scientific reports

Check for updates

OPEN

Bayesian SIR model with change points with application to the Omicron wave in Singapore

Jiaqi Gu & Guosheng Yin✉

The Omicron variant has led to a new wave of the COVID-19 pandemic worldwide, with unprecedented numbers of daily confirmed new cases in many countries and areas. To analyze the impact of society or policy changes on the development of the Omicron wave, the stochastic susceptible-infected-removed (SIR) model with change points is proposed to accommodate the situations where the transmission rate and the removal rate may vary significantly at change points. Bayesian inference based on a Markov chain Monte Carlo algorithm is developed to estimate both the locations of change points as well as the transmission rate and removal rate within each stage. Experiments on simulated data reveal the effectiveness of the proposed method, and several stages are detected in analyzing the Omicron wave data in Singapore.

The ongoing worldwide COVID-19 pandemic caused by the SARS-CoV-2 virus has spread to over 200 countries and areas with more than 523 million confirmed cases and around 7 million deaths by May 2022. Although more than 11 billion doses of vaccine have been administered, several variants of SARS-CoV-2 have appeared with faster transmission and greater virulence and led to diminished effectiveness of developed vaccines, which resulted in multiple COVID-19 pandemic waves. Among all the five variants of concern listed by the World Health Organization, the Omicron variant has a higher level of transmissibility and immune escape capability[1–6], leading to unprecedented numbers of daily confirmed new cases in many countries and areas. Due to the unique characteristics of Omicron (high transmissibility but low mortality or severe cases), existing analysis methods on the spreading of other variants are not applicable to coping with the Omicron wave of COVID-19 pandemic.

In epidemiology, the susceptible-infected-removed (SIR) model[7] is the most popular approach to analyzing the transmission of infectious diseases. Since the beginning of the Omicron wave, a large amount of research on the SIR model or its extensions has been conducted for making inference and prediction. For example, Van Wees et al.[8] applied the standard SIR model to predict the infection rate and hospitalization rate in South African province Gauteng, the United Kingdom and the Netherlands. To incorporate local transmission into the SIR model, Götz[9] developed the global–local SIR model with a locality-adjusted basic reproduction number. The multi-wave SIR model proposed by Ghosh and Ghosh[10] allows researchers to investigate the nonperiodicity of COVID-19 pandemic waves, while Khan and Atangana[11] separated the Omicron variant from other variants in order to estimate its basic reproduction number accurately. However, all the aforementioned approaches assume that throughout a COVID-19 pandemic wave both the transmission rate and removal rate remain unchanged. This assumption is unrealistic because both individual actions for self-protection and government policies in response to the outbreak would decrease the transmission rate. In addition, more medical resources would be allocated to combat the uprise of new cases, leading to an increase of the removal rate during a pandemic wave.

To take the effect of societal changes into consideration, the stochastic SIR model with change points is proposed to analyze the evolvement of the Omicron wave of COVID-19 pandemic. Assuming that both the transmission rate and removal rate are time-varying, binomial models are proposed for the daily reported numbers of confirmed cases and removal cases. A latent indicator vector is introduced to partition a pandemic wave into several stages. Compared to existing works[12–16] that incorporate change points to compartmental models, our model assumes that both the transmission rate and the removal rate are time-varying and can change multiple times during the study period. Instead of being constant between any pair of adjacent change points, time-varying parameters in our model are homogenous in distribution within each stage and thus are more flexible for model fitting. We develop a Markov chain Monte Carlo (MCMC) algorithm to draw posterior samples of parameters and make Bayesian inference of pandemic wave development. Experiments on simulated datasets suggest the

Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong, China. ✉email: gyin@hku.hk

effectiveness of the proposed method in detecting change points of a pandemic wave and the Omicron data in Singapore are analyzed to corroborate the major changes during the Omicron wave.

The rest of this paper is organized as follows. Section "The SIR model" reviews the standard SIR model. In section "Methodology", we propose the stochastic SIR model with change points and develop an MCMC algorithm for making Bayesian inference. Experiments on simulated datasets are conducted in section "Simulations" to illustrate the performance of the proposed method. In section "Analysis of the Omicron wave in Singapore", we analyze the Omicron wave of COVID-19 pandemic in Singapore and find several stages with different transmission rates and removal rates which match the major societal changes in Singapore. We conclude this paper in section "Conclusion".

The SIR model

The susceptible-infected-removed (SIR) model[7] is the most widely used mathematical tool to model the spreading of infectious diseases. Given a closed population of N individuals and three possible states, susceptible (S), infectious (I) and removed (R, either recovered or dead), each individual is assumed to be in one state at any time. As time goes by, the state of each individual would evolve from S to I and then from I to R, implying the process that an individual gets infected and then recovers or dies. Let $S(t)$, $I(t)$ and $R(t)$ denote the number of susceptible, infectious and removed individuals in the population at time point t ($t \geq 0$) respectively. Define $P(t) = I(t)/N$ as the proportion of infectious individuals at time t. Assuming that in-person contacts among individuals follow the Erdős–Rényi model[17,18], the standard SIR model describes the flow of individuals from S to I and then from I to R by a set of ordinary differential equations without the access to individual records of infection and removal,

$$
\begin{cases}
\dfrac{dS(t)}{dt} = -\beta S(t)P(t), \\[2mm]
\dfrac{dI(t)}{dt} = \beta S(t)P(t) - \gamma I(t), \\[2mm]
\dfrac{dR(t)}{dt} = \gamma I(t),
\end{cases}
\tag{1}
$$

where β and γ are the transmission rate parameter and the removal rate parameter respectively. It is clear that

$$
\frac{dS(t)}{dt} + \frac{dI(t)}{dt} + \frac{dR(t)}{dt} = 0,
$$

and thus the population size is fixed as $S(t) + I(t) + R(t) = N$ for $t \geq 0$.

Although the SIR model has been widely used due to its simplicity and interpretability, it cannot fit real Omicron data well. The main reason is that the SIR model assumes that both the transmission rate β and the removal rate γ remain unchanged throughout the whole study period. This assumption is unrealistic for the ongoing COVID-19 pandemic, where the transmission rate would decrease as individual strategies of self-protection or government policies of social distancing are implemented. In addition, the removal rate would also increase as more medical resources are allocated to cope with the pandemic and help to cure patients.

Methodology

The stochastic SIR model with change points. To model the spreading of infectious diseases with time-varying parameters, we propose the stochastic SIR model with change points. Because the numbers of confirmed cases and removed cases are reported only on a daily basis during the COVID-19 pandemic, we assume that the study period is $\{1, \ldots, T\}$, where T is the length of the study period. Let $\Delta \mathbf{I} = (\Delta I_1, \ldots, \Delta I_T)^{\mathsf{T}}$ and $\Delta \mathbf{R} = (\Delta R_1, \ldots, \Delta R_T)^{\mathsf{T}}$ be sequences of daily reported numbers of newly infected (confirmed) cases and removed cases respectively. Given the initial state of the population (S_0, I_0, R_0), we modify the SIR model in (1) and develop the discrete-time stochastic SIR model with change points as follows. For $t = 1, \ldots, T$, we assume

$$
\begin{cases}
\Delta I_t \sim \text{Binomial}(S_{t-1}, 1 - \exp(-\beta_t P_{t-1})), \\
\Delta R_t \sim \text{Binomial}(I_{t-1}, \gamma_t), \\
S_t = S_{t-1} - \Delta I_t, \\
I_t = I_{t-1} + \Delta I_t - \Delta R_t, \\
R_t = R_{t-1} + \Delta R_t,
\end{cases}
\tag{2}
$$

where $P_t = I_t/N$ is the proportion of infectious individuals at day t and (S_t, I_t, R_t) are updated as shown in Fig. 1a. Time-varying parameters $\beta_t \in (0, \infty)$ and $\gamma_t \in (0, 1)$ represent the transmission rate and the removal rate at day t respectively.

