Letters in Spatial and Resource Sciences (2021) 14:79–99
https://doi.org/10.1007/s12076-021-00266-5

ORIGINAL PAPER

Geographically weighted bivariate generalized Poisson regression: application to infant and maternal mortality data

Purhadi[1] · Sutikno[1] · Sarni Maniar Berliana[1,2] · Dewi Indra Setiawan[3]

Received: 7 April 2020 / Accepted: 23 January 2021 / Published online: 3 March 2021

Abstract
Bivariate generalized Poisson regression (BGPR) is an extension of bivariate Poisson regression which deals overdipersion or underdispersion problem. This model gives global regression coefficients for all observations (locations) in the analysis. The BGPR model is then extended to take into account spatial heterogeneity, called geographically weighted bivariate generalized Poisson regression model, that yields varying regression coefficients locally. The regression model is applied to analyse factors affecting number of infant and maternal mortality in East Java, Indonesia.

Keywords Count data · Spatial analysis · Spatial heterogeneity

JEL Classification C10 · C18 · C31 · I10

✉ Sarni Maniar Berliana
sarni@stis.ac.id

Purhadi
purhadi@statistika.its.ac.id

Sutikno
sutikno@statistika.its.ac.id

Dewi Indra Setiawan
dewiindrasetiawan@gmail.com

1 Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

2 Department of Statistics, Politeknik Statistika STIS, Jakarta, Indonesia

3 Research Center for Regional Development and Community Empowerment, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

1 Introduction

Bivariate Poisson regression is a method for modeling a pair correlated count data with several predictor variables that affect both response variables (Karlis and Ntzoufras 2005). This model assumes equidispersion that mean and variance of the response variable are the same. Violation of this assumption occurs when variance is greater than mean (overdispersion) or variance is less than mean (underdispersion). Count data that exhibits underdispersion or overdispersion can be fitted well by applying generalized Poisson regression (Consul and Famoye 1992; Famoye et al. 2004; Wang and Famoye 1997). Bivariate generalized Poisson regression (BGPR) as an extension of bivariate Poisson regression model accommodates overdispersion or underdispersion cases. Zamani et al. (2016) have illustrated this model on health care data.

Furthermore, Poisson regression model has been extended to take into account spatial variation in measuring the relationship between the response variable and explanatory variables, called geographically Poisson regression model. This means that the effect of explanatory variables to the response variable vary spatially. The development of this model followed the work of Brunsdon et al. (1996) in developing statistical technique to analyse data with spatial heterogeneity, called geographically weighted regression (GWR) model.

The GWR model captures spatial heterogeneity as a result of non-stationary process that an attribute show dependency on the location where the attribute was observed. For example, a relatively more developed area would have a relatively lower infant mortality rate than a relatively less developed area. In the case of spatial process, Fotheringham et al. (2002) refer this as spatial non-stationarity. When discussing spatial analysis, another issue that arises is spatial dependency or spatial autocorrelation which describes the similarity of neighbouring observations. Almost all data spatial experience some form of spatial autocorrelation that leads to misleading inference when traditional global model is applied. The usual solution to solve this problem is by applying spatial regression models. However, Fotheringham et al. (2002) has shown that GWR as an alternative model is able to solve the problem of spatially autocorrelated error terms.

Recently, Triyanto et al. (2015) introduced geographically weighted multivariate Poisson regression (GWMPR) model as an extension of geographically weighted Poisson regression (GWPR) model by Nakaya et al. (2005). In their paper, they used maximum likelihood estimation (MLE) method solved by Newton–Raphson algorithm in obtaining the parameter estimates of GWMPR. The GWMPR model was then applied on children health data in 35 locations in Central Java, Indonesia using three response variables concerning underweight, malnutrition, and child mortality. This model takes into account spatial heterogeneity but it does not accommodate underdispersion and overdispersion cases.

In this article, we propose a new count data regression model that takes into account spatial heterogeneity and handles any type of dispersion for correlated response variables, called geographically weighted bivariate generalized Poisson regression (GWBGPR) based on bivariate generalized Poisson distribution by Famoye and Consul (1995). Most recently, Berliana et al. (2020) defined geographically weighted multivariate generalized Poisson regression (GWMGPR) based on multivariate generalized Poisson regression developed by Famoye (2015). The paper presented step-by-step procedure of parameter estimation of the GWMGPR using MLE method and distribution of critical region using likelihood ratio test. Following their steps, we develop GWBGPR model based on different probability mass function than theirs and apply the GWBGPR model on mortality data.

In the next section, the development of GWBGPR model and its parameter estimation and hypothesis testing are presented subsequently. Then, the GWBGPR model is applied on assessing factors affecting infant and maternal mortality in East Java, Indonesia. Conclusions are given in the last section.

2 The GWBGPR model

Following Consul and Shoukri (1985), let a response variable Y, be a generalized Poisson random variable (GPD). Then, the probability function of Y is given by

$$P(y, \lambda, \alpha) = \frac{\lambda\ (\lambda + \alpha y)^{y-1}\ e^{-\lambda-\alpha y}}{y!},\ \ y = 0, 1, 2, \ldots \tag{1}$$

where $\lambda > 0$, $\max(-1, -\lambda/m\ ) \le \alpha < 1$ and $m \ge 4$ is the largest positive integer for which $\lambda + \alpha m > 0$ when $\alpha < 0$. In equation (1), $P(Y = y) = 0$ for $y > m$ when $\alpha < 0$. The model in (1) reduces to the Poisson probability function when $\alpha = 0$ (equidispersion). The GPD is underdispersed when $\alpha < 0$ and overdispersed when $\alpha > 0$.

Famoye and Consul (1995) and Vernic (1997) used trivariate reduction method to construct bivariate generalized Poisson distribution (BGPD). Let Z_0, Z_1, Z_2 be independent generalized Poisson random variables (GPD), $Z_j \sim GPD\left(\lambda_j, \theta_j\right)$, $j = 0, 1, 2$. Let $Y_1 = Z_1 + Z_0$ and $Y_2 = Z_2 + Z_0$. The probability density function of BGPD is given by

$$\begin{aligned} P(y_1, y_2) = {} & \lambda_0\lambda_1\lambda_2 \mathrm{e}^{-(\lambda_0+\lambda_1+\lambda_2)-y_1\alpha_1-y_2\alpha_2} \\ & \sum_{k=0}^{\min(y_1,y_2)} \left[\frac{\left(\lambda_0 + k\alpha_0\right)^{k-1}}{k!} \frac{\left(\lambda_1 + \left(y_1 - k\right)\alpha_1\right)^{y_1-k-1}}{\left(y_1 - k\right)!} \right. \\ & \left. \frac{\left(\lambda_2 + \left(y_2 - k\right)\alpha_2\right)^{y_2-k-1}}{\left(y_2 - k\right)!} e^{k(\alpha_1+\alpha_2-\alpha_0)} \right]. \end{aligned} \tag{2}$$

The probability distribution in (2) reduces to bivariate Poisson distribution when $\alpha_0 = \alpha_1 = \alpha_2 = 0$. The corresponding regression model of (2) is

$$
\begin{aligned}
P(y_1, y_2) = {} & \lambda_0\left(\mu_1 - \lambda_0\right)\left(\mu_2 - \lambda_0\right)\mathrm{e}^{-\left(\mu_1+\mu_2-\lambda_0\right)-y_1\alpha_1-y_2\alpha_2} \\
& \sum_{k=0}^{\min\left(y_1,y_2\right)}\left[\frac{\left(\lambda_0 + k\alpha_0\right)^{k-1}}{k!}\frac{\left(\left(\mu_1 - \lambda_0\right) + \left(y_1 - k\right)\alpha_1\right)^{y_1-k-1}}{\left(y_1 - k\right)!}\right. \\
& \left.\frac{\left(\left(\mu_2 - \lambda_0\right) + \left(y_2 - k\right)\alpha_2\right)^{y_2-k-1}}{\left(y_2 - k\right)!}e^{k(\alpha_1+\alpha_2-\alpha_0)}\right].
\end{aligned} \tag{3}
$$

The relationship between the response variable Y_{ij} and the covariates $\mathbf{x}_i$ are defined as

$$
\mu_{ij} = \exp\left(\mathbf{x}_i^T\boldsymbol{\beta}_j\right), \tag{4}
$$

where $\mathbf{x}_i$ $(i = 1, 2, \dots, n)$ is a vector of covariates and $\boldsymbol{f\beta}_j$ $(j = 1, 2)$ is a vector of regression parameters of Y_j. The model in (4) is a global model that the parameters $\beta_j{}'$s are constant over space.

