Mathematics and Statistics 10(5): 1133-1139, 2022
DOI: 10.13189/ms.2022.100525 http://www.hrpub.org

Geographically Weighted Negative Binomial Regression Modeling using Adaptive Kernel on the Number of Maternal Deaths during Childbirth

Fahimah Fauwziyah*, Suci Astutik, Henny Pramoedyo

Department of Statistics, Faculty of Mathematics and Natural Science, Brawijaya University, Indonesia

Received July 15, 2022; Revised September 15, 2022; Accepted September 29, 2022

Cite This Paper in the Following Citation Styles

(a): *[1] Fahimah Fauwziyah, Suci Astutik, Henny Pramoedyo , "Geographically Weighted Negative Binomial Regression Modeling using Adaptive Kernel on the Number of Maternal Deaths during Childbirth," Mathematics and Statistics, Vol. 10, No. 5, pp. 1133 - 1139, 2022. DOI: 10.13189/ms.2022.100525.*

(b): *Fahimah Fauwziyah, Suci Astutik, Henny Pramoedyo (2022). Geographically Weighted Negative Binomial Regression Modeling using Adaptive Kernel on the Number of Maternal Deaths during Childbirth. Mathematics and Statistics, 10(5), 1133 - 1139. DOI: 10.13189/ms.2022.100525.*

Abstract The standard model that is used for count data is Poisson Regression. In fact, most of the count data is overdispersed, which means that the response variable has greater variance than the mean. So the Poisson Regression cannot be used because overdispersion can cause inaccurate parameter estimators. One of the most widely used methods to overcome overdispersion is Negative Binomial Regression. If there are spatial effects such as spatial heterogeneity that are taken into Negative Binomial model, the appropriate method to analyze is Geographically Weighted Negative Binomial Regression (GWNBR). A spatial weighting matrix is required in the GWNBR model. In this study, three weighting functions were used, that is Adaptive Gaussian Kernel, Adaptive Bisquare Kernel, and Adaptive Tricube Kernel. From the three weighting functions, a model will be formed and the best model will be selected based on the smallest AIC. Count data used in this study is maternal deaths during childbirth in West Java Province, which is the highest case in Indonesia. The results of the analysis show that based on the smallest AIC, the best modeling in maternal deaths during childbirth in West Java is the GWNBR model using the Adaptive Gaussian Kernel weight. The results of the best model were obtained from three groups based on the predictor variables that had a significant effect.

Keywords Adaptive Kernel, Geographically Weighted Negative Binomial Regression, Overdispersion

1. Introduction

Discrete data is data obtained from counting, not measuring. In statistical analysis, the standard model for discrete data or count data is the Poisson Regression model [1]. Poisson Regression is suitable to be applied if it fulfills the assumption of equidispersion, which means the mean of the response variables is equal to the variance [2]. In fact, many equidispersion assumptions are violated because most of the data contain overdispersion. Overdispersion means the variance in the response variable is greater than the mean [3]. Overdispersion can cause parameter estimators to be biased and underestimate the standard error [4].

One of the popular methods of dealing with overdispersion is Negative Binomial Regression which is a combination of Poisson-Gamma distribution. Negative Binomial Regression will produce a model that is applied globally to all regions. Whereas each region has geographical, social, economic, and cultural conditions that are not always the same [5]. So we need a model that is able to consider the effect of the region (spatial).

In spatial modeling, it is necessary to test the assumption of spatial heterogeneity. Spatial heterogeneity is a spatial influence associated with environmental differences and geographical characteristics between observed regions [6].

Spatial heterogeneity can be shown by the different effects of predictor variables on response variables in each region [7]. A method that can effectively accommodate spatial heterogeneity and overdispersion is Geographically Weighted Negative Binomial Regression (GWNBR).

GWNBR is a powerful tool in modeling count data, especially when the data is non-stationary and overdispersion [8]. In addition, this method will produce realistic model predictions and stable parameter estimators [9]. The parameter estimators of the GWNBR method are local, which is each region will have different parameters [10]. In the analysis process, spatial weighting is required which can be formed with the kernel function. The weighting with the adaptive kernel function can produce a weight that adjusts the distribution of the observation point.