The second equation of the proposed model (2) corresponds the term $\gamma I(t)$ in (1), which indicates the number of infectious individuals being removed. Assuming that the removal of different infectious individuals are mutually independent, the number of infectious individuals being removed at day t (ΔR_t) follows $\text{Binomial}(I_{t-1}, \gamma_t)$. However, unlike the term $\beta S(t)P(t)$ in model (1), we make the following assumptions for each susceptible individual at day $t - 1$: (i) The number of in-person contacts with other individuals at day t, $N_{\text{contact},t}$, follows $\text{Poisson}(\lambda_t)$; (ii) Given $N_{\text{contact},t}$, the number of in-person contacts with infectious individuals at day t follows $\text{Binomial}(N_{\text{contact},t}, P_{t-1})$; and (iii) The probability of transmission for each in-person contact with infectious individuals is q_t. As a result, at day t, the probability for a susceptible individual getting infected is $1 - \exp(-\lambda_t q_t P_{t-1})$. For identifiability, we reparameterize $\beta_t = \lambda_t q_t$ as the disease transmission rate at day t.
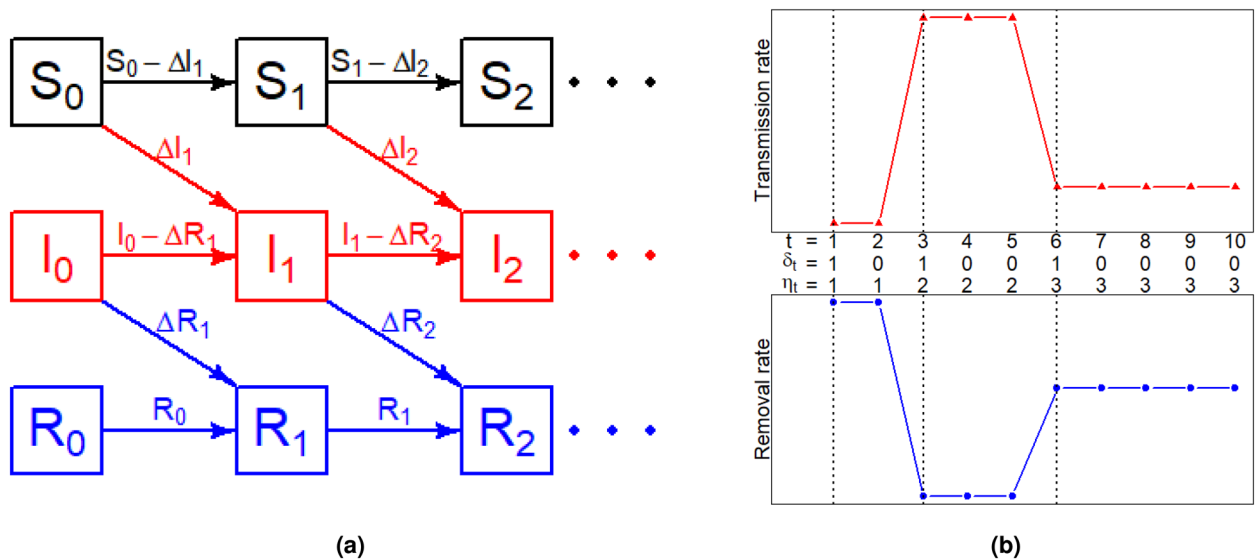
Figure 1. Graphical illustration of (**a**) the update of (S_t, I_t, R_t); (**b**) change-points in the expected transmission rates and expected removal rates of different stages.

Unlike the original SIR model (1) where parameters β and γ are constant throughout the study period, our model (2) assumes that the transmission rate β_t and the removal rate γ_t may change significantly at several change points in the study period. The change points exist because there could be some societal changes, such as the implementation of new social distancing policies, restrictions on individual activities, and the increase of medical resources on patients, which would cause sudden changes in either the transmission rate or the removal rate or both. To depict such dynamic patterns of pandemic evolvement, we introduce a latent binary vector $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_T)^{\mathsf{T}}$, where $\delta_t = 1$ indicates that day t is a change point and $\delta_t = 0$ otherwise, and δ_1 is fixed as 1 for convenience. The index of the stage that contains day t is $\eta_t = \sum_{u=1}^{t} \delta_u$ and the stage index vector is $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_T)^{\mathsf{T}}$, as illustrated in Fig. 1b. Thus, the study period is partitioned into $K = \sum_{t=1}^{T} \delta_t$ stages. Within each stage, parameters (β_t, γ_t) are assumed to be homogeneous with the hierarchical priors,

$$\beta_t | \boldsymbol{\delta}, \mathrm{b} \sim \mathrm{Exp}(b_{\eta_t}),$$

$$\gamma_t | \boldsymbol{\delta}, \mathrm{r} \sim \mathrm{Beta}(r_{\eta_t}, 1),$$

$$b_1, \ldots, b_K \overset{\text{i.i.d.}}{\sim} \mathrm{Gamma}(0.1, 0.1),$$

$$r_1, \ldots, r_K \overset{\text{i.i.d.}}{\sim} \mathrm{Gamma}(0.1, 0.1),$$

$$\delta_1, \ldots, \delta_T \overset{\text{i.i.d.}}{\sim} \mathrm{Bernoulli}(p),$$

(3)

which are denoted as $\pi(\beta_t | \boldsymbol{\delta}, \mathbf{b})$, $\pi(\gamma_t | \boldsymbol{\delta}, \mathbf{r})$, $\pi(b_k)$, $\pi(r_k)$ and $\pi(\boldsymbol{\delta})$ respectively, with $\mathbf{b} = (b_1, \ldots, b_K)^{\mathsf{T}}$ and $\mathbf{r} = (r_1, \ldots, r_K)^{\mathsf{T}}$. The expected transmission rate and the expected removal rate in stage k ($k = 1, \ldots, K$) are $1/b_k$ and $r_k/(1 + r_k)$, respectively. The hyperparameter p expresses one's prior belief on how often a change point occurs in a pandemic wave, because p is the probability of each time point is a change point. We fix $p = 0.01$ in our experiments and also explore the sensitivity analysis with respect to p.

