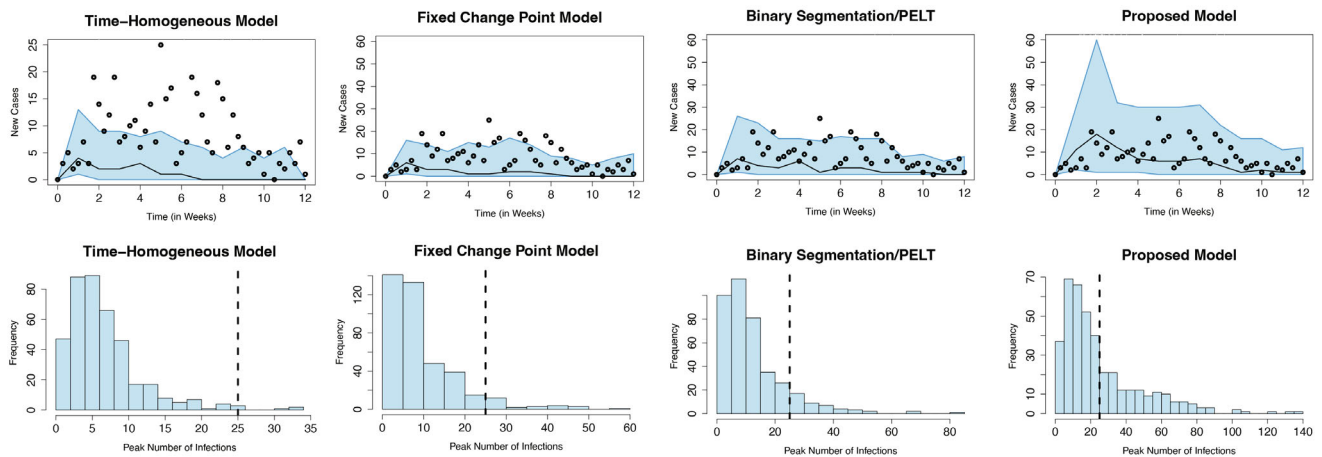


FIGURE 5 Ebola: Posterior predictive distributions of new cases (top row) and peak number of infections (bottom row) for each of the four methods. Observed data are (top row) plotted in black or (bottom row) marked with a dashed line. The proposed model performed best at fitting the observed data.

observed data fall reasonably within the 95% posterior predictive interval for both the Ebola and COVID-19 case studies. We note that there is larger variation in the stochastic trajectories simulated by the COVID-19 model. This is in line with what one would expect, given that the observed data consist of only 15 observation times in the latter case study.

4 | DISCUSSION

This article develops a novel framework for exact Bayesian inference in a stochastic SIR model with a piecewise constant time-varying transmission rate $\beta(t)$. The model is designed to be flexible enough to capture time-heterogeneity in the

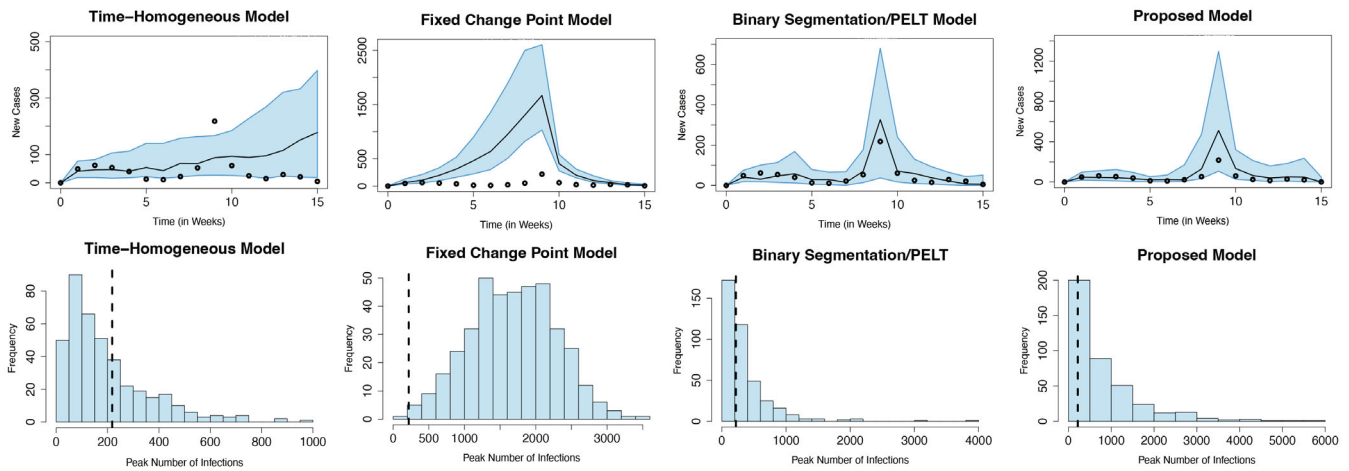


FIGURE 6 COVID-19: Posterior predictive distributions of new cases (top row) and peak number of infections (bottom row) for each of the four methods. Observed data are (top row) plotted in black or (bottom row) marked with a dashed line. The proposed model and the binary segmentation/PELT models performed the best at fitting the observed data.

