

Bayesian Inference for the Transmission Dynamics of Infectious Diseases with a SEIR model

PHP2530

Jialin Liu

May 17, 2024

1. Introduction

Throughout human history, infectious diseases have consistently impacted people's lives and posed significant public health risks, often resulting in severe fatalities. The World Health Organization declared COVID-19, caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), a global pandemic on March 11, 2020 (Cucinotta and Vanelli, 2020). By September 3, 2020, over 26 million cases of COVID-19 were confirmed worldwide, with exceeding 864,000 deaths (Zhou and Ji, 2020). But global pandemics are not recent phenomena - they have occurred throughout recorded history, such as Ebola, plague, and Spanish Flu which resulted in severe financial and psychological disruption (Sampath et al., 2021).

To control outbreaks, countries have started to implement non-pharmaceutical containment, from case isolation and adjustments to educational models to the restrictions of mass gatherings and national lockdowns, to break the transmission path of contagion and thus mitigate the virus's spread. Additionally, the general public's adherence to social distancing and mask-wearing have played significant roles in curbing the spread of COVID-19.

Controlling any growing pandemic and preventing the emergence of a potential new one requires a thorough understanding of the transmission mechanisms behind the emergence of these pandemics. During the long history when people have understood infectious diseases and led to scientific breakthroughs, epidemiological theories, also known as mathematical models of epidemiology were developed (Wu et al., 2022). To mathematically model human-to-human infectious diseases, Kermack and McKendrick (1927) proposed the classic Susceptible-Infected-Recovered (SIR) model as a fundamental compartmental model of stochastic epidemics to explain the number of infected patients in a closed population over time while studying the black death epidemic in London (Wu et al., 2022; Kermack and McKendrick, 1927).

Compartmental modeling approaches involve dividing a target population into different compartments, each representing a possible stage of the disease. The transmission dynamics are characterized by a set of mathematical equations, which are governed by parameters

that describe the flow of individuals through these compartments (Zhou and Ji, 2020). Researchers have spent considerable time evaluating the threats of large-scale infections to society, utilizing the compartmental model to study transmission of the disease, and informing policymakers regarding the effects of interventions. From the perspective of public health policymakers, modeling infectious diseases and inferring parameters of compartmental models play a pivotal role in studying the speed of disease spread. These models help determine whether the inferred parameters accurately reflect the underlying trajectories of disease transmission. For instance, the recovery rate is the parameter that demonstrates the effectiveness of vaccination strategies and transmissibility of virus variants; the spreading rate informs not only immediate responses to the effect of interventions but also preparation for forecasting future infections.

However, missing recorded infections that occurred prior to the first observed data in the initial phase of an outbreak, regardless of prior epidemic duration, poses a major uncertainty in calibrating models and characterizing epidemiological parameters. Due to the lack of observations of infections that occurred before the first recorded data—potentially caused by ineffective containment, limited testing capacity, under-ascertainment, and under-reporting—it is challenging to obtain accurate estimates of key metrics such as the effective reproduction number. Consequently, these estimates are often overestimated at the beginning of a time series.

To address these complexities, we utilize the Bayesian inference framework in a simulated stochastic epidemic to estimate the epidemiological model parameters and link these solutions to another key epidemiological parameter that is widely used to describe the transmissibility of infectious diseases: the basic reproduction number R_0 . In order to make the transmission of infectious diseases as realistic as possible, we extend the classical susceptible-infected-recovered (SIR) model that many existing works focused on to susceptible-exposed-infected-recovered (SEIR) taking into account the complexities of the true disease spread. In the previous work, estimations of compartmental model parameters were conducted. In this paper, we not only perform Bayesian inference for the model's parameters but also for the key parameter R_0 in a comparative analysis of the simulation study. The Bayesian framework quantifies uncertainty for unknown parameters by incorporating information from informative priors and the observed data under a specified likelihood function. Secondly, we perform Bayesian inference on predicted incidence cases and infections, and compare the estimated trajectories with the actual epidemic curves to assess the reliability of the inference.

2. Data Structure and Notation

We introduce a stochastic Susceptible-Exposed-Infected-Recovered (SEIR) model used throughout this paper, with each compartment representing a stage in the projection of the infectious disease.

