Available online at http://scik.org
J. Math. Comput. Sci. 11 (2021), No. 6, 7971-7990
https://doi.org/10.28919/jmcs/6740
ISSN: 1927-5307

GEOGRAPHICALLY WEIGHTED NEGATIVE BIVARIATE BINOMIAL REGRESSION FOR MODELLING THE NUMBER OF DENGUE DISEASES AND THEIR MORTALITY IN EAST NUSA TENGGARA, INDONESIA

ALAN PRAHUTAMA[1,*], ARIEF RACHMAN HAKIM[1], DWI ISPRIYANTI[1], HASBI YASIN[1], BUDI WARSITO[2]
[1]Statistics Department, Faculty of Science and Mathematics, Diponegoro University, Semarang Indonesia
[2]Environment Science, Graduate School Diponegoro University, Semarang Indonesia

Abstract: Dengue disease usually occurs in tropical countries. In 2020, East Nusa Tenggara (NTT), Indonesia was the highest number in the mortality of dengue. For that reason, analysis of spread dengue is indispensable. One of the statistical modelling methods to spread dengue and its factors based on spatial analysis is Geographically Weighted Regression (GWR). GWR was developed to Geographically Weighted Poisson Regression (GWPR) and Geographically Weighted Negative Binomial Regression (GWNB) as handling overdispersion. GWNB has improvement analysis into Geographically Weighted Bivariate Negative Binomial regression (GWBNB) model. This paper applied the GWBNB model for the number of dengue cases and their mortality (as dependent variables) based on population density, percentage of poverty, the number of doctors, the number of health facilities, and the percentage of the area that has good sanitation (as independent variables). It resulted from the BNB model, the factors that impacted the dengue's number and their mortality were poverty, health facilities, and the percentage of areas with good sanitation. In BNB model resulted in 364.4725 of AIC. However, the spatial testing has shown spatial contiguity of dengue's number and their mortality. As a result of the GWBNB model, shown most of the variables were significant and it established the cluster based on its estimation. Besides, with the GWBNB model, the heterogeneity of spatial can be solved. The accuracy of GWBNB estimation was 73% for predict the number of dengue and their mortality in East Nusa Tenggara, Indonesia. The GWBNB model resulted better than the BNB model based on AIC's value.

---

[*]Corresponding author
E-mail address: alan.prahutama@gmail.com
Received September 4, 2021

Keywords: BNB; GWBNB; Dengue's number and their mortality.
2010 AMS Subject Classification: 93A30.

1. INTRODUCTION

Dengue disease is a kind of tropical disease that is caused by the *Aedes aegepty* mosquito. Its spread depends on various factors such as weather, climate, the environment of household, density of population, and health facilities. The government has made many attempts to handle it, such as socializing to keep a clean and healthy environment and increasing the number of health workers in both urban and remote areas. In 2016, the case of dengue was 204,171 cases, with the number of their mortality was 1,598 cases. Meanwhile, in 2017 it decreased significantly which the number of dengue cases was 68,407 cases, with the number of its demise was 493 cases [1]. The region which the highest contributed dengue cases was Java Island., In 2020 the highest dengue cases were in Lampung province, with 3,423 cases and 11 deaths. While, East Nusa Tenggara (NTT), province in Indonesia, had 2,711 cases with 32 death cases (the highest mortality case of dengue in 2020) [2]. Whereas, in 2017, East Nusa Tenggara province had 210 cases of dengue with one death case [1]. It had increased significantly in 2020, so it should be analyzed about the spread of dengue in NTT. Its reach also has dependency among the regions. One method to analyze the spread of disease based on dependency region was spatial analysis. It is one of the statistics branch to model and map the space of variables in the areas [3].

One method in spatial which for modeling between independent variables and dependent variables in regions is GWR. It is also used to map the spread among the regions. GWR results in a regression model in each region, so the heterogeneity of areas can be handled. Global regression was used for modeling spatial data that can results in heterogeneity in its model. So the estimation would be a misperception. Estimate spatial data uses GWR results parameters with different probability of significance in each region [4]. In GWR estimation of the parameter used weighted value which is determined by bandwidth. The Kernel function determines bandwidth. Chosen optimal bandwidth uses Cross-Validation (CV) [4].

Some applications of the GWR method were analysis spread of the average number of smear-positive Tuberculosis (TB) di Xinjiang, China. It used socioeconomic factors, consisting of five variables: population density, the proportion of minorities, the number of infectious disease network reporting agencies, the proportion of the agricultural population, and per capita gross annual domestic product. It resulted that there was a relation between the average number of smear-positive TB and socioeconomic factors. Besides, the GWR model could better be exploited geographically of the average number of smear-positive TB [5]. Similarly, the spread of dengue in Taiwan has been researched by the GWR model with the count of population and population density as predictor variables [6]. Some of the research about spread dengue used the GWR method can be seen in Table 1. In that analysis, the spatial method, especially GWR, could model the dengue spread and its factors. And also, it was able to map the spread based on regions.

On the other hand, the GWR method has been developed in various ways. One of them is GWPR, which is for modeling the count data follows Poisson distribution [7]. Some research about the GWPR model, such as modeled spread of HIV/AIDS, is based on socioeconomic factors in Sub-Saharan Africa. It resulted that, GWPR model obtained 0,6139 of R-square value [8]. In addition, the usage of the GWPR model to model the spread of Covid-19 in Hong Kong was based on the built environment. The GWPR model obtained a close and spatially heterogeneous relationship between the factors and the risk of COVID-19 transmission. The study provided valuable insights that support policymakers in responding to the COVID-19 pandemic and future epidemics [9].

As well as Poisson regression, the GWPR model requires the variance equal as the mean value or model doesn’t contain over or under dispersion. Another method to handle over-dispersion in a geographical context was GWNBR [10]. Researches about the GWNBR model, among others, predicted the wildfire occurrence in the Great Xing’an Mountains. The model fitted and forewarned of the GWNBR model were better than the NB model, produced more precise and stable model parameter estimation, yielded a more realistic spatial distribution of model predictions, and detected the impact hotpots of these predictor variables [11]. Others, analysis of factors that influence maternal mortality in Central Java Province. It resulted that clean and healthy household behavior and the number of community health centers have a significant impact on maternal mortality. And also, there were three classifications based on GWBN results [12].