infection rate while remaining amenable to sampling targeting the exact model posterior. The resulting method yields interpretable parameter estimates favoring model parsimony and comes equipped with corresponding uncertainty estimates. To fit the model to data, we propose a data-augmented MCMC algorithm that samples from the joint posterior distribution of the model parameters and latent variables, performing change point estimation in the process. The framework is user-friendly and allows for prior specification on how strongly one favors constant runs in the transmission rate, leveraging recent developments in efficient proposals for the latent space of unobserved epidemic events when the observed data consist of incidence counts.

Detecting changes in disease spread allows us to gain valuable insight into the efficacy of mitigation policies that were implemented, informing us on where and how quickly a specific policy begins to be effective. While we show that the method is successful in assigning high probability to the true change points in simulations where the ground truth $\beta(t)$ is piecewise constant, our method provides interpretable and flexible approximate estimates even when the ground truth transmission rate is continuously varying. From a practical perspective, the posterior plots of the learned transmission rate are easy to interpret for non-statisticians; not only can one immediately see the uncertainty of each segment, but the most pronounced change points as supported by the data are visually obvious. Finally, because we have a continuous-time generative model, the sampler is ready to be used off-the-shelf for incidence data even when the observation intervals are irregular.

It is important to keep in mind, however, that our method seeks only to describe a change in the transmission rate parameter of the epidemic model. It is important not to over-interpret results of inference—for instance, we cannot make causal claims or attribute changes as directly resulting from a given change in policy. Compared to some prior works that assign a change point time coinciding with the time of a policy change occurring,^{18,19} which make a stronger implicit claim, we improve by instead detecting changes only from incidence data. Moreover, we note that the current model does not regard information on the contact network structure of the epidemic. Similarly, it is therefore possible that changes in overall transmission rate may be attributed to local cluster outbreaks rather than changes in the rate of infectiousness inherent to the disease.

The methodology proposed here leads naturally to several extensions. We have seen success with the piecewise model in empirical studies in this article, and future work may analyze recovery guarantees from a theoretical perspective. The focus of this article is on developing the model framework and enabling efficient Bayesian computation. Not only does this change point model lend a natural interpretation for getting a direct quantity of the drop, or decline, in transmission rate following a policy intervention, but it is also able to segment the epidemic into different stages based on transmission rate. The experiments presented here are focused on a time-varying transmission rate β . However, the proposed change methodology can easily be adapted to model other parameters, such as the recovery rate γ , to be time-heterogeneous as well. We focused on a time-varying transmission rate here as the case studies presented here occurred during time periods in which policy changes were implemented, making it relevant to detect changes in transmission rates. In analyzing

settings where new strains are introduced into the population, for instance, it may be useful to introduce a time-varying γ . Moreover, the method can extend to settings where the true transmission rate need not be piecewise constant. Along these lines, future studies may aim to model a more flexible transmission rate, incorporating time trends such as seasonality in terms beyond piecewise constant building blocks. Past work using more flexible functions such as Gaussian processes have been successful, though they incur significantly higher computational cost and do not scale to large outbreaks due to operations such as matrix inversions.^{13,37} In some cases, these samplers must run for hours for small outbreaks with only a hundred infections.³⁷ Our proposed framework can generate thousands of posterior samples for outbreaks with hundreds of infections in just several minutes on a single machine, suggesting there is room for further model complexity while retaining computational tractability. Additionally, in settings where more information about the outbreak is available, one may consider defining a regression model for $\beta(t)$ based on covariates related to control measures or models of human behavior.³⁸ Levels of mobility, levels of vaccination, and other auxiliary indicators of transmission³⁹ may further be incorporated into the rate function. The broader inferential framework proposed here readily extends to such settings.