Now, the BGPR model in (4) is extended to yield GWBGPR model such that the parameters vary across locations. Suppose that Y_{1i} and Y_{2i} follow BGPD with parameters $\lambda_0,\ \lambda_1,\ \lambda_2,\ \alpha_0,\ \alpha_1,\ \alpha_2$ and $\mathbf{u}_i = \left(\mathbf{u}_{1i}, \mathbf{u}_{2i}\right)$ is a point coordinate in two-dimensional geographic space that represents longitude $\left(\mathbf{u}_{1i}\right)$ and latitude $\left(\mathbf{u}_{2i}\right)$ at location i $(i = 1, 2, \dots, n)$. Thus, the GWBGPR model for location i based on (4) is written as follows

$$
\mu_{ij}\left(\mathbf{u}_i\right) = \exp\left(\mathbf{x}_i^T\boldsymbol{\beta}_j\left(\mathbf{u}_i\right)\right), \tag{5}
$$

where $\boldsymbol{\beta}_j\left(\mathbf{u}_i\right)$ is a vector of parameters of Y_j for location i. To estimate the parameters at locaion i, we approximate (5) by (4) and perform a regression using a subset of data points that are close to i. Thus, the parameter estimates for location i is calculated in the usual way and for the next location i, a new subset of close points is used, and so on.

3 Parameter estimation

Following the work of Berliana et al. (2020), we apply the MLE method in estimating the parameters of the GWBGPR model. The likelihood function for location i corresponds to (3) is written as

$$
L(\Omega) = \prod_{i=1}^{n}\left[\lambda_0\left(\mathbf{u}_i\right)\left(\mu_{i1}\left(\mathbf{u}_i\right) - \lambda_0\left(\mathbf{u}_i\right)\right)\left(\mu_{i2}\left(\mathbf{u}_i\right) - \lambda_0\left(\mathbf{u}_i\right)\right)a\ \sum_{k=0}^{s} bcfg\right] \tag{6}
$$

where

$$
\begin{aligned}
a &= \exp -\left(\mu_{i1}(\mathbf{u}_i) + \mu_{i2}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i)\right) - y_{i1}\alpha_1(\mathbf{u}_i) - y_{i2}\alpha_2(\mathbf{u}_i), \\
b &= \frac{\left(\lambda_0(\mathbf{u}_i) + k\alpha_0(\mathbf{u}_i)\right)^{k-1}}{k!}, \\
c &= \frac{\left(\mu_{i1}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{i1} - k)\alpha_1(\mathbf{u}_i)\right)^{y_{i1}-k-1}}{(y_{i1} - k)!}, \\
f &= \frac{\left(\mu_{i2}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{i2} - k)\alpha_2(\mathbf{u}_i)\right)^{y_{i2}-k-1}}{(y_{i2} - k)!},
\end{aligned}
\tag{7}
$$

$$
g = \exp\left(k\left(\alpha_1(\mathbf{u}_i) + \alpha_2(\mathbf{u}_i) - \alpha_0(\mathbf{u}_i)\right)\right), \tag{8}
$$

and $s = \min(y_{i1}, y_{i2})$. Let Q be the log-likelihood function of equation (6) for location i

$$
\begin{aligned}
Q = \log(L(\Omega)) &= \log\left(L\left(\lambda_0(\mathbf{u}_i), \boldsymbol{\beta}_1(\mathbf{u}_i), \boldsymbol{\beta}_2(\mathbf{u}_i), \alpha_1(\mathbf{u}_i), \alpha_2(\mathbf{u}_i), \alpha_0(\mathbf{u}_i)\right)\right) \\
&= \sum_{i=1}^{n} \log \lambda_0(\mathbf{u}_i) + \sum_{i=1}^{n} \left[\log\left(\mu_{i1}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i)\right)\right] + \sum_{i=1}^{n} \left[\log\left(\mu_{i2}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i)\right)\right] \\
&\quad - \sum_{i=1}^{n} \left[\mu_{1i}(\mathbf{u}_i)\right] - \sum_{i=1}^{n} \left[\mu_{i2}(\mathbf{u}_i)\right] + n\lambda_0(\mathbf{u}_i) - \sum_{i=1}^{n} \left[y_{i1}\alpha_1(\mathbf{u}_i)\right] \\
&\quad - \sum_{i=1}^{n} \left[y_{i2}\alpha_2(\mathbf{u}_i)\right] + \sum_{i=1}^{n} \log B_i
\end{aligned}
\tag{9}
$$

where μ_{ij} is stated in (5) and

$$
\begin{aligned}
B_i &= \sum_{k=0}^{\min(y_{i1}, y_{i2})} B_{i1} B_{i2}, \\
B_{i1} &= \frac{\left(\left(\mu_{i1}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i)\right) + (y_{i1} - k)\alpha_1(\mathbf{u}_i)\right)^{y_{i1}-k-1}}{(y_{i1} - k)!} g,
\end{aligned}
\tag{10}
$$

$$
B_{i2} = \frac{\left(\left(\mu_{i2}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i)\right) + (y_{i2} - k)\alpha_2(\mathbf{u}_i)\right)^{y_{i2}-k-1}}{(y_{i2} - k)!} b, \tag{11}
$$

with b and g are defined in (7) and (8), respectively.

In ordinary regression model, all observations have similar weight in estimating the parameter. In GWBGPR model, we give weight for data contribution in estimating the parameters based on the proximity of the data points to the regression point i. The closer the data points to the regression point, the bigger the contribution of those points in parameter estimation for location i. Several spatial weighting functions are presented in Fotheringham et al. (2002). Here, we show adaptive bisquare weighting function in accordance with what we use in the application later,

$$w_{ii*} = \begin{cases} \left(1 - \left(\frac{d_{ii*}}{h_i}\right)^2\right)^2; & \mathrm{d}_{ii*} \le h_i \\ 0; & \mathrm{d}_{ii*} > h_i \end{cases}$$

where $d_{ii*} = \left(\left(u_{1i} - u_{1i*}\right)^2 + \left(u_{2i} - u_{2i*}\right)^2\right)^{1/2}$, d_{ii*} is Euclidian distance between location i and location $i*$ while h_i is a smoothing parameter or referred to the bandwidth of location i. The selection of optimum bandwidth can be achieved using Generalized Cross Validation (GCV) method (see Fotheringham et al. (2002) for details).

We apply the above weighting function to data points $i*$ $(i* = 1, 2, \dots, n)$ in estimating the parameters for location i $(i = 1, 2, \dots, n)$. Thus, there will be n regression models, each with its own parameters. Then, the weighted log-likelihood function for location i based on (9) is written as

$$\begin{aligned} Q* &= \log\left(L\left(\lambda_0(\mathbf{u}_i), \boldsymbol{\beta}_1(\mathbf{u}_i), \boldsymbol{\beta}_2(\mathbf{u}_i), \alpha_1(\mathbf{u}_i), \alpha_2(\mathbf{u}_i), \alpha_0(\mathbf{u}_i)\right)\right) w_{ii*} \\ &= \sum_{i*=1}^{n}\left[\log \lambda_0(\mathbf{u}_i) w_{ii*}\right] + \sum_{i*=1}^{n}\left[\log\left(\mu_{i*1}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i)\right) w_{ii*}\right] \\ &\quad + \sum_{i*=1}^{n}\left[\log\left(\mu_{i*2}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i)\right) w_{ii*}\right] - \sum_{i*=1}^{n}\left[\mu_{i*1}(\mathbf{u}_i) w_{ii*}\right] \\ &\quad - \sum_{i*=1}^{n}\left[\mu_{i*2}(\mathbf{u}_i) w_{ii*}\right] + n\lambda_0(\mathbf{u}_i) w_{ii*} - \sum_{i*=1}^{n}\left[y_{i*1}\alpha_1(\mathbf{u}_i) w_{ii*}\right] \\ &\quad - \sum_{i*=1}^{n}\left[y_{i*2}\alpha_2(u_i, v_i) w_{ii*}\right] + \sum_{i*=1}^{n}\left[\log B_{i*} w_{ii*}\right]. \end{aligned} \tag{12}$$

The log-likelihood function in (12) is maximized over the parameters $\lambda_0(\mathbf{u}_i)$, $\boldsymbol{\beta}_1(\mathbf{u}_i)$, $\boldsymbol{\beta}_2(\mathbf{u}_i)$, $\alpha_1(\mathbf{u}_i)$, $\alpha_2(\mathbf{u}_i)$, and $\alpha_0(\mathbf{u}_i)$.

