Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2026, 2026:117
https://doi.org/10.28919/cmbn/10288
ISSN: 2052-2541

BAYESIAN PARAMETER ESTIMATION OF SIR AND SIRD MODELS FOR COVID-19 IN INDONESIA USING METROPOLIS–HASTINGS AND HAMILTONIAN MONTE CARLO

XAVIERA NAFTALI, BENNY YONG*

Center for Mathematics and Society, Faculty of Science, Parahyangan Catholic University, Ciumbuleuit 94, Bandung, 40141, Indonesia

Copyright © 2026 the author(s). This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract. Accurate parameter estimation is essential for reliable epidemic modeling and forecasting. This study presents a Bayesian framework for estimating parameters of the SIR and SIRD epidemiological models using two Markov Chain Monte Carlo (MCMC) methods: Metropolis–Hastings (MH) and Hamiltonian Monte Carlo (HMC). Daily COVID-19 surveillance data from Indonesia spanning January to April 2022 are used to estimate the transmission rate (β), recovery rate (γ), and death rate (κ). Model performance is evaluated through convergence diagnostics (Gelman–Rubin statistic), sampling efficiency (effective sample size, ESS), and predictive accuracy (root mean squared error, RMSE). Results show that the MH algorithm exhibits severe convergence failure, with Gelman–Rubin statistics reaching up to 66.63 and ESS values as low as 7.75, indicating highly autocorrelated and inefficient sampling. In contrast, HMC achieves perfect convergence $(\widehat{R} = 1)$ across all parameters with ESS values ranging from about 790 to more than $1,600$, reflecting substantially more efficient posterior exploration. Although both algorithms yield high RMSE under the beta prior for β, adopting a lognormal prior reduces the RMSE dramatically—from over one million to approximately $67,000$ for both the SIR and SIRD models when combined with HMC. Among all configurations tested, the SIRD model estimated via HMC with a lognormal prior for β achieves the lowest RMSE $(67,255.57)$, demonstrating superior predictive performance. These findings suggest that the choice of both the sampling algorithm and the prior distribution critically determines estimation quality

*Corresponding author
E-mail address: benny_y@unpar.ac.id
Received September 15, 2026

in compartmental epidemic models, with HMC and lognormal priors offering a robust and efficient framework for Bayesian epidemiological inference.

Keywords: Bayesian inference; SIR model; SIRD model; Hamiltonian Monte Carlo; Metropolis–Hastings

2020 AMS Subject Classification: 62F15; 65C05; 92D30

1. INTRODUCTION

The COVID-19 pandemic has created an urgent need for reliable models that can describe and predict the spread of infectious diseases. Mathematical epidemiological models play a crucial role in understanding disease transmission dynamics and informing public health decision-making. During the pandemic, governments implemented various policies such as social restrictions and mass vaccination to control the virus's transmission.

Among the most widely used epidemiological models is the Susceptible–Infected–Recovered (SIR) model, which divides the population into three compartments representing individuals who are susceptible to infection, currently infected, and recovered [1]. An extension of this model, the SIRD model, adds a fourth compartment to represent deaths caused by the disease, allowing the death rate to be incorporated alongside transmission and recovery rates. Accurate estimation of these parameters is essential for understanding the course of an epidemic and forecasting its peak and overall impact.

Bayesian inference provides a principled framework for estimating epidemiological parameters by combining prior knowledge with observed data. In this approach, model parameters are treated as random variables with prior distributions that are updated using Bayes' theorem to obtain posterior distributions. However, posterior distributions in epidemic models are often analytically intractable, requiring numerical sampling techniques such as Markov Chain Monte Carlo (MCMC). Traditional algorithms such as Metropolis–Hastings may suffer from slow convergence and high autocorrelation, particularly for nonlinear dynamical systems. This limitation motivates the use of gradient-based methods such as Hamiltonian Monte Carlo, which can explore posterior distributions more efficiently [2].

Previous studies have explored Bayesian methods for parameter estimation in epidemic models. For example, Talawar and Aundhakar [3] used a Metropolis-within-Gibbs algorithm to estimate SIR model parameters from simulated influenza data, while Grinsztajn et al. [4] applied

HMC to real outbreak data from a British boarding school. Recent studies have also investigated mathematical modeling of COVID-19 dynamics and the impact of different virus variants using compartmental epidemic models [5]. However, there is still a lack of research directly comparing the performance of Metropolis–Hastings and Hamiltonian Monte Carlo for estimating parameters in SIR and SIRD models, particularly using real COVID-19 data from Indonesia.

This study aims to address this gap by comparing the accuracy and efficiency of Metropolis–Hastings and HMC in estimating the parameters of SIR and SIRD models for COVID-19 cases in Indonesia. The findings are expected to provide insight into the strengths of each method and help inform future work on epidemic modeling and parameter estimation. The main contributions of this study can be summarized as follows:

1. A comparative Bayesian estimation framework is developed for SIR and SIRD epidemiological models using Metropolis–Hastings and Hamiltonian Monte Carlo methods.

2. Real COVID-19 data from Indonesia are used to evaluate the efficiency and convergence properties of the two MCMC algorithms.

3. The study demonstrates that Hamiltonian Monte Carlo combined with a lognormal prior significantly improves estimation accuracy and sampling efficiency.

2. Research Methodology

This study applies a quantitative approach to estimate parameters in SIR and SIRD epidemiological models using COVID-19 data from Indonesia. It compares the performance of two Bayesian sampling methods, Metropolis–Hastings and Hamiltonian Monte Carlo, focusing on their accuracy and efficiency. This section describes the two sampling algorithms, while the epidemiological models, data, and prior selection are presented in Sections 3–5.

