



# Bayesian modeling of dynamic behavioral change during an epidemic



Caitlin Ward <sup>a,\*</sup>, Rob Deardon <sup>b,c</sup>, Alexandra M. Schmidt <sup>d</sup>

<sup>a</sup> Division of Biostatistics, University of Minnesota, Minneapolis, MN, USA

<sup>b</sup> Faculty of Veterinary Medicine, University of Calgary, Calgary, AB, Canada

<sup>c</sup> Department of Mathematics and Statistics, University of Calgary, Calgary, AB, Canada

<sup>d</sup> Department of Epidemiology, Biostatistics, and Occupational Health, Montreal, QC, Canada

## ARTICLE INFO

### Article history:

Received 13 June 2023

Received in revised form 20 July 2023

Accepted 3 August 2023

Available online 6 August 2023

Handling Editor: Dr Yijun Lou

### Keywords:

Bayesian inference

Compartmental model

SIR

SEIR

Transmission modeling

## ABSTRACT

For many infectious disease outbreaks, the at-risk population changes their behavior in response to the outbreak severity, causing the transmission dynamics to change in real-time. Behavioral change is often ignored in epidemic modeling efforts, making these models less useful than they could be. We address this by introducing a novel class of data-driven epidemic models which characterize and accurately estimate behavioral change. Our proposed model allows time-varying transmission to be captured by the level of "alarm" in the population, with alarm specified as a function of the past epidemic trajectory. We investigate the estimability of the population alarm across a wide range of scenarios, applying both parametric functions and non-parametric functions using splines and Gaussian processes. The model is set in the data-augmented Bayesian framework to allow estimation on partially observed epidemic data. The benefit and utility of the proposed approach is illustrated through applications to data from real epidemics.

© 2023 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

Human behavior is a driving factor in the spread of infectious disease through human populations. In the presence of increasing infection risk, individuals typically engage in protective behaviors to avoid becoming ill. These preventative behavior changes may be imposed by a governing body (e.g., city-wide lockdowns or school closures), or may be the result of personal choices (e.g., social distancing or voluntary masking). Behavioral change can have a substantial impact on the epidemic trajectory by delaying the peak, reducing the total number of individuals that contract the disease, and/or resulting in multiple waves of transmission. Additionally, behavioral change is dynamic; higher disease prevalence tends to result in increased preventative measures, which are subsequently relaxed as prevalence decreases. Understanding these changes in population behavior in response to an epidemic is crucial for public health practitioners and policy makers attempting to stop or slow the spread of the pathogen and allocate valuable resources.

\* Corresponding author.

E-mail address: [ward-c@umn.edu](mailto:ward-c@umn.edu) (C. Ward).

Peer review under responsibility of KeAi Communications Co., Ltd.

Statistical modeling of infectious disease transmission provides a quantitative approach to understanding disease dynamics. The conventional methodology is based on the compartmental SIR model (Kermack & McKendrick, 1927), which segments the population into Susceptible, Infectious, and Removed compartments capturing the important disease states. The model is then parameterized in terms of the rates of flow between compartments. Compartmental models can be implemented deterministically using ordinary differential equations or stochastically, with statistical inference typically carried out using Bayesian methodology. However, the traditional SIR model does not naturally account for transmission dynamics changing in real time as the population reacts to the outbreak, limiting its applicability to real epidemic data.

Previous work incorporating behavioral change into the SIR model framework has been done in the deterministic setting, with models generally falling into one of three categories. One approach adds additional “adherence” or “awareness” compartments to capture reductions in susceptibility and transmissibility for individuals engaging in protective behaviors, with individuals more likely to be adhering to preventative measures when disease prevalence is high (Acuña-Zegarra et al., 2020; Agaba et al., 2017; Del Valle et al., 2005; Perra et al., 2011). Another popular method uses a game-theoretic approach, supplementing the SIR model with a time-varying utility function balancing the costs associated with prevention strategies (monetary, liberty, social) with the benefit of lowering infection risk in the population (Reluga, 2010; Fenichel et al., 2011). The remaining approach allows the transmission rate in the SIR model to be dynamically modified as a function of the recent disease trajectory (Capasso & Serio, 1978; Greenhalgh et al., 2015; Eksin et al., 2019; Weitz et al., 2020; Franco, 2020).

While these deterministic approaches are often easy to describe and simulate from, they can be unrealistic as disease transmission is an inherently stochastic process. Stochasticity is particularly important when describing disease spread in a small population or at the start of an epidemic when there is a small number of infectious individuals, when stochastic events such as so-called “superspreading” are non-negligible (Roberts et al., 2015). The Bayesian implementation incorporates stochasticity and offers several advantages. Epidemic data are often incomplete as infection or recovery times may not be observed. Without such data, deterministic models cannot be properly calibrated. However, Bayesian models implemented with Markov Chain Monte Carlo (MCMC) methods can impute missing data allowing for numerical integration over the probability distributions of the unobserved process. Thus, Bayesian models can perform parameter estimation in the presence of incomplete data, while properly accounting for uncertainty in the estimates. This is known as data-augmentation. Additionally, the Bayesian framework allows for the incorporation of prior information about the disease process, which is often available and can help induce identifiability. This is especially important when data augmentation is required. Finally, the Bayesian approach is also ideally suited for allowed captured parameter uncertainty to be propagated forwards when forecasting, via posterior predictive simulation, giving more realistic uncertainty bounds on forecasts.

However, there are still several unanswered questions. First, can these deterministic behavioral change approaches be translated to the stochastic Bayesian setting? Many of the proposed models include highly complex depictions of behavioral dynamics, however, Bayesian models often require simpler parameterizations than their deterministic counterparts to be computationally tractable (Andersson & Britton, 2012). Second, can disease and behavioral change parameters be estimated from epidemic data? The aforementioned modeling efforts have focused almost entirely on model specification, with results coming from forward simulations using pre-specified parameter values. Notably absent in the literature is any assessment of the statistical properties of these models when fit to real data, due to the aforementioned challenges with calibration deterministic models in the presence of limited data. Finally, how can these models be used in practice to increase understanding of human behavior during an epidemic? Behavioral change is anticipated during any real epidemic, but without any previous work fitting these models to data, their advantage is unknown.

In this article, we advance the field of stochastic infectious disease modeling by answering these important questions. We accomplish this by proposing a novel Bayesian SIR model formulation which captures dynamic behavioral change during an epidemic. The proposed model specifies the transmission rate as a function of recent disease occurrence, and computation is performed via MCMC methods. In our simulation study, we thoroughly investigate the statistical properties of our Bayesian model when fit to data. In particular, we show that behavioral change parameters can be accurately estimated and that posterior predictions from the proposed model can detect subsequent peaks in incidence. To showcase the benefits of considering behavioral change in this class of models, we apply the model to data from an Ebola outbreak in the Democratic Republic of the Congo and the COVID-19 pandemic in New York City.

## 2. Methods

### 2.1. Traditional SIR model

We model transmission using a discrete-time SIR model framework, where susceptible individuals can contract the infection from those who are infectious, and infectious individuals are removed when they no longer transmit the pathogen to others, due to death, isolation, or recovering with immunity. Let  $t = 1, \dots, \tau$  indicate discrete calendar time and  $S_t$ ,  $I_t$ , and  $R_t$  denote the number of individuals in the susceptible, infectious, and removed compartments in the continuous time interval  $[t, t + 1]$ , respectively. Furthermore, define the transition vectors  $I_t^*$  and  $R_t^*$  to represent the number of individuals entering the indicated compartment in this interval. Compartment membership is temporally described by the set of difference equations:

$$S_{t+1} = S_t - I_t^* \quad (1)$$

$$I_{t+1} = I_t + I_t^* - R_t^* \quad (2)$$

$$R_{t+1} = R_t + R_t^*. \quad (3)$$

We assume a closed population, such that  $S_t + I_t + R_t = N$  at all time points, where  $N$  denotes the total population size. Given the population size, a set of initial conditions, and the transition vectors, the compartment membership vectors can be fully determined using Equations (1)–(3).

In the Bayesian framework, we must establish the relationship between data and model parameters using probability distributions. We define the transitions between compartments to be binomially distributed (Lekone & Finkenstädt, 2006; Wilson & Burke, 1942), such that  $I_t^* \sim \text{Bin}(S_t, \pi_t^{(SI)})$  and  $R_t^* \sim \text{Bin}(I_t, \pi^{(IR)})$ . The transition probabilities  $\pi_t^{(SI)}$  and  $\pi^{(IR)}$  describe transmission of the pathogen and the duration of the infectious period, respectively. Assuming an independent Poisson contact process and constant probability of infection given a contact, the transmission probability,  $\pi^{(SI)}$ , takes the form

$$\pi_t^{(SI)} = 1 - \exp\left(-\beta \frac{I_t}{N}\right). \quad (4)$$

The parameter  $\beta$  is interpreted as the transmission rate, which captures both the contact rate and the infection probability, as these are not separately identifiable (Brown et al., 2016). The removal probability,  $\pi^{(IR)}$ , is derived by assuming the length of time an individual is infectious is exponentially distributed with rate  $\gamma$ . In discrete time,  $\pi^{(IR)}$  is the conditional probability of transitioning on day  $s+1$ , given the individual has remained infectious through time  $s$ , resulting in  $\pi^{(IR)} = 1 - \exp(-\gamma)$ . The parameter  $\gamma$  is referred to as the removal rate, however, it is typically more interpretable to consider  $1/\gamma$ , the mean length of the infectious period.