The SEIR model divides the total population into four categories at time t ($t \geq 0$): susceptible individuals, $S(t)$, representing the number of people who are susceptible to the disease

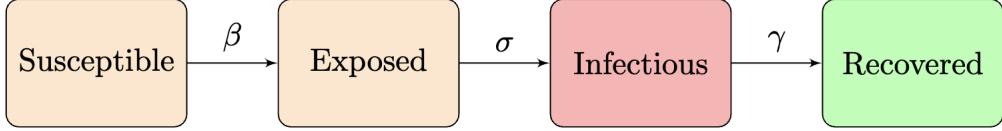


Figure 1: The Deterministic Susceptible (S), Exposed (E), Infected (I) and Recovered (R) model

but not infected at time t or might be infected by the specified type of disease; exposed individuals, $E(t)$, who are exposed to the virus or have been infected but not yet infected to others; infected individuals, $I(t)$, who become infectious and can spread the disease in a wider range through close contact with susceptible populations; and recovered individuals, $R(t)$, referring to those who have recovered and developed immunity and cannot be re-infected. As shown in the transmission mechanism (Figure 1), we approximate changes in these populations at each timestep in terms of the following ordinary differential equations:

$$\begin{aligned} \frac{dS(t)}{dt} &= -\frac{\beta S(t)I(t)}{N}, \\ \frac{dE(t)}{dt} &= \frac{\beta S(t)I(t)}{N} - \sigma E(t), \\ \frac{dI(t)}{dt} &= \sigma E(t) - \gamma I(t), \\ \frac{dR(t)}{dt} &= \gamma I(t). \end{aligned}$$

Each ordinary differential equation in the SEIR model represents a state of the trajectory at time point t , with the model's accuracy depending on the resolution of this system of equations. Here, the parameter β controls the rate of transmission or the spreading rate between susceptible and infectious individuals, which can be expressed as a function of the basic reproduction number R_0 and the infectious period $R_0 \cdot \gamma$. σ is the rate at which exposed individuals become infectious, which is the inverse of the incubation (or latent) period. Recovery rate, γ , is determined by $\gamma = \frac{1}{D}$, where D is the average duration of infectiousness.

This system of equations assumes that the population size is fixed, in other words, any births or deaths caused by other diseases will not be taken into account, and that the disease is transmitted from the susceptible individuals to the infectious individuals, called parental infections. Thus, the total population at time t is summing over all four compartments $N = S(t) + E(t) + I(t) + R(t)$, with $\frac{dN}{dt} = \frac{dS}{dt} + \frac{dE}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$.

3. Methodology

3.1 Statistical Model

Bayesian problems begin with the standard Baye's theorem: $P(\theta|y) = \frac{P(y|\theta)P(\theta)}{P(y)}$. Driven by this, all Bayesian inference follows a consistent framework to model the uncertainty

surrounding parameters of interest, denoted as Ω . This involves deriving posterior functions for Ω which updates our prior belief based on a data-driven likelihood function.

$$\begin{aligned} P(\Omega|D) &= \frac{P(D|\Omega)P(\Omega)}{P(D)} \\ &\propto P(D|\Omega)P(\Omega) \end{aligned}$$

where $P(\Omega)$ follows prior beliefs and $P(D|\Omega)$ is the likelihood, or sampling distribution. The likelihood function describes how the data is generated; the priors allow us to be equipped with existing beliefs before observing the data.

3.2 Sampling distribution

Each compartment has a unique solution at a specific time point, for instance, $I(t)$ in the third ordinary differential equation given all parameters and initial values of each compartment $S(t = 0) > 0; E(t = 0) \geq 0; I(t = 0) \geq 0; R \geq 0$. In realistic cases, we are provided with the observed number of new cases of the diseases reported during a given period, denoted as $Inc(t)$. While the “I” compartment monitors all individuals who are infectious to the susceptible population, $Inc(t)$ captures the newly diagnosed cases within a specific timeframe (Grinsztajn et al., 2021). Constructing the link between current infectious individuals and incident cases, researchers originally often fit the incidence cases into a Poisson likelihood, with the parameter λ which represents the mean rate of the infections I (Cracknell Daniels). However, accounting for the over-dispersion of data with a larger variance than the mean, we suppose the number of incidences follows a negative binomial distribution with $I(t)$ as the number of “successes” until t and ϕ as the “success” probability as follows:

$$Inc(t) \sim NegBinom(I(t), \phi)$$

In particular, the probability mass function of the negative binomial distribution is

$$P(Inc(t) = inc(t)) = \binom{inc(t) + I(t) - 1}{inc(t)} (1 - \phi)^{inc(t)} \phi^{I(t)}$$

3.3 Prior distribution

As an emerging disease and constant exploration of COVID-19 over several years, we have informative domain knowledge about the prior distributions over each of the three parameters $\omega = (\beta, \sigma, \gamma)$. For instance, our prior belief about infectious rate $\gamma \sim N(0.4, 0.5)$ (reported average infectious period of Omicron is around 3) Grinsztajn et al. (2021). One notable thing is that these parameters have to be positive due to non-negative incubation and infectious period D , thus we should truncate the sampling values for those parameters at 0 to avoid any nonsensical inference.