2.1. Metropolis–Hastings. The Metropolis–Hastings algorithm is a Markov Chain Monte Carlo (MCMC) method used to generate samples from a posterior distribution, particularly when the distribution is analytically intractable. In this context, Metropolis–Hastings is used to estimate the parameters of compartmental models such as SIR and SIRD, where direct sampling from the posterior is often not feasible due to complex likelihood functions or high-dimensional parameter spaces.

The algorithm constructs a Markov chain that converges to the target posterior distribution by iteratively proposing candidate samples and accepting or rejecting them based on a probabilistic criterion. The procedure for a single chain is as follows [6]:

1. Select a prior distribution for the parameter θ. Then, randomly draw an initial value θ_c from the prior distribution.
2. Choose a proposal distribution centered at θ_c.
3. Generate a proposed value θ_p from the chosen proposal distribution.
4. Define the proposal distribution centered at θ_p to compute the correction factor.
5. Calculate the correction factor of the proposal distribution using Equation (1).

$$\Lambda = \frac{g(\theta_c|\theta_p)}{g(\theta_p|\theta_c)} \tag{1}$$

6. Compute the posterior probabilities for both θ_c and θ_p, denoted as P(θ_c|data) and P(θ_p|data), respectively.
7. Compute the acceptance ratio using Equation (2).

$$\psi = \min\left(1,\ \frac{P(\theta_p|data)}{P(\theta_c|data)}\Lambda\right) \tag{2}$$

8. Generate a random number r from a uniform distribution on [0,1]. If $\psi \geq r$, accept the proposed value θ_p as the new current value for the next iteration; otherwise, retain the current value θ_c.
9. Repeat steps 2–8 for the desired number of iterations to generate a chain of samples that approximates the target posterior distribution.

Multiple independent chains with different starting points are used in Metropolis–Hastings to explore the parameter space. Convergence diagnostics help verify if the chains have mixed well and reached the target distribution. While the algorithm is versatile, its effectiveness depends on a well-chosen proposal distribution. Poor choices can result in slow convergence, low acceptance rates, or highly correlated samples.

2.2. Hamiltonian Monte Carlo. Hamiltonian Monte Carlo (HMC) is an MCMC method that leverages Hamiltonian mechanics to propose new parameter values. It is particularly efficient in

exploring high-dimensional parameter spaces, which leads to faster convergence and fewer iterations [7]. In contrast to random-walk-based proposals, HMC introduces auxiliary momentum variables and simulates the motion of a particle through parameter space using physical laws. In HMC, the parameter θ is treated as the position of a particle, and an auxiliary variable p represents its momentum. Their joint density is proportional to $e^{-H(\theta,p)}$, where the Hamiltonian function in Equation (3) represents the total energy of the system as the sum of its potential and kinetic energy components:

$$H(\theta,p) = U(\theta) + K(p) \tag{3}$$

The potential energy, denoted as $U(\theta)$ is derived from the negative log posterior, and $K(p)$ is the kinetic energy. Each chain of HMC involves the following steps [8]:

1. Select a prior distribution for the parameter θ, then randomly draw an initial value from the distribution.
2. Sample the momentum variable p from a standard normal distribution.
3. Simulate Hamiltonian dynamics using the leapfrog integrator for L steps with step size ε.
4. Compute the Hamiltonian function before and after the trajectory to evaluate the change in energy.
5. Accept or reject the proposed value using the Metropolis–Hastings acceptance ratio given in Equation (4).

$$\alpha = \min(1, \exp\left(H_{initial} - H_{final}\right)) \tag{4}$$

6. Generate a random number r from a uniform distribution over [0,1]. If $\alpha \geq r$, accept the proposed parameter; otherwise, retain the current parameter value. The accepted value becomes the new current state for the next iteration.
7. Repeat steps 2–6 for the desired number of iterations to generate a chain of samples that approximates the target posterior distribution.

The performance of HMC depends heavily on the step size ε and the number of leapfrog steps L. To improve robustness, this study uses Stan, which automates gradient computation and includes the No-U-Turn Sampler (NUTS). NUTS adaptively selects the number of leapfrog

steps by terminating the simulation when the trajectory begins to reverse direction, avoiding unnecessary computation and improving sampling efficiency [7, 9].

3. Epidemiological Models

Epidemiological models are mathematical tools used to describe the spread of infectious diseases and to analyze the factors that influence disease transmission and control strategies. Two commonly used models in this context are the SIR model and its extended form, the SIRD model.

3.1. SIR Model. The SIR model divides the population into three compartments: susceptible, infected, and recovered/removed. Susceptible individuals, represented by S(t), are those who have not yet been infected and are at risk of contracting the disease. Infected individuals, I(t), are those who are currently carrying the disease and can transmit it to others. Recovered or removed individuals, R(t), are those who are no longer part of the transmission cycle, either because they have recovered with immunity or have been otherwise removed, such as through death [10]. The total population, denoted by N, is assumed to remain constant and equals the sum of all three compartments.

The model includes two main parameters, namely the transmission rate β, which reflects how easily the disease spreads from infected to susceptible individuals, and the recovery rate γ, which indicates how quickly infected individuals recover or are removed from the infectious group. The compartmental representation of the SIR epidemiological model can be seen in Figure 1.