Based on the description above, this study aims to form a GWNBR model and select the model with the best weighting. The data used is the number of maternal deaths during childbirth in West Java province, because the highest cases in Indonesia are located in West Java. This research is expected to provide usefulness for users as a reference in using a method where the response variable is count data that contains overdispersion and has spatial heterogeneity.

2. Materials and Methods

The data used in this study were secondary data regarding the number of maternal deaths during childbirth in West Java. The observation units are 27 regencies/cities which are equipped with information about the latitude and longitude coordinates of each observation. There are six variables involved in this study, which are one response variable and five predictor variables. That is the number of maternal deaths during childbirth (Y), the percentage of pregnant women attending antenatal care at least 4 times (X_1), the percentage of births attended by a midwife (X_2), the number of health centers (X_3), the number of maternity hospitals (X_4), and the percentage of births in health facilities (X_5).

2.1. Poisson Regression

Poisson Regression is a non-linear regression model used to analyze count data. The response variable is based on the characteristic of the Poisson distribution. The Poisson Regression model is presented in equation (1).

$$\mu_i = \exp\left(\beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p\right) \tag{1}$$

2.2. Overdispersion

There is an assumption that must be met in Poisson Regression, that is the variance and mean of the response variable are equal (equidispersion). However, overdispersion is often encountered, where the dispersion parameter varies between observations in the population, so the population depends on the dispersion parameter (θ).

Examination of overdispersion assumption on Poisson Regression can be seen from Pearson Chi-Square divided by degrees of freedom. If $\theta > 1$ then this indicates an overdispersion in the data [11].

$$\theta = \frac{\chi^2}{n-p} \text{ with } \chi^2 = \sum_{i=1}^{n} \frac{\left(y_i - \hat{\mu}_?\right)^2}{\hat{\mu}_i} \tag{2}$$

where:

n : number of observations

p : number of parameters

y_i : value of the ith response variable

$\hat{\mu}_i$: estimator value from Poisson Regression model

2.3. Negative Binomial Distribution Conformity

Distribution testing is carried out to determine whether the response variable (Y) follows a Negative Binomial distribution or not. Distribution testing for discrete data was tested with the Chi-Square test based on the hypothesis:

$H_0 : p = p_i$, response variable follows a Negative Binomial distribution

$H_1 : p \neq p_i$, response variable does not follow a Negative Binomial distribution

Chi-Square test for the Negative Binomial distribution conformity presented in equation (3)[12].

$$\chi^2 = \sum_{i=0}^{m} \frac{\left(n_i - np_i\right)^2}{np_i} \tag{3}$$

The statistics above follows Chi-Square distribution with $m - r - 1$ degrees of freedom. The decision criterion will reject the null hypothesis if $p - value < \alpha$.

2.4. Negative Binomial Regression

Negative Binomial Regression is an alternative solution to overcome overdispersion. The response variable of this regression must follow Negative Binomial distribution. A negative Binomial is a combination of Poisson-Gamma distribution. The presence of Gamma distribution can accommodate overdispersion at Poisson Regression because it does not assume equidispersion in its application [13]. The negative Binomial Regression model is also presented as equation (1).

2.5. Spatial Weighting Matrix

Spatial weighting can be formed by determining the optimum bandwidth and calculating the euclidean distance. Bandwidth is a value that describes the maximum distance (radius) of a region that can influence another region. If the bandwidth is very small, fewer observations are within radius b, so resulting in a larger variance. Conversely, a very large bandwidth will cause a greater bias [14]. Optimum bandwidth is a value that minimizes the Cross Validation score which can be mathematically written as equation (4).

$$CV = \sum_{i=1}^{n} \left(y_i - \hat{y}_{\neq i}(b) \right)^2 \tag{4}$$

where:

b : bandwidth

y_i : observation in the i[th] area

$\hat{y}_{\neq i}(b)$: estimator of y_i at radius b with observation in the i[th] area is omitted from the estimation

Whereas euclidean distance between observation i and observation j presented in equation (5) below [15].