We also assume the prior $\phi \sim Exp(0.5)$ which allows for a wider range of values reflecting the uncertainty. Table 1 lists all prior settings for the model parameter, employing weakly informative priors derived from previous research (Grinsztajn et al., 2021).

Table 1: Prior Distributions for Model Parameters

Parameter	Prior Distribution
β	Normal(2,1)
σ	Normal(0.7, 0.5)
γ	Normal(0.4, 0.5)
ϕ	Expo(0.5)

3.4 Derived quantities and Predicted Cases

In addition to the estimation of compartmental model parameters, understanding the intensity of disease spread over the course of an epidemic is critical for policymakers. The basic reproduction number, R_0 , and effective reproduction number, R_t , are key epidemiological parameters used to describe the transmissibility of infectious diseases. In this paper, we pay extra attention to estimating the basic reproduction number R_0 that measures the average number of secondary infections transmitted from a primary case by a cohort of fully susceptible individuals over a entire period ($0 < t \leq T$) (Sharma et al., 2023; Gostic et al., 2020; Brockhaus et al., 2023). If $R_0 > 1$, incident cases are increasing, potentially leading to an outbreak. If $R_t < 1$, incident cases will decline and the outbreak will eventually disappear as infections die out, and if $R_t = 1$, incident cases will remain stable over time (Barratt and Kirwan, 2018). Generally speaking, R_0 estimates measure epidemic growth and can be used to assess the effectiveness of intervention policies, project future infections, and measure the final size of an epidemic (Gostic et al., 2020; Brizzi et al., 2022). Considering the statistical complexity in accurately estimating R_0 , we conduct a comparative analysis between the standard estimation called EpiEstim, developed by Cori et al.(2013), and Bayesian MCMC estimation.

We are interested in predicting new cases given the posterior distribution of ω to assess whether the predictions fall into a plausible range of true observed incidence cases. The posterior distribution of predicted new incident cases follows $P(Inc_{pred}|Inc) = \int P(Inc_{pred}|\omega)P(\omega|Inc)d\omega$ (Grinsztajn et al., 2021).

4. Computation

4.1 MCMC: Hamiltonian Monte Carlo (a.k.a. Hybrid Monte Carlo)

Hamiltonian Monte Carlo (HMC), an Markov chain Monte Carlo (MCMC) method proposed by Duane et al. (1987); Neal (1996), employs the Hamiltonian physical dynamic system. It calculates the next state using gradient information from the current state and ultimately moving toward a high-probability next state Mbuvha and Marwala (2020). As an advanced version of the Metropolis-Hastings algorithm, this gradient-based and adaptive method excels in high-dimensional problems by better exploring the target distribution

and achieving faster convergence. The HMC algorithm automatically tunes some critical parameters using the No-U-Turn Samplers (NUTS), such as the step size between successive samples. This feature is similar to the Metropolis-Hastings (MH) algorithm, where a smaller step size relates to a higher correlation between successive samples.

4.2 ODE-based Models: Runge-Kutta Method

In this context, it is challenging to determine the posterior distribution up to a proportionality constant using ordinary differential equations (ODEs) in closed form. Stan enables us to code the infectious disease model incorporating ODEs and implements the HMC algorithm, allowing us to sample from our posterior distribution effectively. We can build an ordinary function block for a system of ODEs in Stan by `real[] seir(real time, real[] state, real[] theta, real[] x_r, int[] x_i)` to specify the time, each compartmental state in the ODE system, the parameters we're interested in sampling, etc. To obtain a better approximation, the Runge-Kutta method calculates 4 different slopes within the time interval $(t, \Delta t)$ and results in a weighted average velocity of four slopes obtained at every time step. Stan has a built-in ODE integrator, `integrate_ode_rk45`, to solve ODEs we define in the functions block.

4.3 ODE-based Models: Euler's Method

The simplest algorithm for numerical integration used to calculate the equation of motion is the Euler's Method. It simply calculates the slope at a particular time point and predicts the next value at $t + \Delta t$. This way is too simple to obtain an accurate estimate of the next velocity at $t + \Delta t$, since the result is fully dependent on the length of time interval Δt we use to predict. The shorter time intervals we define, the closer estimates the prediction has. In this paper, we utilize two methods to solve our ODEs and compare which method is more efficient than the other in terms of how many effective draws each method generates.