The traditional SIR model assumes  $\beta$  is constant over time, but this is generally not realistic. More likely, transmission changes as the population responds to the outbreak. Seasonal factors may also contribute, such as the start of the school year or changes in weather. This is incorporated by modifying the transmission probability from Equation (4) as

$$\pi_t^{(SI)} = 1 - \exp\left(-\beta_t \frac{I_t}{N}\right), \quad (5)$$

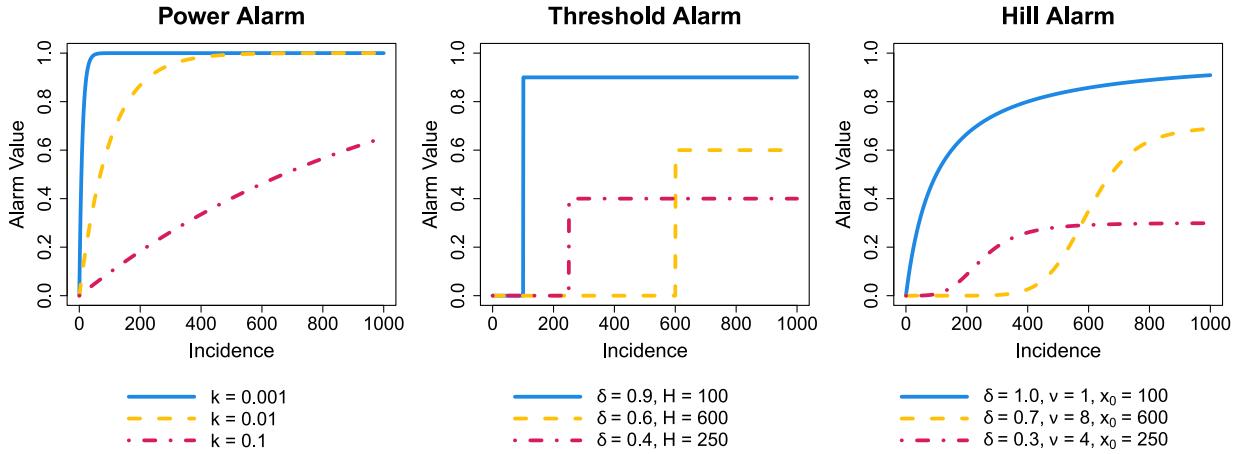
where  $\beta_t$  is the transmission rate at time  $t$ . Changes in transmission can be modeled directly through covariates, such as change points corresponding to the timing of government intervention(s) (Lekone & Finkenstädt, 2006; Ward et al., 2023b) or measures of population mobility (Liu et al., 2020; Sartorius et al., 2021). This allows inference to be made on the relationship between covariate(s) and transmission. Covariates might not capture all important changes in transmission, so more flexible approaches have been proposed, including basis splines (Brown et al., 2016; Hong & Li, 2020), Gaussian processes (Xu et al., 2016), or simple random walks (Irons & Raftery, 2021). All of these approaches are limited when forecasting, as it is difficult to predict the lifting of a government lockdown, and restrictive assumptions must be made to allow any flexibly modeled trajectory of  $\beta_t$  to continue into the future. Thus, an alternate approach accounting for the mechanism of behavioral change is needed.

## 2.2. Behavioral change (BC) model

The proposed behavioral change (BC) model allows for time-varying transmission which captures behavioral change via a dynamically structured dependence on previously observed epidemic trajectory. This is accomplished by allowing a constant transmission rate,  $\beta$ , to be modified by a time-varying level of alarm in the population, denoted  $a_t$ , such that  $\beta_t = \beta(1 - a_t)$ . We consider  $a_t \in [0, 1]$ , such that  $a_t$  corresponds to the proportional reduction in transmission due to the alarm in the population. When  $a_t = 0$ , the population is in its natural ‘unalarmed’ state, and transmission is described only by  $\beta$ . When  $a_t = 1$ , the population is in its maximal alarmed state and transmission is reduced to zero. Plugging this in to Equation (5) yields

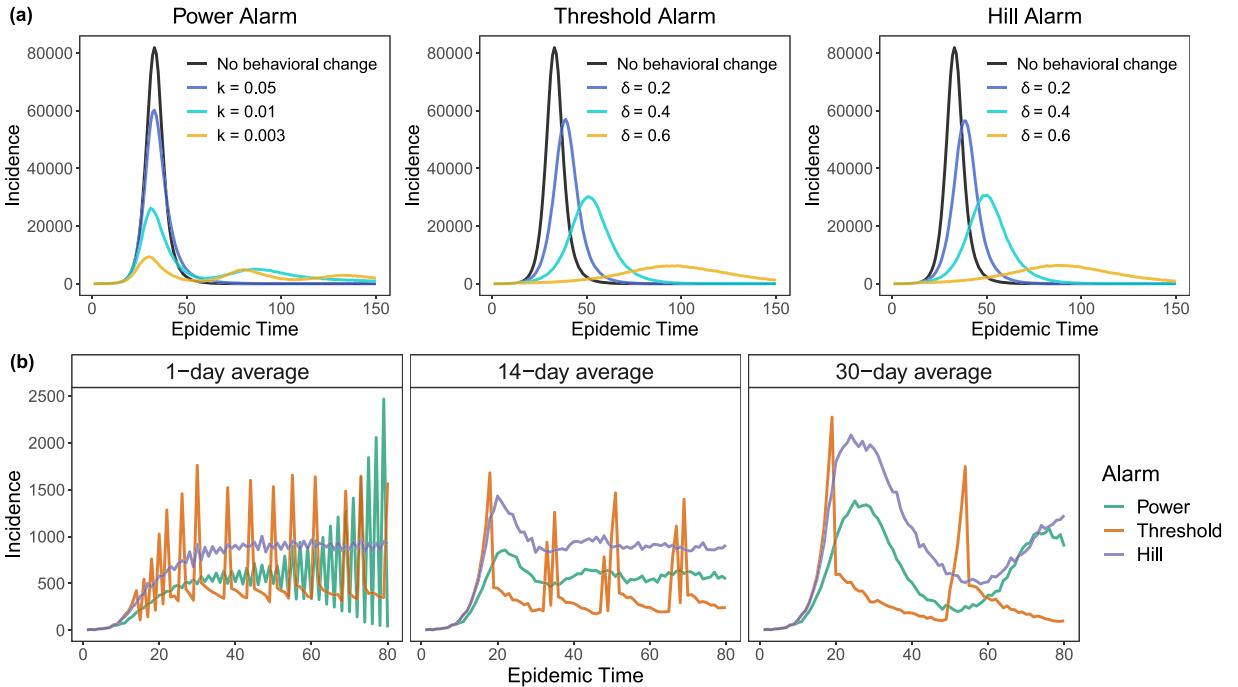
$$\pi_t^{(SI)} = 1 - \exp\left[-\beta(1 - a_t) \frac{I_t}{N}\right]. \quad (6)$$

The structural dependence in the alarm is captured by specifying  $a_t$  as a function of incidence smoothed over the past  $m$  days, such that  $a_t = f\left(\frac{1}{m} \sum_{i=t-m-1}^{t-1} I_i^*\right)$  and the smoothing parameter  $m \in \{1, 2, \dots, t-1\}$ . For  $t < m$ , we use the moving average of the data up until time  $t-1$ , e.g., at time  $t=3$  the alarm is based on the average of the incidence observed at times  $t=1$  and  $2$ . At the start of an epidemic, it is assumed that the alarm is zero. It is possible for the alarm function to depend on other reported metrics of epidemic severity, such as prevalence, hospitalizations, or test positivity rates, but we focus on an incidence-based alarm here. Incidence is advantageous for forecasting, as it is directly generated from the SIR model. Thus, incidence forecasts can be used directly to determine future alarm function values and generate further predictions.



**Fig. 1.** Example alarm functions for various parameter specifications.

Next, we must determine an appropriate functional form for the alarm function. As defined, the alarm must be between 0 and 1, and furthermore we would expect the alarm to be zero when there is no disease in the population, and to monotonically increase as the amount of disease in the population increases. There are many functions that satisfy these characteristics, and we investigate three possibilities of various complexity (Fig. 1). The first function considered is a one-parameter function  $f(x) = 1 - (1 - x/N)^{1/k}$  used in previous deterministic literature (Eksin et al., 2019; Franco, 2020), which we call the “power” alarm. The parameter  $k > 0$  describes the growth rate, with smaller values corresponding with a faster rise in alarm. The next function we consider is a two-parameter constant change point model, which we call the



**Fig. 2.** Example simulated epidemics for various alarm function specifications. (a) Compared to epidemics with no behavioral change, the power alarm with various  $k$  values, the threshold alarm with  $H = 100$  and various  $\delta$  values, and the Hill alarm with  $v = 5, x_0 = 100$ , and various  $\delta$  values. Across all simulations,  $\beta = 0.6$  and  $\gamma = 0.2$ . (b) Using 1-day, 14-day, and 30-day average incidence to inform the alarm function. Across the smoothing levels, the power alarm has  $k = 0.0005$ , the threshold alarm has  $\delta = 0.8$  and  $H = 350$ , and the Hill alarm has  $\delta = 0.85, v = 2$ , and  $x_0 = 450$ . Across all simulations,  $\beta = 0.6$  and  $\gamma = 0.2$ .

"threshold" alarm and specify as  $f(x) = \delta \mathbb{1}(x > H)$ . The threshold alarm is zero until the threshold,  $H$ , is surpassed, at which point it becomes  $\delta$ . Note that  $H$  can take on any value in the observed range of the data informing the alarm function and  $\delta \in [0, 1]$ . The final function analyzed is a modified Hill equation (Geszelyi et al., 2012),  $f(x) = \frac{\delta}{1+(x_0/x)^\nu}$ , which we refer to as the "Hill" alarm. With three parameters, the Hill alarm is the most complex of the three alarm functions considered;  $\delta \in [0, 1]$  describes the asymptote,  $x_0$  is the half occupation value, and  $\nu > 0$  controls the growth rate. The Hill alarm can describe curves similar to the power alarm, as well as sigmoid-shaped curves that resemble a smoothed version of the threshold alarm.

Using these alarm functions in the BC model can generate a multitude of shapes of epidemic curves. Fig. 2a shows that higher levels of alarm reduce peak incidence. In addition, increasing the maximum alarm value in the threshold and Hill alarms delays the peak, while increasing the growth rate of the power alarm does not affect the timing of the initial peak. Fig. 2b illustrates epidemics generated using alarms which reach high values at relatively low levels of incidence, as well as how epidemic trajectory is affected by the amount of data informing the alarm function. When the alarm is based solely on the previous day's incidence, incidence becomes volatile and oscillates between levels producing high and low amounts of alarm. Peaks become more pronounced and spread out when the 14 or 30-day average incidence is used to inform the alarm. Finally, the threshold alarm, which reaches its maximum instantaneously, leads to epidemic curves with very sharp peaks. In contrast, the power and Hill alarm increase gradually, resulting in smoother peaks.

In practice, it may not be obvious which alarm function to use, or a method which does not restrict the shape of the alarm function may be preferred. For this reason, we also investigate the use of more flexible non-parametric approaches to estimating the alarm function using basis splines and Gaussian processes. The spline alarm is modeled using natural cubic splines with estimated knot locations and written as  $f(x) = \mathbf{X}_B' \mathbf{b}$ . The basis matrix,  $\mathbf{X}_B$ , is constructed across the range of observed  $m$ -day average incidence and  $\mathbf{b}$  denotes the associated basis parameters. Constraints were used during estimation to ensure  $f(x) \in [0, 1]$ , but an appropriate link function (e.g., logit) could also be used. For the Gaussian process approach, we assume the logit of the alarm function is a realization from a multivariate normal distribution which is fully specified by its mean  $m(x)$  and covariance  $k(x, x')$ , i.e.,  $\text{logit}[f(x)] \sim \text{MVN}[m(x), k(x, x')]$ . We specify the mean function to start at  $(0, 0)$  and end at  $(\max(x), 1)$  on the logit scale, corresponding with the characteristics we expect of the alarm function. We generally expect the alarm function to be smooth and use the squared exponential covariance function  $k(x, x') = \sigma^2 \exp[-(x - x')^2 / 2l^2]$ , where  $\sigma^2 > 0$  the signal variance controls scaling and  $l > 0$  is the length-scale parameter controlling smoothness. Both alarms are defined across the range of observed  $m$ -day average incidence, so linear interpolation was used to find the value of the alarm on each day.

