The Dengue Hemorrhagic Fever Modeling in Banyumas Regency by Using CAR-BYM, Generalized Poisson, and Negative Binomial

Jajang Jajang[1,*] , Budi Pratikno[1], Mashuri Mashuri[1], Indriani Eko Cahyarini[1]

[1,] *Department of Mathematics, University of Jenderal Soedirman*
[*]*Corresponding author. Email: jajang@unsoed.ac.id*

ABSTRACT
The research studied disease mapping of dengue haemorrhagic fever (DHF) in Banyumas Regency. The generalized Poisson (GP), negative binomial (NB), and CAR-BYM models are then used to modelling the DHF. The predictor variables used in this research are the number of health worker, altitude, and population density. We choose the best model for further analysis and estimate relative risk. The criteria for choosing the best model is root mean square error (RMSE). The result showed that the CAR-BYM is the best model. This is due to the RMSE of CAR-BYS is the lowest (1.65) among others (GP = 3.17 and NB=3.25).. Here, we also presented several analysis of the region case, such as Kembaran District is the highest of DFH. The West Purwokerto is the highest relative risk value. This mean that transmission DHF in West Purwokerto district is highest. We also found that the lowest relative risk value is Gumelar. This mean that transmission DHF in Gumelar district is the lowest.

Keywords: *relative risk, disease mapping, CAR-BYM, generalized Poisson, negative binomial*

1. INTRODUCTION

Dengue fever is a disease that is classified as an infectious disease. Dengue fever is caused by one of the four types of dengue virus. Two mosquitoes that can transmit this virus are *Aedes aegypti* and *Aedes albopictus*. The number of DHF cases as considered as count data and it is a type of spatial data.

Information related to the spread of DHF can be assessed using the disease mapping method. Disease mapping is useful to find geographic distribution of disease burden and disease incident based on risk level.

Mapping of diseases cannot be separated from spatial effects. To overcome this spatial effects in the model, we have to involve a spatial weight matrix in that model. The Spatial weights matrix is non-negative matrix. The element of the matrix, usually, is scored 1 if units are close to each other and 0 if otherwise. However, there are some researcher reconstruct the spatial weights matrix. to determine each element of the matrix, They combine information about geographic distance and interesting variable [1,2].

The number of DHF data can be categorized as count data, so we can use Poisson, generalized Poisson, or negative binomial models for this data. Unfortunately, those model cannot accommodate the spatial aspect. Even though the spatial aspect is important in in disease mapping.

Spatial autoregressive (SAR), spatial error model (SEM), conditional autoregressive-Bessag-York-Mollie (CAR-BYM) are some of popular spatial models. From these three model, the CAR-BYM model can accommodate spatial and non-spatial aspects as the consequence of heterogeneity of cases between regions. The CAR-BYM model can detect areas with relative risk. The relative risk can obtain through interpolation of disease mapping of neighbourhood information. Study on spatial models for area data has been commonly carried out, especially for model with continue response variable type.

Error in the CAR-BYM is dependent, so the popular ordinary least square (OLS) and maximum likelihood (MLE) is not suitable for this case. For this solution, Bayesian estimation method can be used as the right option for this problem. The advantage of this method is flexibility in assumption of the error distribution. To estimate the parameter by using Bayes method, we have to find the integral for high dimensional space. This is because the model involves many parameters. Therefore, the Bayesian method through Markov Chain Monte Carlo (MCMC) is used. The MCMC method has been commonly used by researchers in medical field, such as [3–7].

Banyumas Regency is one of the 35 Regencies/Cities in Central Java Province. According to the Epidemiologic Data and Surveillance Center of the Ministry of Health of the Republic of Indonesia, the causes of increase in and distribution of DHF were, among others, high mobility of the people, urban development, climate change, changes in population density, and population distribution, and other

epidemiologic factors. The number of regional DHF cases such as the case in Banyumas Regency is one type of area data. Interaction is possible since DHF case is of contagious type of case. Therefore, the quantitative measure of a variable that is the attention in an area will be influenced by other area as the consequence of interaction.

