Bayesian Spatial Survival Lognormal 3 Parameter Models for Event Processes Dengue Fever in Tuban

Fetrika Anggraini and Nur Mahmudah

Abstract—**Dengue Haemorrhagic Fever (DHF) is a disease caused by the dengue virus transmitted by the Aedes Aegypti mosquito. This disease often spreads in residential areas every year. DHF is a major health problem because it can affect all age groups and cause death, especially in children. Some of the triggers for the spread of DHF are the area geographical conditions, and people's knowledge and awareness of environmental hygiene. The occurrence of DHF in a certain area is caused by the spread of mosquitoes there, so there is a possibility that other areas will be affected too. Therefore, to determine the factors affecting DHF patients' recovery rate based on the location where the patients seek treatment, the most fitting model is Spatial Survival with Bayesian MCMC method. This study aims to understand the predicting factors of DHF recovery rate based on the patient's residence (W), such as length of hospitalization (Y), sex (X_1), age (X_2), patient participation (X_3), hematocrit level (X_4), thrombocyte count (X_5), haemoglobin count (X_6), body weight (X_7) and patient's medical record (X_8) using lognormal 3 parameter distribution with normal random effect. The result shows that sex (X_1), age (X_2), patient participation (X_3), hematocrit level (X_4), thrombocyte count (X_5) and patient medical record (X_8) are significant factors that affect DHF recovery rate.**

Index Terms—**Bayesian; DBD; Lognormal 3 Parameter; MCMC-Gibbs Sampling; Spatial-Survival.**

I. INTRODUCTION

DENGUE hemorrhagic fever (hereafter abbreviated as DHF) is a disease caused by dengue virus transmitted by Aedes Aegypti mosquito [1]. Mosquito is one of the dangerous animal species in the spread of this virus [2]. Its human to human transmission is a major cause of death [3] mainly in children under 15 years old [4]. The recovery period for DHF usually ranges from 5 to 7 days [5]. Therefore, a model should be used to understand the recovery development [6]. A number of studies on this disease confirm its mortality rate.The rate continues to increase, resulting in higher costs for treatment, management, and medication [7]. There have been many studies on DHF, but it seems that spatial survival analysis is the most fitting method to use [8]. This analysis has been widely applied in the health sector and is known by various terms, such as event history analysis, in other sectors [9]. Survival analysis is a mathematical model for analyzing data in which the response variable is produced by the time until an event occurs [10]. Its aim is to identify risk factors of the incidence [11]. Based on this information, a researcher may aim to determine the predicting factors of a thing or event with risk factors for the occurrence versus time, hence a survival model of a tool will be more adequate [12]. In its development, survival analysis modelling also includes random effects to overcome the heterogeneity/sources of variance that cannot be explained [13]. In survival cases, the timing of an event often depends on the location [14]. Spatial survival analysis is a hazard function to estimate the probability of an object experiencing an event at t-time based on location effects [15]. It is called spatial factor because an event is often related to where it takes place and is influenced by location factor [16]. This factor takes into consideration the closest surrounding areas because they possibly have similar characteristics [13]. To determine the spatial dependence on the random effect of adjacent areas, Bayesian Markov Chain Monte Carlo (MCMC) approach can be used [14]. The research conducted by [16] applied spatial survival model on DHF incidence in Makassar by modeling the patients' length of hospitalization until they discharge for improved condition or recovery, and identified any censored or failed data. Survival model is also used in medical events that trigger death cases which consider spatial effects [17]. A similar study was carried out by [13] who applied spatial survival model into political science. In this field, death does not refer to real death but the survival time of a unit before undergoing a certain political event based on location factors [14]. Previous studies regarding spatial survival models that involve spatial effects assumed that the model would produce good estimates if the survival data in all locations were assumed to have a similar particular distribution [8]. In fact, not all survival data distribution in each location exhibits a clear distribution [18]. Therefore, this study examines the distribution of survival data, which is the lognormal 3 parameter distribution. It is expected that by considering the lognormal 3 parameter distribution and including spatial random effects in the model, the previously unexplained heterogeneity can be explicated. It can be concluded that the result of this study is a model that can be used to determine the factors predicting DHF recovery rate based on where the patients receive treatment. The result can be presented in DHF management dissemination to reduce the number of cases in Tuban Regency.

Manuscript received November 25, 2020; revised June 25, 2021. This work was supported in part by the Penelitian Dosen Pemula (PDP), Research from the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia (KEMENRISTEKDIKTI-BRIN) 2020 under Grant 101/SK/LPPM-UNUGIRI/III/2020.

Fetrika Anggraeni is a Lecturer in the Department of Physics, Universitas Nahdlatul Ulama Sunan Giri Bojonegoro, Jalan Jendral Ahmad Yani No.10, Bojonegoro, Jawa Timur, Indonesia (Corresponding author mail: fetrikaanggraini@gmail.com).

Nur Mahmudah is a Lecturer in the Department of Physics, Universitas Nahdlatul Ulama Sunan Giri Bojonegoro, Jalan Jendral Ahmad Yani No.10, Bojonegoro, Jawa Timur, Indonesia (mahmudahnur437@gmail.com).