Bayesian analysis with MCMC. To make Bayesian inference on parameters $\boldsymbol{\delta}, \mathbf{b}, \mathbf{r}, \boldsymbol{\beta} = (\beta_1, \ldots, \beta_T)^{\mathsf{T}}$ and $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_T)^{\mathsf{T}}$, we develop an MCMC algorithm to draw samples from the posterior distribution,

$$\pi(\boldsymbol{\delta}, \mathrm{b}, \mathrm{r}, \boldsymbol{\beta}, \boldsymbol{\gamma} | S_0, I_0, R_0, \Delta \mathbf{I}, \Delta \mathbf{R}) \propto \pi(\boldsymbol{\delta}) \prod_{k=1}^{K} \pi(b_k) \prod_{t=1}^{T} \left\{ \pi(\beta_t | \boldsymbol{\delta}, \mathbf{b}) \mathbb{P}(\Delta I_t | S_{t-1}, I_{t-1}, \beta_t) \right\}$$

$$\times \prod_{k=1}^{K} \pi(r_k) \prod_{t=1}^{T} \left\{ \pi(\gamma_t | \boldsymbol{\delta}, \mathbf{r}) \mathbb{P}(\Delta R_t | I_{t-1}, \gamma_t) \right\},$$

(4)

where the $\pi(\cdot)$ functions on the right-hand side correspond to the priors in (3), and $\mathbb{P}(\Delta I_t | S_{t-1}, I_{t-1}, \beta_t)$ and $\mathbb{P}(\Delta R_t | I_{t-1}, \gamma_t)$ represent the binomial likelihood functions in (2). At each iteration, parameters are sequentially updated with the Gibbs sampling procedure detailed as follows.

- **Update $\boldsymbol{\delta}$:** Given current values $\boldsymbol{\beta}^{(g)}$ and $\boldsymbol{\gamma}^{(g)}$, $\boldsymbol{\delta}$ is updated via an *add–delete–swap* algorithm. Initializing $\boldsymbol{\delta}^* = \boldsymbol{\delta}^{(g)}$, an operation is selected from {*add*, *delete*, *swap*} with probabilities,
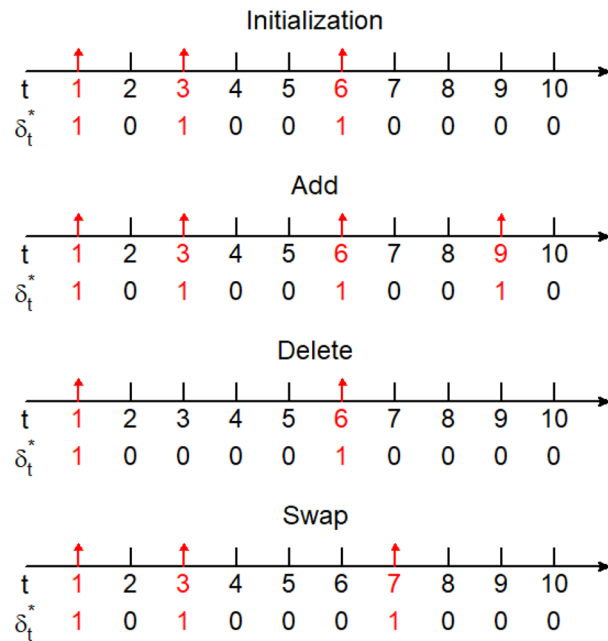
Figure 2. Illustration of the *add–delete–swap* algorithm. Based on the initialization $\boldsymbol{\delta}^* = \boldsymbol{\delta}^{(g)} = (1, 0, 1, 0, 0, 1, 0, 0, 0, 0)$, the candidate $\boldsymbol{\delta}^*$ in the $(g + 1)$-th update can be obtained as: (i) $(1, 0, 1, 0, 0, 1, 0, 0, 1, 0)$ by selecting the *add* operation and updating δ_9^* as 1; (ii) $(1, 0, 0, 0, 0, 1, 0, 0, 0, 0)$ by selecting the *delete* operation and updating δ_3^* as 0; (iii) $(1, 0, 1, 0, 0, 0, 1, 0, 0, 0)$ by selecting the *swap* operation and exchanging values of δ_6^* and δ_7^*.

$$(p_{add}, p_{delete}, p_{swap}) = \begin{cases} (1, 0, 0), & \text{if } K^{(g)} = 1, \\ (0, 1, 0), & \text{if } K^{(g)} = T, \\ (1/3, 1/3, 1/3), & \text{otherwise,} \end{cases} \tag{5}$$

where $K^{(g)} = \sum_{t=1}^T \delta_t^{(g)}$. If the *add* (or *delete*) operation is selected, we randomly select a δ_t^* which is 0 (or 1) and update its value as 1 (or 0). If the *swap* operation is selected, we randomly select a pair of $(\delta_t^*, \delta_{t-1}^*)$ with different values and exchange their values. Examples of the candidate $\boldsymbol{\delta}^*$ obtained by different operations are shown in Fig. 2. Given the candidate $\boldsymbol{\delta}^*$, we compute the Metropolis–Hasting ratio as

$$m_{\text{MH}} = \frac{\pi(\boldsymbol{\beta}^{(g)}, \boldsymbol{\gamma}^{(g)} | \boldsymbol{\delta}^*)}{\pi(\boldsymbol{\beta}^{(g)}, \boldsymbol{\gamma}^{(g)} | \boldsymbol{\delta}^{(g)})} \cdot \frac{\pi(\boldsymbol{\delta}^*)}{\pi(\boldsymbol{\delta}^{(g)})} \cdot \frac{J(\boldsymbol{\delta}^{(g)} | \boldsymbol{\delta}^*)}{J(\boldsymbol{\delta}^* | \boldsymbol{\delta}^{(g)})},$$

where following (3) we can derive

$$\frac{\pi(\boldsymbol{\delta}^*)}{\pi(\boldsymbol{\delta}^{(g)})} = \frac{p^{\sum_{t=1}^T \delta_t^*}(1-p)^{T - \sum_{t=1}^T \delta_t^*}}{p^{\sum_{t=1}^T \delta_t^{(g)}}(1-p)^{T - \sum_{t=1}^T \delta_t^{(g)}}} = \left(\frac{p}{1-p}\right)^{\sum_{t=1}^T (\delta_t^* - \delta_t^{(g)})},$$

$$\pi(\boldsymbol{\beta}^{(g)}, \boldsymbol{\gamma}^{(g)} | \boldsymbol{\delta})$$

$$= \int_{b_1=0}^\infty \cdots \int_{b_K=0}^\infty \int_{r_1=0}^\infty \cdots \int_{r_K=0}^\infty \left\{ \prod_{k=1}^K \pi(b_k) \prod_{t=1}^T \pi(\beta_t^{(g)} | \boldsymbol{\delta}, \text{b}) \prod_{k=1}^K \pi(r_k) \prod_{t=1}^T \pi(\gamma_t^{(g)} | \boldsymbol{\delta}, \text{r}) \right\} dr_K \ldots dr_1 db_K \ldots db_1$$

$$= \prod_{k=1}^K \left\{ \int_0^\infty \pi(b_k) \prod_{t:\eta_t=k} \pi(\beta_t^{(g)} | \boldsymbol{\delta}, \text{b}) db_k \right\} \prod_{k=1}^K \left\{ \int_0^\infty \pi(r_k) \prod_{t:\eta_t=k} \pi(\gamma_t^{(g)} | \boldsymbol{\delta}, \text{r}) dr_k \right\}$$

$$= \prod_{k=1}^K \left\{ \int_0^\infty \frac{0.1^{0.1} b_k^{-0.9} e^{-0.1 b_k}}{\Gamma(0.1)} \times \prod_{t:\eta_t=k} b_k e^{-\beta_t^{(g)} b_k} db_k \right\} \prod_{k=1}^K \left\{ \int_0^\infty \frac{0.1^{0.1} r_k^{-0.9} e^{-0.1 r_k}}{\Gamma(0.1)} \times \prod_{t:\eta_t=k} r_k(\gamma_t^{(g)})^{r_k-1} dr_k \right\}$$

$$\propto \prod_{k=1}^K \left\{ \frac{\Gamma(0.1 + \sum_{t=1}^T \mathbb{I}(\eta_t = k))}{\{0.1 + \sum_{t=1}^T \beta_t^{(g)} \mathbb{I}(\eta_t = k)\}^{0.1 + \sum_{t=1}^T \mathbb{I}(\eta_t=k)}} \right\} \prod_{k=1}^K \left\{ \frac{\Gamma(0.1 + \sum_{t=1}^T \mathbb{I}(\eta_t = k))}{\{0.1 + \sum_{t=1}^T -\log \gamma_t^{(g)} \mathbb{I}(\eta_t = k)\}^{0.1 + \sum_{t=1}^T \mathbb{I}(\eta_t=k)}} \right\},$$

with $\Gamma(\cdot)$ being the gamma function, and the proposal of the Metropolis–Hasting algorithm is