The sampler scales efficiently in per iteration computational cost, allowing applications to much larger outbreaks than those considered in our empirical studies. In the settings we considered, our DA-MCMC algorithm mixes adequately as evident in the trace plots for Δ in Figure S5, which exhibit frequent jumps between 0 and 1. The space of possible change points is already large here, and efficiently sampling over these spaces is a notoriously difficult problem. As we discussed, it may be expected that our MCMC takes a longer time to mix for select parameters that border a change point. For such β_k values, we have observed the posterior distribution can appear bimodal, with the two modes corresponding to two different piecewise regimes. This points to future work that may further improve mixing for these parameters through designing informed proposals not only based on the structure of the disease model, but also on the discrete space of change point configurations. One fruitful line of thought along these lines for high-dimensional discrete problems is the locally balanced proposal described in Reference 40, which uses point-wise informed proposals and balancing functions to bias proposals toward high-probability states.

Our data augmentation approach complements classical changepoint methodology, but open directions remain in further integrating powerful ideas from the changepoint literature within stochastic epidemic models. For instance, References 23 and 35 proposed a set of dynamic programming algorithms to perform direct simulation from the posterior distribution of change point positions. These dynamic programming approaches minimize the cost over all possible segmentations of the data exactly under a given statistical criteria, sidestepping the need to design careful MCMC moves when there is a large space of models to explore. However, to preserve the necessary Markov property that is required for the forward-backward recursions required by these algorithms, these methods require that the observations are conditionally independent on the change points and the parameter values—an assumption that does not hold for SIR case counts explicitly depend on past values of S, I . Similarly, Reference 25 propose a Bayesian variable selection procedure for multiple change point detection not requiring MCMC. Again, these methods are developed in the context of detecting changes in the mean under Gaussian observations, for which they design a sequential procedure for deriving marginal posterior probabilities of change points in closed form. Adapting these methodologies toward parametric change-point detection within mechanistic models such as the SIR remains a promising future direction in a challenging dependent data setting.

ACKNOWLEDGEMENTS

This work was partially supported by NIH Grants R01ES028804 and R01ES027498, and NSF Grants DMS-2230074 and PIPP-2200047. We thank the 2021 DOMath team (Trevor Bowman, Gregory Orme, Pranay Pherwani, Omar Melikechi, and Tao Tang) for contributing valuable conversations during the early stages of this project.

DATA AVAILABILITY STATEMENT

R code is available at https://github.com/JennyHuang19/damcmc_timevarying_beta, including all open-source code and scripts for reproducing the results in this article.

