Comparison of Two Weighting Functions in Geographically Weighted Zero-Inflated Poisson Regression on Filariasis Data

L Amaliana¹, A A R Fernandes¹, and Solimun¹
¹Statistics Department, Brawijaya University, Malang, Indonesia

luthfatul@ub.ac.id

Abstract. Spatial effects are factors to consider in modeling spatial data. These spatial effects can be spatial dependencies and spatial heterogeneity. The purposes of this study are: (1) to form Geographically Weighted Zero-Inflated Poisson (GWZIP) regression model to overcome the problem of spatial heterogeneity and the big enough proportion of zero-inflation in Filariasis case; and (2) to find the best weighting function between the fixed Gaussian kernel and the fixed Bi-square kernel based on the deviance of model. This study uses secondary data covering 35 districts in Central Java Province. The results of this study indicate that there are spatial heterogeneity and the 60% proportion of zero-inflation in Filariasis data. Based on the value of deviance model, it is known that the GWZIP model using fixed Gaussian kernel is better than the GWZIP model using fixed Bi-square kernel.

1. Introduction

Zero-Inflated Poisson (ZIP) regression analysis was initially proposed by [1] to model count data contains a large proportion of zero values in a defect manufacturing case. The number of zeros in the count data can lead to over dispersion. Then, ZIP regression becomes an alternative method to overcome the over dispersion problem in Poisson regression, especially when it comes from a large proportion of zero-inflation. ZIP regression has been studied by [2], [3], and [4] in health, epidemiology, and insurance fields, respectively. While [5] and [6] have been compared ZIP regression and Geographically Weighted Zero-Inflated Poisson (GWZIP) regression in their study. GWZIP regression is a development of ZIP regression which accommodates spatial effects into the model. In contrast to ZIP regression analysis, GWZIP regression can obtains models with local parameter estimators which more represented to every location.

Spatial data contains a lot of information about each of the locations studied. Every site may be different from each other or depend on each other. Spatial effects contained in spatial data can be spatial dependencies and spatial heterogeneity [7]. Spatial heterogeneity can be affected by geographical, economic, and socio-cultural differences of local communities. GWZIP regression is a part of spatial point analysis involving spatial heterogeneity factors into the model. This model is built by including a spatial weighted matrix into the ZIP model. It contains spatial (geographical) information from each location. The information from each location can be accommodated within the model. Entries of spatial weighted matrix are weights of each location relative to a location observed. It was obtained from the Euclidean distance function between locations by using kernel function and a bandwidth. Bandwidth is analogue to the radius of a circle, where if a point in it then it is considered to have an effect. A bandwidth used will affect the accuracy of the model which is associated with the
variance and bias of parameter estimator [8]. The optimum bandwidth plays a role in forming a weighting function.

There are several types of kernel functions used in spatial data modelling. Some of them are Gaussian, Exponential, Bi-square and Tricube kernel functions. Kernel function is used to estimate the parameter in the model if the distance function was a continuous and decreasing monotonic. However, different kernel function may produce parameter estimators that may be different for the same location, as well as for different locations. Gaussian kernel was used by [9] in Geographically Weighted Poisson Regression (GWPR) on infant mortality data.

Filaria data of every district in Central Java Province is an example of spatial data. It is also containing big enough proportion of zero-inflation because Filaria is only occurred in some districts in Central Java. Fixed Gaussian and Bi-square kernel function will be used in this study because that big proportion of zero-inflation. Therefore, this study aims to: (1) obtain GWZIP regression model that can deal with spatial heterogeneity effect and a great proportion of zero-inflation on Filaria case, especially in Central Java Province, and (2) determine the best weighting function between fixed Gaussian and fixed Bi-square kernel functions in GWZIP model.

This paper is structured as follows. Section 2 discusses the literature review about geographically weighted zero-inflated Poisson regression, multicollinearity, spatial heterogeneity and also spatial weighted matrix. Section 3 discusses the method used and section 4 discusses the result of modeling Filaria data by using fixed Gaussian and Bi-square kernel in geographically weighted zero-inflated Poisson regression.

