MethodsX 10 (2023) 101979

Contents lists available at ScienceDirect

MethodsX

journal homepage: www.elsevier.com/locate/mex

Stochastic Bayesian Runge-Kutta method for dengue dynamic mapping

Mukhsar[a,*], Gusti Ngurah Adhi Wibawa[a], Andi Tenriawaru[b], Ida Usman[c], Muhammad Zamrun Firihu[c], Viska Inda Variani[c], Andi Besse Firdausiah Mansur[d], Ahmad Hoirul Basori[d]

[a] Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Halu Oleo Kendari, Indonesia
[b] Department of Computer Science, Faculty of Mathematics and Natural Sciences, Universitas Halu Oleo Kendari, Indonesia
[c] Department of Physics, Faculty of Mathematics and Natural Sciences, Universitas Halu Oleo Kendari, Indonesia
[d] Faculty of Computing and Information Technology in Rabigh, King Abdulaziz University, Jeddah, Saudi Arabia

HIGHLIGHTS

- This study proposes an innovative statistical approach for mapping the dengue dispersion.
- The statistical analysis using the Bayesian Runge-Kutta family can map the region that has high, medium, and low risk.
- The method and application will be helpful for local governments to suppress the dengue in several neighbourhoods with high cases.

ARTICLE INFO

Keywords:
Bayesian
Euler
Four order Runge-Kutta
SIR-SI mode
Stochastic

ABSTRACT

Dengue Hemorrhagic Fever (DHF) is still a threat to humanity that cause death and disability due to changes in environmental and socioeconomic conditions, especially in tropical areas. A critical assessment of the models and methods is necessary. The vital role of stochastic processes of infectious disease research in Bayesian statistical models helps provide an explicit framework for understanding disease transmission dynamics between hosts (humans) and vectors (mosquitoes). This research presents a Bayesian stochastic process for the cross-infection SIR-SI model as a differential equation (susceptibility–infection–recovery for the human population; susceptible–infectious for the vector population) for the dynamic transmission of DHF. Given the difficulties of solving the differential equations precisely, we propose a computational model to approach the solution based on the Runge-Kutta family approach, namely the Euler and the four-order Runge-Kutta methods. The methods used for the discretization process of the SIR-SI model for computational process needs. For comparison purposes, we use a monthly DHF dataset of 10 Kendari, Indonesia, districts from 2019 to 2021. Parameter estimation of the Bayesian SIR-SI model based on the Euler and four-order Runge-Kutta method was updated using Markov Chain Monte Carlo (MCMC) Bayesian. The Euler and four-order Runge-Kutta methods have converged at 10,000 iterations with burn-in 80,000. The numerical simulation results show that the four-order Runge-Kutta approach has the slightest deviance, 106.5. Therefore, this approach is the best one. The relative risk analysis shows that the dynamics of DHF cases fluctuate from January to July every year. However, from January to May, there was a high consistency of DHF cases. Two districts with high case consistency were found, namely Kadia and Wua-Wua. Furthermore, because the

* Corresponding author.
E-mail address: mukhsar.mtmk@uho.ac.id.

https://doi.org/10.1016/j.mex.2022.101979
Received 5 October 2022; Accepted 19 December 2022
Available online 20 December 2022

spread of DHF cases has a spatial effect, the Kadia and Wua-Wua districts need serious attention to suppress the rate of reach of DHF in Kendari City. Intensive observation and intervention in the Kadia and Wua-Wua districts should be carried out in early January to stop the breeding of mosquito larvae. In other neighbourhoods such as Puwatu, Kambu and Kendari Barat, the DHF cases occur temporally as the impacts of DHF cases in the Kadia and Wua-Wua districts. It may facilitate the statisticians to develop further models to adopt a better understanding to control the DHF dynamically. In brief,

Specification Table

Subject Area:	Computer Science/Statistic Analysis
More specific of subject area:	Computational Statistical Analysis
Method Name:	Stochastic Bayesian Runge-Kutta
Name and reference of original method:	NA
Resource availability:	WinBUGS software (MRC Bio Statistic) https://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-winbugs/

Introduction

Dengue Hemorrhagic Fever (DHF) is an ancient disease having a substantial social, economic, and health burden. DHF is not just a disease that remains a public health problem in tropical countries worldwide. Dengue virus has evolved from a sporadic disease to a socially and economically impactful public health problem due to increasing geographic extension, number of cases, and severity. Social impacts are losses in the form of discomfort, family panic, and even death which results in lower life expectancy. The economic impact is in the form of medical and nursing costs, lost work time, school time, transportation and accommodation incurred during treatment, and costs incurred by the government for intervention. Until now, development of drugs to prevent the dengue virus has not been found. For example, early prevention efforts are needed through modeling to provide prevention directions and targeted interventions. Even though the DHF has been investigated for years, it remains a significant public health problem, especially in tropical countries [1–3]. The DHF dynamics have also gained prominence recently since climate change or global warming is predicted to have unexpected effects on its incidence [4,5]. The complexities in the life cycle of the mosquito, highly complex environmental and social interactions, unforeseen consequences of climate change, and population migration between endemic and non-endemic areas continued to contribute to the massive burden of morbidity and mortality accompanying the DHF [6,7]. With no effective vaccine, eradicating mosquito larvae as early as possible is the only advisory prevention.

Many researchers in infectious disease were among the foremost to realize the vital role of statistical models in providing an explicit framework for understanding the disease transmission dynamics between the host (human) and vector (mosquito) [8,9]. The spread of DHF is caused by the influence of human or host factors, aedes mosquitoes or vectors, and the environment, which are spatially correlated. Changes to one component will change the other two components. This change can be in local and global transmission, mainly due to active host mobility. From the merely epidemiological point of view, the empirical data on the propagation of diseases have shown that a nondeterministic behavior is invariably observed. The nondeterministic form explains through a stochastic process. The stochastic processes are also needed to describe the dynamics spatially and temporally. Depending on the type of infectious disease, there are many stochastic variations in the compartment structure. For example, in a SIR model, part of the susceptible population (S) will be infected with the disease so that it enters the infectious population (I) and then recovers or dies from infection and enters the recovery population (R). Because of this, various stochastic models have been proposed to describe the propagation of infectious diseases in human populations [10].

Some authors have developed stochastic models to analyze the dynamic of DHF cases spatially and temporally, such as the Bayesian zero-inflated Poisson spatial-temporal model [6], to investigate the comparative risk of DHF cases. These models require an ecological factor of the environment and two random effects (local and global heterogeneities). But in these models, the dynamics of cross-infection between humans and mosquitoes are not considered. Some authors perform a stochastic process to assess the dynamics of cross-infection between humans and mosquitoes, such as [4,5]. Recent research by [4,5] provided a stochastic process for cross-infection such as susceptible host, infected host, re-covered host and susceptible vector, infective vector, or SIR-SI model. An Euler method was used to discretize the model. The approach to discretizing models built from stochastic processes can use several techniques, for example, four-order Runge-Kutta, Euler Maruyama, and non-standard methods [11]. These methods have varying accuracy. Depending on the stochastic model, the complex process is not necessarily better than the simple method.