Some of NB model has been developed into multivariate analysis which some response variables, were BNB [13]–[15]. NB applied for modeling insurance claim counts in Singapore based on tri-variate responses. The trivariate NB model was proposed to accommodate the dependency among three types of claims: the third party bodily injury, own damage, and third-party property damage [16]. BNB model has been applied in many areas such as modelling the number of hospital stays and non-physician hospital outpatient visits in America. Sixteen variables, including socio-economic variables and insurance and health status variables, were used as independent variables [13]. BNB model has been used in modelling third-party liability claims and comprehensive cover claim which the data from 14,000 automobile policies from a major insurance company in Spain. It resulted that the BNBR model adequately captures the relationship between the two claim counts and the set of explanatory variables [17]. Application BNB model in health aspect has been done in modeling HIV and AIDS in Indonesia. It obtained that the estimation of BNB model was appropiate to model the HIV and AIDS which R-square was 64.9% [18]. The BNB model also developed into spatial analysis which was the GWBNB model [19]. It applied to model the infant mortality in East, Central and West Java, that resulted GWBNB model performed better than BNB model. Some of researches about dengue can be seen in Table 1.

Table 1. Related work of analysis dengue with spatial methods

No	Case	Variables	Methods	Reference
1	Dengue fever in Java island during 2007-2018	Weather effect	Moran’s index, Lisa	Dhewantara et.al [20]
2	dengue fever cases in Kaoshiung City, Taiwan, 2003-2008	Population density, transportation arteries, and water bodies	GWR	Hsueh, et. al [6]
3	Dengue in Jakarta from 2008-2016	Weather	Multiple regression	Fakhruddin et.all [21]
4	Dengue in Swat, Pakistan	Age, gender, and location of the cases were recorded when the patient was admitted to a health facility. Population figures have been extracted from the district health profile	GWR	Atique et. all [22]
5	dengue disease in Lahore, Pakistan	Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), land surface temperature (LST), built-up area, population density	GWR	Tariq and Zaidi [23]
6	dengue fever in Jhapa district, Nepal,	population density; proximity to road proportion of urban area; LST; and NDWI	Semiparametric GWR	Acharya et. all [24]
7	dengue fever in Cali, Colombia	socioeconomic, demographic, and environmental variables	GWR	Delmelle et.all [25]

According to table 1, the analysis of dengue fever has been developed in various spatial methods, especially in the GWR model. Some of researches of dengue uses modelling and spatial analysis such as GWR. Most of resulted of them were spatial model able to capture heterogeneity among the regions. Applied GWBNB, to model the spread dengue, has not ever done before. GWBNB model able to overcome the heterogeneity which is caused by spatial data. This paper proposed the GWBNB model to analyze the spread of dengue in NTT. Its novelty is to apply the GWBNB model for modeling dengue in NTT which was the number of dengue cases and their mortality cases based on population density, environment, and health facilities. Besides, GWBNB can model the dengue spread and know the significant factor of stretch. It is also helpful for mapping spread into some regions.

2. MATERIAL AND METHODS

2.1. NB model

NB model can be used to overcome overdispersion, which is based on the mixed model of Poisson-gamma. Negative binomial regression follows negative binomial distribution [26]:

$$f(x,\mu,\theta)=\frac{\Gamma(x+1/\theta)}{\Gamma(1/\theta)x!}\left(\frac{1}{1+\theta\mu}\right)^{1/\theta}\left(\frac{\theta\mu}{1+\theta\mu}\right)^{x} \tag{1}$$

For $x=0,1,2,\dots,n$ which x is a random variable, and θ and μ are the parameters of the distribution. If $\theta=0$ then negative binomial distribution has variance $V[Y]=\mu$. So that the NB model can be written as:

$$f(x_i)=\mu_i+\varepsilon \tag{2}$$

Based on Eq (2), the estimation of BNB model can be written as:

$$\mu_i=\exp\left(\beta_0+\beta_1X_{1i}+\beta_2X_{2i}+\cdots+\beta_pX_{pi}\right) \tag{3}$$

Which $i=1,2,\dots,n$ as number of observation, p is number of independent variables, μ_i is estimation of BNB model, β is coefficient parameter, X is independent variables, and ε is error of BNB model. Estimation of parameters uses maximum likelihood estimation (MLE). Same as with the regression model, there was hypothesis testing of the model and parameters in the NB model. Hypothesis testing of model uses Chi-square test, while for parameter testing uses Z test.

2.2. GWNBR model

GWNBR model results in local estimate parameters in each location which has different parameters. GWNBR model is as follows [27]:

$$Y_i \sim NB\left[\exp\left(\sum\nolimits_p \beta_p(u_i,v_i)X_{ip}\right), \theta(u_i,v_i)\right]; \tag{4}$$

Where Y_i is an independent variable, X_{ip} is an independent variable, (u_i, v_i) shows location, β_p shows coefficient parameter, θ is dispersion parameter.

2.3. BNB model

The bivariate negative binomial distribution developed the BNB model. Given Y_1 and Y_2 as the random variables which have a negative binomial distribution. Furthermore, they were modeled by regression with independent variables $X_1, X_2, \dots, X_p$, then BNB model as follows [15]:

$$(Y_1, Y_2) \sim BNB\,(\mu_1, \mu_2, \alpha_1, \alpha_2, \psi) \tag{5}$$

Eq. (5) follows the NB distribution in Eq. (1), which Y_1 and Y_2 are dependent variables, μ_1 and μ_2 are estimation model of independent variables, α_1 and α_2 are parameter of BNB distribution, and Ψ is dispersion parameter. Rely on eq. (5) and Eq.(3),the BNB model can be written as:

$$\text{Ln}\,(\mu_{ji}) = \beta_{j0} + \sum\nolimits_{k=1}^{p} \beta_{jk}X_{ki};\, i = 1,2,\dots,n;\, j = 1,2 \tag{6}$$

For i is observation, j follows independent variable, and p is the number of independent variables