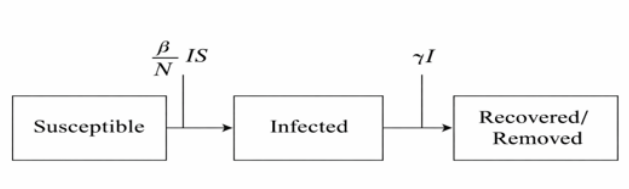


FIGURE 1. Diagram of SIR epidemiological model

These interactions are described using a system of differential equations that track how the number of individuals in each compartment changes over time and can be written as follows:

$$\begin{cases} \dfrac{dS(t)}{dt} = -\dfrac{\beta S(t)I(t)}{N}, \\ \dfrac{dI(t)}{dt} = \dfrac{\beta S(t)I(t)}{N} - \gamma I(t), \\ \dfrac{dR(t)}{dt} = \gamma I(t). \end{cases} \tag{5}$$

Individuals who recover are assumed not to become susceptible again, meaning reinfection is not considered, as shown in Equation (5).

3.2. SIRD Model. The SIRD model extends the SIR model by adding a fourth compartment to represent individuals who have died as a result of the disease. In this model, the population is divided into susceptible, infected, recovered, and deceased individuals, denoted respectively by S(t), I(t), R(t), and D(t). The infected individuals in this model can either recover and move into the recovered compartment or die and move into the deceased compartment. To capture this additional outcome, the model introduces a new parameter, the death rate κ, which represents the per capita rate at which infected individuals die [11]. The compartmental representation of the SIRD epidemiological model can be seen in Figure 2.

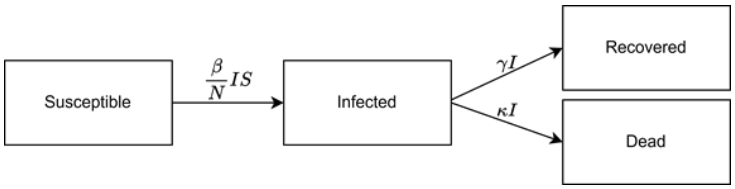


FIGURE 2. Diagram of SIRD epidemiological model

Like the SIR model, the SIRD model assumes a fixed total population, which now includes the deceased. The dynamics of the SIRD model are also described by a system of differential equations that account for the flow of individuals between all four compartments, as shown in Equation (6):

$$\begin{cases} \dfrac{dS(t)}{dt} = -\dfrac{\beta S(t)I(t)}{N}, \\ \dfrac{dI(t)}{dt} = \dfrac{\beta S(t)I(t)}{N} - \gamma I(t) - \kappa I(t), \\ \dfrac{dR(t)}{dt} = \gamma I(t), \\ \dfrac{dD(t)}{dt} = \kappa I(t). \end{cases} \tag{6}$$

By including deaths explicitly, the SIRD model provides a more realistic representation of disease outbreaks that involve significant mortality.

4. Data Description

This study utilizes daily national COVID-19 data from Indonesia covering January 1 to April 30, 2022, as officially reported by the Indonesian COVID-19 Task Force (Satuan Tugas Penanganan COVID-19) [18]. The dataset includes daily counts of new confirmed cases (incidence), recoveries, and deaths, as well as active cases (prevalence), which represent individuals still infected, excluding those who have recovered or died. Table 1 provides a snapshot of the data, which shows the first and last few rows of the daily case records.

TABLE 1. COVID-19 Case Data in Indonesia (January–April 2022)

Date	Recovered	Dead	Incidence	Prevalence
01/01/2022	165	2	274	4399
02/01/2022	190	1	174	4382
03/01/2022	112	5	265	4530
04/01/2022	168	3	299	4658
05/01/2022	180	4	404	4878
⋮	⋮	⋮	⋮	⋮
26/04/2022	4058	30	576	9739
27/04/2022	1178	36	617	9142
28/04/2022	1171	18	412	8365
29/04/2022	572	23	395	8165
30/04/2022	598	17	329	7879

The average and variance of prevalence during this period were 183,043.63 and 36,099,065,094.5, respectively. The variance being significantly higher than the mean indicates overdispersion, thus the data are assumed to follow a negative binomial distribution.

5. Model Fitting

In the Bayesian framework, prior distributions are essential for estimating model parameters. This study uses two models, the SIR model and the SIRD model. The SIR model includes the transmission rate β and recovery rate γ, while the SIRD model adds the death rate κ. To determine suitable priors, daily estimates of these parameters were calculated using processed COVID-19 data. This involved using the cumulative values of the Recovered and Dead columns from Table 1. The estimates for the SIR model parameters were obtained using Equations (7) and (8):

$$\beta_t = \frac{\Delta I(t) + \gamma I(t)}{\frac{S(t)I(t)}{N}} \tag{7}$$

$$\gamma_t = \frac{\Delta R(t)}{I(t)} \tag{8}$$

where $\Delta I(t)$ and $\Delta R(t)$ denote the changes in the number of infected and recovered individuals between day $t+1$ and day t, respectively; that is, $\Delta I(t) = I(t+1) - I(t)$ and $\Delta R(t) = R(t+1) - R(t)$. For the SIRD model, the daily parameter estimates are defined by Equations (9)–(11) below:

$$\beta_t = \frac{\Delta I(t) + \gamma I(t) + \kappa I(t)}{\frac{S(t)I(t)}{N}} \tag{9}$$

$$\gamma_t = \frac{\Delta R(t)}{I(t)} \tag{10}$$

$$\kappa_t = \frac{\Delta D(t)}{I(t)} \tag{11}$$

where $\Delta D(t)$ is the change in the number of deceased individuals between day $t+1$ and day t.

Time-series plots of these daily estimates that are shown in Figure 3 and Figure 4 were used to observe parameter trends and calculate their mean and variance. In the SIR model, β had a mean of 0.0852 and variance of 0.0041, while γ had a mean of 0.0841 and variance of 0.0053. For the SIRD model, β remained the same, γ had a mean of 0.0834 and variance of 0.0052, and κ had a mean of 0.0007 and variance of 0.00000033.