2. Geographically weighted zero-inflated regression
This study uses a regression method and two kernel functions. In this section, spatial heterogeneity problem is explained and spatial weighted matrix included. However, geographically weighted zero-inflated Poisson regression will be discussed first.

2.1. The Model
Zero-Inflated Poisson regression analysis commonly used for count data whose dependent variable contains zero value in large proportion (zero-inflation). ZIP regression is inappropriate when applied to count data involving spatial effects. Then, GWZIPR model becomes an alternative method to deal with. GWZIPR model is a local form of ZIP regression model which result in different local estimator parameter to each location. For each observation on the dependent variable is taken from different locations \((u_i, v_i)\) that are \(y_1, y_2, \ldots, y_n\). In the GWZIPR model the dependent variable \(Y_i\) have different opportunities for \(y_i = 0\) and \(y_i > 0\), that [5]:

\[
P(Y_i = y_i) = \begin{cases} 
\pi_i + (1 - \pi_i)e^{-\mu_i}, & \text{for } y_i = 0 \\
\frac{(1 - \pi_i)e^{-\mu_i} \mu_i^{y_i}}{y_i!}, & \text{for } y_i > 0
\end{cases}
\]

which

\[
\mu_i = e^{x_i^T \beta(u_i, v_i)} \quad \text{and} \quad \pi_i = \frac{e^{x_i^T \gamma(u_i, v_i)}}{1 + e^{x_i^T \gamma(u_i, v_i)}}
\]

In this case, \(\beta(u_i, v_i)\) and \(\gamma(u_i, v_i)\) are parameters located in \((u_i, v_i)\), \(X\) is covariates matrix associated with the probability on zero state \((y_i = 0)\) and mean on Poisson state \((y_i > 0)\). The parameter estimation used in GWZIP regression is Maximum Likelihood (ML) method with In-likelihood function as follows:
\[ \ln L(\mathbf{y}(u_i, v_i), \mathbf{\beta}(u_i, v_i)) = \\
\sum_{i=1}^{n} \left( \ln \left( e^{x_i^T \mathbf{y}(u_i, v_i)} + e^{-x_i^T \beta(u_i, v_i)} \right) - \ln \left( 1 + e^{x_i^T \mathbf{y}(u_i, v_i)} \right) \right) w_{il}(u_i, v_i), \quad y_i = 0 \]
\[ \sum_{i=1}^{n} \left( -e^{x_i^T \beta(u_i, v_i)} + y_i x_i^T \beta(u_i, v_i) \right) - \ln \left( 1 + e^{x_i^T \mathbf{y}(u_i, v_i)} \right) - \ln (y_i!) \right) w_{il}(u_i, v_i), \quad y_i > 0 \]  

(3)

2.2. Multicollinearity
Multicollinearity is the main problem in regression analysis, as well as missing data and outliers. It is a condition in which covariates are highly correlated. The presence of multicollinearity can make parameter estimation being inaccurate. According to [10], the value of Variance-Inflating Factor (VIF) more than 10 indicates the presence of multicollinearity. VIF shows how the variance of parameter estimation increases due to the presence of multicollinearity. VIF values are formulated as follows:

\[ VIF_j = \frac{1}{1 - R_j^2} \]  

(4)

\( R_j^2 \) is coefficient of determination between \( X_j \) and another predictor variable.

2.3. Spatial Heterogeneity
The problem with spatial data other than spatial dependencies is the presence of spatial heterogeneity. Spatial heterogeneity is due to the different characteristics of the observations in one location and the observation at other sites. According to [7], spatial effects in the form of spatial heterogeneity can be identified using Breusch-Pagan test. However, the Breusch-Pagan test is sensitive to the assumption of normality. Therefore, identification of spatial heterogeneity was performed using Koenker-Basset testing [10]. Koenker-Basset (KB) test is based on residual squares (\( \hat{e}_i \)) regression results of response variable against predictor variables used. The residual squares are regressed to the squares of the estimates of the initial model regression. The hypothesis used is \( H_0 \) : there is no heterogeneity and \( H_1 \) : there is heterogeneity. The test statistic used is:

\[ Z_{value} = \frac{\hat{\gamma}_1}{se(\hat{\gamma}_1)} \]  

(7)

which \( e(\hat{\gamma}_1) = \sqrt{var(\hat{\gamma}_1)} \). Reject \( H_0 \) if \( |Z_{value}| > Z_{a/2} \) or \( p-value < \alpha \). If the decision is failed to reject \( H_0 \), then it can be concluded there is no spatial heterogeneity.