This research developed a stochastic process for the SIR-SI model by involving local and global random effects in the model. A combination of both unexpected outcomes leads to a Bayesian stochastic SIR-SI model. A detailed description of local and global heterogeneities can be seen in [12–14]. We use stochastic Euler and four-order Runge–Kutta method to discretize the models for comparison purposes. We compare our model accuracy to those proposed by [9] based on a monthly DHF dataset of 10 districts of Kendari, Indonesia from 2019–2021. Based on the two discretization methods for the SIR-SI model, then are using Bayesian MCMC for parameter estimation of the model. The numerical simulation uses WinBUGS. The advantage of WinBUGS is that besides being open source, it can also generate deviance values. The slightest deviance is the best one. This best model is used to analyze the risk of DHF cases. The model is expected to be used as an effective control to detect and respond to DHF disease quickly.

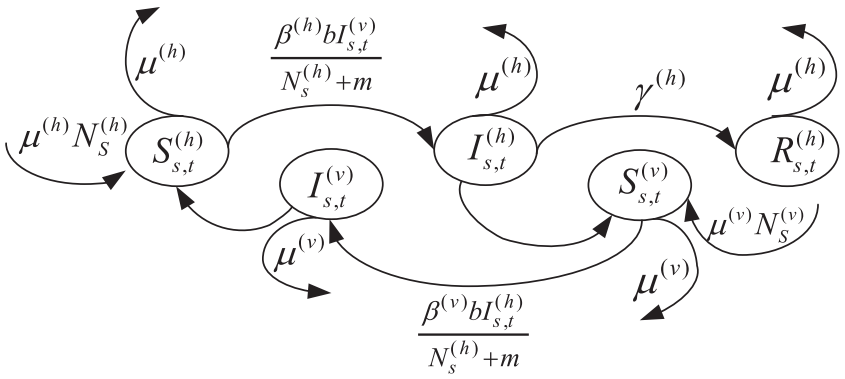

Fig. 1. Compartment of SIR-SI model (adopted from [5]).

This work is organized as follows. We present an epidemiological model of SIR-SI at an early stage involving contact between humans (the host) and a virus-carrying agent or vector (in this study, the virus-carrying agent was mosquitoes). This model is a scheme that organizes the human population into various sub-populations: susceptible individuals (S), infected (I), i.e., individuals who have the virus, show symptoms, and can spread the virus, and recovered individuals (R). Meanwhile, the mosquito population is divided into two sub-populations: those susceptible to dengue virus (S) and those infected with dengue (I). The dengue virus is transmitted to humans through mosquito bites. We will construct a differential equation based on the SIR-SI scheme. The stochastic model will be derived for further investigation in this study. In Part 2, we will present various computational methods for solving the stochastic differential problem under investigation using the Runge-Kutta family method. This section introduces the application of Bayesian and the uncertainty factor as a result of the actual mobility of people. In Section 3, we will use real data and provide some simulations to demonstrate the performance of the computational model, and we will close this work with a conclusion section.

Related works

The compartment model is due to human (host) and mosquito (vector) populations. The human population is referred to as the disease status (for instance, the susceptible host-infected host-recovered host model). In contrast, the mosquito population has two compartments, namely susceptible and infected vectors. The compartment model is expressed in an ordinary differential equation. This system explains that all individuals in the population can interact with the same possibilities [15–17]. However, individual interactions are heterogeneous. Therefore, modeling should consider stochastic processes. The stochastic model reveals an individual in each compartment is regarded as a random variable with specific distribution [16]. The stochastic model can also be applied to the Bayesian paradigm (see, for example, [18–20].

The host in the stochastic model is divided into three classes, namely in the SIR model [21,22]. On the other hand, the vector is parted into two categories, namely the SI model. The mosquitoes infected by the dengue virus will remain the viral carrier for the rest of their lifetimes [23]. Mosquitoes transmit the dengue viruses through their bites, and then a mosquito could be infected by dengue through its bite on an infected human. There is a cross-infection between humans as the host and mosquitoes as the vector. Three assumptions are as follows: (1) susceptible people can become infected and then recover or die due to the infection, (2) susceptible aedes mosquitoes can become infective, but they will not recover or die due to the infection because infective mosquitoes stay infective for the remainder of their lifetimes, (3) the disease spreads in an open environment, with emigration and immigration, and birth and death in a population, 4) the infective rate of an infected individual is proportional to the number of susceptible, and 5) the recovery rate is proportional to the number of infected. The compartment model described in Fig. 1 is a SIR-SI model. It is commonly used in studies of the spread of DHF disease in human populations [5,10].

The susceptible, infected, and recovered host classes are represented by $S_{s,t}^{(h)}$, $I_{s,t}^{(h)}$, $R_{s,t}^{(h)}$, respectively at time $t = 1, 2, \ldots, T$ and region $s = 1, 2, \ldots, S$. The notation $S_{s,t}^{(v)}$ and $I_{s,t}^{(v)}$ represents number of susceptible and infective vectors, respectively. Parameters $\mu^{(h)}$and $\mu^{(v)}$ represent the human birth rates and mosquito death rates, respectively. Furthermore, $\gamma^{(h)}$and b represent the human recovered rate and bthe biting rate per time, respectively. The parameter m represents as blood source of the alternative human available number. Parameters $\beta^{(h)}$ and $\beta^{(v)}$ represent the transmission probability from mosquitoes to humans and from humans to mosquitoes, respectively. The $N_{s,t}^{(h)}$ and $N_{s,t}^{(v)}$ represents the human and mosquito population size, respectively.

The SIR-SI model explained in Fig. 1 can be transformed into the differential equation system [5,25]. The systems are used to provide a link to stochastic process (see Eq. (1)). The $N_{s,t}^{(h)}$ and $N_{s,t}^{(v)}$ are assumed to be constant, such that $S_{s,t}^{(h)} + I_{s,t}^{(h)} + R_{s,t}^{(h)} = N_{s,t}^{(h)}$ and $S_{s,t}^{(v)} + I_{s,t}^{(v)} = N_{s,t}^{(v)}$.