One of the most important quantities estimated by epidemic models is the reproductive number, denoted  $\mathcal{R}_0$ , which quantifies the spread of the pathogen in the population. Various methods of calculating  $\mathcal{R}_0$  exist, and we use the approach of (Ward et al., 2023a). The effective reproductive number is calculated as  $\mathcal{R}_0(t) = S_t \sum_{k=t}^{\infty} [1 - \exp(-\beta_k/N)] \exp(-\gamma)^{k-t}$  and provides the expected number of secondary infectious caused by a single individual that becomes infectious at time  $t$  (for the full derivation, see (Ward et al., 2023a)).  $\mathcal{R}_0$  can be interpreted in relation to the threshold of 1, as  $\mathcal{R}_0 \geq 1$  means the epidemic will continue to propagate through the population and  $\mathcal{R}_0 < 1$  indicates the epidemic will eventually die out.

### 2.3. Bayesian estimation and implementation

The complete log-likelihood for the chain binomial SIR model is

$$\ell(\mathbf{I}^*, \mathbf{R}^* | \Theta) = \sum_{t=0}^{\tau} \left[ \log \binom{S_t}{I_t^*} + I_t^* \log \pi_t^{(SI)} + (S_t - I_t^*) \log (1 - \pi_t^{(SI)}) + \log \binom{I_t}{R_t^*} + R_t^* \log \pi^{(IR)} + (I_t - R_t^*) \log (1 - \pi^{(IR)}) \right], \quad (7)$$

where the parameter vector  $\Theta$  contains  $\beta$  and  $\gamma$  for the traditional model, and includes additional parameters used to estimate  $\beta_t$  for the time-varying transmission models and the BC model. Complete data would provide the time series over the course of the epidemic for the transition vectors  $\mathbf{I}^*$  and  $\mathbf{R}^*$ , the initial conditions  $S_0$  and  $I_0$ , and the population size  $N$ . Often, we do not observe complete information on infectious and removal times. Unless otherwise stated, we assume that incidence ( $\mathbf{I}^*$ ) is observed, and removals ( $\mathbf{R}^*$ ) must be imputed, using data-augmented MCMC methods (Lekone & Finkenstädt, 2006; O'Neill & Roberts, 1999). The R package `nimble` (de Valpine et al., 2017, 2021) was used for computation, as it offers a mechanism for implementing user-defined data-augmented MCMC algorithms.

In the Bayesian framework, the parameter vector must be assigned a prior, with the use and justification of informative priors for any parameters varying by disease application. Often, the gamma distribution is used to specify the prior for  $\beta$  and  $\gamma$  as both parameters must be positive. In the presence of knowledge about the duration of the infectious period, informative priors can be used for  $\gamma$ . Determining informative priors for the parameters of the alarm functions is challenging. We use a vague gamma prior for  $k$  in the power alarm and Uniform(0, 1) priors for  $\delta$  in the threshold and Hill alarms. For  $H$  and  $x_0$  Uniform(min( $x$ ), max( $x$ )) priors are used as we expect the change point or half occupation point to occur during the observed range of incidence. The spline coefficients  $\mathbf{b}$  are given vague  $N(0, 100)$  priors and the knots are given Uniform(min( $x$ ), max( $x$ )) priors. The parameters of the covariance in the Gaussian process model,  $\sigma$  and  $l$  are weakly identified (Zhang, 2004), making the use of informative priors crucial. As the alarm function is estimated on the logit scale, the variability of the function is limited, so we use a gamma(150, 50) prior for  $\sigma$ . The prior for the length-scale parameter was specified as inverse gamma with

the shape and scale parameters determined using the practical range approach (Gelfand et al., 2005). Using this approach, the mean of the prior for  $I$  is specified by finding the value such that the covariance function is 0.05 for two points that are separated by half the maximum distance observed in the data. The prior standard deviation was fixed at two, as that was found to produce reasonable estimation.

### 3. Simulation study

#### 3.1. Simulation set-up

The statistical properties of the BC model are assessed via simulation. The primary goal of the simulation study was determining whether the behavioral change mechanism could be recovered through estimation of the alarm function. The secondary objective was comparing the BC model to the traditional approach without behavioral change and a flexible time-varying transmission model, assessing posterior predictive forecasting and model fit. These aims were addressed by simulating epidemics with behavioral change under the three alarm functions described in Section 2.2. For each of the three data generation scenarios, 50 epidemics were simulated using the initial conditions  $N = 1,000,000$ ,  $S_0 = 999,995$ , and  $I_0 = 5$ . Five models were fit to each simulated epidemic: a BC model using the true alarm function, BC models using the spline and Gaussian process alarms, the model with no behavioral change, and a time-varying transmission model with  $\beta_t$  estimated flexibly using natural cubic splines.

To evaluate posterior prediction, epidemics were simulated over 100 days, with the first 50 days used for model fitting and the subsequent 50 days used to evaluate forecasting accuracy. Simulation parameters were chosen to produce epidemics with a distinguished peak during the first 50 days of the epidemic and with additional peak(s) occurring in the subsequent 50 days. Complete specification of simulation parameters is provided in [Supplementary Table 1](#). Epidemics were generated and BC models were fitted using both 14-day and 30-day average incidence to inform the alarm function. Similar conclusions were found in both settings, so we detail the 30-day average results here and provide the 14-day average results in the [Supplementary Material](#).

Priors were specified as described in Section 2.3. For each model, three MCMC chains were run using various starting values of the parameters. The BC models and the flexible  $\beta_t$  model required more burn-in iterations due to their increased complexity. All models were run for 300,000 iterations post burn-in with samples drawn every 10th iteration. Full descriptions of priors used and MCMC specifications are provided in [Supplementary Table 2](#). Convergence was established by a Gelman and Rubin diagnostic value below 1.1 ([Gelman and Rubin, 1992](#)). A small number of models did not converge despite running for a large number of iterations and have been excluded from the results. More information on these models is provided in the [Supplementary Material](#).

#### 3.2. Simulation results

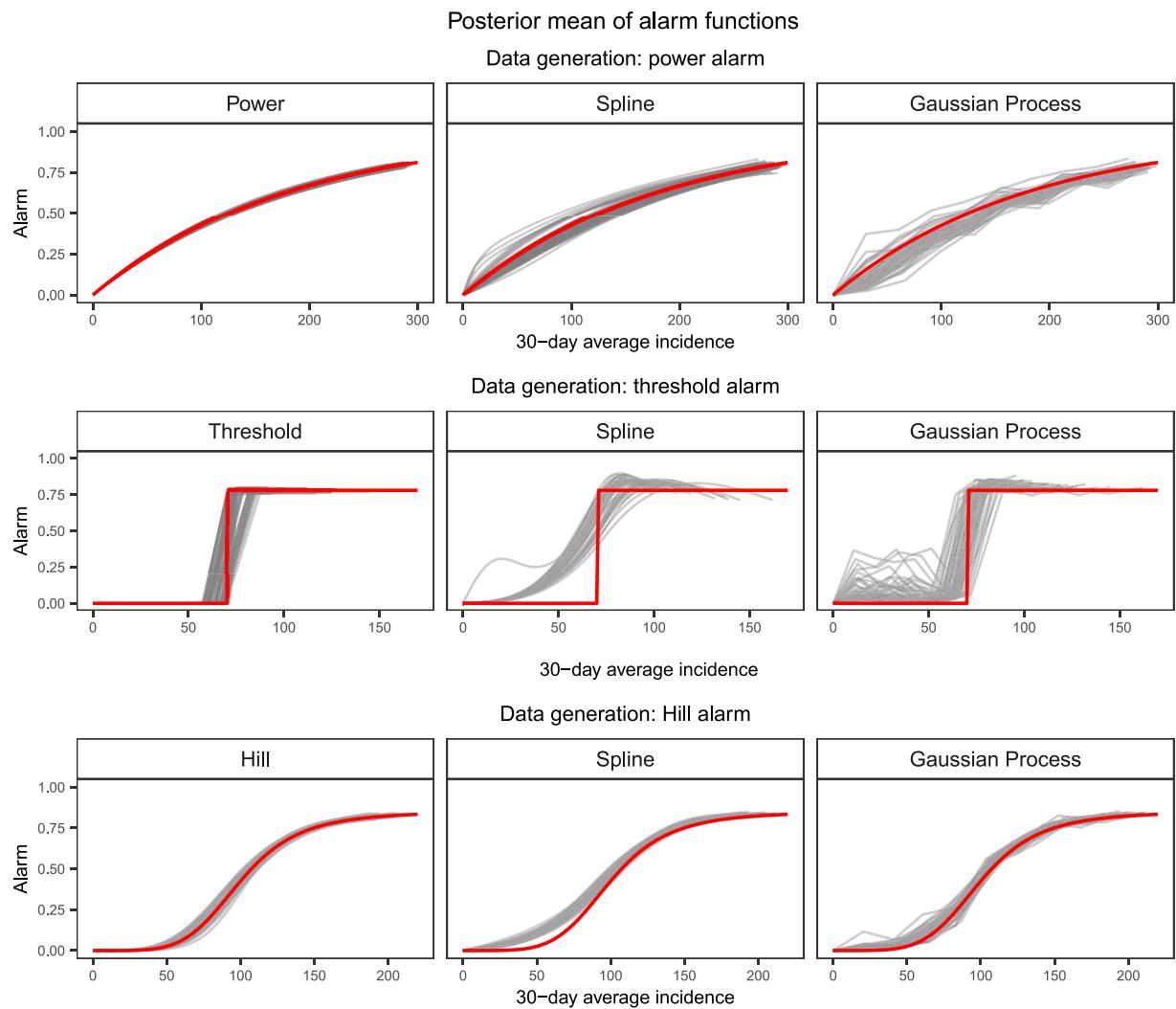
##### 3.2.1. Alarm function estimation

To assess estimation of the alarm function in each data generating scenario, we compare posterior mean alarm function estimates to the true alarm functions ([Fig. 3](#)). The alarm is shown as a function of the 30-day average incidence, ranging from zero to the maximum observed value during the epidemic, which varies between simulations. We find estimation of the alarm function to be excellent when the true functional form of the alarm was used in model fitting. More interestingly, we find the spline and Gaussian process approaches recover the alarm function reasonably well, particularly for the power and Hill alarms. The alarm function recovery is not as precise when the threshold alarm is the true function. However, this is expected as the piecewise constant form of this alarm is generally not well described by splines or Gaussian processes, which are inherently smooth. Despite this, both approaches are able to detect the alarm function rapidly increasing and leveling off quite impressively. These results indicate that when analyzing real data where the true alarm function is unknown, the spline and Gaussian process alarms offer robust and flexible possibilities.