2. RESEARCH METHODOLOGY

2.1 Spatial data and spatial weights Matrix

The area data is one of type of spatial data [8,9]. In the spatial data, suppose D is a region consisting of non-intersecting sub-areas, then *D* is partitioned into finite number of area units [9,10]. Spatial weights matrix (**W**) **is** non-negative matrix that specifies neighborhood set for each observation. Element of spatial matrix ***W is*** 1 for adjacent areas and 0 for the other.

2.2. The Model

In this paper, we focus on discussing we focus Generalized Poisson (GP), negative binomial (NB) and conditional autoregressive-Bessag-York-Mollie (CAR-BYM) models. These model are usually used for count data.

2.2.1 Negative binomial and Generalized Poisson Model

The Poisson model is a common model that is often used for data counts. Suppose Y is random variable Poisson distributed, $Y_i \sim Poi(\mu_i), i = 1,2,\dots,n$, then the probability mass function of Y is

$$p(y_i) = \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!}, y_i = 0,1,2,\dots \quad (1)$$

Expected value and variance of the Y are $\boldsymbol{E(Y_i) = \mu_i}$ and $\boldsymbol{Var(Y_i) = \mu_i}$, respectively. This condition condition is called equidispersion. Poisson's model assumes this equidispersion conditions. However, in many cases, this equidispersion condition is not met, it is called as overdispersion. Common models used in dealing with cases of overdispersion negative binomial (NB) and generalized Poisson (GP) models. Probability mass function (pmf) of NB is

$$f(y,\mu_i,m) = exp\left[y\, ln\left(\frac{m\mu}{1+m\mu}\right) + \frac{1}{m} ln\left(\frac{1}{1+m\mu}\right) + ln\frac{\Gamma\left(y+\frac{1}{m}\right)}{y!\Gamma\left(\frac{1}{m}\right)}\right], \quad (2)$$

where μ is mean of Poisson random variable and m is overdispersion parameter.

In addition to the NB model, the generalized Poisson (GP) model is also commonly used to overcome the problem of overdispersion in the data count. Probability mass function (pmf) of GP is

$$f(y,\mu_i,m) = \left(\frac{\mu_i}{1+m\mu_i}\right)^{y_i} \frac{(1+my_i)^{y_i-1}}{y_i!} \cdot exp\left[\frac{-\mu_i(1+my_i)}{1+m\mu_i}\right]. \quad (3)$$

where μ is mean of Poisson random variable and m is overdispersion parameter.

2.2.2 Poisson-Lognormal Model

The Poisson-Lognormal model was derived from a combination of Poisson distribution with what assuming Poisson heteroscedasticity parameter. Suppose Y_i is a random variable that follows Poisson distribution, $Y_i \sim POI(\mu_i)$, $\mu_i = E_i\theta_i$ and $\theta_i = exp(\eta_i)$. Based on the characteristics of exponential family distribution [11], further, mean Poisson can be stated with

$$log(\mu_i) = log(E_i) + x_i'\beta + v_i \quad (4)$$

with $\boldsymbol{\mu_i}$ is mean of Poisson distributed response variable, $\boldsymbol{log(E_i)}$ adalah *offset*, $\boldsymbol{x'_i}$ is free vector variable, $\boldsymbol{\beta}$ vector parameter, $\boldsymbol{v_i}$ error that follows CAR.

2.2.3 Intrinsic Conditional Autoregressive (ICAR) Model

When an area *n* was given consisting of non-overlapping sub-areas, each adjacent sub-area (have shared borders) was scored 1 and otherwise scored 0. Spatial interaction between area pair i and j could be modeled as conditional normal. $v_i|v_{j\neq i} \sim N\left(\sum_{i\neq j}\varphi_{ij}v_j, \tau_i^2\right)$, $E\left(v_i|v_{j\neq i}\right) = \mu_i + \sum_{J\epsilon N_I}\varphi_{ij}\left(v_j - \mu_i\right)$ and $Var\left(v_i|v_{j\neq i}\right) = \tau_i^2, N_i$ adjacent spatial unit i.