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2} \tag{5}$$

where:

u_i : latitude coordinates on i[th] area

u_j : latitude coordinates on j[th] area

v_i : longitude coordinates on i[th] area

v_j : longitude coordinates on j[th] area

Adaptive Kernel is a function that has a different bandwidth for each observation. In this study, spatial weighting with the Adaptive Kernel function was used because the Adaptive Kernel can produce a weighting amount that adjusts to the distribution of observation. In addition, the Adaptive Kernel is suitable for use in observations that spread out and have irregular patterns. The weight function is presented below [5]:

1. Adaptive Gaussian Kernel

$$w_{ij} = \exp\left[-\frac{1}{2} \left(\frac{d_{ij}}{b_i} \right)^2 \right] \tag{6}$$

2. Adaptive Bisquare Kernel

$$w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{b_i} \right)^2 \right]^2 , d_{ij} \le b_i \\ 0, d_{ij} > b_i \end{cases} \tag{7}$$

3. Adaptive Tricube Kernel

$$w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{b_i} \right)^3 \right]^3 , d_{ij} \le b_i \\ 0, d_{ij} > b_i \end{cases} \tag{8}$$

where:

d_{ij} : euclidean distance

b_i : bandwidth on i[th] area

2.6. Spatial Heterogeneity

The global regression model cannot explain the relationship between variables because the characteristics of the regions are spatially varied. Spatial heterogeneity is a phenomenon in spatial models due to the influence of differences in region characteristics and geographical location between observations. So testing the assumption of spatial heterogeneity is used to determine the existence of relationships between regions that indicate the presence or absence of characteristic differences between one region and another. The breusch-Pagan test is commonly used for detecting spatial heterogeneity.

2.7. Geographically Weighted Negative Binomial Regression (GWNBR)

Geographically Weighted Negative Binomial Regression (GWNBR) is an extension method of GWR that is quite effective in estimating count data that contains overdispersion and has spatial heterogeneity. The GWNBR model can be written as an equation (9)[8].

$$Y_i \sim NB\left[\exp\left(\sum_{k=0}^{p} \beta_k (u_i, v_i) x_{ik} \right), \theta(u_i, v_i) \right] \tag{9}$$

where:

x_{ik} : value of the p[th] predictor variable at observation area (u_i, v_i)

$\beta_k(u_i, v_i)$: regression coefficient of the p[th] predictor variable for each area (u_i, v_i)

$\theta(u_i, v_i)$: dispersion parameter for each area (u_i, v_i)

The parameter estimator of GWNBR uses Maximum Likelihood Estimation (MLE) method with Newton Raphson Iteration which generates local parameter estimation to each region.

Simultaneous parameter testing of the GWNBR model aims to determine the significance of the parameters together. Simultaneous testing is based on the hypothesis:

$H_0 : \beta_1(u_i, v_i) = \cdots = \beta p(u_i, v_i) = 0$

H_1 : at least there is one $\beta_j(u_i, v_i) \neq 0, j = 0, 1, \cdots, p$

$$G = -2 \log \left[\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right] \tag{10}$$

The decision criterion will reject the null hypothesis if $G > \chi_p^2$ or $p - value < \alpha$. Meanwhile the partial parameter testing of the GWNBR model aims to determine the significance of each parameter. Partial testing based on the hypothesis is below:

$H_0 : \beta_j(u_i, v_i) = 0$

$H_1 : \beta_j(u_i, v_i) \neq 0,\ j = 0, 1, \cdots, p$

$$t = \frac{\hat{\beta}_j(u_i, v_i)}{se\left(\hat{\beta}_j(u_i, v_i)\right)} \quad (11)$$

The decision criterion will reject the null hypothesis if $t > t_{\alpha/2}$ or $p - value < \alpha$.

2.8. Best Model Selection

The selection of the best model on GWNBR can be seen from the Akaike Information Criterion (AIC). The best models are determined based on the smallest or most minimum AIC values. The AIC is defined as follows:

$$AIC = -2\log\left(L\left(\hat{\beta}(u_i, v_i), \hat{\theta}(u_i, v_i)\right)\right) + 2k \quad (12)$$

Spatial weighting is an element that influences the value of the log-likelihood, so the greater value of weighting, also greater the log-likelihood and then the value of AIC will be smaller.

3. Result and Discussion

3.1. Overdispersion Testing

Overdispersion testing can be performed by calculating the Pearson Chi-Square divided by its degrees of freedom. From the calculation process obtained, the result of the dispersion parameter is 1.7665 which the value is more than one $(\theta > 1)$, so it could be concluded that there was an overdispersion.