$$\frac{dS_{s,t}^{(h)}}{dt}\frac{dS_{s,t}^{(h)}}{dt} = \mu^{(h)} N_{s,t}^{(h)} - \mu^{(h)} S_{s,t}^{(h)} - \frac{\beta^{(h)} b I_{s,t}^{(v)}}{N_{s,t}^{(h)} + m} S_{s,t}^{(h)}$$

$$\frac{dI_{s,t}^{(h)}}{dt}\frac{dI_{s,t}^{(h)}}{dt} = \frac{\beta^{(h)} b I_{s,t}^{(v)}}{N_{s,t}^{(h)} + m} S_{s,t}^{(h)} - \mu^{(h)} I_{s,t}^{(h)} - \gamma^{(h)} I_{s,t}^{(h)}$$

$$\frac{dR_{s,t}^{(h)}}{dt}\frac{dR_{s,t}^{(h)}}{dt} = \gamma^{(h)} I_{s,t}^{(h)} - \mu^{(h)} R_{s,t}^{(h)}$$

$$\frac{dS_{s,t}^{(v)}}{dt}\frac{dS_{s,t}^{(v)}}{dt} = \mu^{(v)} N_{s,t}^{(v)} - \mu^{(v)} S_{s,t}^{(v)} - \frac{\beta^{(v)} b I_{s,t}^{(h)}}{N_{s,t}^{(v)} + m} S_{s,t}^{(v)}$$

$$\frac{dI_{s,t}^{(v)}}{dt}\frac{dI_{s,t}^{(v)}}{dt} = \frac{\beta^{(v)} b I_{s,t}^{(h)}}{N_{s,t}^{(v)} + m} S_{s,t}^{(v)} - \mu^{(v)} I_{s,t}^{(v)} \tag{1}$$

The numerical approach used in this work is based on the fourth-order Runge–Kutta family to solve systems of ordinary differential equations. To provide a stochastic generalization of the model (1), we define

$$\Phi = \left[S_s^{(h)}(t), I_s^{(h)}(t), R_s^{(h)}(t), S_s^{(v)}(t), I_s^{(v)}(t)\right],\ \forall t \in [0, T]$$

We consider a uniform partition of the interval $[0, T]$ consisting subinterval of $N \in \mathbb{N}$. Define the temporal step size $h = \frac{T}{N}$ and to simplify the analytical process of Eq. (1) is defined as

$$\frac{dS_{s,t}^{(h)}}{dt} \equiv S_s^{*(h)}(t, \Phi),\ \frac{dI_{s,t}^{(h)}}{dt} \equiv I_s^{*(h)}(t, \Phi),\ \frac{dR_{s,t}^{(h)}}{dt} \equiv R_s^{*(h)}(t, \Phi),\ \frac{dS_{s,t}^{(v)}}{dt} \equiv S_s^{*(v)}(t, \Phi),$$

and

$$\frac{dI_{s,t}^{(v)}}{dt} \equiv I_s^{*(v)}(t, \Phi).$$

The Runge–Kutta family is a four-stage technique for stochastic form.

Stage 1

$$K_{11} = hS^{*(h)}\left(t_i, \Phi_i\right) = h\left(\mu^{(h)} N^{(h)} - \mu^{(h)} S^{(h)n} - \frac{\beta^{(h)} b I^{(v)n}}{N^{(h)} + m} S^{(h)n}\right)$$

$$K_{12} = hI^{*(h)}\left(t_i, \Phi_i\right) = h\left(\frac{\beta^{(h)} b I^{(v)n}}{N^{(h)} + m} S^{(h)n} - \mu^{(h)} I^{(h)n} - \gamma^{(h)} I^{(h)n}\right)$$

$$K_{13} = hR^{*(h)}\left(t_i, \Phi_i\right) = h\left(\gamma^{(h)} I^{(h)n} - \mu^{(h)} R^{(h)n}\right)$$

$$K_{14} = hS^{*(v)}\left(t_i, \Phi_i\right) = h\left(\mu^{(v)} N^{(v)} - \mu^{(v)} S^{(v)n} - \frac{\beta^{(v)} b I^{(h)n}}{N^{(v)} + m} S^{(v)n}\right)$$

$$K_{15} = hI^{*(v)}\left(t_i, \Phi_i\right) = h\left(\frac{\beta^{(v)} b I^{(h)n}}{N^{(v)} + m} S^{(v)n} - \mu^{(v)} I^{(v)n}\right)$$

The stage 1 is also known as the Euler method.

Stage 2

$$K_{21} = h\left(\mu^{(h)} N^{(h)} - \mu^{(h)}\left(S^{(h)n} + K_{11}/2\right) - \frac{\beta^{(h)} b\left(I^{(v)n} + K_{11}/2\right)}{N^{(h)} + m}\left(S^{(h)n} + K_{11}/2\right)\right)$$

$$K_{22} = h\left(\frac{\beta^{(h)} b\left(I^{(v)n} + K_{12}/2\right)}{N^{(h)} + m}\left(S^{(h)n} + K_{12}/2\right) - \mu^{(h)}\left(I^{(h)n} + K_{12}/2\right) - \left(\gamma^{(h)} I^{(h)n} + K_{12}/2\right)\right)$$

$$K_{23} = h\left(\gamma^{(h)}\left(I^{(h)n} + K_{13}/2\right) - \mu^{(h)}\left(R^{(h)n} + K_{13}/2\right)\right)$$

$$K_{24} = h\left(\mu^{(v)} N^{(v)} - \mu^{(v)}\left(S^{(v)n} + K_{14}/2\right) - \frac{\beta^{(v)} b\left(I^{(h)n} + K_{14}/2\right)}{N^{(v)} + m}\left(S^{(v)n} + K_{14}/2\right)\right)$$

$$K_{25} = h\left(\frac{\beta^{(v)} b\left(I^{(h)n} + K_{15}/2\right)}{N^{(v)} + m}\left(S^{(v)n} + K_{15}/2\right) - \mu^{(v)}\left(I^{(v)n} + K_{15}/2\right)\right)$$

Stage 3

$$K_{31} = h\left(\mu^{(h)} N^{(h)} - \mu^{(h)}\left(S^{(h)n} + K_{21}/2\right) - \frac{b\left(I^{(v)n} + K_{21}/2\right)}{N^{(h)} + m}\left(S^{(h)n} + K_{21}/2\right)\right)$$

$$K_{32} = h\left(\frac{\beta^{(h)} b\left(I^{(v)n} + K_{22}/2\right)}{N^{(h)} + m}\left(S^{(h)n} + K_{22}/2\right) - \mu^{(h)}\left(I^{(h)n} + K_{22}/2\right) - \left(\gamma^{(h)} I^{(h)n} + K_{22}/2\right)\right).$$

$$K_{33} = h\left(\gamma^{(h)}\left(I^{(h)n} + K_{23}/2\right) - \mu^{(h)}\left(R^{(h)n} + K_{23}/2\right)\right).$$

$$K_{34} = h\left(\mu^{(v)} N^{(v)} - \mu^{(v)}\left(S^{(v)n} + K_{24}/2\right) - \frac{\beta^{(v)} b\left(I^{(h)n} + K_{24}/2\right)}{N^{(v)} + m}\left(S^{(v)n} + K_{24}/2\right)\right)$$