II. BAYESIAN MCMC-*GIBBS SAMPLING*

Bayesian modelling is based on posterior model which combines past data as prior information and observational data as likelihood function construction [19]. Bayesian can overcome the spatial autocorrelation of random effect of time data until an event occurs in adjacent areas [20]. The estimator in the Bayesian approach is the mean or mode of the posterior distribution [21]. Bayesian is highly complicated because of its simulation method which combines Monte Carlo with Markox Chain properties to obtain sample data based on certain sampling scenario [22]. The following are the steps of Monte Carlo Markox Chain simulation method [23]:

1) Determine the initial value.
2) Iterate sample as many as K.
3) Observe the convergence of the sample data.
4) Carry out the burn-in process by removing the first sample as many as B. The period will end when equilibrium condition is reached.
5) Use a parameter as a sample for posterior analysis.
6) Create a posterior distribution plot.
7) Summarize the posterior distribution such as the mean, median, standard deviation, and standard error.

The previously explained posterior distribution is complicated and difficult to be solved manually, so the parameter estimation can be done using Gibbs Sampling [24]. The steps of Gibbs Sampling algorithm process are [25]:

1) Determine the initial value of each parameter.
2) Carry out the simulation process after the initial values are gained.
3) Construct the parameter and save them as a set of values iterated by $(r+1)$from the algorithm.
4) Obtain the result summary of posterior distribution.

III. SPATIAL SURVIVAL ANALYSIS OF LOGNORMAL 3 PARAMETER

Spatial statistics is a statistical method used to analyze spatial data [26]. This method can be used in various fields, such as economics, social, health, meteorology, and climatology [27]. Spatial data are data that contain the "location" information, so it is not only about "what" is measured but also where the data are obtained, and the measurement uses spatial autocorrelation [28].

Survival analysis is a collection of statistical procedures for analyzing data derived from time response variable [29]. It has three functions, those are the survival function, the probability density function or the cumulative of data distribution, and the Hazard function [30]. The Cox survival model is semi parametric because it does not require distribution information which underlies survival time and the baseline hazard function does not have to be determined to estimate the parameter [31].

A spatial survival model is formed by arranging survival data based on adjacent areas in the W_i frailties. Those areas possibly have similar characteristics or same level of risk (hazard) compared to more distant areas [32]. Frailty spatial survival model asserts that random effect has normal distribution which can uncover the spatial autocorrelation cannot be explained in the model [33]. It is loaded with W_i which has normal distribution with smoothing parameter λ [16]. In this research, the distribution of DHF hospitalization period follows lognormal 3 parameter distribution (μ, σ, γ). This probability density function of this distribution is as follows [34]:

$$f(t;\mu,\sigma,\gamma) = \frac{1}{(t-\gamma)\sigma\sqrt{2\phi}} exp\left\{-\frac{[ln(t-\gamma)-\mu]^2}{2\sigma^2}\right\} \tag{1}$$

where $t > \gamma \geq 0, -\infty < \mu < \infty, \sigma > 0$, and γ are location parameters. If t is a response variable with lognormal 3 parameter distribution, then $y = ln(t-\gamma)$ has normal distribution with μ mean and variance. When $\gamma = 0$, the distribution changes into lognormal 2 parameter. Lognormal distribution transformed into standardized normal distribution can be obtained as follows [35]:

$$f(t) = \frac{P\left(Z = \frac{ln(t-\gamma)-\mu}{\sigma}\right)}{\sigma(t-\gamma)} \tag{2}$$

The equation of cumulative distribution function in lognormal 3 parameter distribution or $F(t)$ is as follows [36]:

$$F(t) = \int_0^t \frac{1}{(u-\gamma)\tau\sqrt{2\pi}} exp\left\{-\frac{[ln(u-\gamma)-\mu]^2}{2\sigma^2}\right\} du \tag{3}$$

Based on 3, survival function of lognormal 3 parameter as follows is obtained [37]:

$$\begin{aligned} S(t) &= 1 - F(t) \\ &= 1 - \int_0^t \frac{1}{(u-\gamma)\sigma\sqrt{2\pi}} exp\left\{-\frac{[ln(u-\gamma)-\mu]^2}{2\sigma^2}\right\} du \\ &= P\left[Z > \frac{ln(t-\gamma)-\mu}{\sigma}\right] \end{aligned} \tag{4}$$

Meanwhile, the *hazard* function of lognormal 3 parameter distribution can be determined through [38]:

$$\begin{aligned} h(t) &= \frac{f(t)}{S(t)} \\ &= \frac{P\left(Z = \frac{ln(t-\gamma)-\mu}{\sigma}\right)}{\sigma(t-\gamma)\left\{P\left[Z > \frac{ln(t-\gamma)-\mu}{\sigma}\right]\right\}} \end{aligned} \tag{5}$$

The general Cox regression in 5 could form the following lognormal 3 parameter model [39]:

$$\begin{aligned} h(t,X) &= h_0(t)exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \\ &\quad \beta_p X_p + \mathbf{W}_i) \\ &= \frac{\mu}{\sigma(t-\gamma)} \end{aligned} \tag{6}$$

Next, $h_0(t)$ is a function which value depends on t while is independent from t. Thus, parameter μ can be represented in the following: $\mu = (\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \mathbf{W}_i)$ and baseline hazard in the following [40]:

$$h(t) = \frac{1}{\sigma(t-\gamma)} \tag{7}$$

Therefore, the hazard function is [41]:

$$\begin{aligned} h(t) &= \frac{\mu}{\sigma(t-\gamma)} \\ &= \frac{1}{\sigma(t-\gamma)} exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \\ &\beta_p X_p + \mathbf{W}_i) \end{aligned} \tag{8}$$