To determine the most suitable prior distributions for each model parameter, a distribution fitting process was carried out. This process involved fitting the daily parameter estimates to

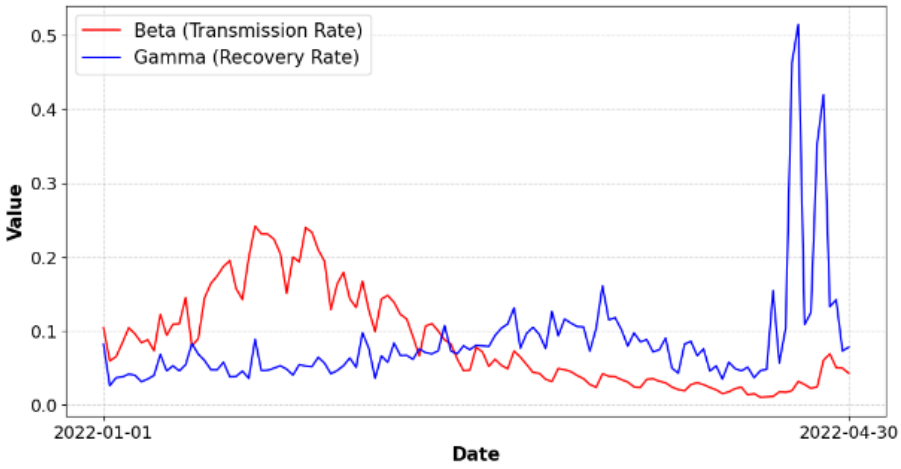


FIGURE 3. Time series of the estimated parameters for the SIR model

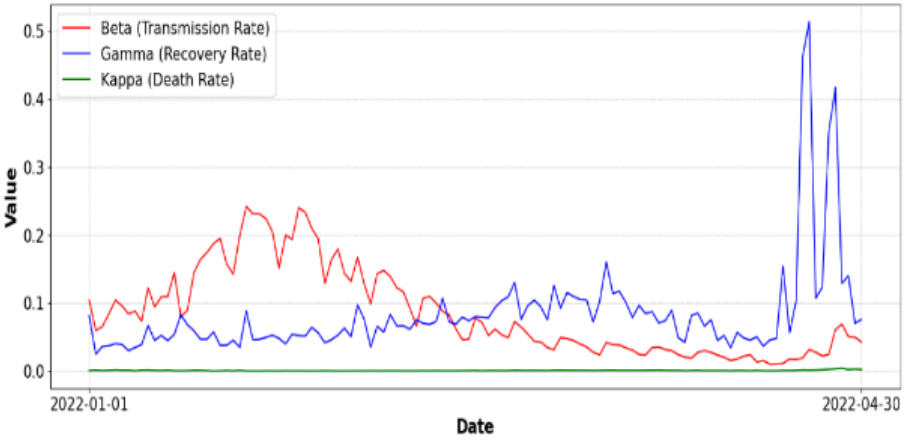
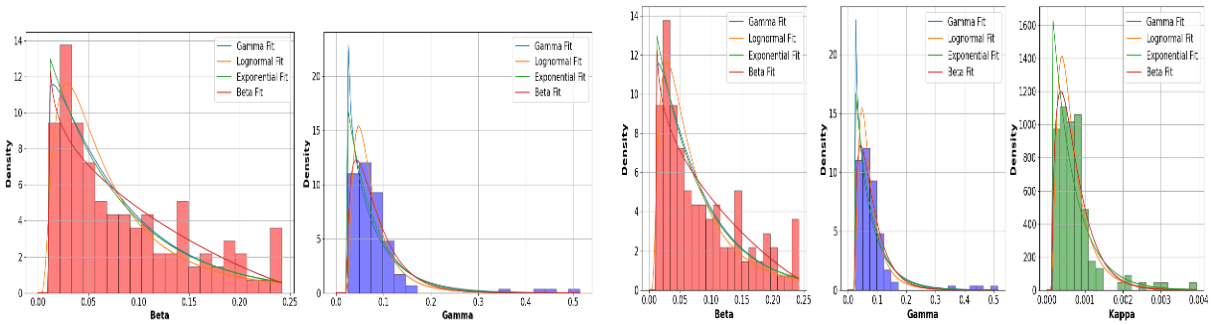


FIGURE 4. Time series of the estimated parameters for the SIRD model

several commonly used continuous probability distributions that are appropriate for strictly positive and potentially skewed data, namely, the gamma, lognormal, exponential, and beta distributions. These distributions were chosen because the estimated parameter values, by nature, cannot be negative and typically do not follow symmetric patterns. The resulting histograms of the daily parameter estimates, overlaid with the fitted distribution curves, are shown in Figure 5.

The estimated parameter distributions were found to be skewed rather than symmetric, making skewed distributions such as gamma, lognormal, exponential, and beta suitable choices. To identify the best-fitting distribution, the Akaike Information Criterion (AIC) was used. AIC evaluates model quality by using the maximum likelihood to estimate how far a model deviates



(a) (b)

FIGURE 5. Fitting distribution plots (a) SIR model, (b) SIRD model

from the true underlying process [12]. The AIC results for the SIR model are shown in Table 2, and for the SIRD model in Table 3.