$$f\left(v_i|v_{j\neq i}\epsilon S\right) = \left(\frac{1}{2\pi\tau_i^2}\right)^{\frac{1}{2}} exp\left(-\frac{1}{2}\sum_{i=0}^{n}\frac{\left(\left(v_j-\mu_i\right)-\rho\sum_{J\epsilon N_I}\varphi_{ij}\left(v_j-\mu_i\right)\right)^2}{\tau_i^2}\right) \quad (5)$$

$\mu_i \epsilon R, \tau_i^2 \epsilon R^+$, $|\rho| < 1$, $\varphi_{ij} = \varphi_{ji}$ and $\varphi_{ii} = 0$. In Spatial Autoregressive, φ_{ij} was the element of weights matrix, the apstial matriks, W, is

$$W = \left(w_{ij}\right), w_{ij} = \begin{cases} 0 & \text{if i} = j \\ \varphi_{ij} & \text{if adjacent } i \neq j. \end{cases}$$

2.2.4 CAR-BYM Model

Conditional Autoregressive –BYM (CAR-BYM) model was Poisson log normal model developed for disease mapping risk. This model covered ICAR component for spatial rarefaction and ordinary random effect component for non-spatial heterogeneity. Poisson regression model was used to estimate relative risk (RR), that is η_i for region i, i=1, 2, …, n, given y_i that was the number of cases.

CAR-BYM model was specified as follows:

$$\eta_i = \mu + x'_i\beta + \phi_i + \theta_i,\ i = 1,2,\dots,n, \quad (6)$$

where

x'_i = observational vector of independent variable i
β = vector of parameter
ϕ_i = ICAR component
μ = average risk level
θ_i = random effect of non-spatial heterogeneity component.

2.2.5 Bayesian Estimation Framework

Suppose Y_i, i=1, 2, …, n are random sample of probability mass function, pmf $P(y|\theta)$, with θ is vector of parameter $\theta = (\theta_i, \dots, \theta_p)$ is [12].

$$P(y|\theta) = \prod_{i=1}^{n} p(y_i|\theta)\, P(\theta). \quad (7)$$

Estimation parameter by using the Bayesian method needed information about parameter θ, called prior distribution. Prior distribution was viewed as introductory knowledge of parameter θ and determined before observation data were given. Here, without losing generality, we take parameter a θ, so we obtain prior distribution $p(\theta)$ and we then create joint pdf/pmf $p(\theta, \boldsymbol{y})$. Based on $p(\theta)$ and $p(\theta, \boldsymbol{y})$ we obtained posterior distribution of the θ, $P(\theta|y)$

$$P(\theta|y) = \frac{P(y,\theta)}{p(y)} = \frac{P(y|\theta)P(\theta)}{p(y)}, \quad (8)$$

where $p(y)$ is marginal probability, $P(y|\theta)$ joint pdf/pmf. Furthermore, based on this posterior distribution, the estimator for parameter θ is $\hat{\theta} = E(\theta|y)$.

The stages of parameter estimation by using Bayes were as follows.

(1) create likelihood function

$$\boldsymbol{l(\beta, \nu)} = \prod_{\boldsymbol{i=1}}^{\boldsymbol{n}} \frac{\boldsymbol{e^{-E_i\theta_i}(E_i\theta_i)^{y_i}}}{\boldsymbol{y_i\ !}} = \boldsymbol{P(y, E, \theta|\beta, \nu)}$$

(2) Determining prior distribution of β and ν

$$p(\beta) = \left(\frac{1}{2\pi}\right)^{\frac{P}{2}} \left(\frac{1}{\tau_\beta}\right)^{P} exp\left(-\frac{1}{2}\sum_{h=0}^{P}\frac{\beta_h^2}{\tau_\beta^2}\right) \boldsymbol{p(\nu_i|\nu_{i\neq j}, \tau_\nu^2)}$$

$$= \left(\frac{\boldsymbol{1}}{\boldsymbol{2\pi}}\right)^{\frac{\boldsymbol{n}}{\boldsymbol{2}}} \boldsymbol{exp}\left(-\frac{\boldsymbol{1}}{\boldsymbol{2}}\sum_{\boldsymbol{i=0}}^{\boldsymbol{n}}\left(\frac{\nu_i - \sum_{i\neq j}^{n}\frac{w_{ij}\cdot\nu_j}{w_{ij}}}{\tau_\beta}\right)^2\right)$$