In order to obtain the estimators of GWBGPR model using the MLE method, we need to calculate the derivatives of Q^* in (12) with respect to each parameter in the model. One can use any optimization technique to obtain the maximum likelihood estimators of the GWBGPR model, such as Newton–Raphson, BHHH, BFGS, etc. Here, we show the derivatives of Q^* until the second order. The first and second order derivatives of $Q^*$with respect to each parameter are presented in the "Appendix".

4 Simultaneous hypothesis testing of GWBGPR

The simultaneous testing on parameters of GWBGPR model was carried out to examine the significance of parameter $\boldsymbol{\beta}_{jl}(\mathbf{u}_i)$ simultaneously using the following hypothesis

$$\begin{aligned} H_0 &: \beta_{j1}(\mathbf{u}_i) = \beta_{j2}(\mathbf{u}_i) = \cdots = \beta_{jm}(\mathbf{u}_i) = 0\,;\ i = 1,2,\ldots,n;\ j = 1,2 \\ H_1 &: \ \text{at least one } \beta_{jl}(\mathbf{u}_i) \neq 0\,;\ l = 1,\ 2,\ \ldots,\ m. \end{aligned} \tag{13}$$

We use maximum likelihood ratio test (MLRT) to obtain critical region of the hypothesis testing for the hypothesis in (13) (see Berliana et al. (2019) for brief procedures of MLRT). Using this method, we need to calculate the maximum likelihood function under the null hypothesis ($L(\omega)$) and the population ($L(\Omega)$). The likelihood function under population have been defined in (6) and its maximum will be achieved when all the estimators are substituted in the equation. The calculation of the log-likelihood function under the null hypothesis $L(\omega)$ is as follows

$$\begin{aligned} L(\omega) = \prod_{i=1}^{n} &\Big[\lambda_{00}(\mathbf{u}_i)\big(\mu_{0i1}(\mathbf{u}_i) - \lambda_{00}(\mathbf{u}_i)\big)\big(\mu_{0i2}(\mathbf{u}_i) - \lambda_{00}(\mathbf{u}_i)\big) \\ &\times a_0 \sum_{k=0}^{s} b_0 c_0 f_0 g_0\Big], \end{aligned} \tag{14}$$

where

$$\begin{aligned} a_0 &= \exp -\big(\mu_{0i1}(\mathbf{u}_i) + \mu_{0i2}(\mathbf{u}_i) - \lambda_{00}(\mathbf{u}_i)\big) - y_{i1}\alpha_{01}(\mathbf{u}_i) - y_{i2}\alpha_{02}(\mathbf{u}_i), \\ b_0 &= \frac{\big(\lambda_{00}(\mathbf{u}_i) + k\alpha_{00}(\mathbf{u}_i)\big)^{k-1}}{k!}, \\ c_0 &= \frac{\big(\mu_{0i1}(\mathbf{u}_i) - \lambda_{00}(\mathbf{u}_i) + (y_{i1} - k)\alpha_{01}(\mathbf{u}_i)\big)^{y_{i1}-k-1}}{(y_{i1} - k)!}, \\ f_0 &= \frac{\big(\mu_{0i2}(\mathbf{u}_i) - \lambda_{00}(\mathbf{u}_i) + (y_{i2} - k)\alpha_{02}(\mathbf{u}_i)\big)^{y_{i2}-k-1}}{(y_{i2} - k)!}, \end{aligned} \tag{15}$$

$$g_0 = \exp\big(k\big(\alpha_{01}(\mathbf{u}_i) + \alpha_{02}(\mathbf{u}_i) - \alpha_{00}(\mathbf{u}_i)\big)\big), \tag{16}$$

with $\mu_{0i1}(\mathbf{u}_i) = \exp\big(\beta_{00}(\mathbf{u}_i)\big)$, $\mu_{0i2}(\mathbf{u}_i) = \exp\big(\beta_{020}(\mathbf{u}_i)\big)$, and $s = \min(y_{i1}, y_{i2})$.

The log-likelihood function corresponds to (14) is calculated as

$$\begin{aligned} Q_0 &= \log(L(\omega)) \\ &= \sum_{i=1}^{n}\big[\log\lambda_{00}(\mathbf{u}_i)\big] + \sum_{i=1}^{n}\big[\log\big(\mu_{0i1}(\mathbf{u}_i) - \lambda_{00}(\mathbf{u}_i)\big)\big] \\ &\quad + \sum_{i=1}^{n}\big[\log\big(\mu_{0i2}(\mathbf{u}_i) - \lambda_{00}(\mathbf{u}_i)\big)\big] - \sum_{i=1}^{n}\big[\mu_{0i1}(\mathbf{u}_i)\big] - \sum_{i=1}^{n}\big[\mu_{0i2}(\mathbf{u}_i)\big] \\ &\quad + \sum_{i=1}^{n}\big[\lambda_{00}(\mathbf{u}_i)\big] - \sum_{i=1}^{n}\big[y_{i1}\alpha_{01}(\mathbf{u}_i)\big] - \sum_{i=1}^{n}\big[y_{i2}\alpha_{02}(\mathbf{u}_i)\big] + \sum_{i=1}^{n}\log B_{0i}, \end{aligned} \tag{17}$$

where

$$B_{0i} = \sum_{k=0}^{\min(y_{i1},y_{i2})} B_{0i1}B_{0i2},$$
$$B_{0i1} = \frac{\left(\left(\mu_{0i1}(\mathbf{u}_i) - \lambda_{00}(\mathbf{u}_i)\right) + (y_{i1}-k)\alpha_{01}(\mathbf{u}_i)\right)^{y_{i1}-k-1}}{(y_{i1}-k)!} g_0, \tag{18}$$
$$B_{0i2} = \frac{\left(\left(\mu_{0i2}(\mathbf{u}_i) - \lambda_{00}(\mathbf{u}_i)\right) + (y_{i2}-k)\alpha_{02}(\mathbf{u}_i)\right)^{y_{i2}-k-1}}{(y_{i2}-k)!} b_0,$$

with b_0 and g_0 are defined in (15) and (16), respectively.

Applying the same procedure presented above, we get the weighted log-likelihood function of (17) for location i as

$$\begin{aligned} Q_0^* &= (\log(L(\omega)))w_{ii*} \\ &= n\log\lambda_{00}(\mathbf{u}_i)w_{ii*} + \sum_{i*=1}^{n}\left[\log\left(\mu_{0i1}(\mathbf{u}_i) - \lambda_{00}(\mathbf{u}_i)\right)w_{ii*}\right] \\ &\quad + \sum_{i*=1}^{n}\left[\log\left(\mu_{0i2}(\mathbf{u}_i) - \lambda_{00}(\mathbf{u}_i)\right)w_{ii*}\right] - \sum_{i*=1}^{n}\left[\mu_{0i1}(\mathbf{u}_i)w_{ii*}\right] \\ &\quad - \sum_{i*=1}^{n}\left[\mu_{0i2}(\mathbf{u}_i)w_{ii*}\right] + n\lambda_{00}(\mathbf{u}_i)w_{ii*} - \sum_{i*=1}^{n}\left[y_{i1}\alpha_{01}(\mathbf{u}_i)w_{ii*}\right] \\ &\quad - \sum_{i*=1}^{n}\left[y_{i2}\alpha_{02}(\mathbf{u}_i)w_{ii*}\right] + \sum_{i*=1}^{n}\left[\log B_{0i}w_{ii*}\right], \end{aligned}$$

where B_{0i} is stated in (18).