2.4. Spatial Weighted Matrix
Weights have an important role in spatial data. It is representing the location where each data is taken. A location can be represented by a coordinate point, such as Latitude and Longitude. Nearby locations should show similar relationships, and vice versa. Distant locations also show spatial heterogeneity. It is indicated by a weighted matrix, \( W(u_i, v_i) \), whose entries are a function of Euclidean distance between locations. Two of kernel functions commonly used are Gaussian and Bi-square kernel. It can describe as follows, respectively:

\[ w_{il}(u_i, v_i) = \exp \left( -\frac{1}{2} \left( \frac{d_{il}}{h} \right)^2 \right) \]  

(8)

and

\[ w_{il}(u_i, v_i) = \begin{cases} 
1 - \left( \frac{d_{il}}{h} \right)^2 & d_{il} < h \\
0 & d_{il} \geq h 
\end{cases} \]  

(9)

which

\[ d_{il} = \sqrt{(u_i - u_l)^2 + (v_i - v_l)^2} \]  

(10)
$d_{il}$ is the Euclidean distance between the $i$-th location and the $l$-th location, while $h$ is bandwidth. Selection of optimum bandwidth can be done by using Cross Validation (CV) method. This method is defined as:

$$CV(h) = \sum_{i=1}^{n}(y_i - \hat{y}_{si}(h))^2$$  
(11)

$\hat{y}_{si}(h)$ is the estimator of $y_i$ if location ($u_i, v_i$) is not included in the estimation and $n$ is number of locations. Minimum CV values can be obtained when the optimum bandwidth is used.

3. Method

Filariasis data used in this study is an empirical data from Central Java Province Health Department [11]. It is a cross-sectional data including 35 districts in Central Java, Indonesia including each of its longitude and latitude. The number of Filariasis case in every district is the dependent variable. Meanwhile, the steps of analysis by using R software 3.3.0 in this study can be described as follows:

1. Do the descriptive analysis of variables
2. Testing the spatial heterogeneity effect
3. Identify proportion of zero-inflation, over dispersion and multicollinearity case
4. Find the optimum fixed bandwidth and Euclidean distance among districts based on the longitude and latitude of each district
5. Find spatial weighted matrix by using Gaussian and Bi-square kernel function
6. Modeling Filariasis case by using GWZIP regression
7. Testing the significance parameter of GWZIP model simultaneously. While the tested statistical used for partially significance parameter is:

$$Z_{value} = \frac{\hat{\beta}_i}{se(\hat{\beta}_i)} \text{ or } Z_{value} = \frac{\hat{y}_i}{se(\hat{y}_i)}$$  
(12)

which $\hat{\beta}_i$ and $\hat{y}_i$ are parameters, $se(\hat{\beta}_i)$ and $se(\hat{y}_i)$ are the standard error of $\hat{\beta}_i$ and $\hat{y}_i$, respectively.
8. Compare and determine the best kernel function.