5. Simulation Study

We generate 500 epidemic scenarios using various parameter configurations. We consider the total population of 10^6 individuals, where the epidemic is seeded by 5 infections $I_0 = 5$ and lasts 104 days ($T = 104$). We assume that the initial values for both the exposed and recovered groups are 0 $E_0 = 0, R_0 = 0$. The latent ($\frac{1}{\sigma}$) and infectious ($\frac{1}{\gamma}$) periods were set to 3.5 and 3 days respectively, consistent with the progression of COVID-19, resulting in a mean generation interval of 6.5 days.

Moreover, we introduce a constant reporting rate $\rho = 1$. This rate reflects the probability that a case is reported and captured in our data. The daily case counts are then generated as binomial distributions of the true incidence, with success probability equal to the reporting rate ρ .

We summarize simulated epidemic trajectories under four different basic reproduction number (R_0) scenarios—1.5, 2, 2.5, and 3—in Figure 2. Each light blue line represents a sim-

ulated scenario, while the dark blue line indicates the average incidence cases across all simulations over time. Under an R_0 of 1.5, the epidemic curve rises slowly and does not reach a peak within the explored timeframe, in contrast to higher R_0 scenarios. As R_0 increases to 2.5 or 3, the incidence sharply spikes, reaching a peak quickly. After reaching the peak, the number of incident infections declines rapidly due to the depletion of susceptible individuals.

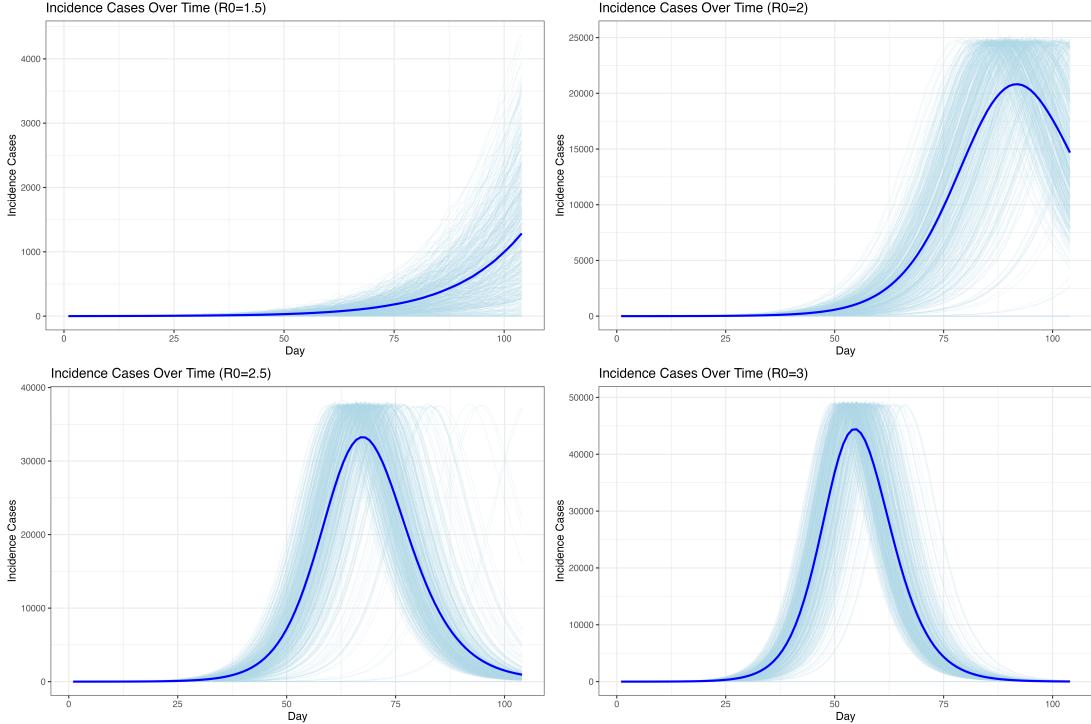
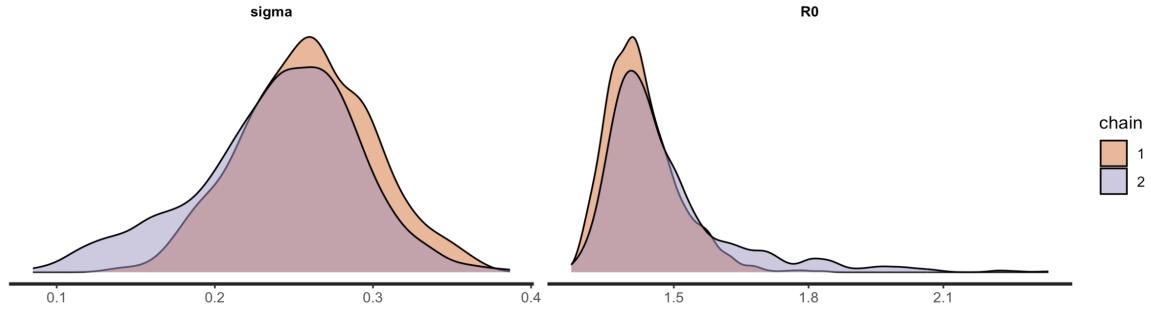