Let G^2 be the test statistic for the hypothesis in (13)

$$G^2 = -2\log\left[\frac{L(\hat{\omega})}{L(\hat{\Omega})}\right] = -2\left[\log\left(L(\hat{\omega}) - L(\hat{\Omega})\right)\right), \tag{19}$$

where $L(\hat{\Omega})$ is the maximum value of (6) and $L(\hat{\omega})$ is the maximum value of (14). The exact distribution of the test statistic in (19) is difficult to achieved. Therefore, we utilized central limit theorem to obtain the distribution of the test statistic (Pawitan 2001). Let the maximum likelihood estimators under population be

$$\underset{2n(p+4)\times 1}{\hat{\boldsymbol{\theta}}} = \left[\hat{\boldsymbol{\theta}}_1\ \hat{\boldsymbol{\theta}}_2\right]^T,$$

where

$$\hat{\boldsymbol{\theta}} = \left[\ \hat{\boldsymbol{\beta}}_1^T(\mathbf{u}_i)\ \hat{\boldsymbol{\beta}}_2^T(\mathbf{u}_i)\ \hat{\lambda}_0(\mathbf{u}_i)\ \hat{\lambda}_1(\mathbf{u}_i)\ \hat{\lambda}_2(\mathbf{u}_i)\ \hat{\alpha}_0(\mathbf{u}_i)\ \hat{\alpha}_1(\mathbf{u}_i)\ \hat{\alpha}_2(\mathbf{u}_i)\ \right]^T,$$
$$\hat{\boldsymbol{\theta}}_1 = [\ \hat{\beta}_{11}(\mathbf{u}_i)\ \hat{\beta}_{12}(\mathbf{u}_i)\ \dots\ \hat{\beta}_{1p}(\mathbf{u}_i)\ \hat{\beta}_{21}(\mathbf{u}_i)\ \hat{\beta}_{22}(\mathbf{u}_i)\ \dots\ \hat{\beta}_{2p}(\mathbf{u}_i)\ ]^T,$$
$$\hat{\boldsymbol{\theta}}_2 = [\ \hat{\beta}_{10}^T(\mathbf{u}_i)\ \hat{\beta}_{20}^T(\mathbf{u}_i)\ \hat{\lambda}_0(\mathbf{u}_i)\ \hat{\lambda}_1(\mathbf{u}_i)\ \hat{\lambda}_2(\mathbf{u}_i)\ \hat{\alpha}_0(\mathbf{u}_i)\ \hat{\alpha}_1(\mathbf{u}_i)\ \hat{\alpha}_2(\mathbf{u}_i)\ ]^T,$$

and the true parameters be

$$\boldsymbol{\theta}_\omega = \left[\boldsymbol{\theta}_{\omega 1}\ \boldsymbol{\theta}_2\right]^T,$$

where the known parameters and the maximum likelihood estimators under the null hypothesis are partitioned as

$$\hat{\boldsymbol{\theta}}_\omega = \left[\boldsymbol{\theta}_{\omega 1}\ \hat{\boldsymbol{\theta}}_{\omega 2}\right]^T,$$

with

$$\underset{2pn\times 1}{\boldsymbol{\theta}_{\omega 1}} = [0\ 0\ \dots\ 0]^T,$$
$$\underset{8n\times 1}{\hat{\boldsymbol{\theta}}_{\omega 2}} = [\hat{\beta}_{\omega 10}(\mathbf{u}_i)\ \hat{\beta}_{\omega 20}(\mathbf{u}_i)\ \hat{\lambda}_{\omega 0}(\mathbf{u}_i)\ \hat{\lambda}_{\omega 1}(\mathbf{u}_i)\ \hat{\lambda}_{\omega 2}(\mathbf{u}_i)\ \hat{\alpha}_{\omega 0}(\mathbf{u}_i)\ \hat{\alpha}_{\omega 1}(\mathbf{u}_i)\ \hat{\alpha}_{\omega 2}(\mathbf{u}_i)]^T.$$

Hence, the hypothesis in (13) can be written as

$$H_0\ :\ \boldsymbol{\theta}_1 = \boldsymbol{\theta}_{\omega 1}$$
$$H_1\ :\ \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_{\omega 1}$$

and the test statistic G^2 in (19) becomes

$$G^2 = \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_{\omega 1}\right)^T (\mathcal{I}^{11})^{-1}\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_{\omega 1}\right),$$

where $\mathcal{I}^{11}$ is the information on $\boldsymbol{\theta}_1$ obtained from the inverse of Fisher information matrix

$$\mathcal{I}^{-1}\left(\hat{\boldsymbol{\theta}}\right) = \begin{bmatrix} \mathcal{I}^{11}_{2pn\times 2pn} & \mathcal{I}^{12}_{2pn\times 8n} \\ \mathcal{I}^{21}_{8n\times 2pn} & \mathcal{I}^{22}_{8n\times 8n} \end{bmatrix}.$$

Thus, under the regularity conditions, see Cameron and Trivedi (2013) for a more detailed information, the distribution of the partioned matrix is

$$\begin{bmatrix} \hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1 \\ \hat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2 \end{bmatrix} \overset{a}{\sim} N\left(\mathbf{0}, \mathcal{I}^{-1}(\boldsymbol{\theta}) \equiv \begin{bmatrix} \mathcal{I}^{11} & \mathcal{I}^{12} \\ \mathcal{I}^{21} & \mathcal{I}^{22} \end{bmatrix}\right),$$

such that

$$\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1\right) \overset{a}{\sim} N\left(\mathbf{0}, \mathcal{I}^{11}\right),$$

$$\left(\mathcal{I}^{11}\right)^{1/2}\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1\right) \overset{a}{\sim} N\left(\mathbf{0}, \mathbf{I}_{2pn}\right),$$

$$G^2 = \left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_{\omega 1}\right)^T \left(\mathcal{I}^{11}\right)^{-1}\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_{\omega 1}\right) \overset{a}{\sim} \chi^2_{2pn}.$$

We reject H_0 when G^2 is more than $\chi^2_{\alpha,2pn}$ where α is the significance level.

As previously mentioned, the GWBGPR model produces a set of local parameter estimates, the spatial distribution of which can then be mapped. We divided each local parameter estimate by the estimate of its standard error and compared it to standard normal distribution to determine each parameter significance.

5 Application

This section presents the application of GWBGPR model using real data about maternal and child health. Maternal and child health are two inseparable issues and therefore they should be jointly analyzed. Level of maternal and child health in a region can be measured by maternal and infant mortality that occur within a period, usually within a year. Infant mortality is death occurring shortly after birth until the baby has not been exactly one year old UNICEF (2015). World Health Organization World (2015) defines maternal death as the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management (from direct or indirect obstetric death), but not from accidental or incidental causes.

In this paper, we model two response variables for GWBGPR model which are number of infant mortality $\left(Y_1\right)$ and number of maternal mortality $\left(Y_2\right)$, both in one year period in 38 regencies/municipalities in East Java, Indonesia. The province of East Java is located on the Java island in the most eastern part where the capital city of Indonesia is also located on this island. Among 38 regions in East Java Province, there are 29 regencies and 9 municipalities where the municipality areas tend to be more developed than regency areas. The explanatory variables included in the model are percentage of deliveries assisted by skilled health worker $\left(X_1\right)$, percentage of pregnant mothers who took iron tablets $\left(X_2\right)$, percentage of pregnancy with complications $\left(X_3\right)$, percentage of women who married at 18 years old or younger $\left(X_4\right)$, and percentage of married women with elementary education or lower $\left(X_5\right)$. We used secondary data that has been published by East Java Provincial Health Office and BPS-Statistics of East Java Province in the 2013 East Java Provincial Health Profile and in the National Economic Social Survey (SUSENAS) publication, respectively.

We assume that the data has bivariate generalized Poisson distribution. We also checked for multicollinearity among predictors because multicollinearity in regression causes problems such as unstable parameter estimate, unintuitive signs of regression coefficients, and higher standard error. Multicollinearity is considered when Variance Inflation Factor (VIF) is more than 10 (Gujarati 2004; Yan and Su 2009). We found no evidence of multicollinearity among the predictors used in this study. The average number of infant mortality is 152.4 deaths with variance 9792.9 and the average number of maternal mortality in East Java is 16.89 deaths with variance 126.2. Thus, both of response variables exhibit overdispersion. The two response variables have positive correlation ($r = 0.74$) with p value $= 0.001$ which means there is a significant correlation between infant and maternal mortality in East Java.

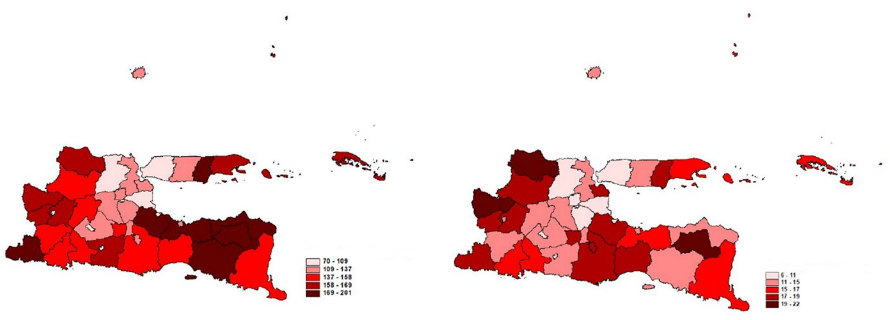
The estimators of the GWBGPR model are obtained using Newton–Raphson iterative algorithm in R software. We chose kernel's bisquare adaptive function to calculate w_{ii*}, the spatial weight of location $i*$ for location i. The simultaneous hypothesis testing on GWBGPR parameters was performed to determine whether one or more parameters have significant effect to the model. We obtained $G^2 = 25,428.05$ and $\chi^2_{0.05;10} = 18.31$. The decision of this test is H_0 is rejected that means at least one parameter have significant effect to infant and maternal mortality.