Estimation parameter of BNB model uses *Maximum Likelihood Estimator* (MLE) with Newton-Raphson iteration. The likelihood function of the BNB model is as follows:

$$(\mu_1,\mu_2,\alpha_1,\alpha_2,\psi) = \prod_{i=1}^{n}\left\{\prod_{j=1}^{2}\prod_{t=0}^{y_j-1}\frac{(\alpha_j{}^{-1}+t)(\mu_{ji})^{y_{ji}}}{y_{ji}!}(\alpha_j{}^{-1})^{\alpha_j{}^{-1}}(\alpha_j{}^{-1}+\mu_{ji})^{-(y_{ji}+\alpha_j{}^{-1})}A\right\} \tag{7}$$

Which $A = [1 + \psi(e^{-y_{1i}} - c_1)(e^{-y_{2i}} - c_2)]$

The testing of BNB model uses *Maximum Likelihood Estimator Test* (MLRT) with hypothesis as follows:

$H_0 : \beta_{j1} = \beta_{j2} = \cdots = \beta_{jp} = 0;\ j = 1,2$

$H_1 :$ at least there is one of $_{jk} \neq 0$; $k = 1,2,\dots,p$

Statistic testing:

$$D = -2\ln\left(\frac{L(\hat{\omega})}{L(\hat{\Omega})}\right) = 2\left(\ln L(\hat{\Omega}) - \ln L(\hat{\omega})\right) \tag{8}$$

for $L(\hat{\omega})$ is maximum likelihood value for model excluded independent variables, while $L(\hat{\Omega})$ is maximum likelihood value for model excluded independent variable. Criteria of testing, rejected H_0 if $D > \chi^2_{[\alpha,(b-a)]}$. In detail, a is the number of parameters in H_0, while b is the number of parameters in H_1. Meanwhile, the parameter testing of the BNB model is as follows:

$H_0: \beta_{jk} = 0$

$H_1: \beta_{jk} \neq 0;\ j = 1,2;\ k = 1,2,\ldots,p$

Statistics testing is the value of $z = \frac{\hat{\beta}_{jk}}{se(\hat{\beta}_{jk})}$, while a criterion of testing is rejected H_0 if $|z| > z_{\alpha/2}$.

2.4. GWBNB model

GWBNB model is developed from the BNB model, which uses location weighted in a parameter estimate. The model of GWBNB is as follows (Ricardo & Carvalho, 2014):.

$$(Y_1, Y_2) \sim BNB(\mu_1(u_i, v_i), \mu_2(u_i, v_i), \alpha_1(u_i, v_i), \alpha_2(u_i, v_i), \psi(u_i, v_i)) \tag{9}$$

So the GWBNB model can be written as:

$$\ln(\mu_{ji}(u_i, v_i)) = \beta_{j0}^* + \sum_{k=1}^{p} \beta_{jk}^* X_{ki}(u_i, v_i) \tag{10}$$

Which $\mu_{ji}(u_i, v_i)$ is estimation of GWBNB model in each location, β_{jk}^* is coefficient parameter of GWBNB model in each location, $X_{ki}(u_i, v_i)$ is independent variable in each location. Same with the BNB model, the estimation parameter of the GWBNB model uses MLE. It results in parameter estimation, which maximizing the function of ln-likelihood. The difference between the estimation parameter of BNB and the GWBNB model was in the GWBNB model, which includes the weighted of location. The testing hypothesis of the GWBNB model uses MLRT as follows:

$H_0: \beta_{j1}^* = \beta_{j2}^* = \cdots = \beta_{jp}^* = 0;\ k = 1,2,\ldots,p;\ j = 1,2$

H_1: at least there is one $\beta_{jk}^* \neq 0$

The statistics testing for GWBNB model is:

$$D_{GW} = -2\ln\left(\frac{L(\hat{\omega}^{GW})}{L(\hat{\Omega}^{GW})}\right) = 2\left(\ln L(\hat{\Omega}^{GW}) - \ln L(\hat{\omega}^{GW})\right) \tag{11}$$

for $L(\hat{\omega}^{GW})$ is maximum likelihood value for model excluded independent variables in GWBNB, while $L(\hat{\Omega}^{GW})$ is maximum likelihood value for model excluded independent variable in GWBNB. Criteria of testing rejected H_0 if $D_{GW} > \chi^2_{[\alpha,(b-a)]}$. Meanwhile, parameter testing of GWBNB model as follows:

$H_0: \beta_{jk}^* = 0$

$H_1: \beta_{jk}^* \neq 0;\ k=1,2,\ldots,p;\ j= 1,2$

Statistics testing as follows:

$$z_{GW} = \frac{\hat{\beta}^*_{jk}}{se(\hat{\beta}^*_{jk})} \tag{12}$$

$\hat{\beta}^*_{jk}$ is estimation parameter of GWBNB model in each location and $se(\hat{\beta}^*_{jk})$ is standard error of coefficient parameter of GWBNB model.

Criteria of testing rejected H_0 if $|z_{GW}| > z_{\alpha/2}$.

2.5. Methods

The area of research was NTT province which consists of 22 regions. The data was collected from Statistics Bureau NTT-Indonesia and data from the ministry of health Indonesia in 2018. Dependent variables were the number of dengue case (Y_1) and the number of their mortality (Y_2) in each region. The independent variables consider the health facilities, population density, and also environmental aspects. Among others, population density (X_1), percentage of poverty (X_2), the number of doctors (X_3), the number of health facilities (X_4), percentage of the area which has good sanitation (X_5). The steps of analysis are as follows:

1 Checked the correlation between two dependent variables. It is needed before modeling two independent variables, because the model is simultaneous and related each other. If the two independent variables were not correlation significantly, they can’t modeling simultaneous.

2 Checked multicollinearity between independent variables. Multicollinearity has impacted bias to the model which high correlation but less of significant variables [29]. If the independent variables contain multicollinearity, it should be excluded or be handled by other methods such principal component analysis.

3 Modeled the number of dengue and their mortality used BNB model. Before the BNB model, we should checked the overdispersion of the model which follows the Poisson regression model. If the model has overdispersion, it should be handled by other methods, one of them is NB model. Checked the overdispersion was implemented by partially, even in this research use two dependent variables.

4 Checked heterogeneity in the number of dengue and their mortality. Checking heterogeneity use Glejser test [29]. If the model contain heterogeneity, then it indicates spatial effects in the model.