Figure 2: Incidence plots over time

After the model using Runge-Kutta method is built in `Stan` code, we pass data from `R` and fit the model by running two independent Markov Chains with a burn-in period that lasts 1000 iterations and another 4000 post-burn-in draws in the sampling phase. Given that our parameters of interest, β (transmission rate) and γ , are linked to R_0 through the expression $R_0 = \frac{\beta}{\gamma}$, we analyze the marginal posterior density for R_0 in one simulation case to determine if the posterior mean converges to the true value of R_0 . Specifically, when R_0 is set to 1.5 during the simulated data generation, the posterior distribution is centered around 1.3, suggesting that the model accurately captures the true transmission dynamics and provides a reliable estimation of the basic reproduction number, effectively addressing the complexity discussed previously. Similarly, for $R_0 = 2$, the marginal posterior density for R_0 averages at 1.85 across two Markov Chains, closely approximating the actual setting of R_0 .

Regarding the marginal posterior density of σ , our model, which employs the Runge-Kutta Method, also demonstrates proficiency in inferring model parameters in both simulation

Marginal Posterior Densities when R_0 of 1.5



Marginal Posterior Densities when R_0 of 2

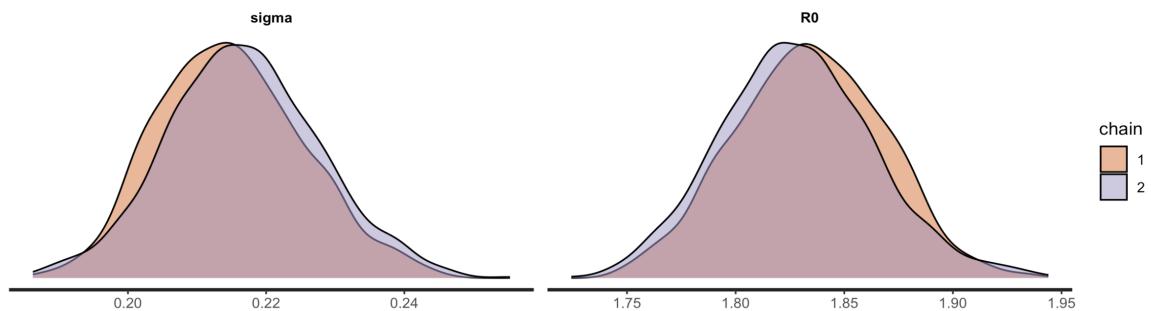


Figure 3: Marginal Posterior Distributions for σ and R_0

scenarios. Our initial parameter setting posits an average incubation period of 3.5 days, corresponding to $\sigma = 0.28$. The marginal posterior distributions, with mean values between 0.21 and 0.26, encompass the true parameter values. Although not always precisely centered, the convergence results are deemed reliable when the procedure is replicated across multiple simulation cases, ultimately leading to more accurate inference of the true parameters.

Once the model is fitted, we aim to assess its efficacy in predicting future behaviors, beyond merely checking parameters. Figure 4 illustrates the predicted incidence cases under various R_0 scenarios compared to the observed epidemic curves. If consistency between the predicted and observed trajectories is achieved, we can conclude that the model accurately captures the transmission dynamics of the simulated outbreak, even with varying model parameters. In Figure 4, the posterior mean of the predicted incidence cases aligns closely with the observed new cases within 90% credible intervals, particularly in scenarios with lower R_0 values. Although the predicted trajectories of new cases show some deviation from the observed data, the model remains reliable. This reliability is supported by the fact that these plots are posterior predictions from a single simulation scenario, and they correspond well with incidence curves averaged across all 500 simulations shown in Figure 2.