4

$$\frac{J(\boldsymbol{\delta}^{(g)}|\boldsymbol{\delta}^*)}{J(\boldsymbol{\delta}^*|\boldsymbol{\delta}^{(g)})} = \begin{cases} 1, & \text{if } \sum_{t=1}^{T}\delta_t^* = \sum_{t=1}^{T}\delta_t^{(g)}, \\ 3/(T-1), & \text{if } (\sum_{t=1}^{T}\delta_t^*, \sum_{t=1}^{T}\delta_t^{(g)}) = (1,2) \text{ or } (T, T-1), \\ (T-1)/3, & \text{if } (\sum_{t=1}^{T}\delta_t^*, \sum_{t=1}^{T}\delta_t^{(g)}) = (2,1) \text{ or } (T-1, T), \\ (\sum_{t=1}^{T}\delta_t^{(g)}-1)/(T-\sum_{t=1}^{T}\delta_t^*), & \text{if } (\sum_{t=1}^{T}\delta_t^*, \sum_{t=1}^{T}\delta_t^{(g)}) \in \{(2,3),(3,4),\ldots,(T-2,T-1)\}, \\ (T-\sum_{t=1}^{T}\delta_t^{(g)})/(\sum_{t=1}^{T}\delta_t^*-1), & \text{if } (\sum_{t=1}^{T}\delta_t^*, \sum_{t=1}^{T}\delta_t^{(g)}) \in \{(3,2),(4,3),\ldots,(T-1,T-2)\}. \end{cases}$$

We then obtain the updated value $\boldsymbol{\delta}^{(g+1)}$ as

$$\boldsymbol{\delta}^{(g+1)} = \begin{cases} \boldsymbol{\delta}^*, & \text{with probability } \min(1, m_{\text{MH}}), \\ \boldsymbol{\delta}^{(g)}, & \text{with probability } 1 - \min(1, m_{\text{MH}}). \end{cases}$$

Correspondingly, we also update $K^{(g+1)} = \sum_{t=1}^{T}\delta_t^{(g+1)}$ and $\eta_t^{(g+1)} = \sum_{u=1}^{t}\delta_u^{(g+1)}$ $(t = 1, \ldots, T)$.

- **Update b and r**: Given current values $\boldsymbol{\delta}^{(g+1)}$, $\boldsymbol{\beta}^{(g)}$ and $\boldsymbol{\gamma}^{(g)}$, we sample $b_k^{(g+1)}$ and $r_k^{(g+1)}$ from

$$b_k \sim \text{Gamma}\left(0.1 + \sum_{t=1}^{T}\mathbb{I}(\eta_t^{(g+1)} = k), 0.1 + \sum_{t=1}^{T}\beta_t^{(g)}\mathbb{I}(\eta_t^{(g+1)} = k)\right),$$

$$r_k \sim \text{Gamma}\left(0.1 + \sum_{t=1}^{T}\mathbb{I}(\eta_t^{(g+1)} = k), 0.1 + \sum_{t=1}^{T}-\log\gamma_t^{(g)}\mathbb{I}(\eta_t^{(g+1)} = k)\right),$$

for $k = 1, \ldots, K^{(g+1)}$, where $\mathbb{I}(\cdot)$ is the indicator function.

- **Update $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$**: Given current values $\boldsymbol{\delta}^{(g+1)}$, $\mathbf{b}^{(g+1)}$ and $\mathbf{r}^{(g+1)}$, we sample $\beta_t^{(g+1)}$ and $\gamma_t^{(g+1)}$ from posterior densities,

$$\pi(\beta_t|S_{t-1}, \Delta I_t, b_{\eta_t^{(g+1)}}^{(g+1)}) \propto \{1 - \exp(-\beta_t P_{t-1})\}^{\Delta I_t} \exp\{-\beta_t P_{t-1}(S_{t-1} - \Delta I_t)\}\pi(\beta_t|b_{\eta_t^{(g+1)}}^{(g+1)}),$$

$$\pi(\gamma_t|I_{t-1}, \Delta R_t, r_{\eta_t^{(g+1)}}^{(g+1)}) \propto \gamma_t^{\Delta R_t}(1 - \gamma_t)^{I_{t-1} - \Delta R_t}\pi(\gamma_t|r_{\eta_t^{(g+1)}}^{(g+1)}),$$

for $t = 1, \ldots, T$.

Using the MCMC algorithm, a set of posterior samples $\{(\boldsymbol{\delta}^{(g)}, \mathbf{b}^{(g)}, \mathbf{r}^{(g)}, \boldsymbol{\beta}^{(g)}, \boldsymbol{\gamma}^{(g)}) : g = 1, \ldots, G\}$ is obtained for inference. Because the main interest lies in detection of change points, we aggregate $\boldsymbol{\delta}^{(1)}, \ldots, \boldsymbol{\delta}^{(G)}$ to obtain a point estimate $\hat{\boldsymbol{\delta}}$ (or the corresponding $\hat{\boldsymbol{\eta}}$). As the indicator vector $\boldsymbol{\delta}$ is a binary vector with 2^{T-1} possible values in total, its posterior mean does not imply a partition of the study period and thus is difficult to interpret. Taking the sequential structure of $\boldsymbol{\delta}$ into consideration, we interpret each $\boldsymbol{\delta}^{(g)}$ (or the corresponding $\boldsymbol{\eta}^{(g)}$) as a cluster of time points $1, \ldots, T$ and obtain $\hat{\boldsymbol{\delta}}$ by solving a clustering aggregation problem as follows.

- For each pair of time points t and t' $(1 \le t < t' \le T)$, we estimate the posterior probability that no change points exist in the time period $\{t + 1, \ldots, t'\}$ as

$$\hat{q}_{tt'} = \frac{1}{G}\sum_{g=1}^{G}\mathbb{I}(\eta_t^{(g)} = \eta_{t'}^{(g)}),$$

where $\mathbb{I}(\cdot)$ is the indicator function.

- The Bayes estimator $\hat{\boldsymbol{\delta}}$ is then obtained as

$$\hat{\boldsymbol{\delta}} = \arg\min_{\boldsymbol{\delta}} \sum_{1 \le t < t' \le T} \left| \mathbb{I}(\eta_t = \eta_{t'}) - \hat{q}_{tt'}\right|,$$

where $\eta_t = \sum_{u=1}^{t}\delta_u$. The Bayes estimator $\hat{\boldsymbol{\eta}}$ can also be obtained correspondingly.