##### 3.2.2. Posterior prediction

The ability of the fitted BC models to predict the epidemic curve was carried out using the posterior distribution for model parameters derived using the first 50 days of the epidemic. For 10,000 posterior draws of the parameters, the future epidemic trajectory was simulated using the model state on day 50 to determine the initial values and proceeding with binomial draws from  $S_t$  and  $I_t$  for  $t = 51, \dots, 100$ . A drawback of the flexible  $\beta_t$  model is there is no obvious mechanism for forecasting, so we compare the model with no behavioral change to the three BC models (true, spline, and Gaussian process alarms). Consistent results were found across simulations. For brevity, the posterior predictive distribution is provided for a single, randomly selected, and typical simulation in [Fig. 4](#).

The model which does not incorporate behavioral change does a poor job of posterior prediction. Assuming constant transmission yields an estimated  $\beta$  which averages over what was observed during the first 50 days. Correspondingly, the posterior predictive trajectories are increasing, but at a slower rate than the growth observed at the start of the epidemic. In contrast, posterior predictions from the BC models are able to detect the intensity and timing of subsequent waves of

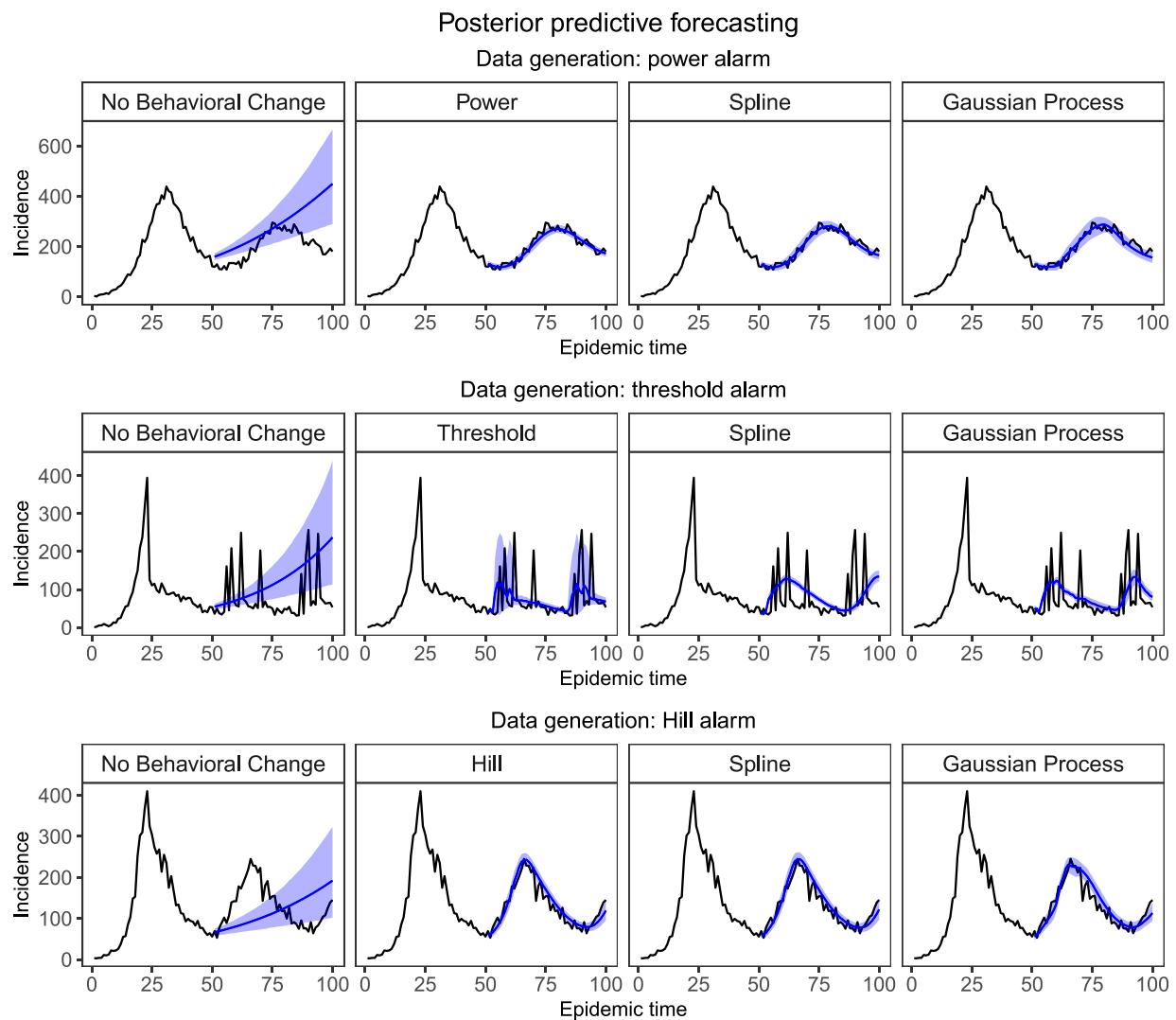


**Fig. 3.** True and posterior mean estimates of alarm functions from 50 simulated epidemics using the correct parametric alarm function, and the spline and Gaussian process alarms for model fitting. True alarm functions shown in red.

transmission. Minimal differences were found between predictions from the parametric and non-parametric approaches for data generated with the power and Hill alarms. For epidemics simulated using the threshold alarm, the non-parametric approaches predict smoother subsequent peaks, due to their failure to capture the abruptness of the change point in the alarm. Finally, the 95% posterior credible intervals for the model with no behavioral change widen as predictions become further from the last observed time point. This is not seen for the BC models, as the structural dependency encoded by the alarm function restricts the shape of the epidemic curve.

### 3.2.3. Comparing model fit

Model fit was assessed with the Widely Applicable Information Criteria (WAIC) (Watanabe & Opper, 2010), for which lower values indicate superior model fit. The distribution of WAIC values and the proportion of epidemics where each model had the lowest WAIC are summarized in Table 1. The true alarm function was selected 82–96% of the time, indicating that WAIC is able to correctly identify the best model. The spline and Gaussian process BC models sometimes provided the best fit. This was less likely when the data was generated from the threshold alarm, where recovery of the alarm function by these methods was poorest. Notably, despite its flexibility, the time-varying  $\beta_t$  model was only selected once and generally had higher average WAIC than any of the BC models. The exception was for data generated from the power alarm, the most gradual alarm function. This indicates that the additional structure imposed by the BC models results in lower WAIC in the simulation setting where data was generated from a model with behavioral change present. The traditional model with no behavioral change performs poorly across all three data generation scenarios.



**Fig. 4.** Mean and 95% credible intervals for posterior predictive forecasts of future incidence compared to the truth for a randomly selected simulation.

**Table 1**

Summaries of WAIC values across 50 simulated epidemics from three data generation scenarios. Models are ordered by mean WAIC.

Data generation	Model fitted	WAIC Mean (SD)	% selected
Power	Power	369.63 (12.80)	92%
	Spline	371.43 (12.73)	8%
	$\beta_t$	375.28 (13.19)	0%
	Gaussian process	381.44 (13.78)	0%
	No Behavioral Change	673.74 (16.25)	0%
Threshold	Threshold	333.19 (12.61)	96%
	Gaussian process	367.06 (48.96)	4%
	Spline	484.39 (65.67)	0%
	$\beta_t$	720.61 (59.87)	0%
	No Behavioral Change	1058.45 (120.73)	0%
Hill	Hill	358.55 (13.61)	82%
	Spline	360.28 (13.33)	10%
	Gaussian process	364.06 (13.95)	6%
	$\beta_t$	375.39 (19.04)	2%
	No Behavioral Change	793.40 (34.11)	0%

## 4. Data applications

### 4.1. Ebola disease

We first illustrate the BC model on a well-studied Ebola outbreak which occurred in 1995 in the Democratic Republic of the Congo (DRC). Ebola is a deadly disease which transmits between humans through direct physical contact with bodily fluids or contaminated clothes or bedding (CDC, 2023b). A person is only infectious once they develop signs and symptoms of Ebola disease, which can occur anywhere between two to 21 days (average eight to 10 days) after initial contact with an ebolavirus (CDC, 2023b,a). Once symptoms have appeared, individuals remain infectious for four to ten days and may either recover or die, with the average case fatality rate around 50% (WHO, 2023). The 1995 DRC epidemic occurred primarily in the city of Kikwit in the Bandundu region, which had a population of 5,363,500 during the outbreak (Lekone & Finkenstädt, 2006). The data used in this analysis are publicly available in the `outbreaks` R package (Jombart et al., 2020) and contain symptom onset date for 291 cases and death date for 236 individuals documented between March and July 1995 (Fig. 5). It is known that 316 infections occurred, but symptom onset date was not recorded for 25 individuals. Ebolavirus was identified as the causative agent of the outbreak on May 9th, after which control measures were immediately introduced. Further details about this epidemic can be found in Khan et al. (1999) (Khan et al., 1999).

Due to the long latent period for Ebola disease, we extend the SIR model introduced in Section 2 to incorporate an additional Exposed compartment describing individuals in the period between contracting the Ebolavirus and having symptoms. The SEIR model is a direct extension to the SIR model for which we now define  $E_t$  as the number of individuals in the exposed compartment and  $E_t^*$  as the number of newly exposed individuals during the continuous time interval  $[t, t + 1]$ . The difference equations become:

$$\begin{aligned}S_{t+1} &= S_t - E_t^* \\E_{t+1} &= E_t + E_t^* - I_t^* \\I_{t+1} &= I_t + I_t^* - R_t^* \\R_{t+1} &= R_t + R_t^*. \end{aligned}$$

The transitions between compartments are still assumed to be binomially distributed as  $E_t^* \sim \text{Bin}(S_t, \pi_t^{(SE)})$ ,  $I_t^* \sim \text{Bin}(E_t, \pi^{(EI)})$  and  $R_t^* \sim \text{Bin}(I_t, \pi^{(IR)})$ . Now,  $\pi_t^{(SE)}$  is the transmission probability of interest, as it describes the probability of an infectious individual transmitting the pathogen to a susceptible individual. As before, the form of the transmission probability can be described by the right hand side of Equations (4) and (5), or 6, depending on whether time-varying transmission is allowed or behavioral change is captured with an alarm function. The other transition probabilities are specified as  $\pi^{(EI)} = 1 - \exp(-\lambda)$  and  $\pi^{(IR)} = 1 - \exp(-\gamma)$ , where  $\lambda$  and  $\gamma$  are the mean lengths of the latent and infectious periods, respectively. Here,  $I^*$  and  $R^*$  are partially defined by the data, and  $E^*$  is completely unobserved. Missing exposure, symptom onset, and removal dates are imputed using data-augmented MCMC methods (Lekone & Finkenstädt, 2006).