3.2. Negative Binomial Distribution Testing

Before analyzing the negative binomial, it is necessary to test whether the response variable already conforms with the Negative Binomial distribution. Distribution testing can be done with the Chi-Square test according to equation (3). The result showed that the statistical value of the Chi-Squared test was 19.0395 with a p-value of 0.0604 which is more significant than 0.05, so it can be concluded that the response variable follows a negative binomial distribution.

3.3. Formation of Spatial Weighting Matrix

In forming a weighting matrix, euclidean distance is required from a region with all regions used in observations and optimum bandwidth in the calculation process. Weighting can be calculated using kernel function, which in this study will use three weighting functions: Adaptive Gaussian Kernel, Adaptive Bisquare Kernel, and Adaptive Tricube Kernel.

For example, one observation in this study was taken namely Bogor Regency. From the information of bandwidth and euclidean distance possessed, the weighting matrix is obtained using the Adaptive Gaussian Kernel as follows:

$W(u_1, v_1) = diag(1, 0.73749, \cdots, 0.01175)$

If using the Adaptive Bisquare Kernel weighting function, the weighting matrix is obtained as follows:

$W(u_1, v_1) = diag(1, 0.81502, \cdots, 0)$

While using the Adaptive Tricube Kernel weighting function according to equation (8), the weighting matrix presented as below:

$W(u_1, v_1) = diag(1, 0.91234, \cdots, 0)$

3.4. Spatial Heterogeneity Testing

Spatial heterogeneity tests are carried out to determine the spatial diversity (heterogeneity) because each region has different conditions. Spatial heterogeneity testing can be done using the Breusch-Pagan test based on the hypothesis:

$H_0 : \sigma_1^2 = \cdots = \sigma_p^2 = \sigma^2$, there is no spatial heterogeneity

H_1: at least there is one $\sigma_i^2 \neq \sigma^2$, there is spatial heterogeneity

The results of the spatial heterogeneity test are presented in Table 1.

Table 1. Spatial Heterogeneity Test Results

Weights	Breusch-Pagan	P-value
Adaptive Gaussian	278.0692	4.069×10^{-57}
Adaptive Bisquare	261.0592	1.773×10^{-53}
Adaptive Tricube	263.563	5.167×10^{-54}

Based on Table 1, it can be seen that the p-value on the three weightings is very small which is less than 0.05, so it can be concluded that there is spatial heterogeneity for all three weightings. This spatial heterogeneity shows that there are differences characteristic of each regency/city in maternal deaths during childbirth in West Java.

3.5. GWNBR Model

Simultaneous testing can be done with the Maximum Likelihood Ratio Test (MLRT). The results of the simultaneous testing are presented in Table 2.

Based on Table 2, it can be seen that the p-value of MLRT tests on the three weightings is very small which is less than 0.05, so it can be concluded that there is at least one predictor variable that has a significant effect on the model in both in the weighting of the Adaptive Gaussian Kernel, Adaptive Bisquare Kernel, and Adaptive Tricube

Kernel.

Table 2. Simultaneous Testing Results

Weights	MLRT	P-value
Adaptive Gaussian	92.7015	8.31×10^{-18}
Adaptive Bisquare	82.9611	8.72×10^{-16}
Adaptive Tricube	84.9215	3.43×10^{-16}

Partial testing of the GWNBR model was carried out to find which predictor variables had a significant effect on the response variable in each region. Parameter significance testing results in different parameters in each region because each weight produces a different significant predictor variable. So regencies/cities that have significant predictor variables in common can be grouped. The grouping based on Adaptive Gaussian Kernel, Adaptive Bisquare Kernel, and Adaptive Tricube Kernel weighting is listed in the table below.

Table 3. Grouping Regency/City Based on Significant Variables with Adaptive Gaussian Kernel

Group	Many Regency/City	Significant Predictor Variables
1	4	X_1, X_2, X_3, X_4
2	4	X_1, X_2, X_3, X_5
3	19	X_1, X_2, X_3, X_4, X_5

Based on Table 3, it shows the grouping of significant predictor variables in regencies/cities with Adaptive Gaussian Kernel weight divided into 3 groups. Group 1 consists of the percentage of pregnant women attending antenatal care at least 4 times, the percentage of births attended by a midwife, the number of health centers, and the number of maternity hospitals. Meanwhile, group 2 consists of the percentage of pregnant women attending antenatal care at least 4 times, the percentage of births attended by a midwife, the number of health centers, and the percentage of births in health facilities. Significant predictor variables in group 3 consisted of the percentage of pregnant women attending antenatal care at least 4 times, the percentage of births attended by a midwife, the number of health centers, the number of maternity hospitals, and the percentage of births in health facilities.