TABLE 2. AIC Values of the Fitted Distributions for the SIR Model Parameters

Distributions	**Transmission Rate β**	**Recovery Rate γ**
Gamma	-372.62	-433.69
Lognormal	-361.11	-462.81
Exponential	-374.51	-434.57
Beta	-386.96	-437.46

TABLE 3. AIC Values of the Fitted Distributions for the SIRD Model Parameters

Distributions	**Transmission Rate β**	**Recovery Rate γ**	**Death Rate κ**
Gamma	-372.62	-432.02	-1535.21
Lognormal	-361.11	-462.57	-1544.05
Exponential	-374.51	-433.08	-1525.63
Beta	-386.96	-436.81	-1533.21

For the SIR model, the beta distribution yielded the lowest AIC for β, while the lognormal distribution had the lowest AIC for γ. Consequently, the initial priors for the SIR model parameters were chosen as a beta distribution for β and a lognormal distribution for γ. Similarly, for the SIRD model, the beta distribution was also the best fit for β, the lognormal for γ, and the lognormal distribution for κ. These selected distributions were thus adopted as the initial

priors for the SIRD model. The parameters of the chosen distributions were derived from the empirical means and variances. For β (beta distribution), shape parameters were $\alpha_1 = 1.53$ and $\alpha_2 = 16.48$. For γ and κ (lognormal distributions), the parameters were $\mu = -2.76,\ \sigma^2 = 0.56$ for γ, and $\mu = -7.521,\ \sigma^2 = 0.515$ for κ. These values were used consistently for both models.

6. Root Mean Squared Error

Root Mean Squared Error (RMSE) is a common evaluation metric used to measure how well a model predicts actual data. It is calculated as the square root of the average of the squared differences between predicted and actual values [13]. RMSE is expressed in the same units as the original data, making it intuitive to interpret. A smaller RMSE indicates a better fit, as it reflects a lower average prediction error. However, since it squares the errors, RMSE is more sensitive to large errors and may not be ideal for datasets with many outliers. Its mathematical formulation is shown in Equation (12):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2} \tag{12}$$

where n is the number of observations, y_i is the actual value, and $\widehat{y}_i$ is the predicted value.

7. Gelman–Rubin Diagnostic

Gelman–Rubin Diagnostic ($\widehat{R}$) is a method used to assess the convergence of Markov chains in MCMC simulations [14]. It compares the variance within each chain to the variance between multiple chains. If the chains have converged, these variances should be similar. Mathematically, the within-chain variance W and between-chain variance B are computed, and the total variance estimate σ^2 is derived by combining both. The Gelman–Rubin statistic is then calculated as shown in Equation (13):

$$\widehat{R} = \sqrt{\frac{\widehat{\sigma}^2}{W}} \tag{13}$$

A value of $\widehat{R}$ close to 1 suggests convergence. Typically, $\widehat{R} < 1.1$ is considered an indicator that the chains have sufficiently converged. This diagnostic helps ensure that MCMC results are reliable and not based on non-stationary chains.

8. Effective Sample Size

Effective Sample Size (ESS) is a metric used in Bayesian parameter estimation to measure how many independent samples are effectively present in an MCMC chain [15]. Since MCMC samples are often autocorrelated, the actual amount of independent information may be much less than the total number of samples. High autocorrelation between samples reduces the ESS, indicating lower sampling efficiency. On the other hand, a higher ESS implies more informative and reliable estimates. ESS is calculated using the formula stated in Equation (14):

$$ESS = \frac{N}{1 + 2\sum_{k=1}^{\infty} \rho_k} \tag{14}$$

where N is the total number of samples, and ρ_k is the autocorrelation at lag k. A lower ESS suggests the presence of redundant information due to correlated samples.

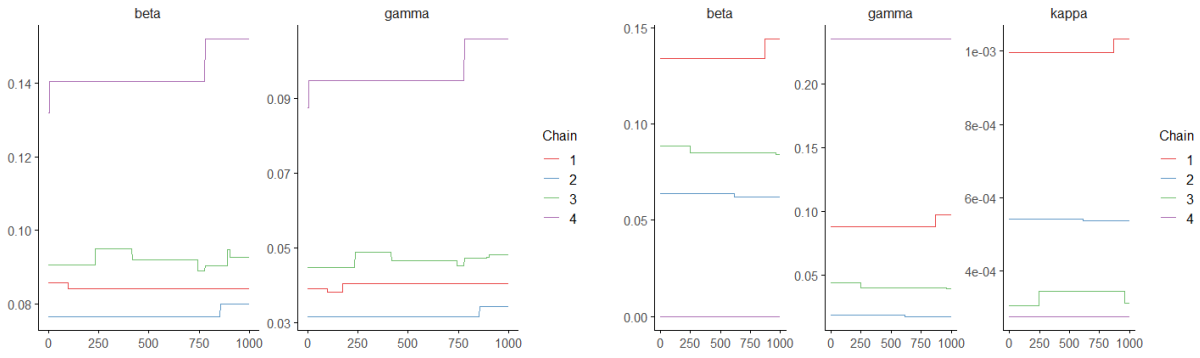
9. Results, Model Validation, and Discussion

This section presents the outcomes of the parameter estimation using the Metropolis–Hastings and Hamiltonian Monte Carlo methods. The performance of each approach is assessed based on the posterior distributions, model fitting, and convergence diagnostics. Furthermore, comparisons between the SIR and SIRD models are discussed to evaluate the suitability and accuracy of each in representing the observed data.

9.1. Parameter Estimation Results Using Metropolis–Hastings. The estimation was carried out using four chains with 2,000 iterations each, where the first 1,000 iterations were discarded as burn-in to reduce the influence of initial values. For the SIR model, the estimated parameters were $\beta = 0.099$ and $\gamma = 0.054$, while for the SIRD model, which includes an additional compartment for deceased individuals, the estimates were $\beta = 0.285$, $\gamma = 0.196$, and $\kappa = 0.0005$.