Using the SEIR framework, various models were fitted. Five BC models using the power, threshold, Hill, spline, and Gaussian process alarms were evaluated. For comparison, we also fit models with no behavioral change, a model with a pre-specified intervention effect, and the flexible  $\beta_t$  model as described in the simulation study. Previous modeling of this

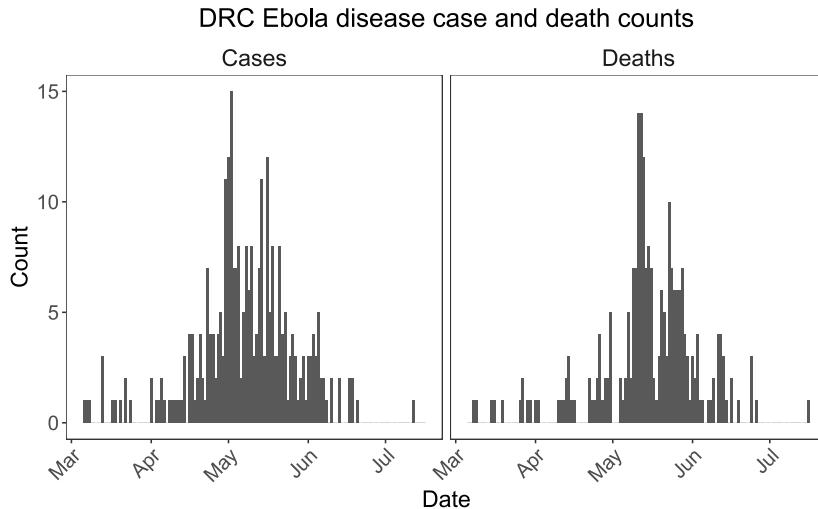


Fig. 5. Democratic Republic of the Congo observed Ebola disease data. Cases are recorded by symptom onset date and deaths by death date.

epidemic has evaluated a pre-specified intervention effect by assuming constant transmission until the time of the intervention, after which transmission decays exponentially (Lekone & Finkenstädt, 2006; Ward et al., 2023a). Mathematically, this is written by allowing  $\beta_t$  in Equation (5) to be equal to  $\exp[\beta_1 + \beta_2(t - t^*)\mathbb{1}(t \geq t^*)]$ , where  $\beta_1$  corresponds to the baseline intensity, and  $\beta_2$  represents the decay in transmission after the intervention was introduced at time  $t^*$  (May 9th).

In the BC models, we let the alarm be informed by cumulative observed incidence as this epidemic was completely observed. Although there were 25 additional infections, as the timings of these are unknown, so we do not allow them to influence the population alarms. The first three observed cases occurred consecutively on March 6, 7, 8, so we start our modeling on March 8th and assume  $E_0 = 2$  and  $I_0 = 1$ . It's known that two deaths occurred prior to this date (Khan et al., 1999), so we set  $R_0 = 2$ . We assume  $N = 5,363,500$  (Lekone & Finkenstädt, 2006). Vague priors for  $\beta$  and the parameters of the alarm function and flexible spline model were specified as in Section 2.3. The priors for  $\lambda$  and  $\gamma$  were informative and specified as in previous modeling of this epidemic (Lekone & Finkenstädt, 2006). Full specification of all priors for model parameters are detailed in Supplementary Table 12. Convergence was established by a Gelman and Rubin diagnostic value below 1.1 (Gelman and Rubin, 1992).

WAIC was used to determine the best fitting models (Table 2). All five BC models had lower WAIC than the standard models used for comparison, indicating the BC models have better fit than the other approaches. Interestingly, the parametric threshold and Hill BC models had lower WAIC than the non-parametric Gaussian process and spline models, with the threshold model having the overall lowest WAIC. Of the standard approaches, WAIC was relatively similar between the flexible  $\beta_t$  model and the model with pre-specified change in transmission after the intervention. The no behavioral change model has the highest WAIC indicating poor fit.

Fig. 6 depicts the estimated alarm functions for all the BC models. The power, spline, and Gaussian process alarms are relatively similar in shape, with steady increase in alarm as more cases were observed during the epidemic, with peak alarm around 0.8–1.0. On the other hand, the threshold and Hill alarms both estimate low alarm values until around 100 total cases had been observed, after which the estimated alarm value shifted to around 0.75. The 100th case was observed on May 2nd, which also corresponds to the peak of the epidemic. It's worth noting, however, that this was one week prior to the discovery of Ebola as the cause of the outbreak and subsequent introduction of control measures.

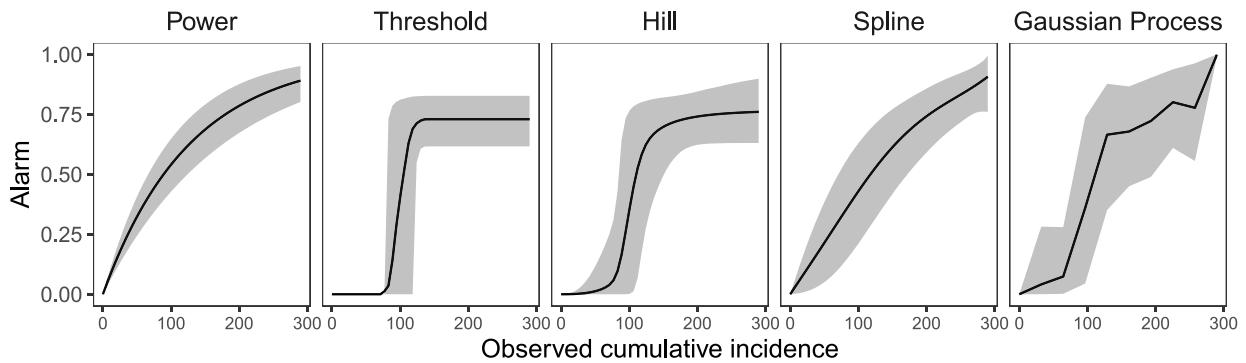
As the threshold BC model provided the best fit per WAIC, we compare estimated reproductive numbers and posterior predictive distributions of cumulative incidence between the BC model and the three standard modeling approaches. The posterior predictive distribution was computed using 10,000 posterior draws of the model parameters, and for each draw, the epidemic trajectory is simulated from the chain binomial model. When fitting the BC models, it was assumed that the alarm was a function of only the 291 observed cases, as the symptom onset date was missing for the other 25. To ensure accuracy in the posterior predictive distribution for the BC models, we assumed only 92% of cases were observed and only observed cases could impact the alarm function.

The BC model, intervention model, and the flexible  $\beta_t$  model all provide relatively similar estimates of  $\mathcal{R}_0(t)$  at the start of the epidemic with posterior means between 1.9 and 2.1 (Fig. 7). However, the intervention model is restricted in shape as the change point in transmission is fixed at May 9th, and the exponential decay form of the intervention effect forces transmission and  $\mathcal{R}_0(t)$  to zero by the end of the epidemic. Conversely, the threshold alarm can detect a change point in transmission which best fits the data, and the alarm level after the change point dictates the reduction in  $\mathcal{R}_0(t)$ , although  $\mathcal{R}_0(t)$  must remain constant after the change point occurs. This flexibility allows the BC model to detect behavioral change resulting in  $\mathcal{R}_0(t)$  dropping below 1 on April 30, around a week prior to the implementation of public health interventions. The flexible  $\beta_t$  model results in reproductive number estimates similar to the intervention model, although with much wider variance at the start of the epidemic and while the intervention model estimates  $\mathcal{R}_0(t) \approx 0$  by the end of May, the flexible  $\beta_t$  models estimates  $\mathcal{R}_0(t)$  to level out at a value of 0.1.

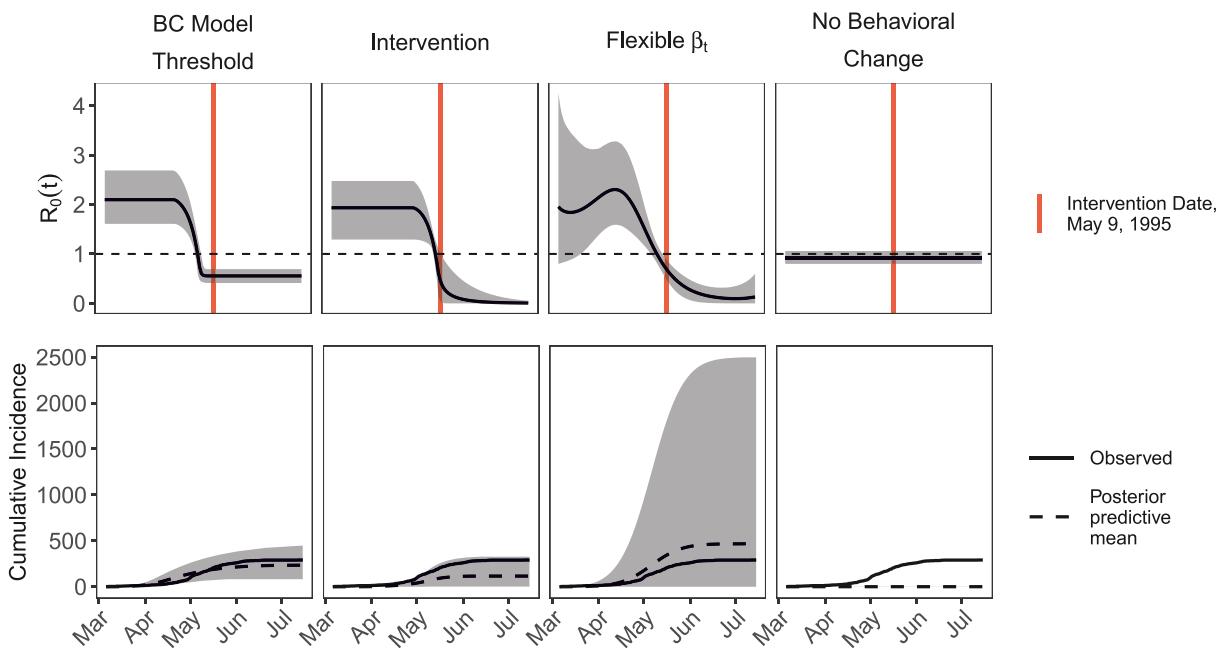
The subtle differences in the trajectory of  $\mathcal{R}_0(t)$  between the modeling approaches lead to larger differences in the posterior predictive distribution of cumulative incidence. As the reproductive number estimated for the intervention model goes quickly to zero, the final size of the epidemic is underestimated with a posterior predictive mean of 116 compared to the observed value of 316. The flexible  $\beta_t$  model overestimates the final size with a posterior predictive mean of 471 and also has huge variability in predicted epidemic trajectory. The threshold BC model has the best posterior predictive distribution

**Table 2**  
WAIC values for all models used in the DRC Ebola disease analysis. Models are ordered by WAIC.

Type	Model fitted	WAIC
BC Model	Threshold	600,676.6
	Hill	603,614.8
	Gaussian Process	605,838.9
	Power	611,096.6
	Spline	616,479.3
Standard Approach	$\beta_t$	622,297.5
	Intervention	622,767.5
	No Behavior Change	675,743.6