(3) create posterior distribution

$$\boldsymbol{p(\beta, \nu, \tau_\beta^2, \tau_\nu^2|y, E, \theta)} =$$

$$\prod_{i=1}^{n}\frac{e^{-E_i\theta_i(E_i\theta_i)^{y_i}}}{y_i\ !}\, x\left(\frac{1}{\sqrt{2\pi}}\right)^{\frac{P}{2}}\left(\frac{1}{\tau_\beta}\right)^{P} \boldsymbol{exp}\left(-\frac{1}{2}\sum_{h=0}^{P}\frac{\beta_h^2}{\tau_\beta^2}\right)$$

$$x\left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \boldsymbol{exp}\left(-\frac{1}{2}\sum_{i=0}^{n}\left(\frac{\nu_i - \sum_{i\neq j}^{n}\frac{w_{ij}\cdot\nu_j}{w_{ij}}}{\tau_\beta}\right)^2\right). \quad \textbf{(9)}$$

2.2.6 Markov Chain Monte Carlo

It is apparent that it is not easy to determine expected value with posterior pdf on (6). Therefore, Bayes parameter estimation was used through Markov Chain Monte Carlo (MCMC). For this, let parameter $\boldsymbol{\theta}$ is *k x 1* vector of the parameters. The Gibb sampler method is as follows

For t=1,2,…,T

- Stage 1. Take θ_1^t from $p(\theta_1 \,|\theta_2^{t-1}, \theta_3^{t-1}, \dots, \theta_k^{t-1}, y)$
- Stage 2. Take θ_2^t from $p(\theta_2 \,|\theta_1^{t-1}, \theta_3^{t-1}, \dots, \theta_k^{t-1}, y)$
- …
- Stage k. Take θ_k^t from $p(\theta_k \,|\theta_1^{t-1}, \theta_2^{t-1}, \dots, \theta_{k-1}^{t-1}, y)$.

If the condition was stable in case of iteration τ_0, estimation of parameter θ_i

$$\widehat{\boldsymbol{\theta}}_{\boldsymbol{k}} = \widehat{\boldsymbol{E}}(\boldsymbol{\theta}_{\boldsymbol{k}}|\boldsymbol{y}) = \frac{\boldsymbol{1}}{\boldsymbol{T}-\boldsymbol{t}_0}\sum_{\boldsymbol{t=t_0+1}}^{\boldsymbol{T}}\boldsymbol{\theta}_{\boldsymbol{k}}^{\boldsymbol{t}}.$$

3. RESULT AND DISCUSSION

3.1 Description analysis

In this paper we used dengue hemorrhagic fever data in Banyumas Regency in 2019. The response variables of the study were the number of dengue cases in Banyumas Regency, with some attributes thought to be related to the number of dengue cases. These variables are the population density of each sub-district, the number of health workers, and the height of the subdistrict area. The scatter plot of theese variables is presented by Figure 1.

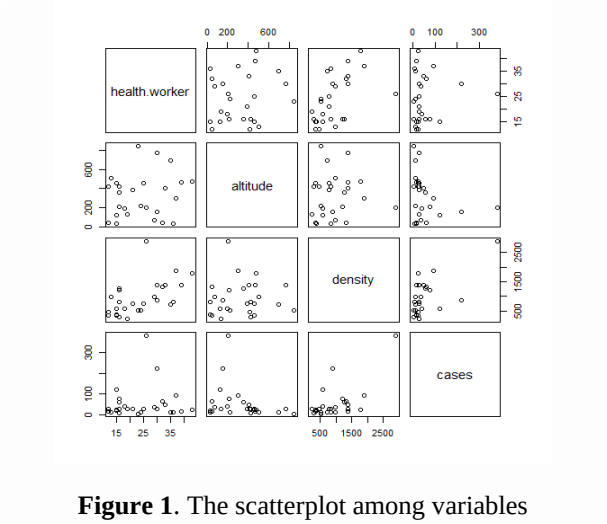

Figure 1. The scatterplot among variables

The Figure 1 shown that relationship among the number of health worker (health worker), altitude, population density (density) and the number of DHF cases. Based on the Figure 1, we cannot see the

relationship between predictor variables and response variable (the number of DHF case) directly. Therefore, we necessary to create model for interpretation the relationship between predictor variables and response variable. From the model, we also can know contribution of each the predictor variables to the response variable.