Estimation of each parameter is obtained from the full conditional distribution of each parameter σ, γ, and β_i with prior distribution determination beforehand [42]. The prior distribution used is the combination between conjugate and informative priors as follows [19]:

$$\begin{aligned} &y \sim Lognormal(\mu, \tau, \gamma) \\ &\mu = \beta^{\mathbf{T}} \mathbf{X}_{ij} + \varepsilon_i, \qquad \varepsilon_i | \varepsilon_{-i}, W_i \sim Normal(a, b), \\ &\beta \sim Normal(v, w) \gamma \sim Gamma(r, s), \\ &\tau \sim Gamma(r, s) \end{aligned} \tag{9}$$

The full conditional distribution of each parameter τ, γ and β_{1+i} and λ is solved by calculating the integrals of the related parameter as follows [8]:

$$\begin{aligned} p(\sigma|\gamma,\lambda,\beta_{1+i}) &\propto \int_{\tau}\int_{\lambda}\int_{\beta_1}\cdots\int_{\beta_{1+p}} I(t|\gamma,\lambda,\beta_1,\ldots\beta_p) \\ &\quad p(\gamma)p(\lambda)p(\beta_1)\ldots p(\beta_p) \\ &\quad d\gamma d\lambda d\beta_1 \ldots d\beta_p \\ p(\gamma|\sigma,\lambda,\beta_{1+i}) &\propto \int_{\gamma}\int_{\lambda}\int_{\beta_1}\cdots\int_{\beta_{1+p}} I(t|\sigma,\lambda,\beta_1,\ldots\beta_p) \\ &\quad p(\sigma)p(\lambda)p(\beta_1)\ldots p(\beta_p) \\ &\quad d\sigma d\lambda d\beta_1 \ldots d\beta_p \\ p(\lambda|\sigma,\gamma,\beta_{1+i}) &\propto \int_{\lambda}\int_{\gamma}\int_{\beta_1}\cdots\int_{\beta_{1+p}} I(t|\sigma,\gamma,\beta_1,\ldots\beta_p) \\ &\quad p(\sigma)p(\gamma)p(\beta_1)\ldots p(\beta_p) \\ &\quad d\sigma d\gamma d\beta_1 \ldots d\beta_p \\ p(\beta_1|\sigma,\gamma,\lambda,\beta_{1+i}\neq 1) &\propto \int_{\tau}\int_{\gamma}\int_{\lambda}\int_{\beta_2}\cdots\int_{\beta_{1+p}} I(t|\sigma,\gamma,\lambda,\beta_2,\ldots\beta_p) \\ &\quad p(\sigma)p(\gamma)p(\lambda)p(\beta_2)\ldots p(\beta_p) \\ &\quad d\sigma d\gamma d\lambda d\beta_2 \ldots d\beta_p \\ &\vdots \\ p(\beta_p|\sigma,\gamma,\lambda,\beta_{1+i}\neq p) &\propto \int_{\tau}\int_{\gamma}\int_{\lambda}\int_{\beta_1}\cdots\int_{\beta_{1+p}} I(t|\sigma,\gamma,\lambda,\beta_1,\ldots\beta_{p-1}) \\ &\quad p(\sigma)p(\gamma)p(\lambda)p(\beta_1)\ldots p(\beta_{p-1}) \\ &\quad d\sigma d\gamma d\lambda d\beta_1 \ldots d\beta_{p-1} \end{aligned} \tag{10}$$

The parameter estimation of Bayesian spatial survival model with lognormal 3 parameter distribution used MCMC Algorithm and Gibbs Sampling [20]. The parameter update process in the model was carried through Gibbs Sampler based on full conditional distribution sample obtained from equation [43]. The elaborated posterior distribution is quite complicated and difficult to be manually solved, so the parameter estimation was done through Gibbs Sampling [44]. The parameter estimation is presented below [45]:

1) Determine the initial value or estimation of each parameter.
   $(\sigma^0, \gamma^0, \lambda^0, \beta_1^0, \ldots \beta_p^0)$
2) Then, a random listing is obtained.
   γ^1 from $p(\gamma|t, \sigma^0, \lambda^0, \beta_1^0, \ldots \beta_p^0)$
   σ^1 from $p(\sigma|t, \gamma^0, \lambda^0, \beta_1^0, \ldots \beta_p^0)$
   λ^1 from $p(\lambda|t, \gamma^0, \sigma^0, \beta_1^0, \ldots \beta_p^0)$
   β_1^1 from $p(\beta_1^1|t, \gamma^0, \sigma^0, \lambda^0, \beta_2^0, \ldots \beta_p^0)$
   $\vdots$
   β_p^1 from $p(\beta_p^1|t, \gamma^0, \sigma^0, \lambda^0, \beta_2^0, \ldots \beta_{p-1}^0)$
3) Iterate the second step until convergence is achieved (sample for model parameter inference is adequate).

IV. METHOD

This study used secondary data of DHF hospitalization record of patients' condition in Koesoma Tuban Hospital which involves spatial/lattice factors. The data taken is the length of hospitalization until the patients were permitted to discharge, which is called the Failure event. The medical recap time was from January 1, 2019 to July 1, 2020. Spatial factors were represented by the proximity of one location to another. Figure 1 presents the map of Tuban Regency.

The variables used were length of hospitalization (Y), sex (X_1), age (X_2), patient participation (X_3), hematocrit level (X_4), thrombocyte count (X_5), hemoglobin count (X_6), body weight (X_7), and medical history (X_8). The Table I presents the description of responses and the predicting variables.