4. Result and discussion

Filariasis is an infectious disease infected by all the species of mosquito, like *Aedes, Anopheles, Culex,* etc. Filariasis also caused by filaria worm that transmitted by mosquito bites. Three species worms caused filariasis disease in Indonesia are *Wuchereria bancrofti, Brugia timori* dan *Brugia malayi.* Data on Filariasis case in every district is a non-negative integer. It is ranging from zero to ten. The covariates which are associated to Filariasis case used in this study are percent of household behave in a clean and healthy way ($X_1$), percent of families with healthy latrines ($X_2$), percent of families with healthy house ($X_3$), and percent of families with healthy drinking water ($X_4$). The descriptive analysis of Filariasis data is summarized in Table 1 as follows.

| Variable | Mean | Standard Deviation | Min | Max |
|----------|------|--------------------|-----|-----|
| Y        | 0.971| 1.963              | 0.00| 10.00 |
| X_1      | 43.50| 14.92              | 8.50| 65.74 |
| X_2      | 76.03| 16.08              | 25.32| 97.46 |
| X_3      | 58.17| 25.72              | 0.00| 89.00 |
| X_4      | 56.36| 25.53              | 0.00| 100.00 |

There are 34 cases of Filariasis in Central Java followed Poisson distribution with $\mu=0.971$. The highest number of Filariasis case is ten cases in which occurred in Demak district, meanwhile it was not occurred in other 21 districts.
The spatial heterogeneity effect on Filariasis data in this study is tested by using Koenker-Basset test. Based on the result of Koenker-Basset test with p-value = 0.0116 and \( \alpha = 0.05 \), it was concluded that there is spatial heterogeneity effect on Filariasis data.

The identification result of zero-inflation, over dispersion and multicollinearity case in this study showed in Table 2, Table 3 and Table 4.

**Table 2. Zero-inflation checking**

| Number of Filariasis Case | Frequency | Percent | Cumulative Percentage |
|---------------------------|-----------|---------|-----------------------|
| 0                         | 21        | 60.00%  | 60.00%                |
| 1                         | 7         | 20.00%  | 80.00%                |
| 2                         | 4         | 11.43%  | 91.43%                |
| 4                         | 1         | 2.86%   | 94.14%                |
| 5                         | 1         | 2.86%   | 97.14%                |
| 10                        | 1         | 2.86%   | 100.00%               |

Table 2 presents the result of zero-inflation checking on Filariasis case. The sample data has a great enough proportion of zero-inflation. It is about 60 percent. According to [1], the proportion of zero-inflation called big if it was more than 50.

**Table 3. Over dispersion checking**

| Deviance | Degrees of Freedom | Deviance / Degrees of Freedom |
|----------|--------------------|------------------------------|
| 67.964   | 30                 | 2.265                        |

Table 3 shows the result of over dispersion checking. The deviance divided by degrees of freedom (df) is greater than 1. It indicates the evidence of over dispersion in the data. It is also indicated by sample mean of dependent variable which smaller than sample variance in Table 1.

**Table 4. Multicollinearity checking**

| Covariates | \( X_1 \) | \( X_2 \) | \( X_3 \) | \( X_4 \) |
|------------|-----------|-----------|-----------|-----------|
| VIF        | 1.040     | 1.077     | 2.174     | 2.200     |

Multicollinearity is a common problem in regression analysis. One of many ways to identify it is by looking at Variance Inflation Factor (VIF). According to [9], the presence of multicollinearity is identified when the value of VIF is greater than 10. Table 4 presents VIF of covariates are smaller than 10. It can conclude that there is no multicollinearity among covariates. Furthermore, all the covariates can be involved to model Filariasis case by using GWZIP regression. The first step to model GWZIP regression on Filariasis data is find the longitude and latitude of each district in Central Java. It was a coordinate point of each district as summarized in Table 5.

According to coordinate point of longitude and latitude in Table 5, then the optimum bandwidth can be obtained by using Cross-Validation (CV) method. The optimum bandwidth used in this study is fixed for all districts and is present in Table 6. Meanwhile, the Euclidean distance between districts relative to a district observed is also calculated based on coordinate points in Table 5. Table 7 summarized the Euclidean distance of each district relative to Cilacap district.
Table 5. Longitude and latitude of districts in Central Java