Despite completing the sampling procedure, the results indicate significant issues with convergence for both models. The Gelman–Rubin diagnostic values were far above the commonly used threshold of 1.1, suggesting that the chains did not converge to the same target distribution. Specifically, for the SIR model, $\widehat{R}$ was 14.16 for β and 14.71 for γ, while for the SIRD model,

the values were even higher, which were 34.21 for β, 66.63 for γ, and 50.12 for κ. These diagnostics clearly indicate that the sampling process failed to reach a stable state, which is further supported by the traceplot visualizations shown in Figure 6 below:



(a) (b)

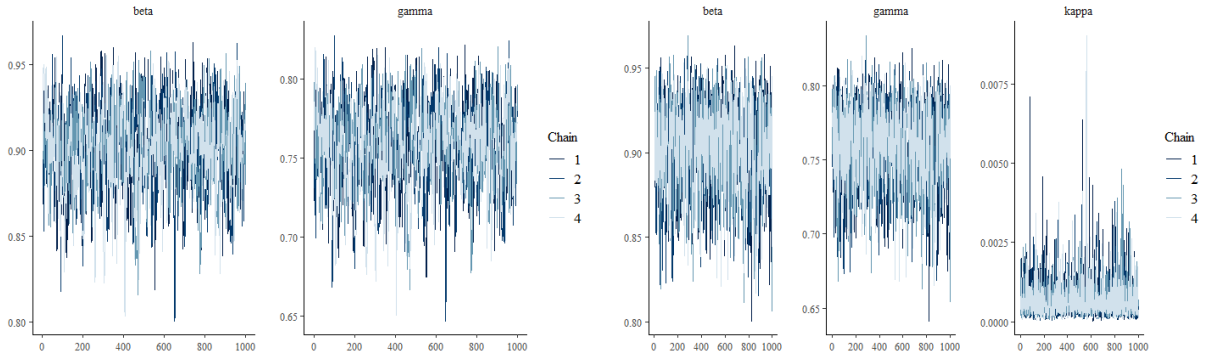
FIGURE 6. Traceplots of parameter estimation using Metropolis–Hastings (a) SIR model, (b) SIRD model

In addition to poor convergence, the sampling efficiency was also notably low. The acceptance rate for proposed values was approximately 0.004 for the SIR model and 0.003 for the SIRD model. Such low acceptance rates imply that most proposed parameter values were rejected, resulting in slow exploration of the posterior space. This inefficiency is further reflected in the effective sample size (ESS) values, which were quite small, 17.85 and 14.24 for β and γ in the SIR model, and 7.87, 7.75, and 8.65 for β, γ, and κ respectively in the SIRD model. Low ESS values suggest high autocorrelation between samples, meaning the chains contained little independent information.

To assess the predictive performance of the models, the discrepancy between the predicted and observed active COVID-19 cases was quantified using the RMSE. The RMSE values were notably high, reaching 6,906,641.89 for the SIR model and 6,224,023.88 for the SIRD model. These large errors indicate a significant difference between the model predictions and actual data, suggesting that the models, as implemented with Metropolis–Hastings, are not reliable for accurate forecasting or inference. Overall, the Metropolis–Hastings method, under the current tuning and prior specifications, failed to produce convergent and accurate posterior estimates

for both the SIR and SIRD models. These results highlight the limitations of the Metropolis–Hastings algorithm in this context and underscore the need for either more informed prior distributions or the use of more efficient sampling techniques, such as Hamiltonian Monte Carlo, which is explored in the next section.

9.2. Parameter Estimation Results Using Hamiltonian Monte Carlo. This section outlines the parameter estimation outcomes obtained via the Hamiltonian Monte Carlo (HMC) method applied to the SIR and SIRD epidemiological models. The sampling process utilized four chains, each running 2,000 iterations, with the initial 1,000 iterations discarded as warm-up to ensure stabilization. For the SIR model, the estimated values for β and γ were 0.900 and 0.760, respectively. In the SIRD model, which additionally accounts for mortality through the parameter κ, the estimates were $\beta = 0.901$, $\gamma = 0.757$, and $\kappa = 0.00071$. Convergence diagnostics indicated strong performance by the HMC algorithm. The Gelman–Rubin statistic was exactly 1 across all parameters, signaling that the chains mixed well and converged effectively to the target posterior distributions. The traceplots presented in Figure 7 show consistent sampling behavior across the four chains.



(a) (b)

FIGURE 7. Traceplots of parameter estimation using Hamiltonian Monte Carlo
(a) SIR model, (b) SIRD model

Regarding sampling efficiency, the HMC method demonstrated notable improvements compared to Metropolis–Hastings. The acceptance rate was very high, approximately 99%, meaning nearly all proposed samples were retained. The effective sample size (ESS) was also significantly greater, with values of 791.32 and 790.39 for β and γ in the SIR model, and 1,164.40,

1,124.45, and 1,657.90 for β, γ, and κ in the SIRD model. These elevated ESS values reflect a lower degree of autocorrelation among samples and thus a richer set of independent draws, increasing estimate reliability.

Although the models demonstrated reasonable parameter estimates and strong convergence diagnostics, the RMSE values remained relatively high, which were 1,039,694.35 for the SIR model and 1,037,925.19 for the SIRD model, indicating a noticeable discrepancy between the predicted and observed COVID-19 case counts. This limitation can be attributed to the structural assumptions of the SIR and SIRD models, which rely on constant transmission and recovery parameters over time. In reality, epidemic dynamics are highly time-dependent and influenced by factors such as public health interventions, behavioral changes, and vaccination rollout. Furthermore, the use of aggregated national-level data may obscure regional heterogeneity, leading to reduced predictive accuracy. These findings suggest that while Bayesian estimation using HMC provides reliable parameter inference, improvements in model structure—such as incorporating time-varying parameters or additional compartments—may be necessary to enhance predictive performance. In addition, the comparison between Metropolis–Hastings and Hamiltonian Monte Carlo highlights fundamental differences in sampling efficiency. The poor performance of Metropolis–Hastings is likely due to its random-walk behavior, which becomes inefficient in high-dimensional and highly correlated parameter spaces. In contrast, HMC leverages gradient information to explore the posterior distribution more effectively, resulting in faster convergence and significantly larger effective sample sizes. However, it should be noted that the performance of Metropolis–Hastings may be improved with better tuning of the proposal distribution or adaptive strategies. Therefore, while HMC is clearly superior in this study, the choice of sampling method should still consider model complexity and computational resources.