5 Modeled the number of dengue and their mortality used GWBNB model with Bi-square Adaptive Kernel. For modeling GWBNB, the first step is count the distance between the observations. It use Euclidean distance. Next, find the optimal bandwidth based on Bi-square

Adaptive Kernel. Actually, there are two kinds of Kernel as weighted namely fix and adaptive. Fix kernel consider the distance with regular distances, while adaptive kernel is fit to disperse distances [4].There are two kinds function of Kernel such as bisquare and Gaussian. Because the number of observation less than 20, we decide to use bisquare function, so as a weighted we use Adaptive Bisquare Kernel function. Based on optimal bandwidth, we find parameter estimate of GWBNB model.

6 Then, interpreting the GWBNB model with make cluster based on estimation result of GWBNB model. In this research we build three cluster among other small, medium, and large. It used to map the spread of the number of dengue and their mortality.

3. RESULTS AND DISCUSSIONS

To begin with the analysis results, we describe the data which consist of the number of dengue and their mortality with its factors in East Nusa Tenggara, Indonesia. It is first analysis to know the distribution of data which it is used to check the data. Table 2 shows statistics descriptive the numberof dengue and their mortality with its factors in East Nusa Tenggara 2018. The average of the number dengue (Y_1) was 72.68 with the highest value was 539 in West Manggarai. The number of dengue mortality (Y_2) has average which was 0.82 (approximate one) with the highest value was 5 in West Manggarai too. The population density (X_1) has the highest value was 2351(thousand per km) in Kupang city as a capital of East Nusa Tenggara province. The percentage of poverty (X_2) has highest value which was 34.85% in Central Sumba, while the lowest value was 9.61% in Kupang city. The highest number of doctors (X_3) and the highest number of health facilities (X_4) were 38 doctors in Kupang city and 36 units respectively. Based on Table 2, the variance of Y_1 is more remarkable than Y_2, and it can be seen from the minimum and maximum values on each variable. And also, population density has high variance. The highest population density was in Kupang city, which capital city of NTT province.

Table 2. Statistics descriptive of the dengue and its factors in NTT 2018

Value	Y_1	Y_2	X_1	X_2	X_3	X_4	X_5
Mean	72.68	0.82	227.14	22.04	18.32	18.36	22.44
Variance	15127.18	2.16	227718.12	52.89	71.37	62.62	839.53
Min	0	0	36	9.61	7	6	0
Max	539	5	2351	34.85	38	36	100

In addition, it requires correlation testing among the independent variables. The correlation between Y_1 and Y_2 variables was 0.853. It shows that Y_1 and Y_2 have high relation. It has been proven that a high case of dengue in certain areas has an increased risk of mortality. Checking significance of the correlation is needed before modeling the bivariate, which is hypothesis testing as follows:

H_0: no correlation between the number of dengue cases (Y_1) and their mortality (Y_2)

H_1: the number of dengue cases (Y_1) and their mortality (Y_2) were correlated

The t-value was 7.3162, and t_{table} was 2.086, so H0 was rejected. There was a correlation between dengue cases (Y_1) and their mortality (Y_2). The next step checked the multi-colinearity among the predictor variables. It can be seen in Table 3 which the value of Variance inflation Factors was less than 5, so they didn't have multi-colinearity.

Table 3. VIF's vale of predictor variables

Variable	VIF
X_1	2.58
X_2	1.53
X_3	3.24
X_4	1.87
X_5	1.79

The following step checked the overdispersion of the Poisson regression model. It can be seen in Table 2 that variance is more remarkable than mean both Y_1 and Y_2. It indicated that there was an over-dispersion of Y_1 and Y_2. Besides, checked of Pearson Chi-square value divided by degree of freedom resulted is greater than 1 (Table 4), which was overdispersion. Checked Person Chi-square was got by Poisson regression modeling. So, in this analysis, used BNB to model the Y_1 and Y_2.

Table 4. Checked Overdispersion

Variable	***Pearson's Chi-square*** **(a)**	**Degree of freedom (b)**	**(a)/(b)**
Y_1	3802.99	16	237.69
Y_2	34.85	16	2.18

3.1 Modeling the number of dengue cases (Y_1) and their mortality (Y_2) use the BNB model

Previously, either the number of dengue or their morality has overdispersion of model. For that reason, we need other model to handle the overdispersion, one of them is NB model. Because there are two dependents variables which have correlation each other, we modeling into BNB model. Firstly, for modeling the number of dengue and their mortality is estimate the model. Table 5 shows the estimation parameter of the BNB model of the number of dengue cases and their mortality in East Nusa Tenggara. Based on table 5, the number of dengue cases was impacted by the number of doctors significantly in 10% (β_3) and the number of health facilities in 5% (β_4). Meanwhile, the number of dengue mortality cases were affected by the percentage of poverty (β_2), the number of health facilities (β_4), and the percentage of areas with good sanitation (β_5).

Table 5. Parameter estimation of BNB model for Y_1 and Y_2 variables

Parameter	The number of dengue cases (μ_1)			The number of dengue mortality cases (μ_2)		
	Coefficient	Z	*p-value*	Coefficient	Z	*p-value*
β_0	-1.0076	-0.7868	0.4314	-0.6061	-0.3447	0.7303
β_1	-0.1651	-0.5239	0.6003	-0.0725	-0.3396	0.7341
β_2	0.0005	1.0954	0.2733	0.0023	104.2034	< 0.0001**
β_3	0.0858	1.8125	0.0699**	-0.0874	-1.3777	0.1683
β_4	0.1083	2.9163	0.0035*	-0.1078	-2.9021	0.0037**
β_5	0.0961	1.3613	0.1734	0.1682	3.2070	0.0013**

*) significant level in 10%; ** in 5%

The model of dengue case and their mortality use BNB model can be written as follows:

$\hat{\mu}_1 = \exp(-1{,}0076 - 0.1651X_1 + 0{,}0005X_2 + 0.0858X_3 + 0.1083X_4 + 0.0961X_5)$

$\hat{\mu}_2 = \exp(-0.6061 - 0.0725X_1 + 0.0023X_2 - 0.0874X_3 - 1.078X_4 + 0.1682X_5)$

For testing of BNB model as follows:

Hypothesis

$H_0 : \beta_{j1} = \beta_{j2} = \cdots = \beta_{j5} = 0$; $j=1,2$

$H_1 :$ at least there is one $\beta_{jk} \neq 0$; $j = 1,2$; $k = 1,2,\ldots,5$

The value of $D(\hat{\boldsymbol{\beta}})$ was 83.4472, which is less than $\chi^2_{(0{,}05;10)} = 18.31$. So, H_0 was rejected, which mean that the BNB model was significant to model the number of dengue and their mortality cases. This model can be interpreted as simultaneous. If each variable assumed value was one unit, the number of dengue cases could be estimated exp(-0.882), which was 2.4 units. For example, the number of doctors has increased in two units, and other variables had assumed constantly, then the number of dengue cases had 0.45 units. Whereas, if the number of health facilities increased two units, dengue cases were exp(0.774)=0.461.