Based on $\hat{\boldsymbol{\delta}}$, we can also compute probability regions of change points. With $\hat{K} = \sum_{t=1}^{T}\hat{\delta}_t$, let $\hat{\tau}_k = \min\{t : \sum_{u=1}^{t}\hat{\delta}_u \ge k\}$ denote the k-th estimated change point $(k = 1, \ldots, \hat{K})$. The $(1 - \alpha)$ highest posterior density (HPD) interval of the k-th change point is computed as $[\hat{\tau}_{lk}, \hat{\tau}_{uk}]$, where

$$\hat{\tau}_{lk}, \hat{\tau}_{uk} = \arg\min_{\tau_l \le \hat{\tau}_k \le \tau_u}(\tau_u - \tau_l), \quad \text{s.t.} \quad \frac{1}{G}\sum_{u=\tau_l}^{\tau_u}\sum_{g=1}^{G}\delta_u^{(g)} \ge 1 - \alpha.$$

As the number of stages varies among posterior samples, we use $G^{-1}\sum_{g=1}^{G}1/b_{\eta_t^{(g)}}^{(g)}$ and $G^{-1}\sum_{g=1}^{G}r_{\eta_t^{(g)}}^{(g)}/(1 + r_{\eta_t^{(g)}}^{(g)})$ as smoothed estimators of the expected transmission rate and expected removal rate at day t respectively.

Simulations

To evaluate the performance of the proposed method in detecting change points of the transmission rate and removal rate during a pandemic wave, we conduct simulated experiments by applying the developed MCMC algorithm for making Bayesian inference.
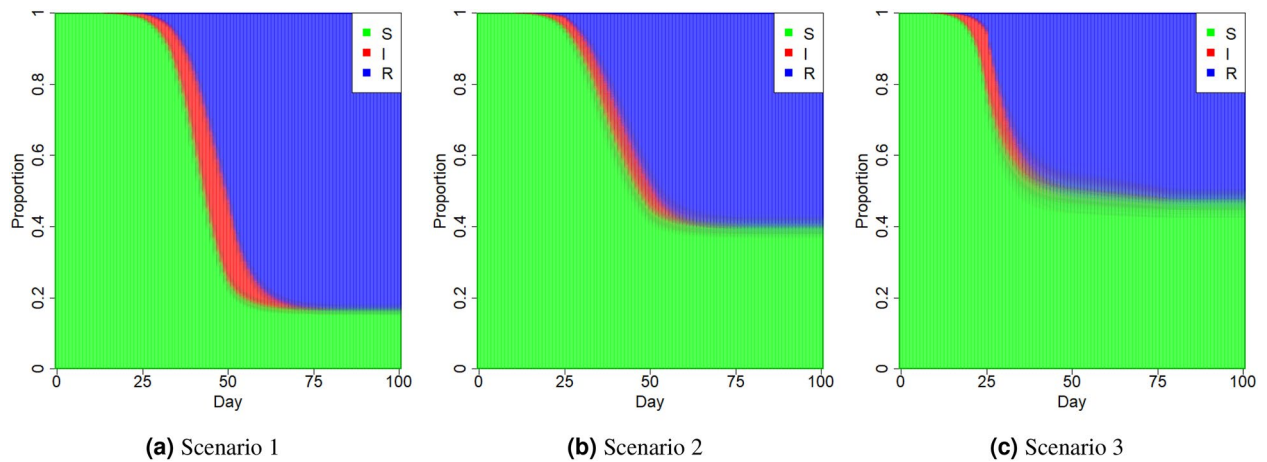
(a) Scenario 1　　　　　　**(b)** Scenario 2　　　　　　**(c)** Scenario 3

Figure 3. Average proportions of susceptible (S), infectious (I) and removed (R) individuals in the population at time t ($t = 1, \ldots, 100$) under different scenarios.

Data generating mechanism. Considering a study period of length $T = 100$ and a fixed population size $N = 1000000$, we partition the study period $\{1, \ldots, T\}$ into four stages evenly. In other words, there are three true change points at $t = 26$, $t = 51$ and $t = 76$. To mimic the dynamics of the transmission rate and removal rate during a pandemic wave, we design three scenarios with the time-varying transmission rate and removal rate as follows,

- Scenario 1:

$$(\beta_t, \gamma_t) = \begin{cases} (0.3, 0.05), & t = 1, \ldots, 25; \\ (0.4, 0.15), & t = 26, \ldots, 50; \\ (0.25, 0.2), & t = 51, \ldots, 75; \\ (0.2, 0.25), & t = 76, \ldots, 100. \end{cases}$$

- Scenario 2:

$$(\beta_t, \gamma_t) = \begin{cases} (0.4, 0.1), & t = 1, \ldots, 25; \\ (0.4, 0.25), & t = 26, \ldots, 50; \\ (0.25, 0.25), & t = 51, \ldots, 75; \\ (0.25, 0.4), & t = 76, \ldots, 100. \end{cases}$$

- Scenario 3:

$$(\beta_t, \gamma_t) = \begin{cases} (0.5, 0.1), & t = 1, \ldots, 25; \\ (0.3, 0.3), & t = 26, \ldots, 50; \\ (0.4, 0.2), & t = 51, \ldots, 75; \\ (0.2, 0.4), & t = 76, \ldots, 100. \end{cases}$$

With initial numbers $(S_0, I_0, R_0) = (N - 50, 50, 0)$, we generate 100 datasets for each of three scenarios. Within each dataset, daily confirmed cases $\Delta \mathbf{I}$ and daily removed cases $\Delta \mathbf{R}$ are generated sequentially as Fig. 1a under the model (2) from $t = 1$ to $t = T$. Given $\Delta \mathbf{I}$ and $\Delta \mathbf{R}$, we exhibit (S_t, I_t, R_t) ($1 \le t \le T$) in Fig. 3. It is clear that the severity of the pandemic wave is the strongest under Scenario 1 and the total number of infections is the smallest under Scenario 3.

Bayesian analysis. We apply the MCMC algorithm proposed in section "Methodology" with hyperparameter $p = 0.01$ to analyze simulated datasets. For each dataset, we discard the first 5000 iterations as burn-in and draw one posterior sample for every 10 iterations in the next 10000 iterations, leading to $G = 1000$ posterior samples $(\boldsymbol{\delta}^{(g)}, \mathbf{b}^{(g)}, \mathbf{r}^{(g)}, \boldsymbol{\beta}^{(g)}, \boldsymbol{\gamma}^{(g)})$ in total.