Based on Table 4, it shows the grouping of significant predictor variables in regencies/cities with Adaptive Bisquare Kernel weight divided into 5 groups. Group 1 consists of the percentage of births attended by a midwife and the number of health centers. Meanwhile, group 2 consists of the percentage of births attended by a midwife, the number of health centers, and the number of maternity hospitals. Group 3 consisted of the percentage of pregnant women attending antenatal care at least 4 times, the percentage of births attended by a midwife, the number of health centers, and the number of maternity hospitals. Group 4 consisted of the percentage of pregnant women attending antenatal care at least 4 times, the percentage of births attended by a midwife, the number of health centers, and the percentage of births in health facilities. While group 5 consisted of the percentage of pregnant women attending antenatal care at least 4 times, the percentage of births attended by a midwife, the number of health centers, the number of maternity hospitals, and the percentage of births in health facilities.

Table 4. Grouping Regency/City Based on Significant Variables with Adaptive Bisquare Kernel

Group	Many Regency/City	Significant Predictor Variables
1	1	X_2, X_3
2	13	X_2, X_3, X_4
3	2	X_1, X_2, X_3, X_4
4	9	X_1, X_2, X_3, X_5
5	2	X_1, X_2, X_3, X_4, X_5

Table 5. Grouping Regency/City Based on Significant Variables with Adaptive Tricube Kernel

Group	Many Regency/City	Significant Predictor Variables
1	1	X_1, X_3, X_4
2	13	X_2, X_3, X_4
3	2	X_1, X_2, X_3, X_4
4	9	X_1, X_2, X_3, X_5
5	2	X_1, X_2, X_3, X_4, X_5

Based on Table 5, it shows the grouping of significant predictor variables in regencies/cities with Adaptive Tricube Kernel weight divided into 5 groups. Group 1 consists of the percentage of pregnant women attending antenatal care at least 4 times, the number of health centers, and the number of maternity hospitals. Meanwhile, group 2 consists of the percentage of births attended by a midwife, the number of health centers, and the number of maternity hospitals. Group 3 consisted of the percentage of pregnant women attending antenatal care at least 4 times, the percentage of births attended by a midwife, the number of health centers, and the number of maternity hospitals. Group 4 consisted of the percentage of pregnant women attending antenatal care at least 4 times, the percentage of births attended by a midwife, the number of health centers, and the percentage of births in health facilities. While group 5 consisted of the percentage of pregnant women attending antenatal care at least 4 times, the percentage of births attended by a midwife, the number of health centers, the number of maternity hospitals, and the percentage of births in health facilities.

3.6. Selection of the Best Model

The selection of the best model can be seen based on the smallest AIC. It means that the parameter estimator is close to the actual parameter value. The AIC calculation is formulated as equation (12). In the calculation process, an element that influences the log-likelihood value is the weighting, where the greater weighting value makes the greater log-likelihood value. Therefore the AIC is getting smaller.

From the results of the GWNBR modeling that has been carried out, the AIC is presented in Table 6.

Based on Table 6, it can be seen that the GWNBR model with Adaptive Gaussian Kernel weighting has the smallest AIC. So the best model for the number of maternal deaths during childbirth in West Java is GWNBR with Adaptive Gaussian Kernel.

Table 6. Best Model Selection Criteria

Weights	AIC
Adaptive Gaussian	44.79975
Adaptive Bisquare	45.09485
Adaptive Tricube	45.04442

3.7. Mapping by the Best Model

Based on the criteria previously discussed, the best model is GWNBR with Adaptive Gaussian Kernel. Grouping map with the best model is shown in Figure 1. The regencies/cities in West Java based on significant predictor variables that influence maternal deaths during childbirth in Figure 1 show that the patterns of regency/city tend to clumped. Regions that close to each other tend to have the same characteristics related to the factors that influence the maternal deaths during childbirth.