| No. | District  | Longitude | Latitude  | No.  | District  | Longitude | Latitude |
|-----|-----------|-----------|-----------|-----|-----------|-----------|----------|
| 1   | Cilacap   | 7.63      | 108.79    | 19  | Kudus     | 7.06      | 111.72   |
| 2   | Banyumas  | 7.44      | 109.05    | 20  | Jepara    | 6.60      | 110.79   |
| 3   | Purbalingga| 7.33      | 108.88    | 21  | Demak     | 6.94      | 110.64   |
| 4   | Banjarnegara| 7.36     | 109.62    | 22  | Semarang  | 7.28      | 110.45   |
| 5   | Kebumen   | 7.64      | 109.69    | 23  | Temanggung| 7.39      | 110.58   |
| 6   | Purworejo | 7.72      | 109.97    | 24  | Kendal    | 6.97      | 109.98   |
| 7   | Wonosobo  | 7.39      | 109.89    | 25  | Batang    | 7.03      | 109.86   |
| 8   | Magelang  | 7.51      | 110.24    | 26  | Pekalongan| 7.04      | 109.64   |
| 9   | Boyolali  | 7.56      | 110.60    | 27  | Pemalang  | 8.11      | 109.48   |
| 10  | Klaten    | 7.63      | 110.63    | 28  | Tegal     | 7.05      | 109.15   |
| 11  | Sukoharjo | 7.68      | 110.40    | 29  | Brebes    | 6.73      | 108.94   |
| 12  | Wonogiri  | 7.89      | 110.99    | 30  | Kota Magelang| 7.47 | 110.21   |
| 13  | Karanganyar| 7.62      | 110.92    | 31  | Kota Surakarta| 7.77 | 110.76   |
| 14  | Sragen    | 7.38      | 110.96    | 32  | Kota Salatiga| 7.34 | 110.50   |
| 15  | Grobogan  | 7.25      | 110.33    | 33  | Kota Semarang| 7.02 | 110.39   |
| 16  | Blora     | 6.57      | 111.41    | 34  | Kota Pekalongan| 6.88 | 109.68   |
| 17  | Rembang   | 7.17      | 111.25    | 35  | Kota Tegal| 6.87      | 109.12   |
| 18  | Pati      | 6.71      | 111.04    |     |           |           |          |

Table 6. Optimum fixed bandwidth

| Kernel function | Bandwidth |
|-----------------|-----------|
| Gaussian        | 1.093     |
| Bi-square       | 2.457     |

Table 7. Euclidean distance and weight of Cilacap district

| District | Euclidean Distance | Gaussian | Bi-square | Gaussian | Bi-square |
|----------|--------------------|----------|-----------|----------|-----------|
| 1        | 0.000              | 1.000    | 1.000     | 0.024    | 0.000     |
| 2        | 0.322              | 0.958    | 0.966     | 0.120    | 0.026     |
| 3        | 0.313              | 0.960    | 0.968     | 0.195    | 0.125     |
| 4        | 0.873              | 0.727    | 0.764     | 0.300    | 0.274     |
| 5        | 0.900              | 0.712    | 0.750     | 0.255    | 0.211     |
| 6        | 1.183              | 0.556    | 0.590     | 0.461    | 0.481     |
| 7        | 1.126              | 0.588    | 0.624     | 0.533    | 0.564     |
| 8        | 1.455              | 0.412    | 0.422     | 0.639    | 0.677     |
| 9        | 1.811              | 0.253    | 0.208     | 0.744    | 0.780     |
| 10       | 1.840              | 0.242    | 0.193     | 0.823    | 0.852     |
| 11       | 1.611              | 0.337    | 0.325     | 0.706    | 0.743     |
| 12       | 2.215              | 0.128    | 0.035     | 0.425    | 0.438     |
| 13       | 2.130              | 0.150    | 0.062     | 0.195    | 0.125     |
| 14       | 2.184              | 0.136    | 0.044     | 0.284    | 0.252     |
| 15       | 1.586              | 0.349    | 0.340     | 0.293    | 0.264     |
| 16       | 2.826              | 0.035    | 0.000     | 0.567    | 0.602     |
| 17       | 2.503              | 0.073    | 0.000     | 0.750    | 0.785     |
| 18       | 2.431              | 0.084    | 0.000     |          |           |
Each district has their spatial weighted matrix. It was a diagonal matrix which is weight of each district relative to a district observed in its main diagonal, such as Cilacap weighted matrix summarized in Table 7. The parameter estimation of GWZIP model for Cilacap district shows in Table 8 as follows.