9.3. Prior Distribution Adjustment. The initial prior distributions for the parameters β, γ, and κ were selected using the AIC, which suggested that a beta distribution was the best fit for β, while a lognormal distribution was more suitable for γ and κ. However, applying these priors in the SIR and SIRD models using the Metropolis–Hastings method resulted in poor parameter

estimation. The Gelman–Rubin diagnostic values were high, which indicates poor convergence, and the RMSE values were also relatively large.

Even though Hamiltonian Monte Carlo (HMC) yielded perfect convergence ($\widehat{R} = 1$) and much lower RMSE than Metropolis–Hastings, the error values were still considered high. To improve the estimation results, a reassessment of the prior distribution was conducted, focusing particularly on the transmission rate parameter β. As an alternative, a lognormal distribution was chosen based on its suitability for modeling positive, right-skewed data. This choice was supported by previous studies on COVID-19 which used lognormal priors for the transmission rate in similar epidemiological models [16, 17]. In this study, the lognormal prior for β was derived as $\mu = -2.69$ and $\sigma^2 = 0.45$. Estimations were then repeated for both SIR and SIRD models using Metropolis–Hastings and HMC, now with the updated lognormal prior for β. The results are summarized in Table 4.

TABLE 4. Performance Comparison of Estimation Methods

Model	Method	Prior β	RMSE	Mean $\widehat{\mathbf{R}}$	Mean ESS	Acceptance Rate
SIR	Metropolis–Hastings	Beta	6,906,641.89	14.43	16.05	0.004
		Lognormal	3,329,720.62	2.55	23.73	0.13
	Hamiltonian Monte Carlo	Beta	1,039,694.35	1	790.85	0.99
		Lognormal	67,619.11	1	947.35	0.99
SIRD	Metropolis–Hastings	Beta	6,224,023.88	50.32	8.09	0.003
		Lognormal	1,094,503.98	2.82	14.19	0.086
	Hamiltonian Monte Carlo	Beta	1,037,925.19	1	1,315.58	0.99
		Lognormal	67,255.57	1	1,717.57	0.99

From the table above, it can be seen that using a lognormal prior for β leads to better estimation results across both models and methods, especially when combined with HMC. In the SIR model using Metropolis–Hastings, the change in prior reduced the RMSE and improved both convergence and acceptance rate, although not sufficiently, where the $\widehat{R}$ value remained above the commonly used threshold of 1.1. Similar trends were observed in the SIRD model, where the lognormal prior substantially reduced the RMSE, but convergence remained poor due to the

increased complexity of the model and the limitations of the Metropolis–Hastings algorithm in higher-dimensional parameter spaces.

In contrast, HMC demonstrated consistent convergence and high efficiency across all experiments. When the lognormal prior was applied, the HMC method produced significantly lower RMSE values, dropping from over one million to under seventy thousand in both SIR and SIRD models. Moreover, the effective sample size (ESS) increased, indicating better sampling quality. These improvements suggest that the lognormal distribution captures the behavior of β more accurately, and that HMC is substantially more robust, especially for complex models like SIRD.

The SIRD model estimated with HMC and a lognormal prior delivered the best results, with the lowest RMSE and highest ESS. This supports the idea that the SIRD model offers a more accurate epidemic representation by separating deaths from recoveries, reducing bias present in the simpler SIR model. HMC's strength lies in its efficient use of gradient information, enabling better convergence and high-quality, weakly correlated samples, unlike Metropolis–Hastings, which struggles in complex models due to its random-walk approach. In summary, the results support the use of a lognormal prior for the transmission rate β, a choice grounded in both empirical performance and literature. Furthermore, HMC proved to be the most effective estimation method, particularly for the more nuanced SIRD model. These findings highlight the importance of careful prior selection and the choice of inference method in achieving accurate and reliable parameter estimation in epidemiological modeling.

9.4. Comparative Analysis of Estimation Methods. To provide a clearer comparison between the estimation approaches, the performance of each method was evaluated using convergence diagnostics, sampling efficiency, and predictive accuracy. The Metropolis–Hastings algorithm consistently exhibited poor convergence behavior, as reflected by extremely large Gelman–Rubin statistics and very low effective sample sizes. These results indicate that the Markov chains were highly autocorrelated and failed to adequately explore the posterior distribution.

In contrast, the Hamiltonian Monte Carlo algorithm demonstrated stable convergence and significantly improved sampling efficiency. The traceplots showed good mixing across chains,

while the Gelman–Rubin statistic remained close to one for all parameters. Furthermore, the effective sample size values obtained using HMC were approximately 40 to 160 times larger than those produced by Metropolis–Hastings.

From a predictive perspective, the SIRD model estimated using HMC with a lognormal prior for the transmission parameter achieved the lowest RMSE among all configurations. This suggests that explicitly modeling disease-induced mortality improves the ability of the model to capture the dynamics of COVID-19 transmission in Indonesia.

These results highlight the importance of using gradient-based MCMC algorithms for parameter estimation in nonlinear epidemiological models. Compared with traditional random-walk methods such as Metropolis–Hastings, Hamiltonian Monte Carlo is able to explore the posterior distribution more efficiently, resulting in faster convergence and higher effective sample sizes.