3.2 Model

In disease mapping, relative risk (RR) often used to measure the risk of a region to others. Relative risk values in model data count depends on observation and expected values. Here, the expected values are obtained from the best model. The best model is selected from generalized Poisson, negative binomial, and CAR-BYM models. The criteria used for model selection is root mean square error (RMSE).

The results of three analysis of variance (ANOVA) model that used for this data are listed in Table 1, Table 2, and Table 3. Based on the three tables (Table 1, Table 2, and Table 3), we can see that coefficient of the population density is positive and significantly correlated with the number of DHF cases. This means that if the population density in a district is increase then the number of DHF cases is also increases. Meanwhile, coefficients of the number health worker and altitude are negative. This mean that if the health worker and altitude are increase then the number of DHF cases are decrease. Furthermore, for calculate the relative risk of DHF for each district in Banyumas regency, we selected the best model by using root mean square error (RMSE).

Table 1. ANOVA of the GP model

Coefficients: Estimate	Std. Error	*z* value	*Pr*(>|*z*|)	
(Intercept):1	3.7305	0.3513	10.616	<2e-16 ***
(Intercept):2	1.1196	0.1208	9.265	<2e-16 ***
density	0.0011	0.0001	8.560	<2e-16 ***
h.worker	-0.0275	0.0131	-2.103	0.0355 *
altitude	-0.0013	0.0005	-2.340	0.0193 *

Table 2. ANOVA of the NB model

Coefficients: Estimate	Std. Error	*z* value	*Pr*(>|*z*|)	
(Intercept)	3.8239	0.3994	9.572	< 2e-16 ***
density	0.0010	0.0002	3.942	8.08e-05 ***
h.worker	-0.0138	0.0172	-0.801	0.423
altitude	-0.0025	0.0006	-4.042	5.30e-05 ***

Table 3. ANOVA of the CAR-BYM model

	Median	2.5%	97.5%	n.effective	eweke diag
(Intercept)	0.2514	-0.4553	1.0594	16.7	0.8
Density	0.0012	0.0007	0.0018	7.9	-1.3
h.worker	-0.0675	-0.0959	-0.0191	7.5	-0.7
Altitude	-0.0011	-0.0030	0.0004	7.7	1.7
tau2	0.0207	0.0036	0.1590	86.2	0.8
sigma2	0.5295	0.2891	1.0312	128.2	1.3

Table 4. The Actual data and prediction

No	DISTRICT	Actual	GP	NB	CAR-BYM
1	Ajibarang	2	2,34	2,14	2,17
2	Banyumas	1	1,23	1,04	0,83
3	Baturraden	6	2,89	2,72	3,48
4	Cilongok	1	2,75	2,58	3,10
5	Gumelar	0	2,24	2,01	0,76
6	Jatilawang	6	4,36	4,52	5,72
7	Kalibagor	3	4,47	4,63	3,96
8	Karanglewas	3	5,07	5,27	4,54
9	Kebasen	6	4,30	4,47	5,42
10	Kedungbanteng	2	3,30	3,25	3,06
11	Kembaran	13	5,96	6,25	11,45
12	Kemranjen	0	3,17	3,15	2,40
13	Lumbir	0	3,40	3,45	2,07
14	Patikraja	10	4,29	4,44	6,01
15	Pekuncen	1	2,83	2,70	1,81
16	Purwojati	1	3,96	4,06	2,22
17	West Purwokerto	9	13,53	14,93	9,41
18	South Purwokerto	11	10,91	12,10	12,12
19	East Purwokerto	2	3,85	3,52	2,52
20	North Purwokerto	7	8,85	9,39	7,07
21	Rawalo	5	3,88	3,96	4,10
22	Sokaraja	8	3,52	3,41	6,08
23	Somaged	3	3,24	3,16	1,42
24	Sumbang	7	3,22	3,03	5,66
25	Sumpiuh	1	3,19	3,18	1,87
26	Tambak	0	3,71	3,78	1,76
27	Wangon	10	3,55	3,57	6,51