On the other hand, the number of dengue mortality can be interpreted as if each variable had one unit, and then the dengue mortality estimation was exp(-1.48990)=0.225 units. If the percentage of poverty, the number of health facilities, and the percentage of areas with good sanitation increased by two units, then the estimated number of dengue mortalitywas 0.239 units.

3.2.Heterogeneity spatial testing

Based on Table 2 about statistics descriptive of the data, the value of variance in each variable shows high value. It means that, in each observation of data indicates heterogeneity of the data. The testing of heterogeneity uses Glesjer testing, while the observation of data contains location or space, so the heterogeneity testing is spatial heterogeneity testing . The testing is as follows:

Hypothesis

H_0 : $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_p^2 = \sigma^2$ (It hasn't heterogeneity spatial)

H_1: at least there is one $\sigma_i^2 \neq \sigma^2$; i = 1,2,...,n (It has heterogeneity spatial)

The statistic of testing is $G = -\left(n - j - 1 - \frac{1}{2}(j-1)\right)\ln\left(\frac{|\Sigma_\Omega|}{|\Sigma_\omega|}\right) \sim \chi^2_{\alpha;v}$, $v = jk$. A criterion of testing is rejected H_0 for $G > \chi^2_{\alpha;v}$. Based on the result, it got the value of G, which was 28.56709. For significant level 5%, the value of was $\chi^2_{(0{,}05;10)} = 18.31$, so H_0 was rejected. It can be concluded that there was heterogeneity spatial between the number of dengue and their mortality. So it would be better to model with the spatial method. This article proposed the GWBNB model to estimate bivariate, which included dengue and mortality cases. In this model has Akaiake Information Criteria (AIC) value, which was 364.4725.

3.3.Modelling the number of dengue cases (Y_1) and their mortality (Y_2) use the GWBNB model

BNB model with uses spatial weighted develops GWBNB model. The first step counts the Euclidean distance among areas. The next step counts the optimal bandwidth using kernel function, based on CV value. The optimum bandwidth of the GWBNB model for modelling the number of dengue and their mortality cases can be seen in Table 6. Both distance and bandwidth are used to determine the weighted spatial. Table 6 shows the bandwidth of the GWBNB model in Y_1 and Y_2, which was used for both of them. The value of bandwidth is based on each location which has various values. The bandwidth is used to determine the weighted which involved kernel function. The function kernel that was used in this research was Adaptive Bisquare Kernel.

Table 6. Optimum bandwidth of GWBNB model

Area	Bandwidth	Area	Bandwidth
Alor	4.852	Nagekeo	3.295
Belu	4.739	Ngada	3.409
Ende	2.779	Rote Ndao	3.491
East Flores	2.956	Sabu Raijua	2.926
Kupang city	3.630	Sikka	2.765
Kupang	3.929	West Sumba	4.945
Lembata	3.567	South West Sumba	5.151
Malaka	9.498	Central Sumba	4.605
Manggarai	3.924	East Sumba	4.078
West Manggarai	4.309	South Timor Tengah	4.582
East Manggarai	3.871	North Timor Tengah	4.517

The estimation of the GWBNB model for the number of dengue cases and their mortality can be seen in Appendix 1. Based on Appendix 1, each location has various estimate value in each variable. All variables in estimation of the number dengue were significant, except X_2 (the percentage of poverty) in Sikka regency was not significant. Meanwhile, estimation parameter for the number mortality dengue also significant, except the intercept's value and percentage of the area which has good sanitation (X_5) in East Flores. The summarize of the estimation parameter of the GWBN model for the number of dengue and their mortality cases can be seen in Table 7. Based on Table 7, can be inferred that the coefficient value of intercept, population density (β_1), the number of doctors (β_3) have positive coefficient, that those variables have contribution in the increasing of the number dengue in East Nusa Tenggara simultaneously. Meanwhile, for estimation the number of mortality dengue has negative impact to decrease the value was area which has good sanitation (β_5) and also the population density (β_1).They have role to decrease the number of mortality which was caused by dengue. The coefficient variables, both of the number dengue and their mortality which have less of variance were the population density and; area which has good sanitation.

Table 7. Statistics descriptive of estimation parameter GWBNB model

Variable	**Value**	β_0	β_1	β_2	β_3	β_4	β_5
Y_1	Min	0.000	0.000	-0.024	0.003	-0.064	-0.047
	Max	3.268	0.019	0.085	0.129	0.071	0.001
	Average	1.730	0.001	0.021	0.051	0.001	-0.014
	Variance	2.612	< 0.0001	0.002	0.003	0.002	< 0.0001
Y_2	Min	-5.804	-0.009	-0.619	-0.431	-0.838	-0.066
	Max	1.991	0.010	0.630	0.626	0.804	0.000
	Average	-0.625	-0.002	0.124	0.072	-0.067	-0.012
	Variance	7.725	< 0.0001	0.126	0.087	0.150	< 0.0001

Furthermore, the estimation number of dengue and their mortality use GWBNB model can be checked by analysis their estimate values of model. As we know that, the data of number dengue and their mortality were count data which have value greater equal than zero. So, if the estimate less than zero was invalid. And also the number of mortality dengue is less equal than the number of dengue in each region. Table 8 shows the number of dengue and their mortality cases estimation. The number of dengue mortality cases in some regions includes Alor, Belu, Lembata, Rote Ndao, Sabu Raijua, South Timor Tengah, and North Timor Tengah, were an invalid model. There was a rule that the number of dengue mortality cases must equal or less than the number of dengue cases. So the percentage of valid model mortality dengue cases was compared to dengue cases which were 73%.