**Table 8. Parameter estimation of GWZIP model for Cilacap district**

| Parameter | Estimate | Std. Error | Z-value | Parameter | Estimate | Std. Error | Z-value |
|-----------|----------|------------|---------|-----------|----------|------------|---------|
| $\hat{\beta}_0$ | -0.019 | 2.435 | -0.008 | $\hat{\beta}_0$ | 0.040 | 2.429 | 0.016 |
| $\hat{\beta}_1$ | 0.031 | 0.005 | 59.041* | $\hat{\beta}_1$ | 0.023 | 0.001 | 42.226* |
| $\hat{\beta}_2$ | -0.021 | 0.004 | -54.742* | $\hat{\beta}_2$ | -0.019 | 0.000 | -47.472* |
| $\hat{\beta}_3$ | 0.015 | 0.004 | 36.009* | $\hat{\beta}_3$ | 0.016 | 0.000 | 36.345* |
| $\hat{\beta}_4$ | -0.005 | 0.003 | -16.811* | $\hat{\beta}_4$ | -0.006 | 0.000 | -18.185* |
| $\gamma_0$ | 1.923 | 9.879 | 0.195 | $\gamma_0$ | 2.162 | 9.814 | 0.220 |
| $\gamma_1$ | 0.122 | 0.021 | 59.041* | $\gamma_1$ | 0.093 | 0.002 | 42.226* |
| $\gamma_2$ | -0.083 | 0.015 | -54.742* | $\gamma_2$ | -0.074 | 0.002 | -47.472* |
| $\gamma_3$ | 0.061 | 0.017 | 36.009* | $\gamma_3$ | 0.063 | 0.002 | 36.345* |
| $\gamma_4$ | -0.021 | 0.012 | -16.811* | $\gamma_4$ | -0.023 | 0.001 | -18.185* |

*) significant at $\alpha = 0.05$ ($Z_{table}=1.96$)

According to Z-value of the first GWZIP model (Gaussian kernel) in Table 8, it was obtained that $X_1$, $X_2$, $X_3$, $X_4$ are significant at $\alpha = 0.05$. The same condition also showed by the second GWZIP model using (Bi-square kernel). It can be concluded that all the covariates are significant partially in the model. The deviance of models is summarized in Table 9 as follows:

**Table 9. Comparison of two weighting function**

| GWZIP Model | Kernel function | Deviance |
|-------------|-----------------|----------|
| (1) | Gaussian | 165.579 |
| (2) | Bi-square | 171.021 |

Table 9 presents the deviance of model (1) is smaller than model (2) with 25 degree of freedom. The smaller deviance is the better model. Therefore, model (1) is better than model (2). In other word, Gaussian kernel function is better fit to Bi-square kernel function in GWZIP model for Filariasis data.

According to Table 9, the deviance of GWZIP model (1) is greater than $\chi^2_{(0.05;8)} = 15.07$. Therefore, GWZIP model (1) is significant simultaneously at $\alpha = 0.05$. The first GWZIP model for Cilacap district based on Table 8 can be written as follows:

$$\ln(\hat{\mu}_t) = -0.019+0.031x_{1t} -0.021x_{2t} +0.015x_{3t} -0.005x_{4t}$$  \hspace{1cm} (i)

and

$$\logit(\hat{\pi}_t) = 1.923+0.112x_{1t} -0.083x_{2t} +0.061x_{3t} -0.021x_{4t}$$  \hspace{1cm} (ii)

which $X_1$ is household behave in a clean and healthy way, $X_2$ is families with healthy latrines, $X_3$ is families with healthy house, and $X_4$ families with healthy drinking water.