ORCID

Jason Xu  <https://orcid.org/0000-0001-5472-3720>

REFERENCES

1. Zelner J, Riou J, Etzioni R, Gelman A. Accounting for uncertainty during a pandemic. *Patterns*. 2021;2(8):100310.
2. Eikenberry SE, Mancuso M, Iboi E, et al. To mask or not to mask: modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infect Dis Model*. 2020;5:293-308.
3. Dehning J, Zierenberg J, Spitzner FP, et al. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science*. 2020;369(6500):eabb9789.
4. Hong HG, Li Y. Estimation of time-varying reproduction numbers underlying epidemiological processes: a new statistical tool for the COVID-19 pandemic. *PLoS One*. 2020;15(7):e0236464.
5. Allen LJ, Driessche PV. Stochastic epidemic models with a backward bifurcation. *Math Biosci Eng*. 2006;3(3):445-458.
6. Andersson H, Britton T. *Stochastic Epidemic Models and their Statistical Analysis*. Vol 151. New York: Springer Science & Business Media; 2012.
7. Dureau J, Kalogeropoulos K, Baguelin M. Capturing the time-varying drivers of an epidemic using stochastic dynamical systems. *Biostatistics*. 2013;14(3):541-555.
8. Fintzi J, Wakefield J, Minin VN. A linear noise approximation for stochastic epidemic models fit to partially observed incidence counts. *Biometrics*. 2021;78:1530-1541.
9. Guttorp P. *Stochastic Modeling of Scientific Data*. Boca Raton, FL: Chapman and Hall/CRC; 2018.
10. Boys RJ, Giles PR. Bayesian inference for stochastic epidemic models with time-inhomogeneous removal rates. *J Math Biol*. 2007;55(2):223-247.
11. Hobolth A, Stone EA. Simulation from endpoint-conditioned, continuous-time Markov chains on a finite state space, with applications to molecular evolution. *Ann Appl Stat*. 2009;3(3):1204-1231.
12. Ho LST, Xu J, Crawford FW, Minin VN, Suchard MA. Birth/birth-death processes and their computable transition probabilities with biological applications. *J Math Biol*. 2018;76(4):911-944.
13. Kypraios T, O'Neill PD. Bayesian nonparametrics for stochastic epidemic models. *Stat Sci*. 2018;33(1):44-56.
14. McKinley TJ, Ross JV, Deardon R, Cook AR. Simulation-based Bayesian inference for epidemic models. *Comput Stat Data Anal*. 2014;71:434-447.
15. King AA, Nguyen D, Ionides EL. Statistical inference for partially observed Markov processes via the R package pomp. arXiv preprint arXiv:1509.00503, 2015.
16. Ionides EL, Nguyen D, Atchadé Y, Stoev S, King AA. Inference for dynamic and latent variable models via iterated, perturbed Bayes maps. *Proc Natl Acad Sci*. 2015;112(3):719-724.
17. McKinley TJ, Vernon I, Andrianakis I, et al. Approximate Bayesian computation and simulation-based inference for complex stochastic epidemic models. *Stat Sci*. 2018;33:4-18.
18. Lekone PE, Finkenstädt BF. Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. *Biometrics*. 2006;62(4):1170-1177.
19. Ward C, Brown GD, Oleson JJ. An individual level infectious disease model in the presence of uncertainty from multiple, imperfect diagnostic tests. *Biometrics*. 2021;79:426-436.
20. Barry D, Hartigan JA. A Bayesian analysis for change point problems. *J Am Stat Assoc*. 1993;88(421):309-319.
21. Chib S. Estimation and comparison of multiple change-point models. *J Econom*. 1998;86(2):221-241.
22. Davis RA, Lee TCM, Rodriguez-Yam GA. Structural break estimation for nonstationary time series models. *J Am Stat Assoc*. 2006;101(473):223-239.
23. Fearnhead P. Exact and efficient Bayesian inference for multiple changepoint problems. *Stat Comput*. 2006;16(2):203-213.
24. Davis RA, Huang D, Yao YC. Testing for a change in the parameter values and order of an autoregressive model. *Ann Stat*. 1995;23:282-304.
25. Cappello L, Padilla OHM, Palacios JA. Scalable Bayesian change point detection with spike and slab priors. arXiv:2106.10383, 2021.
26. Fryzlewicz P. Wild binary segmentation for multiple change-point detection. *Ann Stat*. 2014;42(6):2243-2281.
27. Raftery AE, Akman VE. Bayesian analysis of a Poisson process with a change-point. *Biometrika*. 1986;73(1):85-89.
28. Morsomme R, Xu J. Uniformly ergodic data-augmented MCMC for fitting the general stochastic epidemic model to incidence data. arXiv:2201.09722, 2022.
29. Bailey N. *The Mathematical Theory of Infectious Diseases and its Applications*. London: Charles Griffin & Company Ltd; 1975.
30. Tibshirani R, Saunders M, Rosset S, Zhu J, Knight K. Sparsity and smoothness via the fused lasso. *J R Stat Soc Series B Stat Methodology*. 2005;67(1):91-108.
31. Abraham K, Castillo I, Gassiat E. Multiple testing in nonparametric hidden Markov models: an empirical Bayes approach. *J Mach Learn Res*. 2022;23(94):1-57.
32. Fox EB, Sudderth EB, Jordan MI, Willsky AS. An HDP-HMM for systems with state persistence. *Proceedings of the 25th International Conference on Machine Learning*. New York: Association for Computing Machinery; 2008:312-319.
33. Jefferys WH, Berger JO. Ockham's razor and Bayesian analysis. *Am Sci*. 1992;80(1):64-72.
34. Tanner MA, Wong WH. The calculation of posterior distributions by data augmentation. *J Am Stat Assoc*. 1987;82(398):528-540.
35. Fearnhead P, Rigai G. Changepoint detection in the presence of outliers. *J Am Stat Assoc*. 2019;114(525):169-183.
36. Coltart CE, Lindsey B, Ghinai I, Johnson AM, Heymann DL. The Ebola outbreak, 2013–2016: old lessons for new epidemics. *Philos Trans R Soc Lond B Biol Sci*. 2017;372(1721):20160297.

37. Xu X, Kypraios T, O'Neill PD. Bayesian non-parametric inference for stochastic epidemic models using Gaussian processes. *Biostatistics*. 2016;17(4):619-633.
38. McAdams D, Song Y, Zou D. Equilibrium social activity during an epidemic. *J Econ Theory*. 2023;207:105591.
39. McDonald DJ, Bien J, Green A, et al. Can auxiliary indicators improve COVID-19 forecasting and hotspot prediction? *Proc Natl Acad Sci*. 2021;118(51):e2111453118.
40. Zanella G. Informed proposals for local MCMC in discrete spaces. *J Am Stat Assoc*. 2020;115(530):852-865.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Huang J, Morsomme R, Dunson D, Xu J. Detecting changes in the transmission rate of a stochastic epidemic model. *Statistics in Medicine*. 2024;43(10):1867-1882. doi: 10.1002/sim.10050