The following are the steps to carry out Bayesian spatial survival analysis with lognormal 3 parameter distribution:

1) Assess the lognormal 3 parameter survival model by taking into account the influence of location (spatial) using a Bayesian approach with the following steps: add spatial random effects to the proportional hazard model and determine the prior distribution, or joint distribution, and determine the estimation of model parameters using MCMC and Gibbs sampling.
2) Determine the frailty model with normal distribution of DHF patients in Tuban Regency based on the factors affecting the recovery rate by following this sequence of steps: signify the spatial load by inputting Tuban Regency map into WinBUGS package program, determine the syntax of the random effects, perform "spatial autocorrelation" testing using Moran's I statistical test, and test the distribution of survival time data.
3) Determine the model and survival parameters with a lognormal 3 parameter distribution using Markov Chain Monte Carlo (MCMC) and Gibb Sampling simulations.
4) Determine the model and survival parameters with a lognormal 3 parameter distribution using Markov Chain Monte Carlo (MCMC) and Gibb Sampling simulations.
5) Define the mean and variance of the spatial random effect distribution.
6) Generate T sample θ^1, θ^2, $ldots \ldots \theta^T$ from the posterior distribution $p(\theta|x)$ by updating T as many as n times with sufficient thin so that the Marcov Chain process is fulfilled. The convergent algorithm is described as a state when the algorithm has reached stationary condition in the lognormal 3 parameter posterior distribution.
7) Summarize the posterior distribution (mean, median, standard deviation, MC error, and 95% confidence interval) on the lognormal 3 parameter distribution.

Fig. 1: Map of Tuban Regency

.

TABLE I: Research Variables

No	Variable	Description
1	Time (t)	0 = censored
		1 = uncensored
2	Length of Hospitalization (Y)	Interval
3	Sex (X_1)	1 = male
		0 = female
4	Age (X_2)	0 = $<$ 25 years old
		1 = 25-50 years old
		2 = $>$ 50 years old
5	Patient Participation (X_3)	0 = with health insurance
		1 = general
6	Hematocrit Level (X_4)	0 = Hematocrit level $<$ 42
		1 = Hematocrit level $>$ 42
7	Thrombocyte Count (X_5)	0 = Thrombocyte count $<$ 150.000
		1 = Thrombocyte count $>$ 150.000
8	Hemoglobin Count (X_6)	0 = Hemoglobin count $<$ 15
		1 = Hemoglobin count $>$ 15
9	Body Weight (X_7)	0 = Body weight $<$ 50
		1 = Body weight 50-65
		2 = Body Weight $>$ 65
10	Medical History (X_8)	0 = Have suffered DHF
		1 = Have never suffered from DHF
11	Location (W)	Location of medication or treatment

8) Build and interpret a spatial survival model with a lognormal 3 parameter distribution and determine the predicting factors of DHF recovery rate.
9) Determine the recovery rate (hazard rate) or survival rate of patients in every regency.

V. RESULT AND DISCUSSION

In this study, an analysis of the factors that affect the length of stay of Dengue Hemorrhagic Fever will be analyzed with spatial effects. The first step was carried out by descriptive analysis to determine the characteristics of the length of stay of dengue hemorrhagic fever patients at Koesoma Tuban Hospital, these characteristics can be known based on the time of stay and observed variables.

Based on Table II, it can be seen that of 227 patients, most of the length of stay of DHF patients at Koesoma Tuban Hospital was around 4 days. If seen from the length of hospitalization of the patient until he is in a better or better condition, DHF patients are taken to the hospital for around 2-3 days experiencing fever. The minimum hospitalization for patients is 1 day and the longest is 10 days of hospitalization. Most of the DHF patients in Koesoma Tuban were 19 years old, the youngest was 1 year old and the oldest was 58 years old. This disease not only affects children and adolescents but also adults and even middle age. The average level of hematrokit in Koesoma Tuban Hospital patients was 41.34% with the lowest level was 11% and the

TABLE II: Characteristics of Patients Dengue Hemorrhagic Fever

Variable	Min	Max	Mean	Std.Dev	Var
Length of Hospitalization (Y)	1	10	3.90	1.58	2.50
Sex (X_i)	1	76	19.11	15.7	247.11
Hematrocrit Level (X_4)	11	58	41.34	6.81	46.4
Thrombocyte (X_5)	6000	374000	98470	68.61	4709.1
Hemoglobin Count (X_6)	8	20	14.20	2.208	4.87
Body Weight (X_7)	8	90	40.71	18.01	324.6

highest level was 58%. The higher the patient's hematrokit level, the more severe the patient's condition and the lower the patient's hematrokit level, the patient's condition tends to be better provided that the normal limit of hematrokit levels ranges from 40% to 52%. Increased levels of hematrokit are usually preceded by a decrease in platelets and will continue to increase if bleeding always occurs and will decrease after fluid administration to the patient [16]. It should be noted that the value of the hematrokit is affected by fluid replacement [18]. The average number dengue fever patients platelets of in Tuban Koesoma Hospital for 98470 with the sheer number of platelets minimum 6000 / μl and the largest platelet count 374000 / μl. The less platelet count a person has, the more severe the dengue disease is and the more platelet counts a person has, the better the dengue disease will be provided that the normal limit of platelet count ranges from 150000 / μl to 440000 / μl [16]. So it can be concluded that there are patients who are very unstable because they only have a platelet count of 6000 / μl, due to a lack of public awareness of dengue and treatment is only carried out after the condition is somewhat worse [14]. The mean hemoglobin of dengue hemorrhagic fever patients was 14.20 and most of the body weight of the patients who had been hospitalized for the long period of stay was 41 kg.