Table 8. The value of $\hat{Y}_1$ and $\hat{Y}_2$ of GWBNB model

Area	Y_1	Y_2	**Area**	Y_1	Y_2
Alor	5	7*	Nagekeo	23	1
Belu	13	27*	Ngada	23	1
Ende	25	0	Rote Ndao	10	86*
East Flores	22	1	Sabu Raijua	1	544*
Kupang city	31	0	Sikka	30	5
Kupang	8	5	West Sumba	24	0
Lembata	18	375*	South West Sumba	30	0
Malaka	27	2	Central Sumba	13	0
Manggarai	40	2	East Sumba	30	4
West Manggarai	32	2	South Timor Tengah	29	544*
East Manggarai	23	2	North Timor Tengah	26	357*

*Invalid model

The predicting map of dengue's number and their mortality cases used the GWBNB model can be seen in Figure 1. Based on Figure 1, the number of the cluster was 4 group, which cluster 1 was 1-10; Cluster 2 was 12-23 (11 and 12 excluded, because they were unavailable in the result), cluster 3 was 24-30, and cluster 4 was 31-40 cases. The members of cluster 2 were neighbored each other, i.e., east Manggarai, Ngada, and Nagekeo. And also, in cluster 3, the neighbourhood was South Timor Tengah, North Timor Tengah, and Malaka. The GWBNB model produced AIC's value was 234.932 which it was less than BNB model. So that, GWBNB model resulted better than BNB for modeling the number of dengue cases and their mortality.

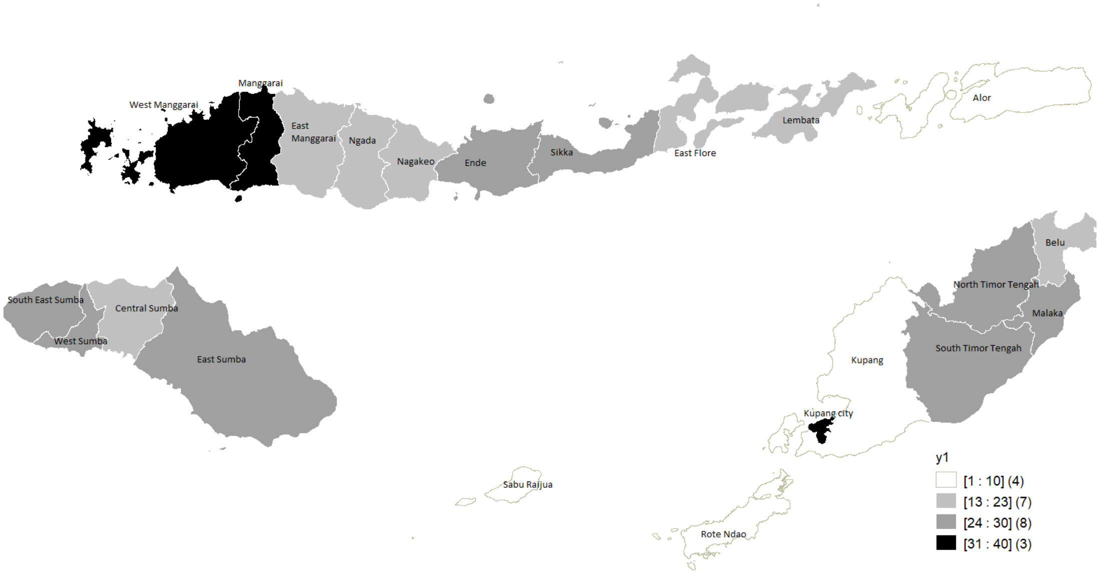

Figure 1. Predicting the number of dengue cases used GWBNB model

Besides, the AIC of GWBNB model less than BNB model which the GWBNB model has good model than BNB model. Moreover, the GWBNB model has more interpretation in spatial analysis. The GWBNB model can complement the BNB model which has homogeneity assumption of its. The GWBNB model can capture the significant variables both of the number of dengue and their mortality in each location. It able to map the locations according to estimation. Furthermore, this results are supported same result by [12] and [11] that analysis of geographically (in this case GWNB model) more effective than global regression (NB model).

4. CONCLUSION

Based on the analysis, it was known that the factors that impacted the number of dengue and their mortality cases. BNB model resulted in the number of dengue cases being affected by the number of health facilities. Meanwhile, the number of dengue mortality cases were affected by the percentage of poverty, the number of health facilities, and the percentage of areas with good sanitation. GWBNB model resulted better than BNB model in this cases. Government can consider the aspect of poverty, health facilities and behavior human in environment to decrease the number of spread dengue, especially in NTT. For further research can include the weather and climate factors such as seasons, temperature, intensity of rains, humidity and intensity of sunlight.

ACKNOWLEDGMENT

This research was funded by Diponegoro University as Development and Applied Research scheme with no of contract 233-103/UN7.6.1/PP/2021.

CONFLICT OF INTERESTS

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

Appendix 1. Table 9. Estimate parameter of GWBNB Model of the number dengue case in NTT