As our main interest lies in the detection of change points, we first investigate the performance of the proposed Bayes estimator $\hat{\boldsymbol{\delta}}$. We present the proportion of 100 simulated datasets where a change point is detected at time point t ($t = 1, \ldots, T$) in Fig. 4a–c. Under all three scenarios, the estimated change points are near the true ones, suggesting that the proposed Bayesian method can accurately locate the change points of a pandemic wave. Compared with the first two scenarios, the estimated change points under Scenario 3 are more concentrated around the truth. We also present the coverage probability of $(1 - \alpha)$ probability regions (i.e., HPD intervals) of change points for each t under different scenarios and different values of α in Fig. 4d–f. The average length of HPD intervals is the shortest under Scenario 3, while intervals of the second and third change points sometimes intersect under Scenarios 1 and 2.
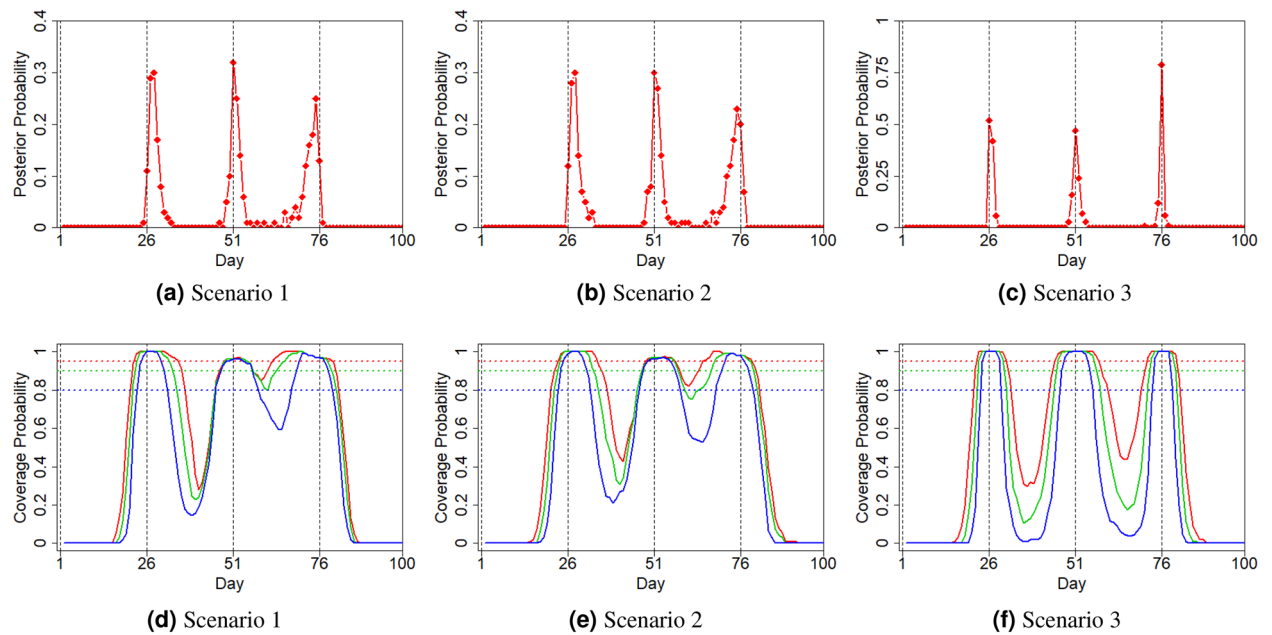
6

Figure 4. (**a**–**c**) Proportions that a change point is estimated at time t ($1 \leq t \leq T$) under different scenarios; (**d**–**f**) Coverage probability of the $(1 - \alpha)$ HPD intervals of change points at each t under different scenarios ($1 - \alpha = 95\%$ in red; 90% in green; 80% in blue).

To quantitatively measure the agreement between the true $\boldsymbol{\delta}$ (or $\boldsymbol{\eta}$) and the estimator $\hat{\boldsymbol{\delta}}$ (or $\hat{\boldsymbol{\eta}}$), we transfer the stage allocation of time points $\{1, \ldots, T\}$ to a clustering problem and adopt the adjusted Rand index[19] and the mutual information[20] as evaluation metrics.

- Adjusted Rand index (ARI):

$$\text{ARI} = \frac{(\text{TP} + \text{TN}) - \{(\text{TP} + \text{FP})(\text{TP} + \text{FN}) + (\text{TN} + \text{FP})(\text{TN} + \text{FN})\}}{1 - \{(\text{TP} + \text{FP})(\text{TP} + \text{FN}) + (\text{TN} + \text{FP})(\text{TN} + \text{FN})\}},$$

where the proportions of true positives (TP), false positives (FP), false negatives (FN) and true negatives (TN) in stage allocation are given by

$$\begin{cases} \text{TP} = \frac{2}{T(T-1)} \sum_{1 \leq t < t' \leq T} \mathbb{I}(\eta_t = \eta_{t'}, \hat{\eta}_t = \hat{\eta}_{t'}), \\ \text{FP} = \frac{2}{T(T-1)} \sum_{1 \leq t < t' \leq T} \mathbb{I}(\eta_t \neq \eta_{t'}, \hat{\eta}_t = \hat{\eta}_{t'}), \\ \text{FN} = \frac{2}{T(T-1)} \sum_{1 \leq t < t' \leq T} \mathbb{I}(\eta_t = \eta_{t'}, \hat{\eta}_t \neq \hat{\eta}_{t'}), \\ \text{TN} = \frac{2}{T(T-1)} \sum_{1 \leq t < t' \leq T} \mathbb{I}(\eta_t \neq \eta_{t'}, \hat{\eta}_t \neq \hat{\eta}_{t'}). \end{cases}$$

The range of ARI is $[0, 1]$, where a larger ARI suggests that $\hat{\boldsymbol{\delta}}$ is more similar to $\boldsymbol{\delta}$ and the maximum possible value is obtained when $\hat{\boldsymbol{\delta}} = \boldsymbol{\delta}$.

- Mutual information (MI):

$$\text{MI} = \sum_{k=1}^{K} \sum_{k'=1}^{\hat{K}} \frac{n_{kk'}}{T} \log \left(\frac{n_{kk'} T}{n_{k \cdot} n_{\cdot k'}} \right),$$

where

$$\begin{cases} \hat{K} = \sum_{t=1}^{T} \hat{\delta}_t, \\ n_{kk'} = \sum_{t=1}^{T} \mathbb{I}(\eta_t = k, \hat{\eta}_t = k'), \\ n_{k \cdot} = \sum_{t=1}^{T} \mathbb{I}(\eta_t = k), \\ n_{\cdot k'} = \sum_{t=1}^{T} \mathbb{I}(\hat{\eta}_t = k'). \end{cases}$$

The range of MI is $[0, \sum_{k=1}^{K} n_{k \cdot}/T \times \log(T/n_{k \cdot})]$, where a larger MI suggests that $\hat{\boldsymbol{\delta}}$ is more similar to $\boldsymbol{\delta}$ and the maximum possible value is obtained when $\hat{\boldsymbol{\delta}} = \boldsymbol{\delta}$. Under our three scenarios, the maximum possible value of MI is

	ARI ∈ [0, 1]			MI ∈ [0, 1.386]		
	Scenario 1	Scenario 2	Scenario 3	Scenario 1	Scenario 2	Scenario 3
$p = 0.01$	0.855 (0.092)	0.862 (0.094)	0.960 (0.035)	1.198 (0.107)	1.206 (0.107)	1.327 (0.049)
$p = 0.05$	0.954 (0.039)	0.953 (0.037)	0.977 (0.023)	1.319 (0.054)	1.316 (0.052)	1.351 (0.035)
$p = 0.2$	0.625 (0.040)	0.618 (0.042)	0.709 (0.050)	1.144 (0.057)	1.146 (0.051)	1.213 (0.058)