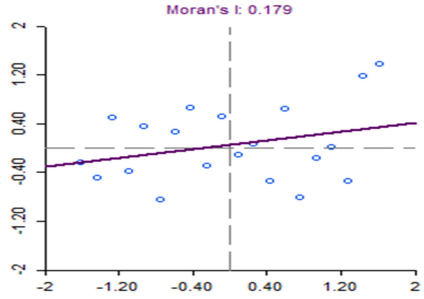

Fig. 2: Moran's I Indeks for the number of patients with dengue hemorrhagic fever

.

Second step on This study analyzed the spatial survival of lognormal 3 parameter to determine the variables which predict survival time of DHF patients during hospitalization in Koesoma Tuban Hospital until they discharged by considering spatial factors with autocorrelation. Spatial autocorrelation in cases of dengue hemorrhagic fever is a condition where there are significant similarities or differences between regions based on the ratio between the number of deaths (people with dengue hemorrhagic fever) [38] that can last up to a certain time in each district in Tuban. This study uses the statistical calculation of Moran's I global test with the aim of knowing whether or not there is a relationship or relationship between the number of dengue hemorrhagic fever sufferers in a district with neighboring areas can be seen in Figure 2.

Figure 2 shows the Moran's I index of 0.179, which is in the range 0 and 1, so it can be concluded that the resulting autocorrelation is Positive Spatial autocorrelation, which means that there is relationship or number of dengue hemorrhagic fever. This positive autocorrelation identifies that adjacent sub-districts have different characteristic values and the ratio of the number of deaths to the number of people affected by dengue hemorrhagic fever in each district in Tuban City. So then an assumption arises to model survival or the rate of recovery of dengue hemorrhagic fever patients from death by considering the spatial autocorrelation, so that it is expected that the model obtained is able to explain the heterogeneity of the data.

In addition to using the Moran's I index, it is necessary to test significant spatial autocorrelation through hypotheses in determining the presence or absence of spatial autocorrelation in the incidence of length of stay of dengue hemorrhagic fever in Tuban district [14]. The test was performed using permutations as much as 999 times can be seen in Figure 3 with the following hypothesis:

H_0 : I = 0 (No Autocorrelation)
H_1 : I $\neq$ 0 (No Autocorrelation)

Based on Figure 3 the p value (0.0380) is smaller than α = 0.05 (5%) so reject H_0 which means sufficient evidence to say that there is spatial autocorrelation in the incidence of dengue fever in Tuban Regency. The survival model which has a significant spatial effect will be tested whether the spatial effect has a CAR (Conditional Autoregrresive) or Normal frailty distribution. The following is a comparison of the spatial survival model in the 3-parameter lognormal distribution using the DIC value.

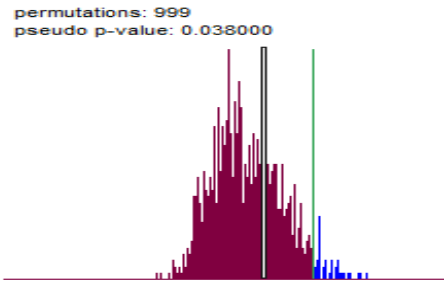

Fig. 3: permutations as much as 999 times to the Moran's I index

.

TABLE III: DIC Value of Spatial Survival

Model Spatial Survival	Random Effect	DIC
Lognormal 3 Parameter	CAR	8826.240
Lognormal 3 Parameter	Normal	7582.980

Based on Table III shows that the model used with consideration of location effects and taking into account the proximity matrix between locations is Spatial Survival in the 3-parameter Lognormal distribution. with Normal frailty because it has the smallest DIC value. This shows that the spatial survival model that gives rise to heterogeneity can be explained by the survival model having a normal frailty distribution. After selecting the model, the next step is to determine the factors that affect the hazard or the rate of survival of dengue hemorrhagic fever hospitalization from death. The following are the results of the parameter estimation of the spatial survival model with a lognormal distribution of 3 parameters with Normal frailty which is presented in Table IV.

Table IV presents the significant factors in DHF recovery rate if the values in the range of 2.5% to 97.5% do not

contain a 0. The variable column lists the factors assumed to affect the recovery rate, the mean column shows the size of the model parameters, and the next four columns show the estimated values at the 97.5% confidence interval. The predicting factors were sex (X_1), age (X_2), patient participation (X_3), hematocrit level (X_4), thrombocyte count (X_5) and patient medical history (X_8). The value of the normal frailty parameter distribution $\lambda = \sqrt{\sigma}$ where σ is significant at 1.287. It means that in this survival lognormal 3 parameter distribution model, spatial effects were found. The normal frailty distribution occurs between units of observation in one group while frailty between groups will be mutually independent.