	Estimation of parameter							Probability of each parameter					
Area	theta	Intercept	X_1	X_2	X_3	X_4	X_5	Intercept	X_1	X_2	X_3	X_4	X_5
Alor	0.3679	0.0069	-0.0004	0.0844	0.1287	-0.0643	-0.0053	0.00	0.00	0.00	0.00	0.00	0.00
Belu	0.3717	0.0066	-0.0004	0.0792	0.1259	-0.0581	-0.0061	0.00	0.00	0.00	0.00	0.00	0.00
Ende	0.3813	3.1590	0.0006	-0.0134	0.0054	0.0267	-0.0153	0.00	0.00	0.00	0.00	0.00	0.00
East Flores	0.3608	2.6726	-0.0004	0.0069	0.0532	-0.0387	0.0014	0.00	0.00	0.00	0.00	0.00	0.00
Kupang city	0.3810	0.0054	0.0000	0.0661	0.1101	-0.0273	-0.0106	0.00	0.00	0.00	0.00	0.00	0.00
Kupang	0.3762	0.0062	-0.0003	0.0730	0.1217	-0.0477	-0.0074	0.00	0.00	0.00	0.00	0.00	0.00
Lembata	0.3684	0.0065	-0.0005	0.0850	0.1272	-0.0620	-0.0041	0.00	0.00	0.00	0.00	0.00	0.00
Malaka	0.1969	0.0001	0.0194	0.0019	0.0030	0.0024	0.0000	0.00	0.00	0.00	0.00	0.00	0.00
Manggarai	0.3874	3.2243	0.0007	-0.0168	0.0032	0.0290	-0.0167	0.00	0.00	0.00	0.00	0.00	0.00
West Manggarai	0.3878	3.2380	0.0007	-0.0179	0.0030	0.0298	-0.0176	0.00	0.00	0.00	0.00	0.00	0.00
East Manggarai	0.3873	3.2222	0.0007	-0.0167	0.0032	0.0289	-0.0166	0.00	0.00	0.00	0.00	0.00	0.00
Nagekeo	0.3857	3.1970	0.0006	-0.0152	0.0036	0.0279	-0.0155	0.00	0.00	0.00	0.00	0.00	0.00
Ngada	0.3862	3.2022	0.0006	-0.0156	0.0035	0.0282	-0.0158	0.00	0.00	0.00	0.00	0.00	0.00
Rote Ndao	0.3872	0.0037	0.0006	0.0504	0.0802	0.0235	-0.0182	0.00	0.00	0.00	0.00	0.00	0.00
Sabu Raijua	0.3895	0.0029	0.0017	0.0389	0.0779	0.0713	-0.0467	0.00	0.00	0.00	0.00	0.00	0.00
Sikka	0.3732	3.0895	0.0005	-0.0008	0.0153	0.0092	-0.0140	0.00	0.00	0.28	0.00	0.00	0.00
West Sumba	0.3882	3.2632	0.0009	-0.0223	0.0033	0.0346	-0.0219	0.00	0.00	0.00	0.00	0.00	0.00
South West Sumba	0.3882	3.2682	0.0009	-0.0218	0.0030	0.0338	-0.0213	0.00	0.00	0.00	0.00	0.00	0.00
Central Sumba	0.3880	3.2505	0.0009	-0.0218	0.0035	0.0343	-0.0215	0.00	0.00	0.00	0.00	0.00	0.00
East Sumba	0.3878	3.2238	0.0009	-0.0236	0.0046	0.0375	-0.0236	0.00	0.00	0.00	0.00	0.00	0.00
South Timor Tengah	0.3743	0.0064	-0.0003	0.0752	0.1227	-0.0517	-0.0069	0.00	0.00	0.00	0.00	0.00	0.00
North Timor Tengah	0.3728	0.0065	-0.0003	0.0779	0.1251	-0.0560	-0.0064	0.00	0.00	0.00	0.00	0.00	0.00

Appendix 2. Estimate parameter of GWBNB Model of the number dengue mortality cases in NTT

	Estimation of parameter							Probability of each parameter					
Area	**theta**	**Intercept**	**X_1**	**X_2**	**X_3**	**X_4**	**X_5**	**Intercept**	**X_1**	**X_2**	**X_3**	**X_4**	**X_5**
Alor	10.080	-2.982	-0.007	0.521	0.252	-0.342	-0.010	0.660	0.000	0.000	0.000	0.000	0.000
Belu	11.213	-4.301	-0.009	0.604	0.378	-0.439	-0.012	0.515	0.000	0.000	0.000	0.000	0.000
Ende	-0.275	1.032	0.003	-0.136	-0.137	0.204	-0.007	0.857	0.000	0.000	0.000	0.000	0.000
East Flores	4.196	-0.172	-0.002	0.147	0.033	-0.113	0.000	0.982	0.000	0.000	0.020	0.000	0.933
Kupang city	8.833	-5.804	-0.007	0.522	0.443	-0.450	-0.015	0.449	0.000	0.000	0.000	0.000	0.000
Kupang	10.154	-5.147	-0.008	0.572	0.413	-0.451	-0.012	0.427	0.000	0.000	0.000	0.000	0.000
Lembata	6.934	-1.542	-0.004	0.327	0.114	-0.215	-0.002	0.804	0.000	0.000	0.000	0.000	0.006
Malaka	2.420	0.000	0.003	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Manggarai	-0.031	1.544	0.002	-0.124	-0.136	0.183	-0.011	0.770	0.000	0.000	0.000	0.000	0.000
West Manggarai	-0.085	1.675	0.002	-0.129	-0.143	0.189	-0.011	0.763	0.000	0.000	0.000	0.000	0.000
East Manggarai	-0.027	1.526	0.002	-0.124	-0.135	0.182	-0.011	0.772	0.000	0.000	0.000	0.000	0.000
Nagekeo	0.008	1.278	0.002	-0.119	-0.126	0.177	-0.011	0.805	0.000	0.000	0.000	0.000	0.000
Ngada	0.006	1.336	0.002	-0.120	-0.128	0.178	-0.011	0.795	0.000	0.000	0.000	0.000	0.000
Rote Ndao	8.347	-1.463	-0.008	0.363	0.430	-0.520	-0.009	0.000	0.000	0.000	0.000	0.000	0.000
Sabu Raijua	7.861	-0.724	-0.008	0.395	0.626	-0.838	-0.066	0.969	0.000	0.000	0.000	0.000	0.000
Sikka	-7.603	1.912	0.010	-0.619	-0.431	0.804	-0.009	0.774	0.000	0.000	0.000	0.000	0.000
West Sumba	-0.391	1.991	0.003	-0.148	-0.180	0.226	-0.007	0.774	0.000	0.000	0.000	0.000	0.000
South West Sumba	-0.350	1.990	0.002	-0.146	-0.175	0.221	-0.008	0.776	0.000	0.000	0.000	0.000	0.000
Central Sumba	-0.350	1.917	0.003	-0.145	-0.175	0.221	-0.008	0.765	0.000	0.000	0.000	0.000	0.000
East Sumba	-0.521	1.902	0.003	-0.152	-0.193	0.242	-0.005	0.772	0.000	0.000	0.000	0.000	0.000
South Timor Tengah	11.370	-5.153	-0.009	0.630	0.453	-0.488	-0.015	0.452	0.000	0.000	0.000	0.000	0.000
North Timor Tengah	11.105	-4.555	-0.009	0.605	0.393	-0.449	-0.012	0.480	0.000	0.000	0.000	0.000	0.000

REFERENCES

[1] Ministry-Health-Indonesia, Situasi Penyakit Demam Beradrah di Indonesia tahun 2017, (2018), Jakarta, Indonesia.