Figure 1 shows the estimate number of infant and maternal deaths based on the GWBGPR model with five covariates. It appears that infant and maternal mortality tends to be higher in the eastern part of East Java Province. This may be due to a higher difficulties and a more limited access to qualified health facilities, considering that their locations are further away from both the provincial capital and the national capital.

We obtained 38 GWBGPR models for each response variable in accordance with the number of regencies and municipalities in East Java Province. For example purpose, we choose one region to be presented here, which is Pacitan Regency. (The complete 38 models of GWBGPR are available from the authors.) The parameter coefficients and their significances are presented in Table 1. It shows that the variables are overdispersed since α_0, α_1, and α_2 are significantly different from zero. It is also evident that the response variables have positive correlation showed by positive λ_0. The GWBGPR model on number of infant mortality for Pacitan is

$$\hat{\lambda}_1 = e^{5.0543 - 0.0262X_1 + 0.0206X_2 - 0.0011X_3 - 0.0023X_4 + 0.0542X_5}. \tag{20}$$

Iron tablet consumption (X_2) and low education level (X_5) have positive effect on infant mortality, while skilled birth attendant (X_1) and pregnancy complication (X_3) give negative effect. Equation (20) shows that every additional 1% of delivery



(a) Infant mortality **(b)** Maternal mortality

Fig. 1 Estimate number of infant and maternal mortality attended by skilled health worker $\left(X_1\right)$ will decrease the average number of infant mortality $e^{-0.0262} = 0.9741$ times assuming other variables constant. In the same manner, we found that every additional 1% of pregnant women took iron tablets during pregnancy $\left(X_2\right)$ and of married women with low level education will increase the average number of infant mortality assuming other variables constant. Meanwhile, every additional 1% of pregnancy with complications $\left(X_3\right)$ will decrease the average number of infant mortality.

Next, we have the GWBGPR model on number of maternal mortality for Pacitan as follows

$$\hat{\lambda}_2 = \mathrm{e}^{0.1277-0.0141X_1+0.0372X_2+0.0008X_3-0.0023X_4+0.0545X_5}. \tag{21}$$

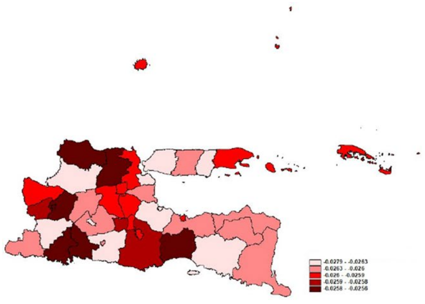
From the model in (21), each additional 1% of delivery attended by skilled health worker $\left(X_1\right)$ will decrease the average number of maternal mortality by $e^{-0.0141} = 0.986$ times assuming other variables constant. Meanwhile, each additional 1% of iron tablets consumption during pregnancy $\left(X_2\right)$ and of married women with low level education $\left(X_5\right)$ will increase the average number of maternal mortality.

We describe the maps of local parameter etimates in pairs for infant mortality and maternal mortality model in Figs. 2, 3, 4, 5 and 6. The distributions of the local parameter estimates for the two models (infant mortality and maternal mortality) show different pattern within the study area. The figures show that the covariates that can explain of high mortality event vary across space. Some covariates may have no power in explaining high mortality in a less developed regions and some may have high power in a more developed regions.

Table 1 Parameter estimation of GWBGPR model for Pacitan

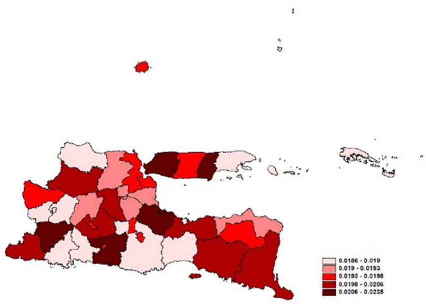
Parameter	Estimation	SE	Z	p value
$\beta_{1.0}$	5.0543	0.0821	61.556	0.000*
$\beta_{1.1}$	– 0.0262	0.0009	– 28.436	0.000*
$\beta_{1.2}$	0.0206	0.0009	23.948	0.000*
$\beta_{1.3}$	– 0.0011	0.0004	– 3.005	0.003*
$\beta_{1.4}$	– 0.0023	0.0013	– 1.820	0.069
$\beta_{1.5}$	0.0542	0.0022	24.957	0.000*
$\beta_{2.0}$	0.1277	0.2383	0.536	0.592
$\beta_{2.1}$	– 0.0141	0.0026	– 5.337	0.000*
$\beta_{2.2}$	0.0372	0.0025	14.712	0.000*
$\beta_{2.3}$	0.0008	0.0011	0.730	0.465
$\beta_{2.4}$	– 0.0023	0.0039	– 0.581	0.561
$\beta_{2.5}$	0.0545	0.0065	8.445	0.000*
λ_0	1.3355	0.0066	202.038	0.000*
α_0	0.0091	0.0003	27.726	0.000*
α_1	72.5624	0.2269	319.841	0.000*
α_2	10.4593	0.1113	94.017	0.000*

*p value < 0.05



(a) Infant mortality model **(b)** Maternal mortality model

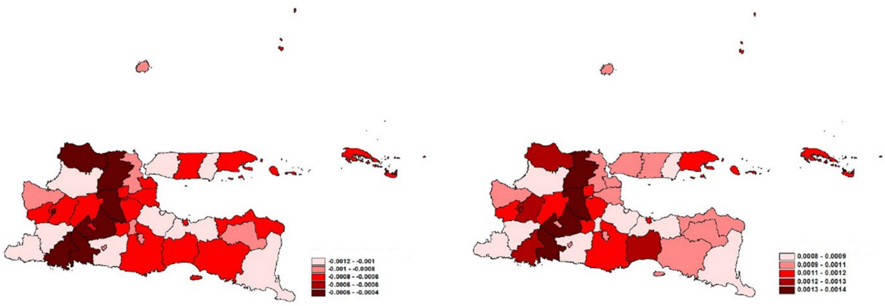
Fig. 2 Local estimates of deliveries by skilled health worker parameter



(a) Infant mortality model **(b)** Maternal mortality model

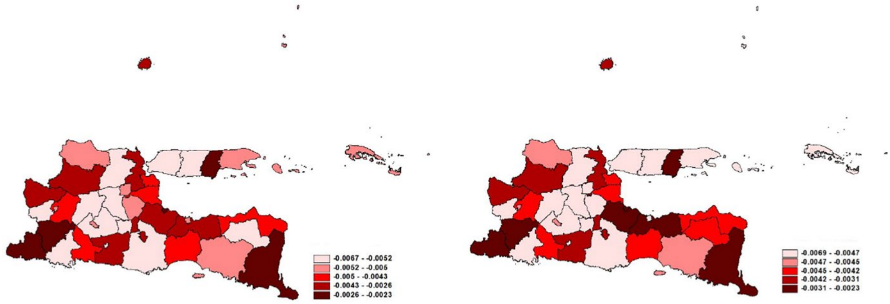
Fig. 3 Local estimates of iron tablets during pregnancy parameter

The coefficient of deliveries assisted by skilled health worker and iron tablet consumption should have negative signs on infant and maternal mortality and the rest of explanatory variables should have positive signs on infant and maternal mortality (McCarthy and Maine 1992; Mosley and Chen 1984; Muldoon et al. 2011). Unfortunately, however, the results show that our models encounter sign reversal in several regression coefficients. As mentioned earlier, reversal signs may happen due to multicollinearity problem. Nevertheless, Fotheringham and Oshan (2016) stated that GWR model is robust to multicollinearity except for extreme cases. Their research showed that under moderate-to-high levels of multicollinearity, the model still produce accurate local parameter estimates. Therefore, the explanation of the results regarding sign reversal in this study may need further research. The problems associated with geographically weighted regression framework, especially in sign reversal has been discussed by Farber and Paez (2007).