Table V shows that all significant W_i* values affected the patients recovery rate because they did not contain 0 value within 2.5% to 97.5% interval. This indicates that DHF patients in all districts of Tuban Regency had different recovery rates. This difference can be seen in the significant interval width of the recovery rate caused by the normal random effect parameter (λ). Thus, it can be said that these DHF cases had spatial dependence on the variance and mean components, meaning that the different variance and mean values in each district resulted in the difference of confidence interval in each district too. Furthermore, to determine the risk level /tendency of a particular factor, odds ratio was used. Odds ratio is the comparison of individual odds in certain factor/predictor (x) condition in the expected category with the factor/predictor (x) in the comparison category. Based on the posterior parameters obtained in Table V, the recovery rate (hazard) of DHF patients in each district in Tuban Regency could be modeled as follows:

$$h(t) = \frac{1}{\sigma_i(t-1.216)} * e^{(3.198 - 0.597X_{1.0} - 0.265X_{1.1} + \ldots - 0.409X_{8.1} + \mathbf{W}_i)} \quad (11)$$

The model can be interpreted as sex (X_1) with($\hat{\beta} = -0.597$) value significantly affected recovery rate by exp(-0.597) = 0.550. This shows that female DHF patients tend to recover 0.55 times more slowly than the male patients. It explains why the death rate among female patients is higher, which is because their bodies are more susceptible to dengue virus [38]. The same interpretation applies for all variables.

The results of the spatial survival analysis with normal random effects distribution to uncover unexplained heterogeneity/sources of variance in the model with the Bayesian MCMC method showed that the predicting factors of DHF were sex (X_1), age (X_2), patient participation (X_3), hematocrit level (X_4), thrombocyte count (X_5) and patient medical history (X_8). The random effect with normal distribution in this study showed that the districts in Tuban Regency were significant for the DHF recovery rate. This means that in these DHF cases, there was indeed a spatial effect in the lognormal 3 parameter distribution survival model. This result can be used as a basis for Tuban District Health Department in formulating strategic steps to accelerate DHF recovery rate. Spatial random effect in DHF is also addressed in [16]. Study by [14] using the Weibull distribution with random Conditional Autoregressive (CAR) effect was able to explain the heterogeneity that previously could not be explained, suggesting that sex, age, and thrombocyte count are significant factors. However, this study found that locations were not significant DHF recovery rate, particularly in Pamekasan. Research on DHF conducted by [16] concluded that the predicting factors of DHF patients recovery rate are sex, age, thrombocyte count, and hemoglobin count. It can be concluded that the large number of DHF cases can be epidemiologically important for community-focused health programs because they involve spatial effects. Lognormal 3 parameter spatial survival model with normal random effect can provide information relevant for DHF management dissemination so that the cases number in Tuban Regency can be reduced.

VI. Conclusion

Based on the analysis results, it can be concluded that the survival time lognormal 3 parameter distribution can be applied to spatial survival model which indicates six predicting factors of DHF recovery rate in Tuban Regency. Those factors are sex (X_1), age (X_2), patient participation (X_3), hematocrit level (X_4), thrombocyte count (X_5) and patient medical history (X_8). The spatial dependence on random effects with normal distribution was because of the significant relationship between the patients well being and environmental discomfort at 1.28. The value of normal random effect shows that in DHF cases, there is a spatial dependence of variance and mean differences from spatial random effects in each district, resulting in different confidence intervals of recovery rate in each district in Tuban City.

References

[1] S. Thangavelu, C. Jenil Dhas and M. Zeya Ansari, "Dengue fever complicated by hemophagocytosis-a difficult to manage case," *The Open Hematology Journal*, vol. 7, no. 1, pp. 1-5, 2019.

[2] A. Ghosh, N. Chowdhury and G. Chandra, "Plant extracts as potential mosquito larvicides Indian," *The Indian journal of medical research*, vol. 135, no. 5, pp. 581-98, 2012.

[3] S. B. Halstead, "Dengue," *The lancet*, vol. 370, no. 9599, pp. 1644-1652, 2012.

[4] J. M. T. Bezerra, J. P. Miranda, J. P. N. Neto, A. C. R. Cruz, W. P. Tadei and V. C. S. Pinheiro, "Occurrence of Aedes aegypti (*Diptera, Culicidae*) in a Dengue Transmission Area at Coastal Maranhão State, Brazil," *The open tropical medicine journal*, vol. 6, no. 1, pp. 5-10, 2013.

[5] P. C. Austin, "A tutorial on multilevel survival analysis: Methods, models and applications," *International Statistial Review*, vol. 85, no. 2, pp. 185-203, 2017.

[6] A. Riebler, S. H. Sorbye, D. Simpson and H. Rue, "An intutive bayesian spatial model for disease mapping that accounts for scaling," *Statistical method in medical research*, vol. 25, no. 4, pp. 1145-1165, 2016.

[7] K. A. Chaudhry, F. Jamil, M. Razzaq and B. F. Jilani, "Survival analysis of dengue patients of Pakistan," *International Journal of Mosquito Research*, vol. 5, no. 6, pp. 5-9, 2018.

[8] N. Mahmudah, N. Iriawan and S. W. Purnami,"Bayesian Spatial Survival Models For HIV/AIDS Event Processes In East Java," *Indian Journal of Public Health Research and Development*, vol. 9, no. 11, pp. 1586-1591, 2018.

[9] Jr, R. G Miller *Survival analysis* New York: John Wiley and Sons, 2011

[10] D. G. Kleinbaum and M. Klien,*Survival Analysis. A Self-Learning Text*, New York: Springer, 2012.

[11] J. Ribeiro, *Business Survival Analysis Using SAS*, USA: SAS Institute, 2017.

[12] L. Handayani, M. Fatekurohman and D. Anggraeni, "Survival Analysis in Patients with Dengue Hemorrhagic Fever (DHF) Using Cox Proportional Hazard Regression," *International Journal of Advanced Engineering Research and Science*, vol. 4, no. 7, pp. 138-145, 2017.