Stage 4

$$K_{41} = h\left(\mu^{(h)} N^{(h)} - \mu^{(h)}\left(S^{(h)n} + K_{31}/2\right) - \frac{\beta^{(h)} b\left(I^{(v)n} + K_{31}/2\right)}{N^{(h)} + m}\left(S^{(h)n} + K_{31}/2\right)\right)$$

$$K_{42} = h\left(\frac{\beta^{(h)} b\left(I^{(v)n} + K_{32}/2\right)}{N^{(h)} + m}\left(S^{(h)n} + K_{32}/2\right) - \mu^{(h)}\left(I^{(h)n} + K_{32}/2\right) - \left(\gamma^{(h)} I^{(h)n} + K_{32}/2\right)\right)$$

$$K_{43} = h\left(\gamma^{(h)}\left(I^{(h)n} + K_{33}/2\right) - \mu^{(h)}\left(R^{(h)n} + K_{33}/2\right)\right)$$

$$K_{44} = h\left(\mu^{(v)} N^{(v)} - \mu^{(v)}\left(S^{(v)n} + K_{34}/2\right) - \frac{\beta^{(v)} b\left(I^{(h)n} + K_{34}/2\right)}{N^{(v)} + m}\left(S^{(v)n} + K_{34}/2\right)\right)$$

$$K_{45} = h\left(\frac{\beta^{(v)} b\left(I^{(h)n} + K_{35}/2\right)}{N^{(v)} + m}\left(S^{(v)n} + K_{35}/2\right) - \mu^{(v)}\left(I^{(v)n} + K_{35}/2\right)\right)$$

The spread of DHF cases can be transmitted with local and cross site due to social interaction of people. Consequently, an additional component in Eq. (1) is required, namely the random effect of local ($u_{s,t}^{(h)}$) and global ($v_{s,t}^{(h)}$) heterogeneity. Random global heterogeneity represents the actual mobility of people cross locations. Therefore, location weighting is required. In the Bayesian paradigm, the CAR model involves weighting between locations represented by the neighbourhood matrix written.

$v_{s,t}|v_{j,j\neq s} \sim N(\rho\sum_{j\in\varepsilon(i)}\frac{v_j}{P_s}, \frac{1}{\tau_v P_s})$, $\varepsilon(i)$ represents the number of neighbours at location s, and $-1 \leq \rho \leq 1$. Based on the analytical description related to the numerical approach using the Runge-Kutta family, the discretization using the Euler method (or stage 1) for model (1) is described in Eq. (2).

$$S_{s,\ t+1}^{(h)} = \mu^{(h)} N_{s,t}^{(h)} + \left(1 - \mu^{(h)}\right) S_{s,t}^{(h)} - A_{s,t}^{(h)}, \quad A_{s,t}^{(h)}\ \text{Poisson}\left(\lambda_{s,t}^{(h)}\right)$$

$$\lambda_{s,t+1}^{(h)} = \exp\left(\beta_0^{(v)} + u_{s,t}^{(h)} + v_{s,t}^{(h)}\right)\frac{\beta^{(h)} b I_{s,t}^{(v)}}{N_{s,t}^{(h)} + m} S_{s,t}^{(h)}$$

$$I_{s,t+1}^{(h)} = \left(1 - \mu^{(h)}\right) A_{s,t}^{(h)} - B_{s,t}^{(h)}, \quad B_{s,t}^{(h)} = \gamma^{(h)} I_{s,t}^{(h)}$$

$$R_{s,t+1}^{(h)} = \left(1 - \mu^{(h)}\right) R_{s,t}^{(h)} + C_{s,t}^{(h)},\ C_{s,t}^{(h)} = \mu^{(h)} I_{s,t}^{(h)}$$

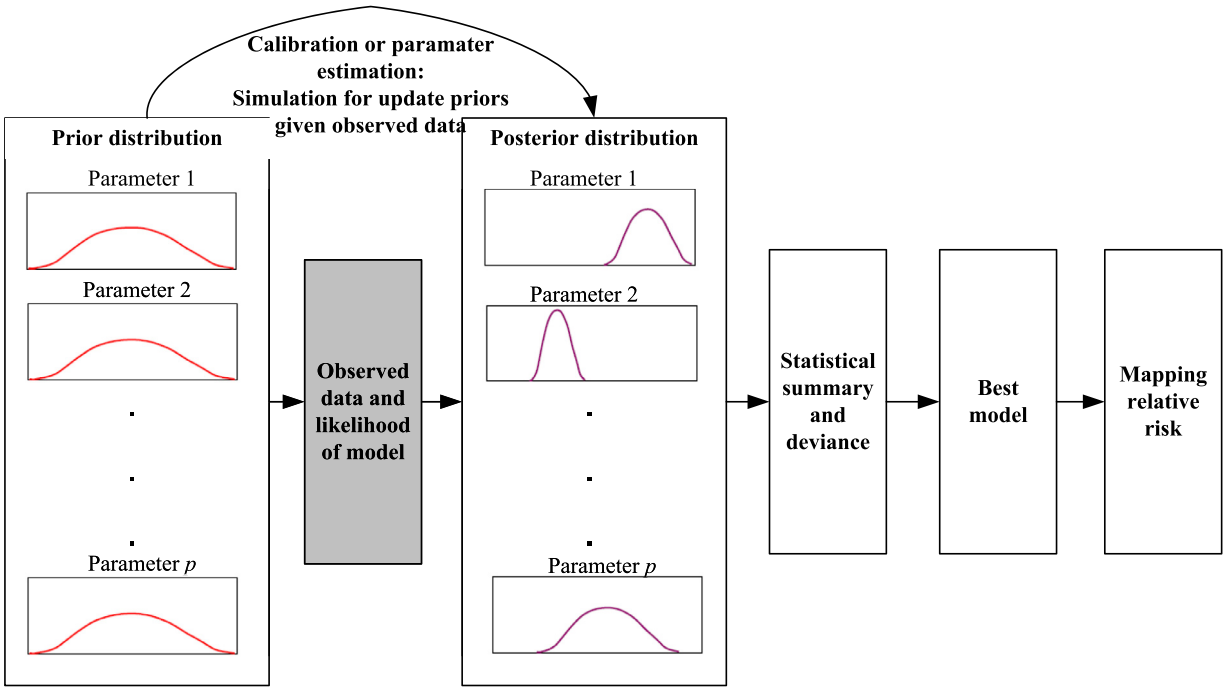

Fig. 2. Bayesian approach to model calibration.

$$S_{s,t+1}^{(v)} = \left(1-\mu^{(v)}\right)S_{s,t}^{(v)} + \mu^{(v)}N_{s,t}^{(v)} - \frac{\beta^{(v)}bI_{s,t}^{(h)}}{N_{s,t}^{(v)}+m}S_{s,t}^{(v)}$$

$$I_{s,t+1}^{(v)} = \left(1-\mu^{(v)}\right)I_{s,t}^{(v)} + \frac{\beta^{(v)}bI_{s,t}^{(h)}}{N_{s,t}^{(v)}+m}S_{s,t}^{(v)} \tag{2}$$

Random effect represents the actual mobility of people to provide procedures of variations in the relative risk. The terms $A_{s,t}^{(h)}$ and $B_{s,t}^{(h)}$ in the model (2) describe the newly infective humans and mosquitoes. In the fundamental model, the count data use the Poisson distribution for new invectives. The Eq. (2) of the human new infective is assumed to be independent of Poisson distributions. The parameter $\beta_0^{(v)}$ represents a constant term to describe the overall rates of human populations.