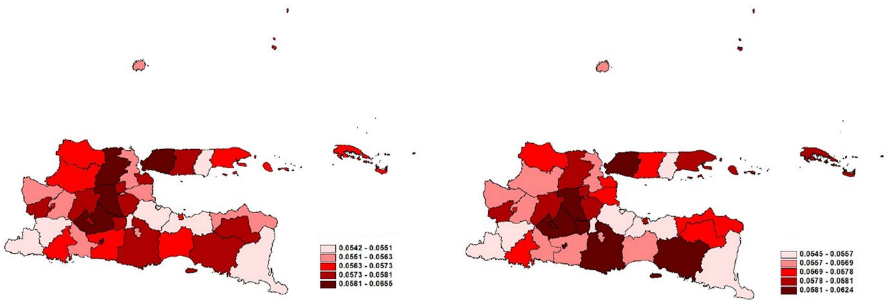
(a) Infant mortality model **(b)** Maternal mortality model

Fig. 4 Local estimates of pregnancy with complications parameter



(a) Infant mortality model **(b)** Maternal mortality model

Fig. 5 Local estimates of early married parameter



(a) Infant mortality model **(b)** Maternal mortality model

Fig. 6 Local estimates of low education parameter

6 Conlusions

This paper proposed GWBGPR, a regression model for correlated count data that takes into account spatial heterogeneity for any type of dispersion. The application of GWBGPR model on mortality data shows spatial variation in distribution of mortality events and in explaining power of each explanatory variable of mortality events across locations in East Java Province. The results also shows reversal signs in regression coefficients that need further study. The basic GWBGPR model can be extended into mixed GWBGPR model in analyzing infant and maternal mortality in which some parameters are global and some parameters are local.

Appendix

First and second partial derivatives of Q^* in (12) with respect to the parameters are shown below.

First partial derivatives

The first partial derivative of $Q^*$with respect to $\lambda_0(\mathbf{u}_i)$ is

$$\frac{\partial Q^*}{\partial \lambda_0(\mathbf{u}_i)} = \sum_{i*=1}^{n}\left[-\frac{w_{ii*}}{\lambda_0(\mathbf{u}_i)} - 3n\,w_{ii*} + \frac{1}{B_{i*}}\frac{\partial B_{i*}}{\partial \lambda_0(\mathbf{u}_i)}w_{ii*}\right],$$

where

$$\frac{\partial B_{i*}}{\partial \lambda_0(\mathbf{u}_i)} = \sum_{k=0}^{\min(y_{i*1},y_{i*2})}\left[\frac{\partial B_{i*1}}{\partial \lambda_0(\mathbf{u}_i)}B_{i*2} + \frac{\partial B_{i*2}}{\partial \lambda_0(\mathbf{u}_i)}B_{i*1}\right].$$

Differentiating B_{i*1} and B_{i*2} in (10) and (11) with respect to $\lambda_0(\mathbf{u}_i)$, we get

$$\frac{\partial B_{i*1}}{\partial \lambda_0(\mathbf{u}_i)} = \left[-(y_{i*1}-k-1)\frac{\mu_{i*1}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{i*1}-k)\alpha_1(\mathbf{u}_i)}{(y_{i*1}-k)!}g\right],$$

$$\frac{\partial B_{i*2}}{\partial \lambda_0(\mathbf{u}_i)} = \frac{\left((\mu_{i*2}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i)) + (y_{i*2}-k)\alpha_2(\mathbf{u}_i)\right)^{y_{i*2}-k-1}}{(y_{i*2}-k)!}\,g\,h,$$

with g is defined in (8) and

$$h = \frac{-(y_{i*2}-k-1)}{\mu_{i*2}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{i*2}-k)\alpha_2(\mathbf{u}_i)} + \frac{k-1}{\lambda_0(\mathbf{u}_i) + k\,\alpha_0(\mathbf{u}_i)}.$$

The first partial derivative of $Q^*$with respect to $\boldsymbol{\beta}_1(\mathbf{u}_i)$ is

$$\frac{\partial Q^*}{\partial \boldsymbol{\beta}_1(\mathbf{u}_i)} = \sum_{i*=1}^{n} \frac{\mu_{i*1}(\mathbf{u}_i)\mathbf{x}_{i*}}{\mu_{i*1}(\mathbf{u}_i)} w_{ii*} + \sum_{i*=1}^{n} \mu_{i*1}(\mathbf{u}_i)\mathbf{x}_{i*} w_{ii*} + \sum_{i*=1}^{n} \frac{1}{B_{i*}} \frac{\partial B_{i*}}{\partial \boldsymbol{\beta}_1(u_i, v_i)} w_{ii*},$$

where

$$\frac{\partial B_{i*}}{\partial \boldsymbol{\beta}_1(\mathbf{u}_i)} = \sum_{k=0}^{\min(y_{i*1}, y_{i*2})} \left[\frac{\partial B_{i*1}}{\partial \boldsymbol{\beta}_1(\mathbf{u}_i)} B_{i*2} + \frac{\partial B_{i*2}}{\partial \boldsymbol{\beta}_1(\mathbf{u}_i)} B_{i*1} \right].$$

Differentiating B_{i*1} in (10) with respect to $\boldsymbol{\beta}_1(\mathbf{u}_i)$, we get

$$\frac{\partial B_{i*1}}{\partial \boldsymbol{\beta}_1(\mathbf{u}_i)} = \frac{\mu_{i*1}(\mathbf{u}_i)(y_{i*1} - k - 1)}{(y_{i*1} - k)!} \times \left(\mu_{i*1}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{i*1} - k)\alpha_1(\mathbf{u}_i)\right)^{y_{i*1}-k-2} g,$$

with g is defined in (8).

The first partial derivative of $Q^*$with respect to $\boldsymbol{\beta}_2(\mathbf{u}_i)$ is

$$\frac{\partial Q^*}{\partial \boldsymbol{\beta}_2(\mathbf{u}_i)} = \sum_{i*=1}^{n} \frac{\mu_{i*2}(\mathbf{u}_i)\mathbf{x}_{i*}}{\mu_{i*2}(\mathbf{u}_i)} w_{ii*} + \sum_{i*=1}^{n} \mu_{i*2}(\mathbf{u}_i)\mathbf{x}_{i*} w_{ii*} + \sum_{i*=1}^{n} \frac{1}{B_{i*}} \frac{\partial B_{i*}}{\partial \boldsymbol{\beta}_2(\mathbf{u}_i)} w_{ii*},$$

where

$$\frac{\partial B_{i*}}{\partial \boldsymbol{\beta}_2(\mathbf{u}_i)} = \sum_{k=0}^{\min(y_{1i*}, y_{2i*})} \left[\frac{\partial B_{1i*}}{\partial \boldsymbol{\beta}_2(\mathbf{u}_i)} B_{2i*} + \frac{\partial B_{2i*}}{\partial \boldsymbol{\beta}_2(\mathbf{u}_i)} B_{1i*} \right].$$

Differentiating B_{2i*} in (11) with respect to $\boldsymbol{\beta}_2(\mathbf{u}_i)$, we have

$$\frac{\partial B_{2i*}}{\partial \boldsymbol{\beta}_2(\mathbf{u}_i)} = \frac{\mu_{2i*}(\mathbf{u}_i)\mathbf{x}_{i*}(y_{2i*} - k - 1)}{(y_{2i*} - k)!} \times \left(\mu_{2i*}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{2i*} - k)\alpha_2(\mathbf{u}_i)\right)^{y_{1i*}-k-2} b,$$

with b is defined in (7).