[13] D. Darmofal, *Bayesian Spatial Survival Models for Political Event Processes*, Columbia: Gambrell Hall, 2008.

TABLE IV: Posterior Summaries of Spatial Survival Normal Frailty Model

Variable	Parameter	Mean	2.50%	Median	97.50%	Description
$X_{1.0}$	b_1	-0.597	-2.677	-0.612	-2.107	Significant
$X_{1.1}$	b_2	-0.265	-1.756	-0.058	-1.224	Significant
$X_{2.0}$	b_3	0.011	-1.739	-0.117	-2.250	Significant
$X_{2.1}$	b_4	0.758	-2.633	-0.755	-1.384	Significant
$X_{2.2}$	b_5	0.725	-2.429	-0.740	-1.011	Significant
$X_{3.0}$	b_6	0.316	-1.739	-0.516	-1.740	Significant
$X_{3.1}$	b_7	-0.342	-1.727	-0.375	-1.405	Significant
$X_{4.0}$	b_8	-0.231	-2.351	-0.185	-1.282	Significant
$X_{4.1}$	b_9	0.325	-1.528	-0.335	-0.548	Significant
$X_{5.0}$	b_{10}	0.966	-2.870	-0.736	-0.345	Significant
$X_{5.1}$	b_{11}	0.806	-2.546	-0.871	-0.908	Significant
$X_{6.0}$	b_{12}	-1.157	-2.479	-1.282	-0.836	Significant
$X_{6.1}$	b_{13}	-0.780	-2.129	-0.734	0.621	Insignificant
$X_{7.0}$	b_{14}	0.210	-1.469	0.161	2.179	Insignificant
$X_{7.1}$	b_{15}	0.225	-1.399	0.219	1.676	Insignificant
$X_{7.2}$	b_{16}	-0.209	-1.948	-0.205	1.420	Insignificant
$X_{8.0}$	b_{17}	-1.138	-3.739	-1.113	-0.890	Significant
$X_{8.1}$	b_{18}	-0.409	-1.509	-0.371	0.534	Insignificant
Constanta	b_0	3.198	1.494	3.159	5.204	Significant
lambda	λ	1.287	0.508	1.033	3.650	Significant
sigma	σ	1.216	0.075	0.937	3.884	Significant

TABLE V: Spatial Effect in Each District

Location	mean	2.5%	median	97.5%
Bancar	-1.259	-4.808	-0.915	-0.505
Bangilan	0.090	1.910	0.129	1.689
Grabakan	-0.615	-3.158	-0.432	-0.853
Jatirogo	-0.554	-3.393	-0.359	-1.255
Jenu	-0.003	-3.038	-0.013	-3.050
Kenduran	0.016	-3.031	-0.010	-3.109
Kerek	0.002	-3.029	-0.009	-3.092
Merauke	-0.011	-3.081	-0.001	-3.079
Montong	0.005	-2.981	0.008	-3.220
Palang	0.005	-3.052	0.002	-3.122
Parengan	-0.005	-2.985	-0.006	-3.016
Plumpang	-0.009	-3.123	-0.009	-3.015
Rengel	-0.013	-3.160	-0.009	-2.991
Semanding	0.023	-3.020	-0.024	-3.112
Senori	-0.017	-3.153	-0.013	-3.065
Singgahan	0.020	-3.008	-0.013	-3.050
Soko	-0.016	-3.093	-0.008	-3.053
Tambak boyo	-0.011	-3.169	-0.014	-3.088
Tuban	-0.005	-2.984	-0.011	-3.097
Widang	-0.003	-3.016	-0.003	-2.997

[14] N. Iriawan, S. Astutik and D. D. Prastyo, "Markov Chain Monte Carlo – Based Approaches for Modeling the Spatial Survival with Conditional Autoregressive (CAR) Frailty," *International Journal of Computer Science and Network Security*, vol. 10, no. 12, pp. 211-216, 2010.

[15] A. S. Thamrin and I. Taufik, *Spatial Random Effects Survival Models to Assess Geographical Inequalities in Dengue Fever Using Bayesian Approach: a Case Study* Ristol: Institute of Physics Publishing, 2018.

[16] A. Aswi, S. Cramb, E. Duncan, W. Hu, G. White and K. Mengersen, "Bayesian Spatial Survival Models for Hospitalization of Dengue : A Case Study of Wahidin Hospital," *International Journal of Environmental Research and Public Health*, vol. 17, no. 3, pp. 1-12, 2020.

[17] H. Zhou, T. Hanson and J. Zhang, "spBayesSurv: Fitting Bayesian Spatial Survival Models Using R," *Journal of Statistical Software*, vol. 5, no. 2, pp. 1-32, 2018.

[18] N. Mahmudah and H. Pramoedyo, "Pemodelan Spasial Survival Weibull-3 Parameter dengan Frailty Berdistribusi Conditional Autoregrressive (CAR)," *Natural B*, vol. 3, no. 1, pp. 93-102, 2015.

[19] Y. Fang, "A Bayesian Approach to Inference and Prediction for Spatially Correlated Count Data Based on Gaussian Copula Model." *IAENG International Journal of Applied Mathematics*, vol. 44, no. 3, 126-133, 2014.

[20] J. Kruschke, *Doing Bayesian Data Analysis*, USA: Elsevier Science, Academic Press, 2014.

[21] K. Motarjem, M. Mohammadzadeh and A. Abyar, "Bayesian Analysis of Spatial Survival Model with Non-Gaussian Random Effect," *Journal of Mathematical Sciences*, vol. 237, no. 5, pp. 692–701, 2019.