The numerical process for the four-order Runge-Kutta is carried out the same as the Euler method or stage 1. Therefore, discretizing the model (1) using the four-order Runge Kutta is described in Eq. (3) below.

$$S_{s,t+1}^{(h)} = S_{s,t}^{(h)} + \frac{1}{6}\left(K_{11}+2K_{21}+2K_{31}+K_{41}\right)$$

$$I_{s,t+1}^{(h)} = I_{s,t}^{(h)} + \frac{1}{6}\left(K_{12}+2K_{22}+2K_{32}+K_{42}\right)$$

$$R_{s,t+1}^{(h)} = R_{s,t}^{(h)} + \frac{1}{6}\left(K_{13}+2K_{23}+2K_{33}+K_{43}\right)$$

$$S_{s,t+1}^{(v)} = S_{s,t}^{(v)} + \frac{1}{6}\left(K_{14}+2K_{24}+2K_{34}+K_{44}\right)$$

$$I_{s,t+1}^{(v)} = I_{s,t}^{(v)} + \frac{1}{6}\left(K_{15}+2K_{25}+2K_{35}+K_{45}\right) \tag{3}$$

The Bayesian approach estimates the parameters of models (2) and (3). The Bayesian paradigm is widely used to analyze complex statistical models [19]. In the Bayesian paradigm, the observation data is assumed to have a distribution. The parameters of models in the Bayesian issue have uncertain properties. Therefore, the parameters will have a prior distribution [16]. Obtaining posterior distribution requires previous distribution information and the likelihood function. The work of [19] has described some prior distributions known in Bayesian, namely conjugate, non-conjugate, informative, and non-informative prior.

Table 1
Summary of parameter estimation for both models, 80,000 burn-in 10,000 iterations.

Node	Mean	Standard Deviation	MC error	2.5%	Median	97.5%	Sample
Euler Method							
betaH0	1.077	0.454	0.027	0.257	1.046	2.094	10,000
betaH	0.500	0.526	1.180	0.412	0.480	0.512	10,000
betaV	0.003	0.001	0.002	0.417	0.003	0.005	10,000
muH	0.004	0.002	0.025	0.430	0.004	0.020	10,000
muV	0.515	0.035	0.003	0.460	0.540	0.590	10,000
deviance	463.4	3.350	0.220	458.5	462.9	470.9	10,000
Four order Runge-Kutta							
betaH0	0.979	0.462	0.018	0.126	0.968	1.917	10,000
betaH	0.457	0.536	0.502	0.378	0.464	0.508	10,000
betaV	0.002	0.001	0.024	0.319	0.002	0.005	10,000
muH	0.003	0.002	0.025	0.330	0.003	0.006	10,000
muV	0.520	0.040	0.004	0.450	0.520	0.580	10,000
deviance	106.5	3.660	0.042	99.40	106.6	113.9	10,000

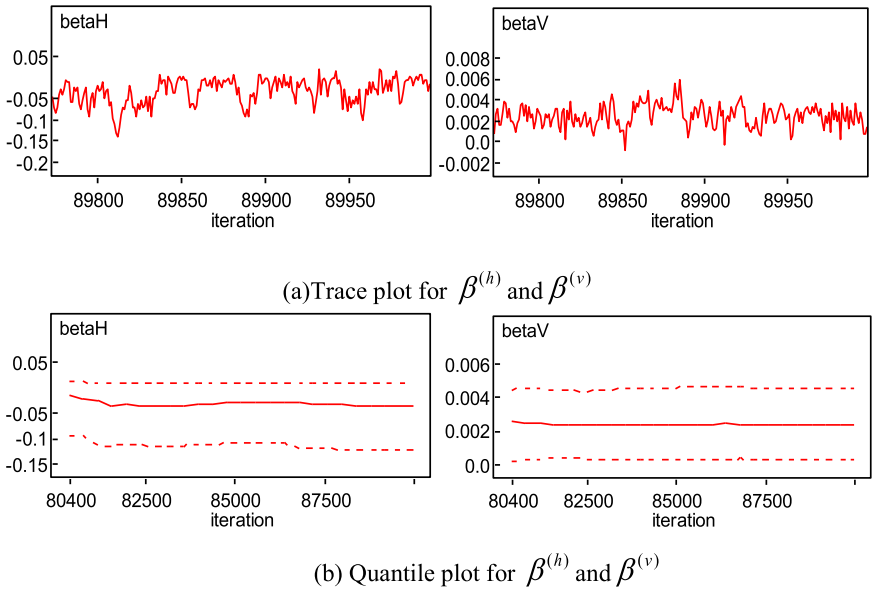

Fig. 3. Samples of parameter estimation of the model (1) for four-order RungeKutta, 10,000 iterations and burn-in 80,000.

Materials and methods

DHF data was used as a feeder to measure the approximation accuracy of Euler and the four-order Runge-Kutta. DHF data used as feeder to measure the approximation accuracy of Euler and the four-order Runge-Kutta. The data set used in this study is the monthly DHF cases in 10 Kendari, Indonesia districts, which are recorded every day from 2019 to 2021. Mandonga, Baruga, Puuwatu, Kadia, Wua-Wua, Poasia, Abeli, Nambo, Kambu, Kendari, and Kendari Barat. The population density of Kendari is approximately 1.364 people per square kilometer (km^2). The highest rainfall rate ranges from 3000 mm^3 to 3030 mm^3, usually from January–April. Therefore, the number of infective mosquitoes is expected to increase during the season. The rainfall data is obtained from BMKG Kendari. Rainfall is an essential element in supporting the breeding of mosquito larvae [25]. The incidence of dengue was positively associated with increased rainfall [20].

Infected mosquito data are calculated from newly infected based on the SIR-SI model. To estimate the total mosquito population, we use the research conducted by [24,25] $N_{s,t}^{(v)} \approx 8.6892 N_{s,t}^{(h)}$. The estimate is based on the SIR-SI model for dengue disease transmission where the starting values are set. Here, some of the estimations are based on information taken from an article by [25], $I_{s,0}^{(v)} = 0.0557 \times S_{s,0}^{(h)}$. The number of alternative hosts available as the blood source m is assumed to be zero and b is 2.33 [4].

Numerical computations are performed using WinBUGS software. This package is open source to carry out Markov Chain Monte Carlo (MCMC) analysis of Bayesian models [16]. The observations random sample of (model 2), $\lambda_{s,t,k}^{(h)}$ for $k = 1, 2, \ldots, n$ is generated from the posterior distribution of infected humans $\lambda_{s,t}^{(h)}$. The posterior expected mean number is approximated by $\tilde{\lambda}_{s,t}^{(h)} = \frac{1}{n} \sum_{k=1}^{n} \lambda_{s,t,k}^{(h)}$.