**Fig. 6.** Posterior means and 95% credible intervals for all estimated alarm functions for the DRC Ebola outbreak.

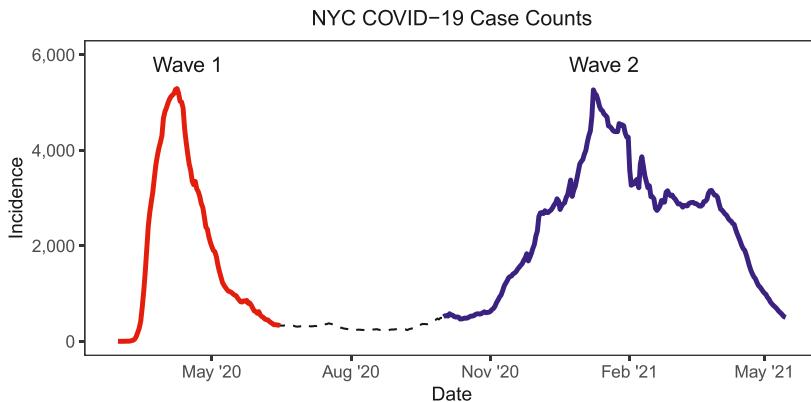


**Fig. 7.** Posterior means and 95% credible intervals for the effective reproductive number over time and the posterior predictive distribution of cumulative incidence for the DRC Ebola outbreak. Results are presented for the BC model using the threshold alarm, the model with an exponential decay intervention, the flexible  $\beta_t$  model, and the model with no behavioral change.

although it still underestimates the final epidemic size with the mean at 235. The model with no behavioral change is forced to average transmission over the entire epidemic and estimates  $R_0(t)$  just below 1 over the entire time period, which results in posterior predictive distribution over epidemics that only infect one or two individuals before dying out.

#### 4.2. COVID-19

After the first cases of COVID-19 were identified in China in December 2019, the SARS-CoV-2 virus spread rapidly across the globe, being declared a pandemic by the World Health Organization in March 2020 ([WHO, 2020](#)). We illustrate the use of the BC model to evaluate change in behavior over time using two waves of COVID-19 data from New York City (NYC) during the period March 2020–May 2021. The data used for this study are publicly available as part of the NYC Department of Health and Mental Hygiene Github repository ([NYC Health, 2022](#)). Based on the observed case counts, we define two waves of COVID-19 in NYC to be analyzed (Fig. 8).



**Fig. 8.** New York City COVID-19 data and waves defined for the analysis. Wave 1 occurs between Mar 1 and Jun 15, 2020 and Wave 2 between Oct 1, 2020 and May 15, 2021.

To illustrate the differences in the various BC model specifications, the power, threshold, Hill, spline, and Gaussian process alarms were all used to model both COVID-19 waves. As in the simulation study, the model with no behavioral change and the flexible  $\beta_t$  model were also fitted for comparison purposes. Prior to model fitting, the reported counts of new cases over time were smoothed to the 7-day average to account for weekly fluctuations in reporting. Although COVID-19 is known to have a short latent period around 5–6 days (Xin et al., 2022), this is difficult to incorporate in a fully Bayesian model due to the lack of reliable data arising from the presence of asymptomatic infections and lag between infectiousness and testing (Subramanian et al., 2021), and therefore the SIR model is commonly used (de Oliveira et al., 2020; Irons & Raftery, 2021; Lawson and Kim, 2021, 2022). For this analysis, we assume the smoothed incidence provides the time-series  $I^*$ , and use data-augmented MCMC to estimate  $R^*$ . While we know this is not true because of under-reporting, this assumption reduces computational expense. Additionally, as we anticipate behavioral change to be informed by observed epidemic trajectory, we do not expect the absence of undetected cases in our model to impact the estimation of our alarm function.

The population size was set as  $N = 8,804,190$ , the recorded population from the 2020 census. For each wave, we allow the initial conditions  $S_0$  and  $I_0$  to be estimated, using strong priors based on the past epidemic trajectory. Vague priors for  $\beta$  and the parameters of the alarm function and flexible spline model were specified as in Section 2.3. The prior for  $\gamma$  was specified to correspond with a mean infectious period of three days and 80% prior probability of the mean between 2 and 4 days. This corresponds with the typical length of time between an individual becoming contagious and testing positive and subsequently isolating. Results of a sensitivity analysis on this prior indicated little impact to the main conclusions across various priors and are provided in the Supplementary Material. Full specification of all priors for model parameters are detailed in Supplementary Table 21. In this example, the epidemic was not fully observed, so the BC models were fit using both 30-day and 60-day average incidence to inform the alarm functions. Convergence was established by a Gelman and Rubin diagnostic value below 1.1 (Gelman and Rubin, 1992).

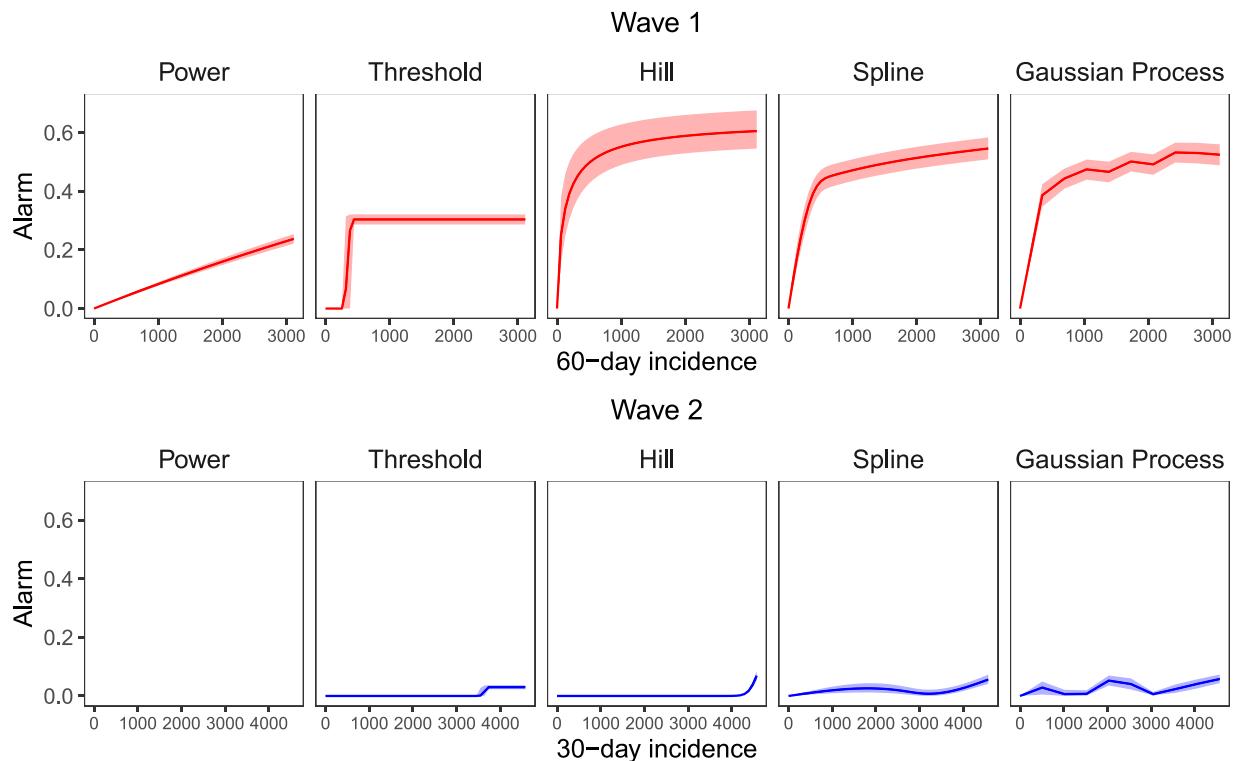
To determine the best fitting models for the two COVID-19 waves, WAIC values were compared (Table 3). Modeling  $\beta_t$  flexibly with splines provided the best fit to both waves, indicating that the additional structure imposed by the alarm function does not fully capture the observed epidemic trajectory. The Gaussian process and Hill alarms were among the BC models with the lowest WAIC across both waves. For the first wave, using 60-day average incidence to inform the alarm function offered better model fit, while the 30-day average performed better for the second wave. This is likely due to the different shape of the epidemic curve during Wave 2, which peaked in the beginning of January 2021, but leveled off between March–April 2021 before incidence was truly driven down. In contrast, Wave 1 showed steady decline in incidence post-peak, indicating the behavioral change in the population continued until the wave died out. The model with no behavioral change had the highest WAIC for Wave 1 and for Wave 2 had the second highest WAIC, indicating the importance of incorporating of behavioral change when considering real epidemic data.

Based on the WAIC results, we present the estimated alarm functions for each BC model based on 60-day average incidence for Wave 1 and 30-day average incidence for Wave 2 (Fig. 9). Despite the restricted shapes of the parametric functions, the estimated alarm functions were generally similar. The spline and Gaussian process alarms were very alike, which is not surprising due to the relationship between Gaussian processes and splines (Wahba, 1978). Comparing the estimated alarm functions between the two waves can be used to evaluate changes in pandemic response over time. During the first wave, the alarm reached very high levels at relatively low observed incidence. The second wave started while many restrictions from the first wave were still in place, and correspondingly the BC models estimate very slight increases in alarm at higher levels of observed incidence. In the second wave, the estimated Gaussian process alarm is not monotonically increasing which allows it to capture raised alarm during the large winter peak of the epidemic and during the subsequent mini-peak occurring in the spring. As the Gaussian process alarm achieved the lowest WAIC, it seems this flexibility allows for better model fit.