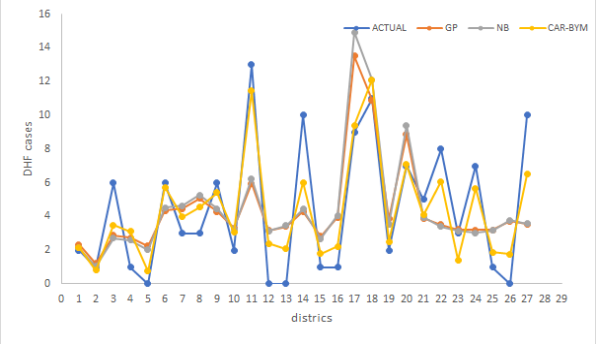

Figure 2. The accuracy of GP, NB and CAR-BYM.

Figure 2 showed accuracy of prediction by using plot between DHF actual and DHF prediction. Based on the Figure 2, we can see that accuracy of the CAR.BYM prediction for DHF is good (green color line close to black color line). Based on generalized Poisson, negative binomial, and CAR-BYM model, we then computed root mean square error (RMSE) of their errors. The RMSE of are 40.95, 40.64 and 1.29, respectively (Figure 3).

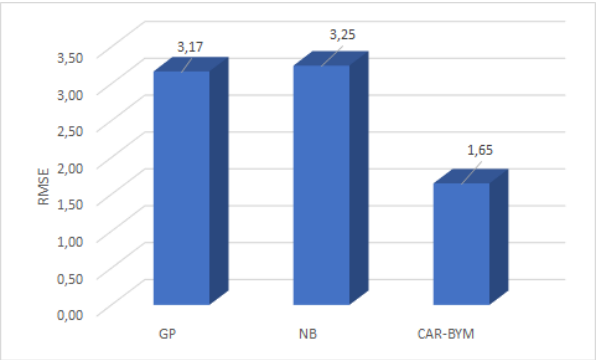

Figure 3. The RMSE GP, NB and CAR-BYM model

Due to this RMSE, we concluded that the CAR-BYM model is the best model for this DHF modeling. Therefore, we used CAR-BYM model to calculate and prediction of relative risk (RR). The relative risk (RR) values for each district resulting from CAR-BYM model is presented in Table 5. Table 5 is the detail of RR value in each district, while Figure 4 is the presentation of RR value in Map.

Table 5. The RR values for each districts

No	District	RR
1	Ajibarang	0,306
2	Banyumas	0,198
3	Baturraden	0,873
4	Cilongok	0,365
5	Gumelar	0,175
6	Jatilawang	1,213
7	Kalibagor	1,049
8	Karanglewas	1,040
9	Kebasen	1,162
10	Kedungbanteng	0,728
11	Kembaran	2,063
12	Kemranjen	0,485
13	Lumbir	0,579
14	Patikraja	1,431
15	Pekuncen	0,330
16	Purwojati	0,878
17	West Purwokerto	2,621
18	South Purwokerto	2,447
19	East Purwokerto	0,600
20	North Purwokerto,	2,191
21	Rawalo	1,080
22	Sokaraja	0,976
23	Somaged	0,526
24	Sumbang	0,883
25	Sumpiuh	0,446
26	Tambak	0,476
27	Wangon	1,087

The values of RR for each district are presented by Figure 4.