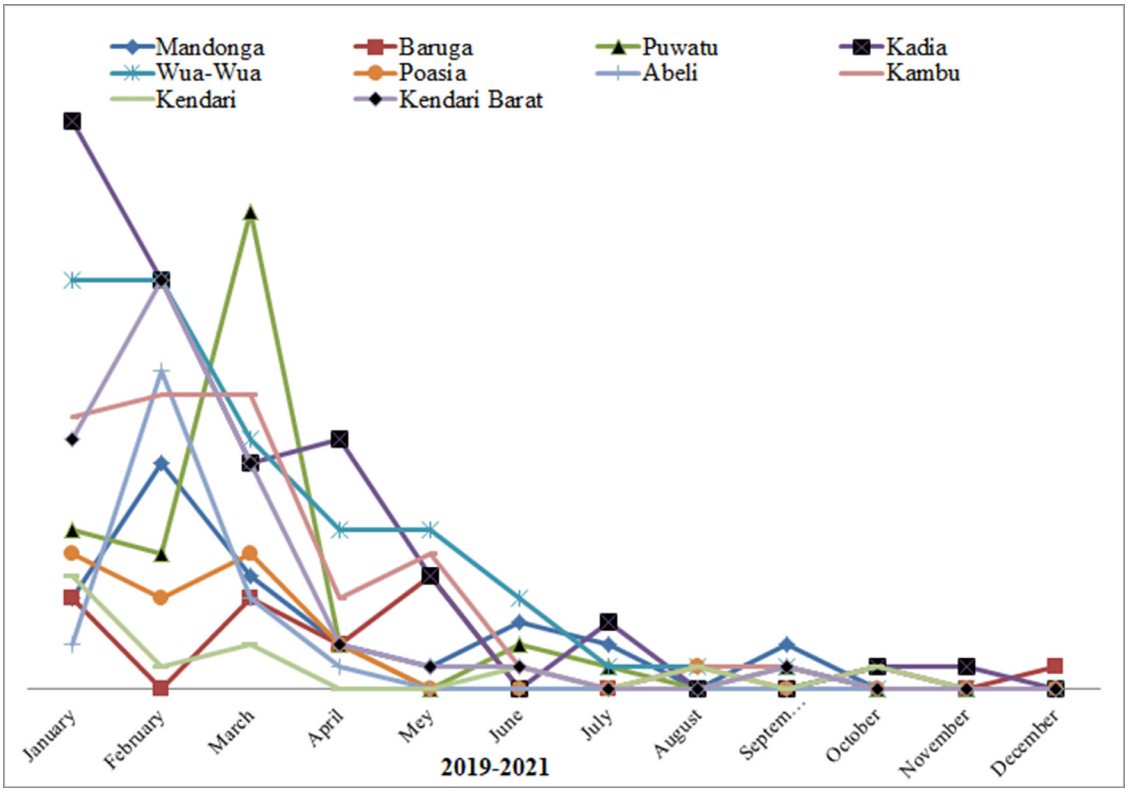

Fig. 4. Monthly incident rate.

The relative risk parameter or $\theta_{s,t}^{(h)}$ is written $\theta_{s,t}^{(h)} = \frac{\lambda_{s,t}^{(h)}}{e_{s,t}^{(h)}}$. Therefore, the posterior expected of DHF relative risk is defined as

$$\tilde{\theta}_{s,t}^{(h)} = \frac{1}{n}\sum_{k=1}^{n}\theta_{s,t,k}^{(h)} = \frac{1}{n}\sum_{k=1}^{n}\frac{\lambda_{s,t,k}^{(h)}}{e_{s,t}^{(h)}} = \frac{1}{n}\sum_{k=1}^{n}\frac{\tilde{\lambda}_{s,t}^{(h)}}{e_{s,t}^{(h)}}.$$

The Bayesian approach to model calibration and simulation is described in Fig. 2. The prior distribution is updated by conditioning on observed data and the likelihood of the model. The resulting conditional distribution is the posterior distribution which is generally approximated via Markov chain Monte Carlo (MCMC) sampling. MCMC methods are tools for simulating data from an arbitrary distribution of Bayesian frames. The interest in MCMC methods lies in that the preceding samplers are usable for sampling from a wide range of distributions. Despite this quasi-universality, problems can arise during the algorithm implementation, in particular numerical overflows, slowness of the runs, or difficulties in convergence assessment. The principle of MCMC methods is to create a random walk in the parameters space, which converges to a stationary distribution, the joint posterior distribution. Whatever the method used, an important feature is to check that the algorithm has been run long enough to ensure the simulated sample is representative of the target distribution. Consequently, the first iterations are usually not used for inference and are considered burning time. Moreover, the simulated values are not independent, thus be needed to obtain acceptable accuracy.

Results and discussion

The Parameter estimation of the model (1) based on the Euler and four order Runge-Kutta method was updated using Markov Chain Monte Carlo (MCMC) Bayesian. The Euler and four-order Runge-Kutta methods have converged at 10,000 iterations with burn-in 80,000. The history of the simulation process is cut at 80,000 iterations, and then the simulation process is carried out 10,000 times. This convergence process has fulfilled the properties of MCMC and is reinforced by historical plots. For example, trace and quantile plots show that the historical method for 10,000 times is in the same zone (see Fig. 3). A summary of parameter estimation for model (1) is presented in Table 1. The numerical simulation results show that the four-order Runge-Kutta approach has the slightest deviance, 106.5. The deviance is widely used in Bayesian modeling of goodness-of-fit measures in statistics to compare fitted models. Several methods can be used to assess goodness-of-fit, including chi-square statistics, Akaike information criterion, Bayesian information criterion, deviance, and posterior predictive loss. In this study, we use the deviance because it is readily available in WinBUGS software and identifies weaknesses with other measures, particularly for models with several random effects. The model with the smallest deviance is the best one. The DHF relative risk analysis for the next step is referring the estimation results of the