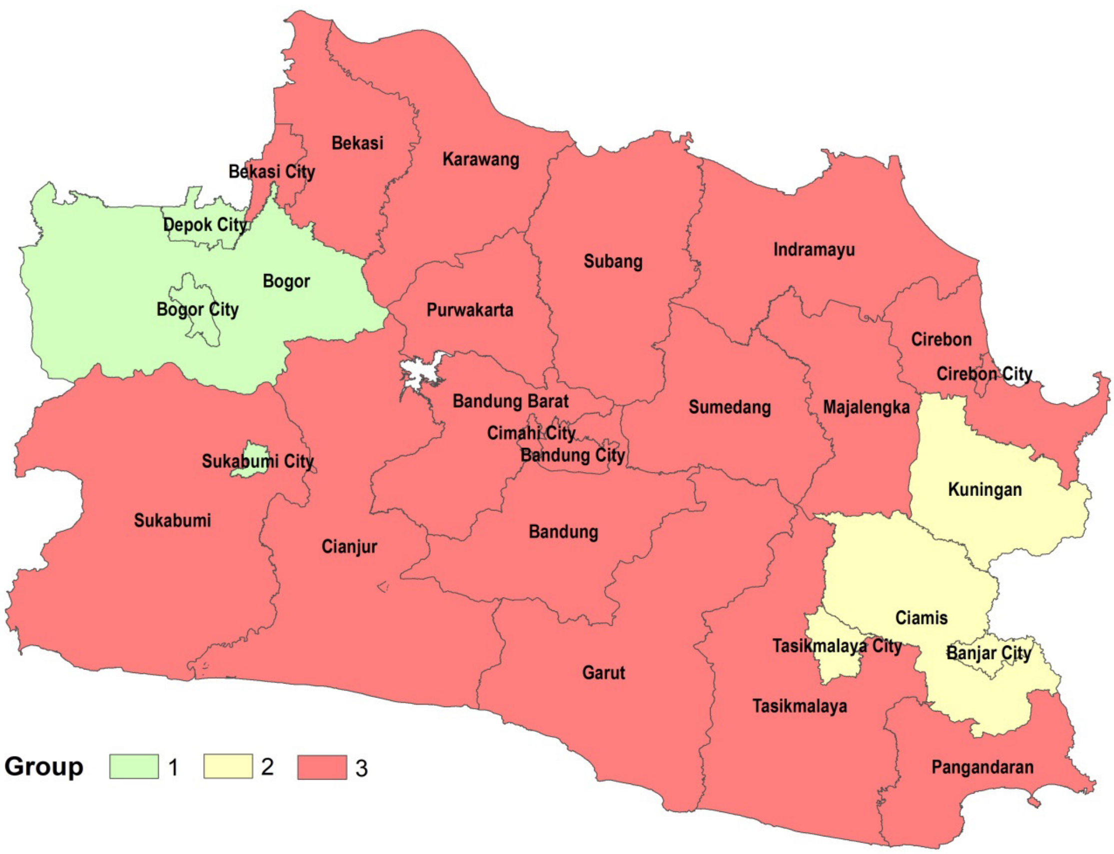

Figure 1. Grouping Map of the Best Model

4. Conclusions

The conclusion based on the results of the analysis is that the best modeling of maternal deaths during childbirth in West Java is the GWNBR model with the Adaptive Gaussian Kernel, which is based on the smallest AIC. The best model results were obtained from three groups based on the predictor variables that had a significant impact. The significant predictor variables in group 1 were the percentage of pregnant women attending antenatal care at least 4 times (X1), the percentage of births attended by a midwife (X2), the number of health centers (X3), and the number of maternity hospitals (X4). Meanwhile, group 2 was the percentage of pregnant women attending antenatal care at least 4 times (X1), the percentage of births attended by a midwife (X2), the number of health centers (X3), and the percentage of births in health facilities (X5). And group 3 was the percentage of pregnant women attending antenatal care at least 4 times (X1), the percentage of births attended by a midwife (X2), the number of health centers (X3), the number of maternity hospitals (X4), and the percentage of births in health facilities (X5).

REFERENCES

[1] C. Cameron and P. K. Trivedi, *Regression Analysis of Count Data*, 2nd ed. New York: Cambridge University Press, 2013.