Table 1. Average values of accuracy measures of the Bayes estimator $\hat{\boldsymbol{\delta}}$ with standard deviations in parentheses.
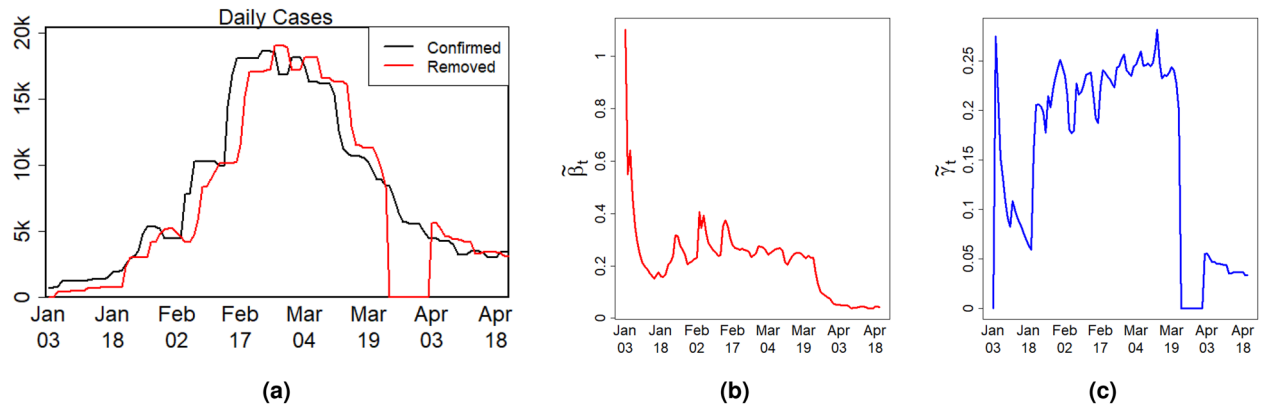
Figure 5. (**a**) The numbers of daily confirmed and removed COVID-19 cases, (**b**) the naive estimator of the time-varying transmission rate, and (**c**) the naive estimator of the time-varying removal rate during the Omicron wave in Singapore.

$$\sum_{k=1}^{K} \frac{n_{k\cdot}}{T} \times \log\left(\frac{T}{n_{k\cdot}}\right) = \sum_{k=1}^{4} \frac{25}{100} \times \log\left(\frac{100}{25}\right) = \log 4 = 1.386,$$

because $n_{k\cdot} = 25$ for $k = 1, \ldots, 4$.

The performance of the Bayes estimator $\hat{\boldsymbol{\delta}}$ under different scenarios is summarized in the first row of Table 1. It is clear that our estimator can accurately estimate the underlying $\boldsymbol{\delta}$, especially under Scenario 3 with large changes of the transmission rate and removal rate between adjacent stages.

We also vary the hyper-parameter p to investigate how the performance of $\hat{\boldsymbol{\delta}}$ is sensitive to the choice of prior probability. Both the ARI and MI of $\hat{\boldsymbol{\delta}}$ are larger under $p = 0.05$ than those under $p = 0.01$ but with no significant gaps. This is because $p = 0.05$ is closer to the true proportion of change points $K/T = 0.04$ than 0.01 and thus true change points can be selected with more certainty, even though the signal strengths of true change points are still strong enough to be detected using $p = 0.01$. When $p = 0.2$, the ARI of $\hat{\boldsymbol{\delta}}$ becomes smaller due to more false positives in change point detection. In general, the performance of $\hat{\boldsymbol{\delta}}$ is stable as long as the prior probability p is not far larger than the true proportion of change points. A guideline to choose p is first checking the naive estimators of the transmission rate and removal rate (as discussed in the next section) and then taking the ratio of the number of change points and the study length T.

Analysis of the Omicron wave in Singapore

At the beginning of 2022, Singapore witnessed the outbreak of the Omicron wave of COVID-19 pandemic. To investigate change points of the Omicron wave, we collect the daily reported number of confirmed cases and removed cases in Singapore from January 3 to April 21, 2022, during which 893854 confirmed cases and 453 deaths of COVID-19 were reported in $T = 109$ days. Let $N = 5930134$ be the population size. The daily reported number of confirmed and removed cases during the study period is presented in Fig. 5a.

To gain insight into the trend of β_t and γ_t, we compute the naive estimator of model (2),

$$\tilde{\beta}_t = -\frac{1}{P_{t-1}} \log\left(\frac{\Delta I_t}{S_{t-1}}\right), \quad \tilde{\gamma}_t = \frac{\Delta R_t}{I_{t-1}}, \quad t = 1, \ldots, T.$$

As shown in Fig. 5b,c, the transmission rate was very high in the first several days and then became stable between January 18 and March 19, 2022. After March 19, the transmission rate decreased to almost zero. The removal rate was low before January 18, 2022. Between January 18 and March 23, the removal rate increased stably, which then had a sharp drop followed with a sudden increase because recovery cases were not reported from March 24 to April 1, 2022.
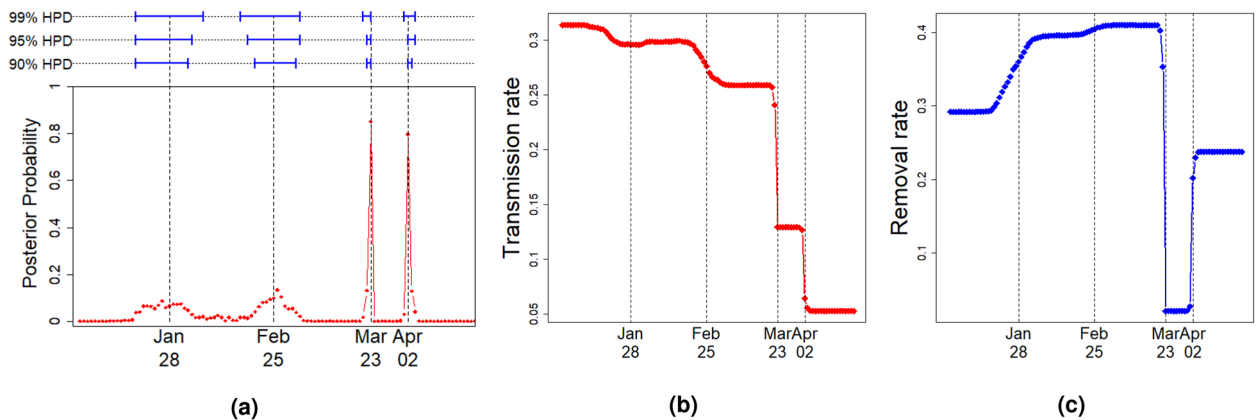
Figure 6. The Bayes estimator of (**a**) $\boldsymbol{\delta}$ and 95% probability regions (HPD intervals) for change points; (**b**) the expected transmission rate; and (**c**) the expected removal rate.

Although the naive estimator suggests possible existence of change points in the study period, it cannot provide the estimates of change points with quantitative uncertainty measurement and thus we apply the proposed MCMC algorithm with hyperparameter $p = 0.01$ to draw posterior samples of parameters for Bayesian inference. We discard the first 5000 iterations as burn-in and draw one posterior sample for every 10 iterations in the next 10000 iterations, leading to $G = 1000$ posterior samples $(\boldsymbol{\delta}^{(g)}, \mathbf{b}^{(g)}, \mathbf{r}^{(g)}, \boldsymbol{\beta}^{(g)}, \boldsymbol{\gamma}^{(g)})$ in total. The Bayes estimator $\hat{\boldsymbol{\delta}}$ is presented in Fig. 6a, with estimated posterior probabilities $\sum_{g=1}^{G} \delta_t^{(g)}/G \ (t = 1, \ldots, T)$ and the corresponding probability regions of change points. Four change points are detected in the Omicron wave in Singapore. The probability regions of the third (March 23, 2022) and fourth (April 2, 2022) change points are shorter, while uncertainties of the first two change points (January 28 and February 25, 2022) are greater. With the smoothed estimators of the expected transmission rate and expected removal rate on day $t \ (t = 1, \ldots, T)$ shown in Fig. 6b,c, we can interpret different stages as follows.