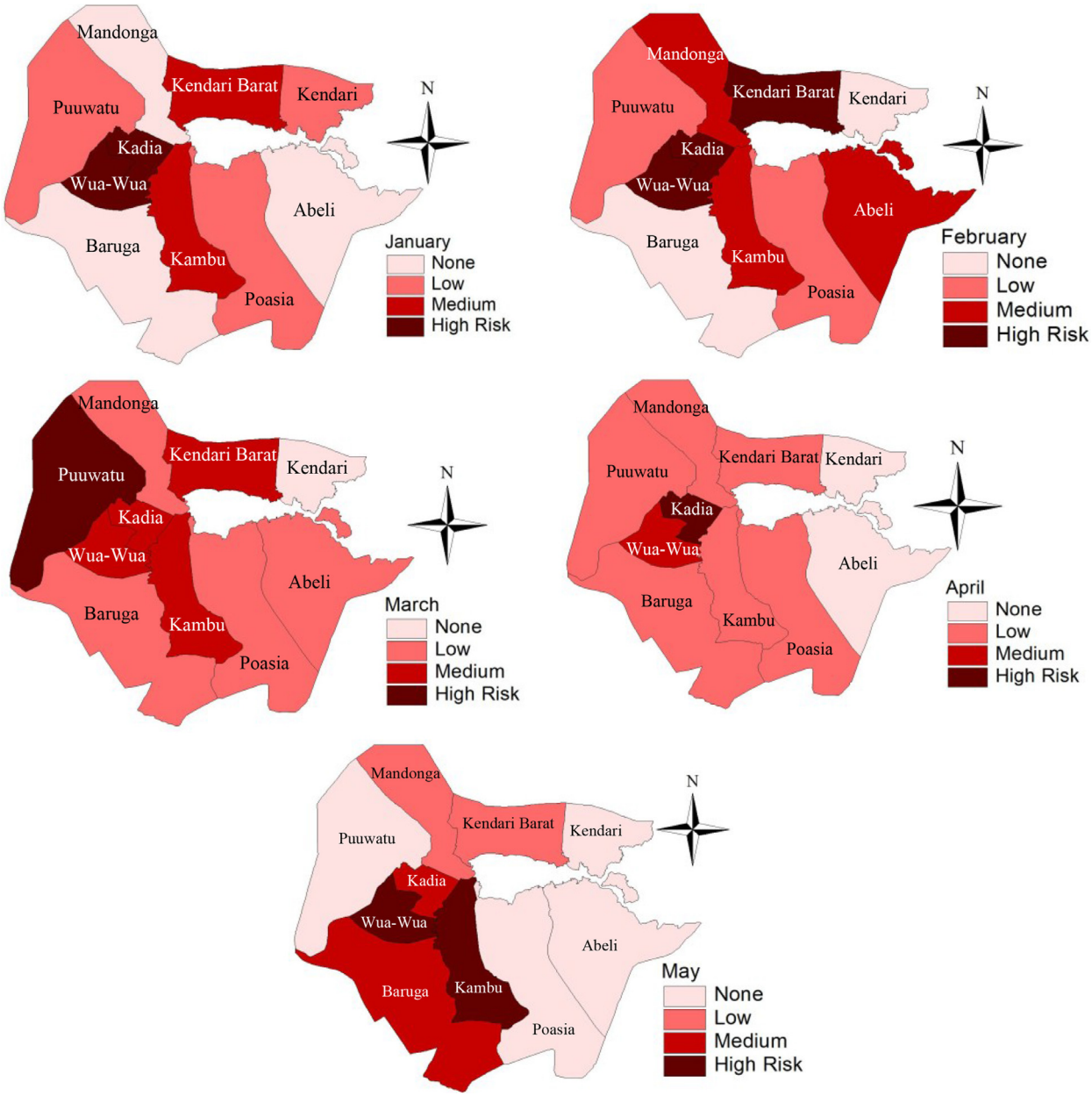

Fig. 5. The relative risk maps using four order Runge-Kutta of the SIR-SI Model.

four-order Runge-Kutta method. Figure 3 shows samples of parameter estimation for $\beta^{(h)}$and $\beta^{(v)}$. Other parameter estimates,$\mu^{(h)}$ and $\mu^{(v)}$, represent the human birth rates or human birth and mosquito death proportions, respectively. The ${\beta_0}^{(h)}$ and ${\beta_0}^{(v)}$ represent intercepts of humans and mosquitoes are given in Table 1. Because all parameters are proportions means that these parameters have no units or, in other words, are constants.

The relative risk analysis shows that the dynamics of DHF cases fluctuate from January to July every year. An incidence rate of monthly aggregate data from 2019 to 2021 confirms the dynamics of DHF from January to July (see Fig. 4). This whole time represents the dynamics of the spread of dengue each year. However, from January to May, there was a high consistency of DHF cases. In this study, two districts with high case consistency were found, namely Kadia and Wua-Wua districts. The incidence rate in Figure 4 is the ratio of the number of cases to the total time of the population per 10,000 persons. This incidence rate measures the dynamics and risk of DHF occurrence in each district in Kendari City. As an illustration, in Kadia District in January, there were a dynamic average of 25 cases per 10,000 persons during the 2019-2021 period. Then the dynamics of DHF issues sloped closer to 0 in December. In contrast, in the Puwatu District for January, there was an average dynamic of 7 cases per 10,000 persons during

the 2019-2021 period. But instead, the dynamics of DHF increased strictly in March by around 21 cases per 10,000 persons, then decreased to 0 from April to December. For ten districts in Kendari City from January to May, DHF cases fluctuated during the 2019-2021 period. The incidence rate only captures the dynamics or trends of DHF cases and cannot describe the DHF risk. Therefore, the incidence rate is used as initial information to analyze DHF risk.

Furthermore, because the spread of DHF cases has a spatial effect, the Kadia and Wua-Wua districts need serious attention to suppress the rate of reach of DHF in Kendari City. Intensive observation and intervention in the Kadia and Wua-Wua districts should be carried out in early January to stop the breeding of mosquito larvae. In other neighbourhoods such as Puuwatu, Kambu and Kendari Barat, the DHF cases occur temporally as the impacts of DHF cases in the Kadia and Wua-Wua districts (see Fig. 5).

Conclusion

Dengue Haemorrhagic Fever (DHF) is an ancient disease having a vast social, economic, and health burden. Even though the DHF has been investigated for years, it remains a significant public health problem, especially in tropical countries. Many researchers in infectious disease were among the foremost to realize the crucial role of statistical models in providing an explicit framework for understanding the disease transmission dynamics between the host (human) and vector (mosquito). Depending on the type of infectious disease, there are many stochastic variations in the compartment structure. This research provides a Bayesian stochastic process for cross-infection or the Bayesian SIR-SI model. The model is involved local and global random effects. A Runge-Kutta family methods are used to discretize the model, namely Euler and four-order Runge-Kutta methods. For comparison purposes, we use a monthly DHF dataset of 10 Kendari, Indonesia, districts from 2019–2021. Parameter estimation of the Bayesian SIR-SI model based on the Euler and four-order Runge-Kutta method was updated using Markov Chain Monte Carlo (MCMC) Bayesian. The Euler and four-order Runge-Kutta methods have converged at 10,000 iterations with a burn-in of 80,000. The numerical simulation results show that the four-order Runge-Kutta approach has minor deviance, 106.5. Therefore, this approach is the best one. The relative risk analysis shows that the dynamics of DHF cases fluctuate from January to July every year.