[2] Kompas, Pasien DBD Capai 17.820, Ini 10 Provinsi dengan Jumlah Kasus Tertinggi, (2020), Jakarta, Indonesia.

[3] L. Anselin, Spatial Econometric: Methods and Models, (1988), Netherlands: Springer Netherlands.

[4] A. S. Fotheringham, C. Brunsdon, and M. Charlton, Geographically weighted Regression: The Analysis of Spatially Varying Relathionships, (2002), USA: John Wiley & Sons, Inc.

[5] W. Wei, J. Yuan-Yuan, Y. Ci, A. Ahan, and C. Ming-Qin, Local spatial variations analysis of smear-positive tuberculosis in Xinjiang using Geographically Weighted Regression model, BMC Public Health 16 (2016), 1–9.

[6] Y. H. Hsueh, J. Lee, and L. Beltz, Spatio-temporal patterns of dengue fever cases in Kaoshiung City, Taiwan, 2003–2008, Appl. Geogr. 34 (2012), 587–594.

[7] C. Xu, Y. Wang, W. Ding, and P. Liu, Modeling the Spatial Effects of Land-Use Patterns on Traffic Safety Using Geographically Weighted Poisson Regression, Networks Spat. Econ. 20 (2020), 1015–1028.

[8] S. Sun, Z. Li, H. Zhang, H. Jiang, and X. Hu, Analysis of HIV/AIDS epidemic and socioeconomic factors in sub-Saharan Africa, Entropy 22 (2020), 1–17.

[9] J. Huang, M. P. Kwan, Z. Kan, M. S. Wong, C. Y. T. Kwok, and X. Yu, Investigating the relationship between the built environment and relative risk of COVID-19 in Hong Kong, ISPRS Int. J. Geo-Information 9 (2020), 23-35.

[10] R. Fitriani and I. Gede Nyoman Mindra Jaya, Spatial modeling of confirmed COVID-19 pandemic in East Java Province by geographically weighted negative binomial regression, Commun. Math. Biol. Neurosci. 20 (2020), 58.

[11] Z. Su, H. Hu, M. Tigabu, G. Wang, A. Zeng, and F. Guo, Geographically weighted negative Binomial regression model predicts wildfire occurrence in the Great Xing'an Mountains better than negative Binomial Model, Forests, 10 (2019), 1–16.

[12] F. W. R. Fadilah, S. S. Handajani, E. Zukhronah, and H. Pratiwi, Geographically weighted negative binomial regression model to analysis of factors that influence on maternal mortality in Central Java Province, AIP Conf. Proc. 2202 (2019), 020105.

[13] S. So, D. H. Lee, and B. C. Jung, An alternative bivariate zero-inflated negative binomial regression model using a copula, Econ. Lett. 113 (2011), 183–185.

[14] C. M. Ng, S. H. Ong, and H. M. Srivastava, A class of bivariate negative binomial distributions with different index parameters in the marginals, Appl. Math. Comput. 217 (2010), 3069–3087.

[15] F. Famoye, On The Bivariate Negative Binomial Regression Model, J. Appl. Stat. 37 (2010), 969–981.

[16] P. Shi and E. A. Valdez, Multivariate negative binomial models for insurance claim counts, Insur. Math. Econ. 55 (2014), 18–29.

[17] F. Liu and D. Pitt, Application of bivariate negative binomial regression model in analysing insurance count data, Ann. Actuar. Sci. 11 (2017), 390–411.

[18] A. Tohari, N. Chamidah, and Fatmawati, Modeling of HIV and AIDS in Indonesia Using Bivariate Negative Binomial Regression, IOP Conf. Ser. Mater. Sci. Eng. 546 (2019), 1–7.

[19] Y. S. Dewi, Purhadi, Sutikno, and S. W. Purnami, Evaluation of geographically weighted multivariate negative Binomial method using multivariate spatial infant mortality data, J. Phys. Conf. Ser. 1397 (2019), 1-12.

[20] P. Dhewantara, M. Riandi, T. Wahono, M. Mujiyanto, and S. Hidayat, Geographical and temporal dynamics of dengue fever in Java island, Indonesia (2007–2018): Evidence base to guide effective interventions, Int. J. Infect. Dis. 101 (2020), 346-356.

[21] M. Fakhruddin, Assessing the interplay between dengue incidence and weather in Jakarta via a clustering integrated multiple regression model, Ecol. Complex. 39 (2019), 268-275.

[22] S. Atique, Investigating spatio-temporal distribution and diffusion patterns of the dengue outbreak in Swat, Pakistan, J. Infect. Public Health 11 (2018), 550–557.

[23] B. Tariq and A. Z. Zaidi, Geostatistical modeling of dengue disease in Lahore, Pakistan, SN Appl. Sci. 1 (2019), 1–12.

[24] B. K. Acharya, C. X. Cao, T. Lakes, W. Chen, S. Naeem, and S. Pandit, Modeling the spatially varying risk factors of dengue fever in Jhapa district, Nepal, using the semi-parametric geographically weighted regression model, Int. J. Biometeorol. 62 (2018), 1973–1986.

[25] E. Delmelle, M. Hagenlocher, S. Kienberger, and I. Casas, A spatial model of socioeconomic and environmental determinants of dengue fever in Cali, Colombia, Acta Trop. 164 (2016), 169–176.

[26] J. Hilbe, Negative Binomial Regression, Second Edition. New York: Cambridge University Press, (2011).

[27] A. Ricardo and T. Carvalho, Geographically Weighted Negative Binomial Regression-Incorporating Overdispersion. New York: Springer Science, (2013).

[28] A. R. da Silva, T. C. V. Rodrigues, Geographically Weighted Negative Binomial Regression—Incorporating Overdispersion, Stat. Comput. 24 (2014), 769–783.

[29] W. H. Greene, Econometric Analysis, 5th ed. New Jersey USA: Prentice Hall, (2002).