Model (i) explained that: (1) every 1% of change in household behave in a clean and healthy way can increasing the probability of the average of Filariasis case 1.031 times than previously; (2) every 1% of change in families with healthy latrines can decreasing the probability of the average of Filariasis case 1.021 times than previously; (3) every 1% of change in families with healthy house can increasing the probability of the average of Filariasis case 1.015 times than previously; and (4) every 1% of change in families with healthy drinking water can decreasing the probability of the average of Filariasis case 1.005 times than previously, if another covariates are constant. While model (ii) shows the probability of Filariasis case not occurred in every district are associated to household behave in a
clean and healthy way, families with healthy latrines, families with healthy house, and families with healthy drinking water.

Every district has different weighted matrix, then a district can have different significant covariates than other districts. Some districts can form some groups based on the significant variable. However, GWZIP model using Gaussian kernel function obtained in this study gives the same significant covariates for all districts. So, there only one group formed and contained of all districts. It was caused by proportion of zero-inflation which is dominates the dependent variable. Other research on infant mortality data by [9] showed GWPR model by using Gaussian kernel function exactly form three groups of districts, while Bi-square kernel function can form seven groups.

5. Conclusion
This study proposed a part development of statistical spatial point modeling that can handle the over dispersion caused by a great proportion of zero-inflation and the spatial heterogeneity effect. The Filariasis data shows 60% proportion of zero values and significant spatial heterogeneity effect. The two GWZIP regression model using two kernel functions indicate all the covariates have a significant influence to the number of Filariasis case. Unfortunately, there is no group of districts formed according to their significant covariates. So, all the covariates in this study are global parameter for GWZIP model in every district. While, spatial model commonly forms some groups of districts, have global and local parameters. Comparing two GWZIP model based on the value of deviance presents the model using Gaussian kernel is better than the model using Bi-square kernel.

6. Acknowledge
The authors thank to: (1) the Faculty of Mathematics and Natural Science, Brawijaya University; and (2) Central Java Province Health Department for providing Filariasis data.

References
[1] Lambert, D 1992 Zero-Inflated Poisson Regression, With an Application to Defects in Manufacturing Technometrics 34 1
[2] Bohning, D, Dietz, E, Schlattmann, P, Mendonca, L, and Kirchner, U 1999 The Zero-Inflated Poisson Model and The Decayed, Missing, and Filled Teeth Index in Dental Epidemiology Journal of Royal Statistical Society A 162 195
[3] Lestari, A, Purhadi and Ratna, M 2009 Pemodelan Regresi Zero Inflated Poisson (Aplikasi pada Data Pekerja Seks Komersial di Klinik Reproduksi Putat Jaya Surabaya) Phytagoras: Jurnal Pendidikan Matematika 5 57
[4] Mouatassim, Y and Ezzahid, E H 2012 Poisson Regression and Zero-Inflated Poisson Regression : Application to Private Health Insurance Data European Actuary Journal 2 187
[5] Purhadi, Dewi, Y S and Amaliana, L 2015 Zero Inflated Poisson and Geographically Weighted Zero-Inflated Poisson Regression Model: Application to Elephantiasis (Filariasis) Counts Data Journal of Mathematics and Statistics 11 52
[6] Kalogirou, S 2016 Destination Choice of Athenians: An Application of Geographically Weighted Versions of Standard and Zero Inflated Poisson Spatial Interaction Models Geographical Analysis 48 191
[7] Anselin, L 1988 Spatial Econometrics: Methods and Models (Dordrecht: Kluwer Academic Publishers)
[8] Nakaya, T, Fotheringham, A S, Brunsdon, C and Charlton, M 2005 Geographically Weighted Poisson Regression for Disease Association Mapping Statistics in Medicine 24 2695
[9] Aulele, S N 2018 Model Geographically Weighted Poisson Regression dengan Pembobot Fungsi Kernel Gauss Barekeng: Jurnal Ilmu Matematika dan Terapan 5 25
[10] Gujarati, D N 2004 Basic Econometrics, 4th Edition (New York: Mc.Graw-Hill Companies)
[11] Central Java Province Health Department 2017 Profil Kesehatan Provinsi Jawa Tengah (Semarang: (Central Java Province Health Department)