However, from January to May, there was a high consistency of DHF cases. Two districts with high case consistency were found, namely Kadia and Wua-Wua. Furthermore, because the spread of DHF cases has a spatial effect, the Kadia and Wua-Wua districts need serious attention to suppress the rate of a stretch of DHF in Kendari City. Intensive observation and intervention in the Kadia and Wua-Wua districts should be carried out in early January to suppress the breeding of mosquito larvae. In other neighbourhoods such as Puwatu, Kambu, and Kendari Barat, the DHF cases occur temporally as the impacts of DHF cases in the Kadia and Wua-Wua districts. Further research will apply Euler-Maruyama and nonstandard methods for discretization. In general, the fitting of these models has been possible because of the availability of different computational techniques, the most notable being Markov chain Monte Carlo (MCMC). For large models or big data sets, MCMC can be tedious, and reaching the required samples can take a very long time. Also, autocorrelation may occur, and more iterations may be required. Alternatively, the posterior distributions of the parameters might be approximated in some way. However, most models are highly multivariate, and resembling the full rear distribution may not be possible in practice. The integrated nested Laplace approximation (INLA) focuses on the posterior margins for latent Gaussian models. This also means that INLA will be handy when only marginal inference on the model parameters is required.

Funding

Authors wishes to acknowledge the financial support from the KEMENDIKBUD RISTEK of Indonesia through grant SP-DIPA-023.17.1.690523/2022.

Declaration of competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We thank also to the Department of Health, BPS, and BMKG of Kendari for their permission to use their observation data in this study. Also, we wish to thank the anonymous reviewers for their comments and criticisms. All of their comments were taken into account in the revised version of the paper, resulting in a substantial improvement with respect to the original submission.

References

[1] X.Deparis A. Tran, P. Dussart, J. Morvan, P. Rabarison, F. Remy, L. Polidori, J. Gardon, Dengue spatial and temporal patterns, French Guiana 2001, Emerg. Infect. Dis. 10 (4) (2004) 615–621 2024.
[2] A. Rohani, I. Asmaliza, S. Zainah, H.L. Lee, Detection of dengue from feld aedes aegypti and aedes albopictus adults and larvae, Southeast Asian J. Trop. Med. Public Health 28 (1) (1997) 138–142.
[3] D.J. Gubler, Epidemic dengue haemorrhagic fever as a public health, social and economic problem in the 21st century, Trends Microbiol. 10 (2) (2002) 100–103.

[4] N.A. Samat, D.F. Percy, Vector-borne infectious disease mapping with stochastic difference equations: an analysis of dengue disease in Malaysia, J. Appl. Stat. 39 (9) (2012) 2029–2046.
[5] A.S.Ahmar Mukhsar, M.A. El Safty, H. El-Khawaga, M. El Sayed, Bayesian convolution for stochastic epidemic model, Intell. Autom. Soft Comput. 34 (2) (2022) 1175–1186 Tech Science Press, vol.
[6] Mukhsar, The Bayesian zero-inflated negative binomial (t) spatio-temporal model to detect an endemic DHF location, Far East J. Math. Sci. 109 (2) (2018) 357–372.
[7] B.Abapihi Mukhsar, A. Sani, E. Cahyono, P. Adam, F.A. Abdullah, Extended convolution model to Bayesian spatio-temporal for diagnosing DHF endemic locations, J. Interdiscip. Math. 19 (2) (2016) 233–244.
[8] A. Sani, B. Abapihi, Mukhsar, Kadir, Relative risk analysis of dengue cases using convolution extended into spatio-temporal model, J. Appl. Stat. 42 (11) (2015) 2509–2519.
[9] L. Esteva, C. Vargas, Analysis of a dengue disease transmission model, Math. Biosci. 150 (2) (1998) 131–151.
[10] A. Raza, M.S. Arif, M. Rafiq, A reliable numerical analysis for stochastic dengue epidemic model with incubation period of virus, Adv. Differ. Equ. 2019 (32) (2019) 1–19.
[11] K.E. Macías-Díaz, A. Raza, N. Ahmed, M. Rafiqe, Analysis of a nonstandard computer method to simulate a nonlinear stochastic epidemiological model of coronavirus-like diseases, Comput. Methods Programs Biomed. 2021 (204) (2021) 1–10.
[12] N.Iriawan Mukhsar, B.S.S. Ulama, Sutikno, New look for DHF relative risk analysis using Bayesian poisson-lognormal 2-level spatio-temporal, Int. J. Appl. Math. Stat. 47 (17) (2013) 39–46.
[13] L. Knorr-Held, J. Besag, Modelling risk from a disease in time and space, Stat. Med. 17 (18) (1998) 2045–2060.
[14] D. Clancy, P.D. O'Neill, Bayesian estimation of the basic reproduction number in stochastic epidemic models, Bayesian Anal. 3 (4) (2008) 737–758.
[15] U. Fatima, D. Baleanu, N. Ahmed, S. Azam, A. Raza, M. Rafiq, M.A. Rehman, Numerical study of computer virus reaction diffusion epidemic model, Comput. Mater. Contin. 66 (3) (2021) 3183–3194.
[16] B.A. Lawson, Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology, CRC Press, Chapman&Hall, 2008 [online] https://www.routledge.com/Bayesian-Disease-Mapping-Hierarchical-Modeling-in-Spatial-Epidemiology/Lawson/p/book/9780367781224 .
[17] D. Clancy, A stochastic SIS infection model incorporating indirect transmission, J. Appl. Probab. 42 (3) (2005) 726–737.
[18] J.D. Sanders, J.L. Talley, A.E. Frazier, B.H. Noden, Landscape and anthropogenic factors associated with adult aedes aegypti and aedes albopictus in small cities in the Southern Great Plains, Insects 11 (699) (2020) 1–20.
[19] A.B. Lawson, W.J. Browne, V.C.L. Rodeiro, Disease Mapping With WinBUGS and MLwiN, John Wiley &Sons, Chichester, UK, 2003 [online] https://onlinelibrary.wiley.com/doi/book/10.1002/0470856068 .
[20] L.A. Waller, B.P. Carlin, H. Xia, A.E. Gelfand, Hierarchical spatio temporal mapping of disease rates, J. Am. Stat. Assoc. 92 (438) (1997) 607–617.
[21] K.A. Kabir, K. Kuga, J. Tanimoto, Analysis of SIR epidemic model with information spreading of awareness, Chaos Solit. Fractals 119 (2019) 118–125.
[22] F. Li, X. Meng, X. Wang, Analysis and numerical simulations of a stochastic SEIQR epidemic system with quarantine-adjusted incidence and imperfect vaccination, Comput. Math. Methods Med. 2018 (2018).
[23] A. Miao, T. Zhang, J. Zhang, C. Wang, Dynamics of a stochastic sir model with both horizontal and vertical transmission, J. Appl. Anal. Comput. 8 (4) (2018) 1108–1121.
[24] N. Gao, Y. Song, X. Wang, J. Liu, Dynamics of a stochastic sis epidemic model with nonlinear incidence rates, Adv. Differ. Equ. 2019 (1) (2019) 41.
[25] H. Nishiura, Mathematical and statistical analysis of the spread of dengue, Dengue Bull. 30 (2006) 51–67.