[2] P. Tlhaloganyang and R. M. Sakia, "Zero Inflated Poisson Distribution in Equidispersed Data with Excessive Zeros," *Res. J. Math. Stat. Sci.*, vol. 8, no. 1, pp. 31–34, 2020.

[3] K. F. Sellers, A. W. Swift, and K. S. Weems, "A Flexible Distribution Class for Count Data," *J. Stat. Distrib. Appl.*, vol. 4, no. 22, pp. 1–21, 2017, doi: 10.1186/s40488-017-0077-0.

[4] A. Melliana, Y. Setyorini, H. Eko, S. Rosi, and Purhadi, "The Comparison of Generalized Poisson Regression and Negative Binomial Reression Methods in Overcoming Overdispersion," *Int. J. Sci. Technol. Res.*, vol. 2, no. 8, pp. 255–258, 2013, [Online]. Available: www.ijstr.org.

[5] K. P. Sinaga, "Poverty Data Modeling in North Sumatera Province Using Geographically Weighted Regression (GWR) Method," *Int. J. Sci. Res.*, vol. 4, no. 2, pp. 1738–1742, 2015, [Online]. Available: https://www.ijsr.net/archive/v4i2/SUB151577.pdf.

[6] C. Nurhayati, B. Warsito, H. Yasin, and A. Rusgiyono, "Robust Geographically Weighted Regression with Least Absolute Deviation (Case Study: The Percentage of Diarrhea Occurrence in Semarang 2015)," *J. Phys Conf. Ser*, vol. 1217, no. 012099, 2019, doi: 10.1088/1742-6596/1217/1/012099.

[7] R. Afifah, Y. Andriyana, and I. G. N. M. Jaya, "Robust Geographically Weighted Regression with Least Absolute Deviation Method in Case of Poverty in Java Island," in *AIP Conference Proceedings*, 2017, vol. 1827, no. 020023, doi: 10.1063/1.4979439.

[8] A. R. da Silva and T. C. V. Rodrigues, "Geographically Weighted Negative Binomial Regression-Incorporating Overdispersion," *Stat Comput*, vol. 24, pp. 769–783, 2014, doi: 10.1007/s11222-013-9401-9.

[9] 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*, vol. 10, no. 377, pp. 1–16, 2019, doi: 10.3390/f10050377.

[10] H. Yasin, I. Suryani, and P. Kartikasari, "Graphical Interface of Geographically Weighted Negative Binomial Regression (GWNBR) Model using R-Shiny," *J. Phys Conf. Ser*, vol. 1943, no. 012155, 2021, doi: 10.1088/1742-6596/1943/1/012155.

[11] A. Rahayu, "Model-model Regresi untuk Mengatasi Masalah Overdispersi pada Regresi Poisson," *J. Peqguruang Conf. Ser.*, vol. 2, no. 1, pp. 1–5, 2020.

[12] K. Mutaqin and Y. P. Adri, "Prediksi Klasemen Akhir Kompetisi Sepakbola Indonesia Menggunakan Metode Perluasan Ekspektasi Phytagoras," *J. Data Anal.*, vol. 2, no. 1, pp. 22–30, 2019, doi: 10.24815/jda.v2i1.14123.

[13] N. M. R. Keswari, I. W. Sumarjaya, and N. L. P. Suciptawati, "Perbandingan Regresi Binomial Negatif dan Regresi Generalisasi Poisson dalam Mengatasi Overdispersi (Studi Kasus: Jumlah Tenaga Kerja Usaha Pencetak Genteng di Br. Dukuh, Desa Pejaten)," *E-Jurnal Mat.*, vol. 3, no. 3, pp. 107–115, 2014, doi: 10.24843/mtk.2014.v03.i03.p072.

[14] Tizona et al, "Pemodelan Geographically Weighted Regression (Gwr) dengan Fungsi Pembobot Adaptive Kernel Bisquare untuk Angka Kesakitan Demam Berdarah di Kalimantan Timur Tahun 2015," *J. Eksponensial*, vol. 8, no. 1, pp. 87–94, 2017.

[15] A. S. Fotheringham, C. Brunsdon, and M. Charlton, *Geographically Weighted Regression the analysis of spatially varying relationship*. England: John Wiley and Sons Ltd, 2002.