Building on the approach used in predicting incidence cases, we aim to assess the model's predictive performance concerning the number of infections at specific time points. Across

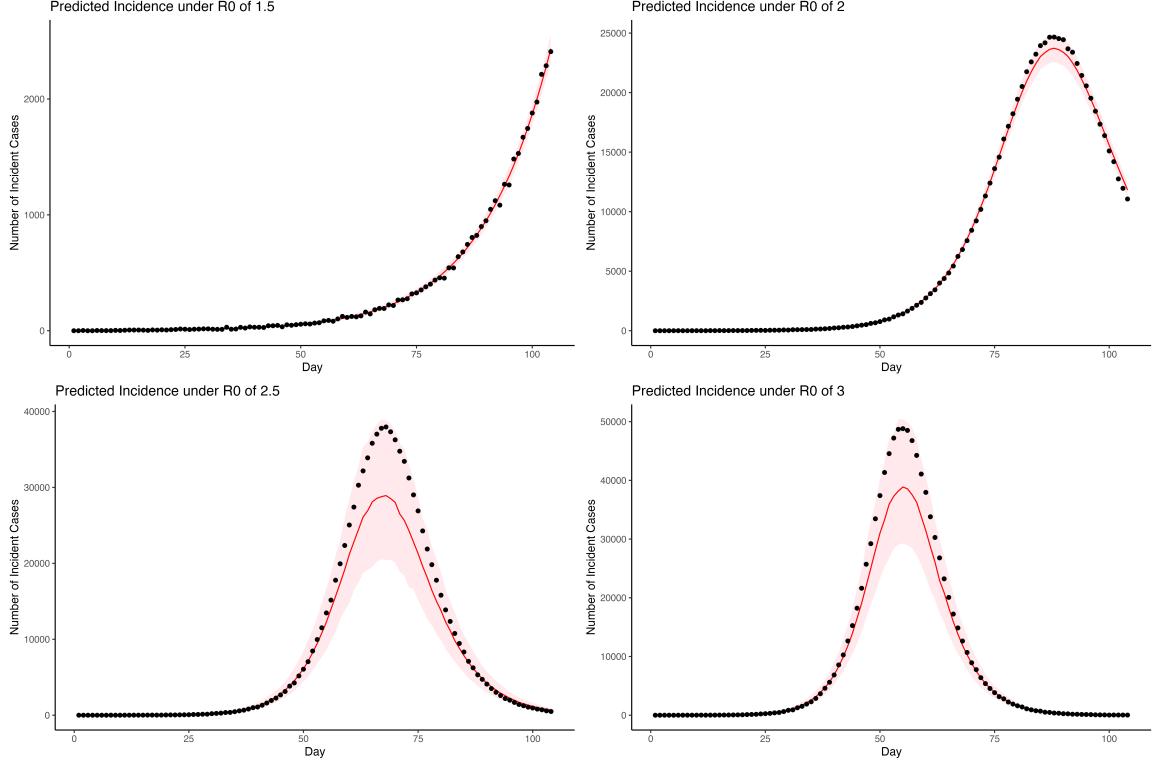


Figure 4: Plot of Predicted Incident Cases over Time

all simulation scenarios, our model demonstrates an ability to accurately quantify uncertainty and predict infections, maintaining only a small discrepancy with the observed data (Figure 5). It is unrealistic and unnecessary for defined models to perfectly align with observed epidemic curves; perfect alignment is not the goal of estimations. More crucially, the model provides valuable insights into the transmission dynamics, particularly the timing and speed of reaching the peak of infections and the point at which infections die out or susceptible depletion reach.

We define bias as the difference between the estimated value and the true R_0 . In our analysis, we compare estimations generated by Bayesian MCMC inference with those obtained from the popular EpiEstim method developed by Cori et al. (2013). These methods measure the intensity of infectious disease spread and compute the basic (and effective) reproduction numbers to inform public health professionals. Across all R_0 scenarios, the Bayesian MCMC method consistently shows lower absolute bias and percentage of bias compared to EpiEstim. This confirms the model's ability to quantify uncertainty about missing cases prior to the first observed data at the initial phase of an epidemic, thus eliminating bias that could lead to an overestimation of the basic reproduction number at the beginning stage.