10. Conclusion

This study investigated Bayesian parameter estimation for SIR and SIRD epidemiological models using two MCMC sampling methods: Metropolis–Hastings and Hamiltonian Monte Carlo. The empirical analysis based on COVID-19 data from Indonesia shows that the Metropolis–Hastings algorithm suffers from poor convergence, low acceptance rates, and very small effective sample sizes. As a result, the parameter estimates obtained from this method are unreliable for modeling epidemic dynamics.

In contrast, Hamiltonian Monte Carlo demonstrated excellent convergence properties and significantly higher sampling efficiency. The method consistently produced Gelman–Rubin statistics close to one and large effective sample sizes, indicating reliable exploration of the posterior distribution. Among all configurations, the SIRD model estimated using HMC with a lognormal prior for the transmission rate achieved the best predictive performance.

These results highlight the importance of selecting both an appropriate epidemiological model and an efficient Bayesian inference algorithm. The findings suggest that Hamiltonian Monte Carlo provides a robust framework for parameter estimation in compartmental epidemic models and can improve the reliability of epidemic modeling studies.

Despite the promising results, this study has several limitations. The models assume a constant population and do not account for time-varying transmission rates, vaccination effects, or

changes in public behavior. In addition, the analysis is based on reported case data, which may be affected by underreporting or reporting delays. These limitations may influence the accuracy of the estimated parameters and the predictive performance of the models. Future research may incorporate time-varying transmission rates, vaccination effects, or mobility data to further improve the predictive accuracy of compartmental epidemic models.

ACKNOWLEDGMENT

The authors would like to thank Andreas Parama Wijaya and Robyn Irawan for their useful inputs.

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

REFERENCES

[1] F. Brauer, C. Castillo-Chávez, Z. Feng, Mathematical Models in Epidemiology, Springer New York, 2019. https://doi.org/10.1007/978-1-4939-9828-9.

[2] J. Andrade, J. Duggan, A Bayesian Approach to Calibrate System Dynamics Models Using Hamiltonian Monte Carlo, Syst. Dyn. Rev. 37 (2021), 283–309. https://doi.org/10.1002/sdr.1693.

[3] A.S. Talawar, U.R. Aundhakar, Parameter Estimation of SIR Epidemic Model Using MCMC Methods, Glob. J. Pure Appl. Math. 12 (2016), 1299–1306. https://www.ripublication.com/gjpam16/gjpamv12n2_11.pdf.

[4] L. Grinsztajn, E. Semenova, C.C. Margossian, J. Riou, Bayesian Workflow for Disease Transmission Modeling in Stan, Stat. Med. 40 (2021), 6209–6234. https://doi.org/10.1002/sim.9164.

[5] S. Gupta, Rupali, Y.K. Rajoria, G.P. Sahu, Mathematical Modelling on Dynamics of Multi-variant SARS-CoV-2 Virus: Estimating Delta and Omicron Variant Impact on COVID-19, IAENG Int. J. Appl. Math. 55 (2025), 180–188.

[6] T. Donovan, R.M. Mickey, Bayesian Statistics for Beginners: A Step-by-Step Approach, Oxford University Press, Oxford, 2019. https://doi.org/10.1093/oso/9780198841296.001.0001.

[7] M.D. Hoffman, A. Gelman, The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo, J. Mach. Learn. Res. 15 (2014), 1593–1623. http://jmlr.csail.mit.edu/papers/v15/hoffman14a.html.

[8] R.M. Neal, MCMC Using Hamiltonian Dynamics, in: Handbook of Markov Chain Monte Carlo, Chapman and Hall/CRC, pp. 113–162, (2011). https://doi.org/10.1201/b10905-6.

[9] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, et al., Bayesian Data Analysis, Chapman and Hall/CRC, 2013. https://doi.org/10.1201/b16018.

[10] M. Martcheva, An Introduction to Mathematical Epidemiology, Springer US, 2015. https://doi.org/10.1007/978-1-4899-7612-3.

[11] D. Sen, D. Sen, Use of a Modified SIRD Model to Analyze COVID-19 Data, Ind. Eng. Chem. Res. 60 (2021), 4251–4260. https://doi.org/10.1021/acs.iecr.0c04754.

[12] M. Mazerolle, Improving Data Analysis in Herpetology: Using Akaike's Information Criterion (AIC) to Assess the Strength of Biological Hypotheses, Amphibia-Reptilia 27 (2006), 169–180. https://doi.org/10.1163/156853806777239922.

[13] C.J. Willmott, K. Matsuura, Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in Assessing Average Model Performance, Clim. Res. 30 (2005), 79–82. https://doi.org/10.3354/cr030079.

[14] D. Vats, C. Knudson, Revisiting the Gelman–Rubin Diagnostic, Stat. Sci. 36 (2021), 518–529. https://doi.org/10.1214/20-sts812.

[15] J.K. Kruschke, Doing Bayesian Data Analysis, Elsevier, 2015. https://doi.org/10.1016/C2012-0-00477-2.

[16] R. Mbuvha, T. Marwala, Bayesian Inference of COVID-19 Spreading Rates in South Africa, PLoS One 15 (2020), e0237126. https://doi.org/10.1371/journal.pone.0237126.

[17] M. Mandal, S. Mandal, Detection of Transmission Change Points During Unlock-3 and Unlock-4 Measures Controlling COVID-19 in India, J. Drug Deliv. Ther. 11 (2021), 76–86. https://doi.org/10.22270/jddt.v11i2.4600.

[18] Satuan Tugas Penanganan COVID-19, Peta Sebaran COVID-19, 2022. https://web.archive.org/web/20220330235109/https://covid19.go.id/peta-sebaran.