The first partial derivative of Q^* with respect to $\alpha_1(\mathbf{u}_i)$ is

$$\frac{\partial Q^*}{\partial \alpha_1(\mathbf{u}_i)} = -\sum_{i*=1}^{n} \left[y_{1i*} w_{ii*} + \frac{1}{B_{i*}} \frac{\partial B_{i*}}{\partial \alpha_1(\mathbf{u}_i)} w_{ii*} \right],$$

where

$$\frac{\partial B_{i*}}{\partial \alpha_1(\mathbf{u}_i)} = \sum_{k=0}^{\min(y_{1i*}, y_{2i*})} \left[\frac{\partial B_{1i*}}{\partial \alpha_1(\mathbf{u}_i)} B_{2i*} + \frac{\partial B_{2i*}}{\partial \alpha_1(\mathbf{u}_i)} B_{1i*} \right].$$

Differentiating B_{1i*} in (12) with respect to $\alpha_1(\mathbf{u}_i)$, we get

$$\begin{aligned}\frac{\partial B_{1i*}}{\partial \alpha_1(\mathbf{u}_i)} &= \frac{\left(\mu_{1i*}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{1i*} - k)\alpha_1(\mathbf{u}_i, v_i)\right)^{y_{1i*}-k-1}}{(y_{1i*} - k)!} \\ &\quad \times g\left(\frac{(y_{1i*} - k - 1)(y_{1i*} - k)}{\mu_{1i*}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{1i*} - k)\alpha_1(\mathbf{u}_i, v_i)} + 1 \right).\end{aligned}$$

The first partial derivative of Q^* with respect to $\alpha_2(\mathbf{u}_i)$ is

$$\frac{\partial Q^*}{\partial \alpha_2(\mathbf{u}_i)} = -\sum_{i*=1}^{n} \left[y_{2i*} w_{ii*} + \frac{1}{B_{i*}} \frac{\partial B_{i*}}{\partial \alpha_2(\mathbf{u}_i)} w_{ii*} \right],$$

where

$$\frac{\partial B_{i*}}{\partial \alpha_2(\mathbf{u}_i)} = \sum_{k=0}^{\min(y_{i*1}, y_{i*2})} \left[\frac{\partial B_{i*1}}{\partial \alpha_2(\mathbf{u}_i)} B_{i*2} + \frac{\partial B_{i*2}}{\partial \alpha_2(\mathbf{u}_i)} B_{i*1} \right].$$

Differentiating B_{i*1} and B_{i*2} in (10) and ((11) with respect to $\alpha_2(\mathbf{u}_i)$, we get

$$\begin{aligned}\frac{\partial B_{i*1}}{\partial \alpha_2(\mathbf{u}_i)} &= \frac{\left(\mu_{i*1}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{i*1} - k)\alpha_1(\mathbf{u}_i)\right)^{y_{i*1}-k-1}}{(y_{i*1} - k)!} g, \\ \frac{\partial B_{i*2}}{\partial \alpha_2(\mathbf{u}_i)} &= \frac{(y_{i*2} - k - 1)(y_{i*2} - k)}{(y_{i*2} - k)!} \\ &\quad \times \left(\mu_{i*2}(\mathbf{u}_i) - \lambda_0(u_i, v_i) + (y_{i*2} - k)\alpha_2(\mathbf{u}_i)\right)^{y_{i*1}-k-2} b,\end{aligned}$$

with b and g are defined in (7) and (8), respectively.

The first partial derivative of Q^* with respect to $\alpha_0(\mathbf{u}_i)$ is

$$\frac{\partial Q^*}{\partial \alpha_0(\mathbf{u}_i)} = \sum_{i*=1}^{n} \left[\frac{1}{B_{i*}} \frac{\partial B_{i*}}{\partial \alpha_0(\mathbf{u}_i)} w_{ii*} \right],$$

where

$$\frac{\partial B_{i*}}{\partial \alpha_0(\mathbf{u}_i)} = \sum_{k=0}^{\min(y_{i*1}, y_{i*2})} \left[\frac{\partial B_{i*1}}{\partial \alpha_0(\mathbf{u}_i)} B_{i*2} + \frac{\partial B_{i*2}}{\partial \alpha_0(\mathbf{u}_i)} B_{i*1} \right].$$

Differentiating B_{i*1} and B_{i*2} in (10) and (11) with respect to $\alpha_0(\mathbf{u}_i)$, we have

$$\frac{\partial B_{i*1}}{\partial \alpha_0(\mathbf{u}_i)} = -\frac{\left(\mu_{i*1}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{i*1} - k)\alpha_1(\mathbf{u}_i)\right)^{y_{i*1}-k-1}}{(y_{i*1} - k)!} g,$$

$$\frac{\partial B_{i*2}}{\partial \alpha_2(\mathbf{u}_i)} = \frac{\left(\mu_{i*2}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{i*2} - k)\alpha_2(\mathbf{u}_i)\right)^{y_{i*1}-k-2}}{(y_{i*2} - k)!} b\, k\,(k-1),$$

with b and g are defined in (7) and (8), respectively.

Second partial derivatives

Let

$$f_1 = \frac{y_{i*1} - k - 1}{\left(\mu_{i*1}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{i*1} - k)\alpha_1(\mathbf{u}_i)\right)^2}, \tag{22}$$

$$f_2 = \frac{y_{i*2} - k - 1}{\left(\mu_{i*2}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{i*2} - k)\alpha_2(\mathbf{u}_i)\right)^2}, \tag{23}$$

$$f_3 = \frac{k - 1}{\left(\lambda_0(\mathbf{u}_i) + k\alpha_0(\mathbf{u}_i)\right)^2}, \tag{24}$$

$$f_4 = \frac{y_{i*1} - k - 1}{\mu_{i*1}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{i*1} - k)\alpha_1(\mathbf{u}_i)}, \tag{25}$$

$$f_5 = \frac{y_{i*2} - k - 1}{\mu_{i*2}(\mathbf{u}_i) - \lambda_0(\mathbf{u}_i) + (y_{i*2} - k)\alpha_2(\mathbf{u}_i)}. \tag{26}$$

Then, we have

$$\frac{\partial^2 Q^*}{\partial\left(\lambda_0\left(\mathbf{u}_i\right)\right)^2} = \sum_{i*=1}^{n}\left[\left(\frac{1}{\lambda_0\left(\mathbf{u}_i\right)} + \sum_{k=0}^{\min\left(y_{i*1},y_{i*2}\right)}\left(-f_1 - f_2 + f_3\right)\right)w_{ii*}\right],$$

$$\frac{\partial^2 Q^*}{\partial\lambda_0\left(\mathbf{u}_i\right)\partial\boldsymbol{\beta}_1\left(\mathbf{u}_i\right)} = \sum_{i*=1}^{n}\left[\left(\sum_{k=0}^{\min\left(y_{i*1},y_{i*2}\right)} f_1\,\mu_{i*1}\left(\mathbf{u}_i\right)\mathbf{x}_{i*}\right)w_{ii*}\right],$$

$$\frac{\partial^2 Q^*}{\partial\lambda_0\left(\mathbf{u}_i\right)\partial\boldsymbol{\beta}_2\left(\mathbf{u}_i\right)} = \sum_{i*=1}^{n}\left[\left(\sum_{k=0}^{\min\left(y_{i*1},y_{i*2}\right)} f_2\,\mu_{i*2}\left(\mathbf{u}_i\right)\mathbf{x}_{i*}\right)w_{ii*}\right],$$

$$\frac{\partial^2 Q^*}{\partial\lambda_0\left(\mathbf{u}_i\right)\partial\alpha_1\left(\mathbf{u}_i\right)} = \sum_{i*=1}^{n}\left[\left(\sum_{k=0}^{\min\left(y_{i*1},y_{i*2}\right)} f_1\left(y_{i*1} - k\right)\right)w_{ii*}\right],$$

$$\frac{\partial^2 Q^*}{\partial\lambda_0\left(\mathbf{u}_i\right)\partial\alpha_2\left(\mathbf{u}_i\right)} = \sum_{i*=1}^{n}\left[\left(\sum_{k=0}^{\min\left(y_{i*1},y_{i*2}\right)} f_2\left(y_{i*2} - k\right)\right)w_{ii*}\right],$$

$$\frac{\partial^2 Q^*}{\partial\lambda_0\left(\mathbf{u}_i\right)\partial\alpha_0\left(\mathbf{u}_i\right)} = \sum_{k=0}^{\min\left(y_{i*1},y_{i*2}\right)}\left(-f_3\,k\,w_{ii*}\right),$$

$$\frac{\partial^2 Q^*}{\partial\boldsymbol{\beta}_1\left(\mathbf{u}_i\right)\partial\boldsymbol{\beta}_1^T\left(\mathbf{u}_i\right)} = \sum_{i*=1}^{n}\left[\left(\sum_{k=0}^{\min\left(y_{i*1},y_{i*2}\right)} \mathbf{x}_{i*}\,\mu_{i*1}\,\mathbf{x}_{i*}\left(1 - f_1\,\mu_{i*1}\left(\mathbf{u}_i\right) + f_4\right)\right)w_{ii*}\right],$$

$$\frac{\partial^2 Q^*}{\partial\boldsymbol{\beta}_1\left(\mathbf{u}_i\right)\partial\boldsymbol{\beta}_2^T\left(\mathbf{u}_i\right)} = 0,$$

$$\frac{\partial^2 Q^*}{\partial\boldsymbol{\beta}_1\left(\mathbf{u}_i\right)\partial\alpha_1\left(\mathbf{u}_i\right)} = \sum_{i*=1}^{n}\left[\left(\sum_{k=0}^{\min\left(y_{i*1},y_{i*2}\right)} -f_4\,\mu_{i*1}\left(\mathbf{u}_i\right)\mathbf{x}_{i*}\left(y_{i*1} - k\right)\right)w_{ii*}\right],$$