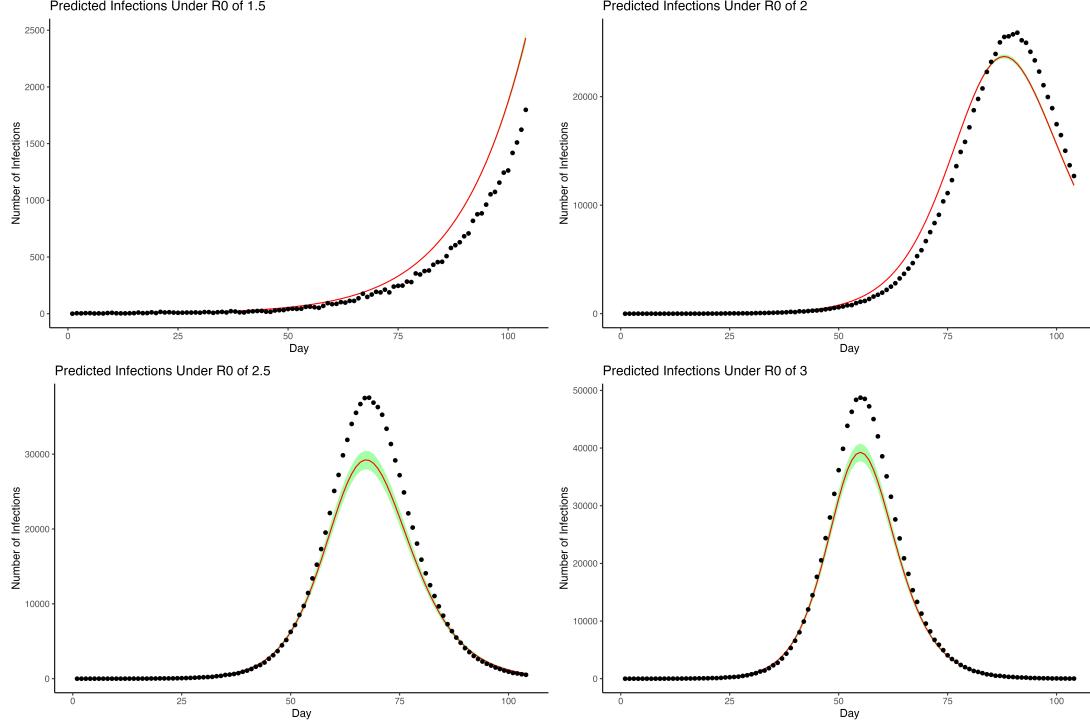


Figure 5: Plot of Predicted Infections over Time

Table 2: Absolute and percentage forms of bias and standard deviation of R_0 Estimators under different R_0 scenarios

R_0 Value	Methods	Bias	SD	% of Bias
1.5	\hat{R}_0 (Bayes MCMC)	0.05	0.02	3
	EpiEstim	0.53	0.11	35
2.0	\hat{R}_0 (Bayes MCMC)	0.17	0.03	9
	EpiEstim	0.43	0.04	21
2.5	\hat{R}_0 (Bayes MCMC)	0.4	0.05	19
	EpiEstim	0.45	0.23	20
3.0	\hat{R}_0 (Bayes MCMC)	0.5	0.07	23
	EpiEstim	0.76	0.39	53

Based on Table 2, Bayesian MCMC inference for R_0 strongly outperforms EpiEstim estimates in terms of lower standard deviation, suggesting that Bayesian MCMC estimates are more precise and less variable. Our model, which employs the Runge-Kutta solver, effectively addresses bias resulting from unobserved parent cases. Specifically, our proposed model tends to provide estimates closer to the true R_0 value, with generally smaller biases

and standard deviations compared to the EpiEstim method.

However, the bias for Bayesian MCMC estimations appears to increase under high R_0 settings. Additionally, the model with the Runge-Kutta method shows improvement over the Euler's method, achieving a larger effective sample size and \hat{R} being close to 1. Although the bias and standard deviation trends suggest that the proposed inference method may not always provide a more accurate estimate of R_0 as R_0 increases, the Bayesian MCMC method consistently outperforms other methods in terms of bias and standard deviation, even in higher R_0 scenarios.

Appendix A.

Software in the form of **R** and **Stan** code, together with data used in the simulation study and results including figures and tables is available at <https://github.com/Jialin-pluszero/Bayesian-Analysis-of-Infectious-Diseases-Model>.