1. **January 3 to January 27**: The transmission rate was high while the removal rate was low, indicating the outbreak of the Omicron wave and medical resources in Singapore were not ready to cope with the Omicron variant yet.
2. **January 28 to February 24**: The transmission rate decreased slightly, indicating that people in Singapore had realized the outbreak of the Omicron wave and begun self-protection activities. In addition, the removal rate increased to a high level as the Ministry of Health in Singapore begun deploying medical resources to COVID-19 patients.
3. **February 25 to March 22**: The Omicron wave reached the peak and then passed it due to the sharp decrease in the transmission rate.
4. **March 23 to April 1**: An abnormal period was detected when recovery cases were not reported for several days, leading to a sharp drop in the estimated removal rate.
5. **April 2 to April 21**: The Omicron wave became stabilized and under control.

Conclusion

The stochastic SIR model with change points is developed to quantify the evolvement of the Omicron wave of COVID-19 pandemic. The proposed model assumes that the numbers of daily confirmed cases and removed cases follow binomial distributions and the probability of a susceptible individual getting infected is determined by the proportion of currently infectious individuals and the time-varying transmission rate. We develop an MCMC algorithm to draw posterior samples and provide both the Bayes estimators and probability regions of change points, which are corroborated by experiments on simulated data. By analyzing the Omicron wave in Singapore, four change points are detected, corresponding to the increase of medical resources, decrease of the transmission rate, the beginning and the end of an abnormal period with no removal cases reported. The proposed model is applicable to a pandemic wave of any infectious disease under the impact of social changes. It is also possible to make inference on the spreading of COVID-19 and compare the effect of disease-control policies in different countries, as long as both daily numbers of confirmed cases and removal cases are available. In addition, our model can be easily extended to incorporate other interpersonal contact patterns into disease transmission. Although our model (2) is also applicable to sequential (real-time) change point detection in an ongoing epidemic, the goals for real-time detection are to minimize the detection delay as well as false detections of change points[21]. Timely detecting changes in the transmission rate and removal rate is critically important for policy making and identifying any possible new variants of the virus. To accommodate such objectives, construction of the posterior loss and Bayes estimators of change points warrant further investigation.

Data availability

All data generated and analyzed during this study are included in the supplementary file "Code.zip".

References

1. Lyngse, F. P. *et al.* Household transmission of the SARS-CoV-2 Omicron variant in Denmark. *Nat. Commun.* **13**, 5573. https://doi.org/10.1038/s41467-022-33328-3 (2022).
2. Pearson, C. A. B. *et al.* Bounding the levels of transmissibility and immune evasion of the Omicron variant in South Africa. *medRxiv*https://doi.org/10.1101/2021.12.19.21268038 *(2021)*.
3. Lu, L. *et al.* Neutralization of severe acute respiratory syndrome coronavirus 2 omicron variant by sera from BNT162b2 or Coro-naVac vaccine recipients. *Clin. Infect. Dis.*https://doi.org/10.1093/cid/ciab1041 *(2021)*.
4. Zhang, L. *et al.* The significant immune escape of pseudotyped SARS-CoV-2 variant Omicron. *Emerg. Microbes Infect.* **11**, 1–5. https://doi.org/10.1080/22221751.2021.2017757 (2021).
5. Gozzi, N. *et al.* Preliminary modeling estimates of the relative transmissibility and immune escape of the Omicron SARS-CoV-2 variant of concern in South Africa. *medRxiv*https://doi.org/10.1101/2022.01.04.22268721 *(2022)*.
6. Viana, R. *et al.* Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature* **603**, 679–686. https://doi.org/10.1038/s41586-022-04411-y (2022).
7. Kermack, W. O. & McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. A Contain. Pap. Math. Phys. Character* **115**, 700–721. https://doi.org/10.1098/rspa.1927.0118 (1927).
8. van Wees, J.-D. *et al.* SIR model for assessing the impact of the advent of Omicron and mitigating measures on infection pressure and hospitalization needs. *medRxiv*https://doi.org/10.1101/2021.12.25.21268394 *(2021)*.
9. Götz, T. Analysis of an SIR-model with global and local infections. arXiv (2022).
10. Ghosh, K. & Ghosh, A. K. Study of COVID-19 epidemiological evolution in India with a multi-wave SIR model. *Nonlinear Dyn.*https://doi.org/10.1007/s11071-022-07471-x *(2022)*.
11. Khan, M. A. & Atangana, A. Mathematical modeling and analysis of COVID-19: A study of new variant Omicron. *Physica A Stat.l Mech. Appl.* **599**, 127452. https://doi.org/10.1016/j.physa.2022.127452 (2022).
12. Dehning, J. *et al.* Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* **369**, eabb9789. https://doi.org/10.1126/science.abb9789 (2020).
13. Kim, Y.-J., Seo, M. H. & Yeom, H.-E. Estimating a breakpoint in the pattern of spread of COVID-19 in South Korea. *Int. J. Infect. Dis.* **97**, 360–364. https://doi.org/10.1016/j.ijid.2020.06.055 (2020).
14. Dass, S. C. *et al.* A data driven change-point epidemic model for assessing the impact of large gathering and subsequent movement control order on COVID-19 spread in Malaysia. *PLoS One* **16**, e0252136. https://doi.org/10.1371/journal.pone.0252136 (2021).
15. Jiang, S., Zhou, Q., Zhan, X. & Li, Q. BayesSMILES: Bayesian segmentation modeling for longitudinal epidemiological studies. *J. Data Sci.* **19**, 365–389. https://doi.org/10.6339/21-jds1009 (2021).
16. Perakis, G., Singhvi, D., Lami, O. S. & Thayaparan, L. COVID-19: A multiwave SIR-based model for learning waves. *Prod. Oper. Manag.*https://doi.org/10.1111/poms.13681 *(2022)*.
17. Erdös, P. & Rényi, A. On random graphs I. *Publ. Math. Debrecen* **6**, 290 (1959).
18. Gilbert, E. N. Random graphs. *Ann. Math. Stat.* **30**, 1141–1144. https://doi.org/10.1214/aoms/1177706098 (1959).
19. Hubert, L. & Arabie, P. Comparing partitions. *J. Classif.* **2**, 193–218. https://doi.org/10.1007/bf01908075 (1985).
20. Steuer, R., Kurths, J., Daub, C. O., Weise, J. & Selbig, J. The mutual information: Detecting and evaluating dependencies between variables. *Bioinformatics* **18**, S231–S240. https://doi.org/10.1093/bioinformatics/18.suppl_2.s231 (2002).
21. Polunchenko, A. S. & Tartakovsky, A. G. State-of-the-art in sequential change-point detection. *Methodol. Comput. Appl. Probab.* **14**, 649–684. https://doi.org/10.1007/s11009-011-9256-5 (2011).

Acknowledgements

Author contributions

Both J.G. and G.Y. wrote and revised the manuscript, and J.G. conducted all experiments and drew all plots using R-3.6.1.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-25473-y.

Correspondence and requests for materials should be addressed to G.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.