[22] A. Rabie and J. Li, "E-Bayesian estimation for Burr-X distribution based on Type-I hybrid censoring scheme." *IAENG International Journal of Applied Mathematics*, vol. 48, no. 3, pp. 244-250, 2018.

[23] P. E. Jacob, "Unbiased Markov chain Monte Carlo methods with couplings," *Journal of the Royal Statistical Society (Statistical Methodology)*, vol. 82, no. 3, pp. 543-600, 2020.

[24] K. M. Banner, K. M. Irvine and T. J. Rodhouse, "The Use of Bayesian Priors In Ecology : The Good, The Bad and not great," *Methods In Ecology And Evolution*, vol. 11, no. 8, pp. 882-889, 2020.

[25] Q. Feng, S. Sha and L. Dai, "Bayesian Survival Analysis Model for Girth Weld Failure Prediction," *Applied Sciences*, vol. 9, no. 6, pp. 1-11, 2019.

[26] M. Blangiardo and M. Cameletti,*Spatial and Spatio-Temporal Bayesian Models with R-INLA*, New York: Wiley, 2015.

[27] R. P. Haining and G. Li, *Modelling Spatial and Spatial-Temporal Data: A Bayesian Approach*, New York: Chapman and Hall/CRC, 2020.

[28] S. Banerjee, M. M. Wall and B. P. Carlin, "Frailty modeling for spatially correlated survival data, with application to infant mortality in Minnesota," *Biostatistics*, vol. 4, no. 1, pp. 123-142, 2003.

[29] P. D. Allison, *Survival Analysis Using SAS®A Practical Guide Second Edition*, USA: SAS Institute Inc, Cary, 2010.

[30] S. Guo,*Survival Analysis*, New York: Oxford University Press, 2010.

[31] K. Bogaerts, A. Komarek and E. Lesaffre, *Survival Analysis with Interval-Censored Data: A Practical Approach with Examples in R, SAS, and BUGS*, New York: Chapman and Hall/CRC, 2018.

[32] S. Banerjee, "Spatial data analysis," *The annual review of public health*, vol. 37, pp. 47-60, 2016.

[33] S. Momenyan, A. Kavousi, T. Baghfalaki and J. Poorolajal, "Bayesian modeling of clustered competing risks survival times with spatial random effects," *Epidemiology Biostatistics and Public Health*, vol. 17, no. 2, pp. 47-60, 2020.

[34] M. A. Erango, "Bayesian Joint Modeling of Longitudinal and Survival Time Measurement of Hypertension Patients," *Risk Management and Healthcare Policy*, vol. 13, pp. 73-81, 2020.

[35] D. Darmofal,"Bayesian spatial survival models for political event processes." *American Journal of Political Science*, vol. 53, no. 1, pp. 241-257, 2009.

[36] J. G. Ibrahim, M. H. Chen, and D. Sinha,*Bayesian Survival Analysis*, Singapore: Springer, 2010.

[37] B. Yu, "A Bayesian MCMC approach to survival analysis with doubly-censored data,"*Computational Statistics and Data Analysis*, vol. 54, pp. 1921-1929, 2010.

[38] S. Amalia, N. Iriawan and D. D. Prastyo,"Survival Analysis and Factor Influencing The Recovery Of Dengue Hemorrhagic Fever Patient By Using Bayesian Mixture Survival," *Proceedings of the Third International Conference on Mathematics and Natural Sciences*, pp. 91-97, 2010.

[39] S. Selvin, *Survival Analysis for Epidemiologic and Medical Research*, New York: C ambridge University Press, 2008.

[40] T. Amir, F. Farshad and Y. Meh, "Spatial Survival Analysis of Initiation Age and Prevalence of Smoking in Iran; Results from a Population Based Study," *Archives of Iranian Medicine*, vol. 7, no. 23, pp. 462-468, 2020.

[41] M. E. S. Hannah, *Spatial survival analysis* Berlin: Spinger, 2015.

[42] P. Congdon, *Applied Bayesian Modelling*, USA: John Wiley & Sons, 2003.
[43] J. C. Gelman, J. B. Carlin, H. S. Stern,D. B. Dunson, A. Vehtari, and D. B. Rubin, "*Bayesian data analysis* New York: CRC, 2013.
[44] I. Ntzoufras, *Bayesian Modeling Using WinBUGS*, New Jersey: John Wiley & Sons, Inc., 2009.
[45] A. Aswi, S. M. Cramb, P. Moraga and K. Mengersen, "Bayesian spatial and spatio-temporal approaches to modelling dengue fever: a systematic review," *Epidemiology and Infection*, vol. 147, pp. 1-14, 2018.

Fetrika Anggraeni received the B.S. and M.S. degrees from Universitas Negeri Yogyakarta, Yogyakarta, Indonesia, in 2012 and 2015, respectively. At present, she is currently a lecturer in the Department of Statistics, Universitas Nahdlatul Ulama Sunan Giri, Bojonegoro, Indonesia. Her research interests include modeling, economics and statistics.

Nur Mahmudah received the B.S. degree in Statistics, from Brawijaya University, Malang, Indonesia, in 2010 and the M.S. degree in Statistics from ITS, Surabaya, Indonesia, in 2018, respectively. At present, she is currently a lecturer in the Department of Statistics, Universitas Nahdlatul Ulama Sunan Giri, Bojonegoro, Indonesia. Her research interests include modeling, computational statistics, Bayesian and Stochastic modeling.