References

- Helen Barratt and Maria Kirwan. Epidemic theory (effective & basic reproduction numbers, epidemic thresholds) & techniques for analysis of infectious disease data (construction & use of epidemic curves, generation numbers, exceptional reporting & identification of significant clusters), 2018. URL <https://www.healthknowledge.org.uk/public-health-textbook/research-methods/1a-epidemiology/epidemic-theory>.
- Andrea Brizzi, Megan O'Driscoll, and Ilaria Dorigatti. Refining reproduction number estimates to account for unobserved generations of infection in emerging epidemics. *Clinical Infectious Diseases*, 75(1):e114–e121, February 2022. ISSN 1537-6591. doi: 10.1093/cid/ciac138. URL <http://dx.doi.org/10.1093/cid/ciac138>.
- Elisabeth K. Brockhaus, Daniel Wolfram, Tanja Stadler, Michael Osthege, Tanmay Mitra, Jonas M. Littek, Ekaterina Krymova, Anna J. Klesen, Jana S. Huisman, Stefan Heyder, Laura M. Helleckes, Matthias an der Heiden, Sebastian Funk, Sam Abbott, and Johannes Bracher. Why are different estimates of the effective reproductive number so different? a case study on covid-19 in germany. *PLOS Computational Biology*, 19(11):e1011653, November 2023. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1011653. URL <http://dx.doi.org/10.1371/journal.pcbi.1011653>.
- Bethan Naomi Cracknell Daniels. Bayesian inference for SARS-CoV-2 transmission modelling, using Stan. URL https://github.com/ImperialCollegeLondon/ReCoDE_IDMS.
- Domenico Cucinotta and Maurizio Vanelli. Who declares covid-19 a pandemic. *Acta Bio Medica Atenei Parmensis*, 91(1):157–160, March 2020. ISSN 25316745, 03924203. doi: 10.23750/abm.v91i1.9397. URL <https://doi.org/10.23750/abm.v91i1.9397>.
- Simon Duane, A.D. Kennedy, Brian J. Pendleton, and Duncan Roweth. Hybrid monte carlo. *Physics Letters B*, 195(2):216–222, September 1987. ISSN 0370-2693. doi: 10.1016/0370-2693(87)91197-x. URL [http://dx.doi.org/10.1016/0370-2693\(87\)91197-X](http://dx.doi.org/10.1016/0370-2693(87)91197-X).
- Katelyn M. Gostic, Lauren McGough, Edward B. Baskerville, Sam Abbott, Keya Joshi, Christine Tedijanto, Rebecca Kahn, Rene Niehus, James A. Hay, Pablo M. De Salazar, and et al. Practical considerations for measuring the effective reproductive number, rt. *PLOS Computational Biology*, 16(12), 2020. doi: 10.1371/journal.pcbi.1008409.
- Léo Grinsztajn, Elizaveta Semenova, Charles C. Margossian, and Julien Riou. Bayesian workflow for disease transmission modeling in stan. *Statistics in Medicine*, 40(27): 6209–6234, September 2021. ISSN 1097-0258. doi: 10.1002/sim.9164. URL <http://dx.doi.org/10.1002/sim.9164>.
- William Ogilvy Kermack and Anderson G McKendrick. 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(772):700–721, 1927.
- Rendani Mbuvha and Tshilidzi Marwala. Bayesian inference of covid-19 spreading rates in south africa. *PLOS ONE*, 15(8):e0237126, August 2020. ISSN 1932-6203. doi: 10.1371/journal.pone.0237126. URL <http://dx.doi.org/10.1371/journal.pone.0237126>.

Radford M. Neal. *Monte Carlo Implementation*, page 55–98. Springer New York, 1996. ISBN 9781461207450. doi: 10.1007/978-1-4612-0745-0_3. URL http://dx.doi.org/10.1007/978-1-4612-0745-0_3.

Shrikanth Sampath, Anwar Khedr, Shahraz Qamar, Aysun Tekin, Romil Singh, Ronya Green, and Rahul Kashyap. Pandemics throughout the history. *Cureus*, September 2021. ISSN 2168-8184. doi: 10.7759/cureus.18136. URL <http://dx.doi.org/10.7759/cureus.18136>.

Vijay Sharma, Rajnish Sharma, and Balbir B. Singh. Etymologia: Reproduction number. *Emerging Infectious Diseases*, 29(8), August 2023. ISSN 1080-6059. doi: 10.3201/eid2908.221445. URL <http://dx.doi.org/10.3201/eid2908.221445>.

Yichun Wu, Yaqi Sun, and Mugang Lin. Sqeir: An epidemic virus spread analysis and prediction model. *Computers and Electrical Engineering*, 102:108230, September 2022. ISSN 0045-7906. doi: 10.1016/j.compeleceng.2022.108230. URL <http://dx.doi.org/10.1016/j.compeleceng.2022.108230>.

Tianjian Zhou and Yuan Ji. Semiparametric bayesian inference for the transmission dynamics of covid-19 with a state-space model. *Contemporary Clinical Trials*, 97:106146, October 2020. ISSN 1551-7144. doi: 10.1016/j.cct.2020.106146. URL <http://dx.doi.org/10.1016/j.cct.2020.106146>.