**Table 3**

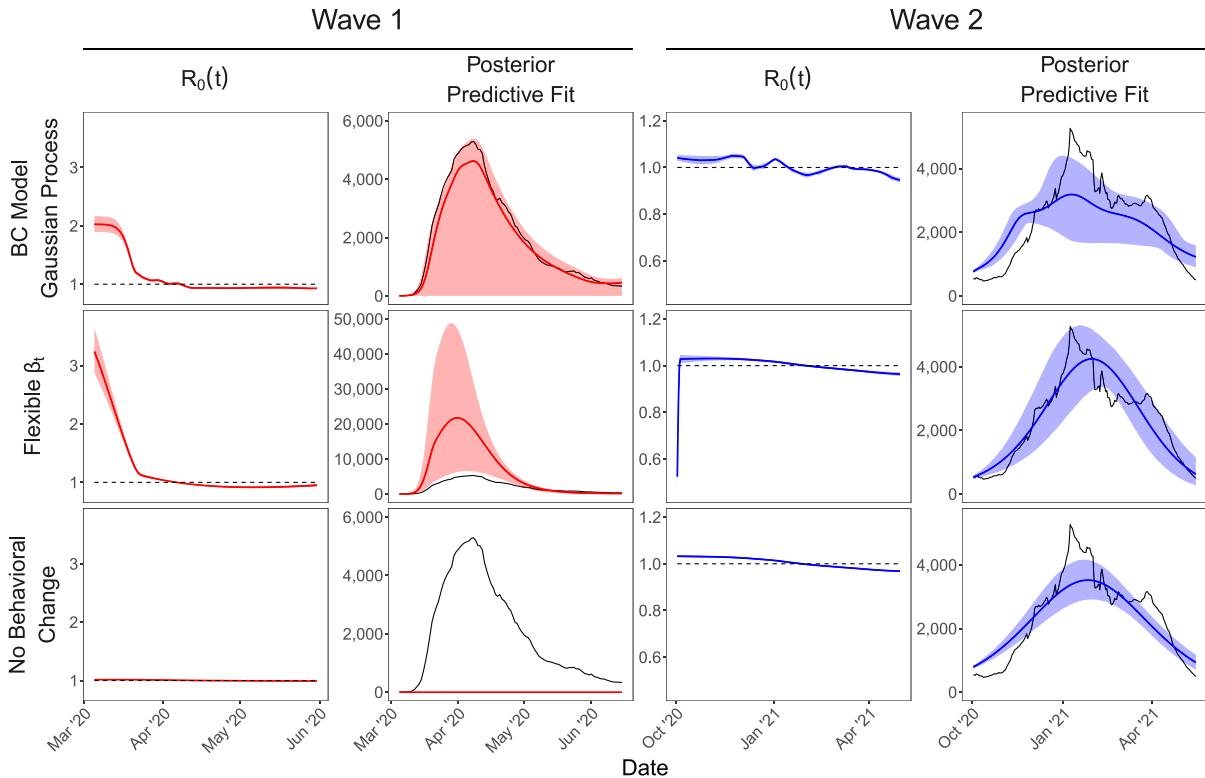
WAIC values for all converged models used in the NYC COVID-19 analysis. Within each wave, models are ordered by WAIC. The BC models chosen for the final results are indicated in italics. The power alarm models did not converge for Wave 2 and are therefore excluded from these results.

Wave	Model fitted	Smoothing	WAIC
Wave 1	$\beta_t$	None	1089.83
	Gaussian Process	60-day	1092.92
	Spline	60-day	1099.02
	Hill	60-day	1125.36
	Spline	30-day	1157.33
	Gaussian Process	30-day	1157.89
	Hill	30-day	1190.37
	Threshold	30-day	1466.06
	Threshold	60-day	1466.34
	Power	60-day	1540.83
Wave 2	Power	30-day	1753.43
	No Behavior Change	None	2263.96
Wave 2	$\beta_t$	None	2865.70
	Gaussian Process	30-day	3025.92
	Threshold	30-day	3041.86
	Hill	30-day	3045.90
	Spline	30-day	3051.68
	Threshold	60-day	3078.72
	Hill	60-day	3080.96
	Gaussian Process	60-day	3091.50
	No Behavior Change	None	3102.84
	Spline	60-day	3105.28



**Fig. 9.** Posterior means and 95% credible intervals for all estimated alarm functions from each wave of the NYC COVID-19 epidemic. The power alarm model did not converge for Wave 2 and is therefore excluded from these results.

Finally, we compare estimated reproductive numbers and posterior predictive distributions between the BC model using the Gaussian process alarm, the flexible  $\beta_t$  model, and the model with no behavioral change (Fig. 10). The Gaussian process



**Fig. 10.** Posterior means and 95% credible intervals for the effective reproductive number of time and the posterior predictive distribution from each wave of the NYC COVID-19 epidemic. Results are presented for the BC model using the Gaussian process alarm, the flexible  $\beta_t$  model and the model with no behavioral change.

alarm was chosen as it had the lowest WAIC of the BC models for both waves. The posterior predictive distribution was computed using 10,000 posterior draws of the model parameters, which includes  $S_0$  and  $I_0$ . For each draw, the epidemic trajectory is simulated from the chain binomial model and is compared to the true observed epidemic curve. For the first COVID-19 wave, the trajectory of  $\mathcal{R}_0(t)$  is quite different between the three models, particularly between March and April 2020. The BC model estimates  $\mathcal{R}_0(t)$  starting around 2, while the flexible  $\beta_t$  model estimates a higher value around 3, and the model with no behavioral change estimating  $\mathcal{R}_0(t) \approx 1$  over the entire wave. This leads to vastly different posterior predictive distributions, due to the high variability in stochastic epidemic models at the start of an epidemic. Interestingly, the flexible  $\beta_t$  model has poor posterior predictive fit, despite having the lowest WAIC. This is likely because WAIC weights each time point equally in calculating the log predictive density, whereas the posterior predictive epidemic trajectory is highly influenced by the estimated  $\mathcal{R}_0(t)$  at time one. Additionally, the flexible  $\beta_t$  model has a very wide 95% posterior predictive interval, estimating the peak incidence to be anywhere between 6000 and 50,000 cases per day. The BC model has lower uncertainty in the posterior predictive distribution, but the 95% interval does include the scenario where the epidemic dies out immediately. The high variability in these distributions is not surprising as forecasting at the very beginning of an epidemic is extremely difficult with many unknowns.

In Wave 2, the estimated  $\mathcal{R}_0(t)$  for the BC model and the no behavioral change model is just above 1 until mid-January 2021. The flexible  $\beta_t$  model follows a similar trajectory, except  $\mathcal{R}_0(t)$  is below 1 for the first two days, which allows it to better capture the slow growth of the epidemic at the beginning of October 2020 in the posterior predictive trajectory. The BC model struggles with the shape of the epidemic curve in Wave 2, as incidence stops declining between March and April 2021. However, the posterior predictive distribution appears slightly better than that of the model with no behavioral change.

## 5. Discussion

There is a critical need to understand the dynamics of population behavior changing in response to an infectious disease outbreak. Guided by the previous deterministic literature, we developed a novel Bayesian epidemic model framework which characterizes behavioral change dynamics at the population level while remaining simple enough to be computationally feasible. We showed that the proposed BC model can accurately estimate the mechanism of behavioral change across a wide range of scenarios, including when flexible non-parametric methods are used. The practical implications and usefulness of

the proposed approach were illustrated with two relevant case studies using Ebola and COVID-19 data, although the model could be applied to any communicable disease.

Our simulation study conducted a thorough investigation of the BC model and made several notable findings. First, when behavioral change impacts epidemic trajectory, the BC model is able to accurately estimate the mechanism of behavioral change. We considered three different alarm functions to describe behavioral change, and although the functions impact the epidemic trajectory differently, we were able to reasonably capture all three functions using non-parametric splines and Gaussian processes. This is hugely beneficial, as it may be difficult to choose an appropriate functional form of the alarm when analyzing real epidemic data. We also showed that posterior predictive forecasts from the BC model can accurately detect a second peak, something that is not feasible with the traditional SIR model. Finally, WAIC provided an accurate metric for selecting the best fit model and we found the additional structure in the BC models generally lead to lower WAIC values than the more flexible approach of estimating  $\beta_t$  directly, even though both approaches are able to capture changes in transmission over time.

Our analysis of the Ebola and COVID-19 epidemics offers numerous insights into behavioral change. In both analyses, we found the incorporation of behavioral change offers superior model fit compared to an approach without behavioral change. In the analysis of the Ebola outbreak, we illustrated the ability of the BC model to detect behavioral change which occurs separately from a government intervention. Allowing for structured behavioral change without restricting the timing or impact offered superior model fit and provided additional insight on the population engaging in protective behaviors a week prior to Ebola being identified as the cause of the outbreak. The use of the SEIR model in this analysis also shows how the BC model can be incorporated into more complex compartmental models when additional data is available. In the analysis of the first two waves of the COVID-19 pandemic in New York City, we presented the use of the BC model for comparing behavioral change over time.

Although illustrating the proposed model on real epidemic is extremely valuable, these analyses are not without limitations. In particular, the COVID-19 pandemic has led to many modeling challenges arising from the large presence of asymptomatic cases, lack of testing availability, changes in disease severity over time, and waning immunity. By modeling the COVID-19 waves separately, we mitigate issues with changing disease severity and waning immunity, not to mention changes in the way the population/government reacted to increasing cases between waves. However, it might be desirable to model multiple waves in one analysis, in which case we would want to extend our model to allow for time-varying alarm functions and/or changes in pathogen transmissibility. We have not incorporated undetected infections in the model, however, some preliminary simulations (not shown) have found that the presence of undetected infections which are not accounted for in the model does not impact the estimation of the alarm function when the alarm is based on observed cases, a realistic assumption. While this model misspecification does likely lead to underestimation of the reproductive number, our primary goal in this work is to illustrating the BC models ability to estimate behavioral change and we believe this is accomplished with the presented model. Various methods to account for undetected infections exist, including the addition of vaccinated or asymptomatic compartments (Angeli et al., 2022) or incorporating sero-prevalence surveys (Irons & Raftery, 2021). The BC model could be directly incorporated into these more complex structures when needed to achieve the analysis goals.

In this work we considered a population-average model which assumes homogeneous mixing and equal susceptibility for all members of the population. These assumptions may not be realistic, as factors like age are known to impact contact rates and susceptibility. Many extensions to the Bayesian SIR model have been introduced to relax these assumptions, including the use of a stratified population structure (Brown et al., 2016; Porter & Oleson, 2016), the addition of spatial random effects in the transmission rate (Lawson & Kim, 2021; Mahsin et al., 2022), or allowing the transmission rate to be impacted by individual-level covariates (Deardon et al., 2010). Establishing the BC model in the population-averaged framework is important as data are often limited, but ongoing work seeks to incorporate behavioral change into models with individual covariates and spatial structure. In a more precise model where susceptibility and contact patterns are estimated separately, the alarm function could modify either factor. The alarm function itself may also be modeled as a function of individual-level covariates. Additionally, one may consider both the transmission rate and the alarm function to vary spatially, and models incorporating spatial structure in the alarm function could estimate region-specific alarms.

Across all alarm functions considered, we assume the alarm level is zero when there have been no cases observed. This assumption is realistic for new or localized outbreaks, but may be restrictive when considering the scale of a global pandemic. For example, we may expect alarm in a rural region to be elevated prior to the first incidence, based on knowledge of disease occurrence in a nearby urban area. While one could consider adding a parameter to estimate a baseline alarm level, without strong priors this would be unidentifiable with the baseline transmissibility  $\beta$  in a single population model. However, in a multi-region spatial model with some global baseline transmissibility this type of “nugget effect” could be estimated for each region.

The proposed framework models behavioral change implicitly by allowing transmission to depend only on the perceived amount of disease in the population. However, time-varying measures of population behavior (e.g., Google mobility reports) exist and could be used to incorporate behavioral change directly in the specification of the transmission rate (Vanni et al., 2021; Hu et al., 2021). One disadvantage of this approach is it requires simulation of the behavioral change metric, which may be challenging. In contrast, transmission in the BC model is based on previous incidence, which is generated directly from the SIR model. Despite this issue, comparing an approach using mobility metrics to the BC model is a promising avenue for future work. Developments including both measures of mobility and an alarm function based on epidemic trajectory in the

transmission rate would be particularly interesting. In this setting, one could examine whether the alarm function is able to detect “finer grain” behavioral change which may not be captured by mobility data, e.g., voluntary masking.