$$\frac{\partial^2 Q^*}{\partial\boldsymbol{\beta}_1\left(\mathbf{u}_i\right)\partial\alpha_2\left(\mathbf{u}_i\right)} = 0,$$

$$\frac{\partial^2 Q^*}{\partial\boldsymbol{\beta}_1\left(\mathbf{u}_i\right)\partial\alpha_0\left(\mathbf{u}_i\right)} = 0,$$

$$\frac{\partial^2 Q^*}{\partial\boldsymbol{\beta}_2\left(\mathbf{u}_i\right)\partial\boldsymbol{\beta}_2^T\left(\mathbf{u}_i\right)} = \sum_{i*=1}^{n}\left[\left(\sum_{k=0}^{\min\left(y_{i*1},y_{i*2}\right)} \mathbf{x}_{i*}\,\mu_{i*2}\left(\mathbf{u}_i\right)\mathbf{x}_{i*}\left(1 - f_2\,\mu_{i*2}\left(\mathbf{u}_i\right) + f_5\right)\right)w_{ii*}\right],$$

$$\frac{\partial^2 Q^*}{\partial\boldsymbol{\beta}_2\left(\mathbf{u}_i\right)\partial\alpha_1\left(\mathbf{u}_i\right)} = 0,$$

$$\frac{\partial^2 Q^*}{\partial\boldsymbol{\beta}_2\left(\mathbf{u}_i\right)\partial\alpha_2\left(\mathbf{u}_i\right)} = \sum_{i*=1}^{n}\left[\left(\sum_{k=0}^{\min\left(y_{i*1},y_{i*2}\right)} -f_2\left(y_{i*2} - k\right)\mu_{i*2}\left(\mathbf{u}_i\right)\mathbf{x}_{i*}\right)w_{ii*}\right],$$

$$\frac{\partial^2 Q^*}{\partial\boldsymbol{\beta}_2\left(\mathbf{u}_i\right)\partial\alpha_0\left(\mathbf{u}_i\right)} = 0,$$

$$\frac{\partial^2 Q^*}{\partial\left(\alpha_1\left(\mathbf{u}_i\right)\right)^2} = \sum_{i*=1}^{n}\left[\left(\sum_{k=0}^{\min\left(y_{i*1},y_{i*2}\right)} -f_1\left(y_{i*1} - k\right)^2\right)w_{ii*}\right],$$

$$\frac{\partial^2 Q^*}{\partial\alpha_1\left(\mathbf{u}_i\right)\partial\alpha_2\left(\mathbf{u}_i\right)} = 0,$$

$$\frac{\partial^2 Q^*}{\partial\alpha_1\left(\mathbf{u}_i\right)\partial\alpha_0\left(\mathbf{u}_i\right)} = 0,$$

$$\frac{\partial^2 Q^*}{\partial\left(\alpha_2\left(\mathbf{u}_i\right)\right)^2} = \sum_{i*=1}^{n}\left[\left(\sum_{k=0}^{\min\left(y_{i*1},y_{i*2}\right)} -f_2\left(y_{i*2} - k\right)^2\right)w_{ii*}\right],$$

$$\frac{\partial^2 Q^*}{\partial\alpha_2\left(\mathbf{u}_i\right)\partial\alpha_0\left(\mathbf{u}_i\right)} = 0,$$

$$\frac{\partial^2 Q^*}{\partial\left(\alpha_0\left(\mathbf{u}_i\right)\right)^2} = \sum_{i*=1}^{n}\left[\left(\sum_{k=0}^{\min\left(y_{i*1},y_{i*2}\right)} \frac{-k^2(k-1)}{\lambda_0\left(\mathbf{u}_i\right) + k\alpha_0\left(\mathbf{u}_i\right)}\right)w_{ii*}\right],$$

where f_1,f_2,f_3,f_4, andf_5 are defined in (22) until (26), respectively.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by P, S, SMB, and DIS. The first draft of the manuscript was written by SMB and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Berliana, S.M., Purhadi, Sutikno, Rahayu, S.P.: Multivariate generalized Poisson regression model with exposure and correlation as a function of covariates: parameter estimation and hypothesis testing. In: AIP Conference Proceedings, vol. 2192, p. 090001 (2019)

Berliana, S.M., Purhadi, Sutikno, Rahayu, S.P.: Parameter estimation and hypothesis testing of geographically weighted multivariate generalized Poisson regression. Mathematics **8**(1523), 1–14 (2020)

Brunsdon, C., Fotheringham, A.S., Charlton, M.E.: Geographically weighted regression: a method for exploring spatial nonstationarity. Geogr. Anal. **28**(4), 281–298 (1996)

Cameron, A.C., Trivedi, P.K.: Regression Analysis of Count Data, 2nd edn. Cambridge University Press, New York (2013)

Consul, P., Famoye, F.: Generalized Poisson regression model. Commun. Stat. Theory Methods **21**(1), 89–109 (1992)

Consul, P.C., Shoukri, M.M.: The generalized Poisson distribution when the sample mean is larger than the sample variance. Commun. Stat. Simul. Comput. **14**(3), 667–681 (1985)

Famoye, F.: A multivariate generalized Poisson regression model. Commun. Stat. Theory Methods **44**, 497–511 (2015)

Famoye, F., Consul, P.C.: Bivariate generalized Poisson distribution with some applications. Metrika **42**, 127–138 (1995)

Famoye, F., Wulu, J.T., Singh, K.P.: On the generalized Poisson regression model with an application to accident data. J. Data Sci. **2**, 287–295 (2004)

Farber, S., Paez, A.: A systematic investigation of cross-validation in GWR model estimation: empirical analysis and Monte Carlo simulation. J. Geogr. Syst. **9**(4), 371–396 (2007)

Fotheringham, A.S., Oshan, T.M.: Geographically weighted regression and multicollinearity: dispelling the myth. J. Geogr. Syst. **18**, 303–329 (2016)

Fotheringham, A.S., Brundson, C., Charlton, M.: Geographically Weighted Regression: The Analysis of Spatially Varying Relationship. Wiley, Hoboken (2002)

Gujarati, D.N.: Basic Econometrics, 4th edn. McGraw Hill, New York (2004)

Karlis, D., Ntzoufras, I.: Bivariate Poisson and diagonal inflated bivariate Poisson regression models in R. J. Stat. Softw. **14**(10), 1–36 (2005)

McCarthy, J., Maine, D.: A framework for analyzing the determinants of maternal mortality. Stud. Fam. Plan. **23**(1), 23–33 (1992)

Mosley, W.H., Chen, L.C.: An analytical framework for the study of child survival in developing countries. Popul. Dev. Rev. **10**(Supplement: Child Survival: Strategies for Research), 25–45 (1984)

Muldoon, K.A., Galway, L.P., Nakajima, M., Kanters, S., Hogg, R.S., Bendavid, E., Mills, E.J.: Health system determinants of infant, child and maternal mortality: a cross-sectional study of UN member countries. Glob. Health **7**(42), 1–10 (2011)

Nakaya, T., Fotheringham, A.S., Brunsdon, C., Charlton, M.: Geographically weighted Poisson regression for disease association mapping. Stat. Med. **24**, 2695–2717 (2005)

Pawitan, Y.: In All Likelihood. Clarendon Press, Oxford (2001)

Triyanto, Purhadi, Otok, B.W., Purnami, S.W.: Parameter estimation of geographically weighted multivariate Poisson regression. Appl. Math. Sci. **9**(82), 4081–4093 (2015)

UNICEF: Levels and Trends in Child Mortality. United Nations Children's Fund, New York (2015)

Vernic, R.: On the bivariate generalized Poisson distribution. Astin Bull. **27**(01), 23–32 (1997)

Wang, W., Famoye, F.: Modeling household fertility decisions with generalized Poisson regression. J. Popul. Econ. **10**, 273–283 (1997)

World Health Organization: Trends in Maternal Mortality: 1990 to 2015: Estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division. World Health Organization, Geneva (2015)

Yan, X., Su, X.G.: Linear Regression Analysis: Theory and Computing. World Scientific Publishing Co. Pte. Ltd., Singapore (2009)

Zamani, H., Faroughi, P., Ismail, N.: Bivariate generalized Poisson regression model: applications on health care data. Empir. Econ. **51**(4), 1607–1621 (2016)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.