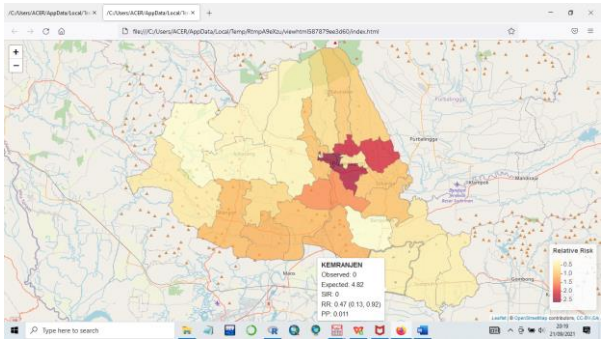

Figure 4. RR of DHF mapping in Banyumas

Based on figure 4 we find the areas that have RR from the highest to the lowest. The RR values from high to low of an area are represented by dark to light colors. The dark color indicates a high risk area of DHF and light color indicate a low risk area of DHF.

The DHF cases of Kembaran is the highest. Although, the RR of West Purwokerto, South Purwokerto and North Purwokerto are higher than Kembaran. Based on Table 5 most districts generally have an RR < 1. The districts that have RR > 1 show that the transmition of DHF in the area is greater than average RR in the Banyumas. The highest RR value is

that of West Purwokerto District and the lowest relative risk value is that of Gumelar District. Gumelar district is at the lowest risk of transmission and West Purwokerto district is at the highest risk of transmission.

4. CONCLUSION

The CAR-BYM model is the best model that can be used for DHF case modeling in Banyumas Regency. Based on factors studied has information that if both the number of health workers and altitude are increase then the DHF cases is decrease. Based on the studied, we also obtained information that if the population density is increase then the DHF case is also increase.

The best model for modeling the DHF case in Banyumas Regency is the CAR-BYM model. The relative risk for each area is determined by using the CAR-BYM model. The relative risk is directly proportional to the number of DHF cases.

The highest relative risk value is that of West Purwokerto District and the lowest relative risk value is that of Gumelar District. Gumelar district is at the lowest risk of transmission and West Purwokerto district is at the highest risk of transmission.

AUTHORS' CONTRIBUTIONS

JJ, DESIGNED and PERFORMED the experiments, DERIVED the MODELS and ANALYSED the data, B.P and MM VERIFIED the background and analytical methods. I.E.C PROVIDED data.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to the Ministry of Research and Technology and Higher Education for their funding for this research through Research Grant BLU UNSOED year 2021.

REFERENCES

[1] J. Aldstadt, A. Getis, Using AMOEBA to create a spatial weights matrix and identify spatial clusters, Geographical Analysis, vol. 38, no. 4, 2006, pp. 327–343.
[2] Jajang, A. Saefuddin, I. W. Mangku, H. Siregar, Comparing Performances of WG, WGnew and WC on Dynamic Spatial Panel Model By Monte Carlo Simulation, Far East Journal Of Mathematical Sciences, vol. 80, no. 2, 2014. pp. 155–167.
[3] Y.J. Cheng, Geographical information systems-based spatial analysis and implications for syphilis interventions in Jiangsu province, People's Republic of China, Geospatial Health, vol. 7, no. 1, 2012, pp. 63–72.
[4] X. Han, L. Lee, Bayesian estimation and model selection for spatial Durbin error model with finite distributed lags, Regional Science and Urban Economics, vol. 43, no. 5, 2013, pp. 816–837.
[5] J. Hendricks, C. Neumann, A Bayesian approach for the analysis of error rate studies in forensic science, Forensic Science International, vol. 306, 2020, p. 110047.
[6] R. Srinivasan, P. Venkatesan, Bayesian random effects model for disease mapping of relative risks, Ann Biol Res, vol. 5, no. 1, 2014, pp. 23–31.
[7] H.S. Stern, N. Cressie, Posterior predictive model checks for disease mapping models, Statistics in Medicine 19 (17-18), 2000, pp. 2377–2397.
[8] C. Gaetan, X. Guyon, Spatial statistics and modeling, Springer, 2010.
[9] N.A. Cressie, Statistics for spatial data John Willy and Sons, Inc., New York, 1993.
[10] S. Banerjee, B.P. Carlin, A.E. Gelfand, Hierarchical modeling and analysis for spatial data, CRC press, 2014.
[11] P. McCullagh, J.A. Nelder, Generalized Linear Models, CRC Press, 1989.
[12] J.G. Ibrahim, Bayesian survival analysis, Springer, 2001.