## Software

Software in the form of R code, together with data used and complete documentation is available at <https://github.com/ceward18/epidemicBCM>.

## Declaration of competing interest

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

Funding for the project was provided by the Canadian Statistical Sciences Institute (CANSSI) Distinguished Postdoctoral Fellowship. This research was enabled in part by computational resources provided by the University of Calgary and the University of Minnesota.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.idm.2023.08.002>.

## References

- Acuña-Zegarra, M. A., Santana-Cibrian, M., & Velasco-Hernandez, J. X. (2020). Modeling behavioral change and COVID-19 containment in Mexico: A trade-off between lockdown and compliance. *Mathematical Biosciences*, 325, Article 108370.
- Agaba, G. O., Kyrychko, Y. N., & Blyuss, K. (2017). Mathematical model for the impact of awareness on the dynamics of infectious diseases. *Mathematical Biosciences*, 286, 22–30.
- Andersson, H., & Britton, T. (2012). *Stochastic epidemic models and their statistical analysis*, ume 151. Springer Science & Business Media.
- Angeli, M., Neofotistos, G., Mattheakis, M., & Kaxiras, E. (2022). Modeling the effect of the vaccination campaign on the COVID-19 pandemic. *Chaos, Solitons & Fractals*, 154, Article 111621.
- Brown, G. D., Oleson, J. J., & Porter, A. T. (2016). An empirically adjusted approach to reproductive number estimation for stochastic compartmental models: A case study of two Ebola outbreaks. *Biometrics*, 72, 335–343.
- Capasso, V., & Serio, G. (1978). A generalization of the kermack-mckendrick deterministic epidemic model. *Mathematical Biosciences*, 42, 43–61.
- CDC. (2023a). *Ebola Disease signs and symptoms*. URL: <https://www.cdc.gov/vhf/ebola/symptoms/index.html>.
- CDC. (2023b). *Ebola disease transmission*. URL: <https://www.cdc.gov/vhf/ebola/transmission/index.html>.
- 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, 239.
- Del Valle, S., Hethcote, H., Hyman, J. M., & Castillo-Chavez, C. (2005). Effects of behavioral changes in a smallpox attack model. *Mathematical Biosciences*, 195, 228–251.
- Eksin, C., Paarporn, K., & Weitz, J. S. (2019). Systematic biases in disease forecasting—the role of behavior change. *Epidemics*, 27, 96–105.
- Penichel, E. P., Castillo-Chavez, C., Cuddia, M. G., Chowell, G., Parra, P. A. G., Hickling, G. J., Holloway, G., Horan, R., Morin, B., Perrings, C., et al. (2011). Adaptive human behavior in epidemiological models. *Proceedings of the National Academy of Sciences*, 108, 6306–6311.
- Franco, E. (2020). A feedback SIR (fSIR) model highlights advantages and limitations of infection-based social distancing. arXiv preprint arXiv:2004.13216 28.
- Gelfand, A. E., Kottas, A., & MacEachern, S. N. (2005). Bayesian nonparametric spatial modeling with Dirichlet process mixing. *Journal of the American Statistical Association*, 100, 1021–1035.
- Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7, 457–472.
- Gesztesy, R., Zsuga, J., Kemeny-Beké, A., Varga, B., Juhasz, B., & Tosaki, A. (2012). The hill equation and the origin of quantitative pharmacology. *Archive for History of Exact Sciences*, 66, 427–438.
- Greenhalgh, D., Rana, S., Samanta, S., Sardar, T., Bhattacharya, S., & Chattopadhyay, J. (2015). Awareness programs control infectious disease—multiple delay induced mathematical model. *Applied Mathematics and Computation*, 251, 539–563.
- Hong, H. G., & Li, Y. (2020). Estimation of time-varying reproduction numbers underlying epidemiological processes: A new statistical tool for the COVID-19 pandemic. *PLoS One*, 15, Article e0236464.
- Hu, T., Wang, S., She, B., Zhang, M., Huang, X., Cui, Y., Khuri, J., Hu, Y., Fu, X., Wang, X., et al. (2021). Human mobility data in the COVID-19 pandemic: Characteristics, applications, and challenges. *International Journal of Digital Earth*, 14, 1126–1147.
- Irons, N. J., & Raftery, A. E. (2021). Estimating SARS-CoV-2 infections from deaths, confirmed cases, tests, and random surveys. *Proceedings of the National Academy of Sciences*, 118.
- Jombart, T., Frost, S., Nouvellet, P., Campbell, F., & Sudre, B. (2020). *Outbreaks: A collection of disease outbreak data*. URL: <https://CRAN.R-project.org/package=outbreaks>. r package version 1.9.0.
- Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London - Series A: Containing Papers of a Mathematical and Physical Character*, 115, 700–721.
- Khan, A. S., Tshioko, F. K., Heymann, D. L., Le Guenno, B., Nabeth, P., Kerstiëns, B., Fleerackers, Y., Kilmarx, P. H., Rodier, G. R., Nkuku, O., et al. (1999). The reemergence of Ebola hemorrhagic fever, Democratic Republic of the Congo, 1995. *The Journal of infectious diseases*, 179, S76–S86.
- Lawson, A. B., & Kim, J. (2021). Space-time COVID-19 bayesian SIR modeling in South Carolina. *PLoS One*, 16, Article e0242777.
- Lawson, A. B., & Kim, J. (2022). Bayesian space-time SIR modeling of COVID-19 in two US states during the 2020–2021 pandemic. *PLoS One*, 17, Article e0278515.
- Lekone, P. E., & Finkenstädt, B. F. (2006). Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. *Biometrics*, 62, 1170–1177.
- Liu, L., Vikram, S., Lao, J., Ben, X., D'Amour, A., O'Banion, S., Sandler, M., Saorous, R. A., & Hoffman, M. D. (2020). *Estimating the changing infection rate of COVID-19 using Bayesian models of mobility*. medRxiv.
- Mahsin, M., Deardon, R., & Brown, P. (2022). Geographically dependent individual-level models for infectious diseases transmission. *Biostatistics*, 23, 1–17.

- NYC Health. (2022). *Nyc coronavirus disease 2019 COVID-19 data*. <https://github.com/nychealth/coronavirus-data>.
- de Oliveira, A. C. S., Morita, L. H. M., da Silva, E. B., Zardo, L. A. R., Fontes, C. J. F., & Granzotto, D. C. T. (2020). Bayesian modeling of COVID-19 cases with a correction to account for under-reported cases. *Infectious Disease Modelling*, 5, 699–713.
- O'Neill, P. D., & Roberts, G. O. (1999). Bayesian inference for partially observed stochastic epidemics. *Journal of the Royal Statistical Society: Series A*, 162, 121–129.
- Perra, N., Balcan, D., Gonçalves, B., & Vespignani, A. (2011). Towards a characterization of behavior-disease models. *PLoS One*, 6, Article e23084.
- Porter, A. T., & Oleson, J. J. (2016). A spatial epidemic model for disease spread over a heterogeneous spatial support. *Statistics in Medicine*, 35, 721–733.
- Reluga, T. C. (2010). Game theory of social distancing in response to an epidemic. *PLoS Computational Biology*, 6, Article e1000793.
- Roberts, M., Andreasen, V., Lloyd, A., & Pellis, L. (2015). Nine challenges for deterministic epidemic models. *Epidemics*, 10, 49–53.
- Sartorius, B., Lawson, A., & Pullan, R. (2021). Modelling and predicting the spatio-temporal spread of COVID-19, associated deaths and impact of key risk factors in England. *Scientific Reports*, 11, 1–11.
- Subramanian, R., He, Q., & Pascual, M. (2021). Quantifying asymptomatic infection and transmission of COVID-19 in New York City using observed cases, serology, and testing capacity. *Proceedings of the National Academy of Sciences*, 118, Article e2019716118.
- de Valpine, P., Paciorek, C., Turek, D., Michaud, N., Anderson-Bergman, C., Obermeyer, F., Wehrhahn Cortes, C., Rodríguez, A., Temple Lang, D., Paganin, S., & Hug, J. (2021). *NIMBLE: MCMC, particle filtering, and programmable hierarchical modeling*. <https://doi.org/10.5281/zenodo.1211190>. URL: <https://cran.r-project.org/package=nimble>.
- de Valpine, P., Turek, D., Paciorek, C., Anderson-Bergman, C., Temple Lang, D., & Bodik, R. (2017). Programming with models: Writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational & Graphical Statistics*, 26, 403–417.
- Vanni, F., Lambert, D., Palatella, L., & Grigolini, P. (2021). On the use of aggregated human mobility data to estimate the reproduction number. *Scientific Reports*, 11, 1–10.
- Wahba, G. (1978). Improper priors, spline smoothing and the problem of guarding against model errors in regression. *Journal of the Royal Statistical Society: Series B*, 40, 364–372.
- Ward, C., Brown, G. D., & Oleson, J. J. (2023a). Incorporating infectious duration-dependent transmission into Bayesian epidemic models. *Biometrical Journal*, 65, Article 2100401.
- Ward, C., Brown, G. D., & Oleson, J. J. (2023b). An individual level infectious disease model in the presence of uncertainty from multiple, imperfect diagnostic tests. *Biometrics*, 79, 426–436.
- Watanabe, S., & Opper, M. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11.
- Weitz, J. S., Park, S. W., Eksin, C., & Dushoff, J. (2020). Awareness-driven behavior changes can shift the shape of epidemics away from peaks and toward plateaus, shoulders, and oscillations. *Proceedings of the National Academy of Sciences*, 117, 32764–32771.
- WHO. (2020). *WHO Director-General's opening remarks at the media briefing on COVID-19*, 11 March 2020. URL: <https://www.who.int/director-general-speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19--11-march-2020>.
- WHO. (2023). *Ebola virus disease overview*. URL: <https://www.who.int/health-topics/ebola>.
- Wilson, E. B., & Burke, M. H. (1942). The epidemic curve. *Proceedings of the National Academy of Sciences of the United States of America*, 28, 361.
- Xin, H., Li, Y., Wu, P., Li, Z., Lau, E. H., Qin, Y., Wang, L., Cowling, B. J., Tsang, T. K., & Li, Z. (2022). Estimating the latent period of coronavirus disease 2019 (covid-19). *Clinical Infectious Diseases*, 74, 1678–1681.
- Xu, X., Kypraios, T., & O'Neill, P. D. (2016). Bayesian non-parametric inference for stochastic epidemic models using Gaussian processes. *Biostatistics*, 17, 619–633.
- Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *Journal of the American Statistical Association*, 99, 250–261.