Parameter estimation and hypothesis testing of geographically and temporally weighted bivariate generalized Poisson regression

Setyorini Indah Purwanti1*, Sutikno1 and Purhadi1

1 Department of Statistics, Faculty of Sciences, Institut Teknologi Sepuluh Nopember (ITS). Arief Rahman Hakim Street, Surabaya 60111 Indonesia

Corresponding E-mail: setyoriniindahpurwanti@gmail.com

Abstract. Poisson regression is used to model the data with the response variable in the form of count data. This modeling must meet the equidispersion assumption. That is, the average value is the same as the variance. However, this assumption is often violated. Violation of the equidispersion assumption in Poisson regression modeling will result in invalid conclusions. These violations are an overdispersion and an underdispersion of the response variable. Generalized Poisson Regression (GPR) is an alternative if there is a violation of the equidispersion assumption. If there are two correlated response variables, modeling will use the Bivariate Generalized Poisson Regression (BGPR). However, in the panel data with the observation unit in the form of an area, BGPR is not quite right because there is spatial and temporal heterogeneity in the data. Geographically and Temporally Weighted Bivariate Generalized Poisson Regression (GTWBGPR) is a method for modeling spatial and temporal heterogeneity data. GTWBGPR is a development of GWBGPR. In GTWBGPR, besides accommodating spatial effects, it also accommodates temporal effects. This research will discuss the parameter estimation and test statistics for the GTWBGPR model. Parameter estimation uses Maximum Likelihood Estimation (MLE), but the result is not closed-form, so it is solved by numerical iteration. The numerical iteration used is Newton-Raphson. The test statistic for simultaneous testing uses the Maximum Likelihood Ratio Test (MLRT). With large samples, then this test statistic has a chi-square distribution approximation. So the test statistic for the partial test uses the Z test statistic.

Keywords: GTWBGPR, MLE, MLRT, spatial heterogeneity, temporal heterogeneity

Track Name: Advanced Technology and Renewable Energy

1. Introduction
Poisson regression is used to see the relationship between response variables in count data with predictor variables. In modeling the data using Poisson regression, it must meet the equidispersion assumption. That is, the average value is the same as the variance. The variance value greater than the mean in the Poisson distribution is called overdispersion, while the variance value that is less than the mean is called underdispersion [1]. If this assumption is violated, the standard error will be
underestimated, leading to incorrect conclusions about the significance of the independent variables [2]. The equidispersion condition is often not fulfilled. Generalized Poisson Regression (GPR) is an alternative for modeling data in case of overdispersion or underdispersion. In some studies, GPR is better at modeling data than Poisson regression [1,2,3], negative binomial regression [1,2,3], and Poisson Inverse Gaussian Regression [1]. However, if there are two correlated response variables, use the Bivariate Generalized Poisson Regression (BGPR) [4,5,6].

This regression resulted in invalid estimation results. One alternative to modeling data with spatial heterogeneity is to use Geographically Weighted Regression (GWR). GWR can accommodate spatial variations in data using geographic weighting to obtain different parameter estimates for different region points. Thus, using GWR on spatial data can accommodate Simpson's paradox. Simpson's paradox is the difference in parameter estimates in the data model in the aggregate and modeled data separately [7]. BGPR is unsuitable for spatial data because it assumes the same weights for all regions. Therefore Purhadi et al. [8] developed Geographically Weighted Bivariate Generalized Poisson Regression (GWBGPR).

Panel data is an example of data with a time dimension. Panel data is data obtained from more than one observation. This data's advantages over cross-section and time-series data are more variability, more informative, and less collinearity between variables [9]. GWBGPR is not suitable for analyzing panel data. When using GWBGPR in analyzing panel data, it is assumed that each period's parameters are the same. This assumption is certainly not realistic since data varies between periods. Geographically and Temporally Weighted Regression (GTWR) is an alternative for modeling data containing spatial and temporal heterogeneity [10,11,12]. Therefore, this time we will propose Geographically and Temporally Weighted Bivariate Generalized Poisson Regression (GTWBGPR). GTWBGPR can be used to model panel data with the observation unit in the form of an area, where the response variable is in the form of count data and correlates with each other. In modeling GTWR, it is usually done simultaneously for all periods. In GTWBGPR, parameter estimation and hypothesis testing will be carried out gradually in each period.

Parameters estimators of the regression model are obtained using the parameter estimation method. Several parameter estimation methods that often use in regression models with Generalized Poisson distributed response variables are Maximum Likelihood Estimation (MLE), Method of Moment (MM), and Generalized Method of Moment (GMM). The methods used to estimate parameters for GPR are MLE [1,2,3], MM [1], and GMM [13]. In BGPR, used MM [4] and MLE [6] to estimate parameters. GWBGPR [8] and GWMGPR [14] use MLE to estimate parameters. After getting the parameter estimator, the next step is hypothesis testing. The test statistic for testing hypotheses with a Generalized Poisson distributed response variable is obtained through the Maximum Likelihood Ratio Test (MLRT) method. Therefore, in GTWBGPR, the parameter estimator uses MLE, and hypothesis testing uses MLRT because most of the models with Generalized Poisson distributed response variables use MLE as a parameter estimation method and MLRT to obtain test statistic in hypothesis testing.

2. Materials and Method

2.1 Geographically Weighted Bivariate Generalized Poisson Regression (GWBGPR)

Suppose the response variable \( Y_{i1} \) dan \( Y_{i2} \) distributed by BGP, the probability density function of \( Y_{i1} \) and \( Y_{i2} \) from GWBGPR is as follows [8]:

\[
f(y_{i1}, y_{i2}) = \lambda_0(u_i, v_i)\lambda_1(u_i, v_i)\lambda_2(u_i, v_i)
\]

\[
\exp\left\{-(\lambda_0(u_i, v_i) + \lambda_1(u_i, v_i) + \lambda_2(u_i, v_i)) - y_{i1}\phi_1(u_i, v_i) - y_{i2}\phi_2(u_i, v_i)\right\}
\]
\[
\left(\begin{array}{c}
\sum_{k=0}^{\min(\eta_i, \eta_j)} \left( \hat{\lambda}_i (u, v_i) + k \phi_i (u, v_i) \right)^{k-1} \frac{1}{k!} \left( \hat{\lambda}_i (u, v_i) + (\eta_i - k) \phi_i (u, v_i) \right)^{\eta_i-k-1} \\
\left(\begin{array}{c}
\hat{\lambda}_2 (u, v_i) + (\eta_2 - k) \phi_2 (u, v_i) \right)^{\eta_2-k-1} \frac{1}{(\eta_2-k)!} \left(\exp \left( k \left[ \phi_1 (u, v_i) + \phi_2 (u, v_i) - \phi_0 (u, v_i) \right] \right) \right)
\end{array}\right)
\right)
\]

(1)

where \((u, v_i)\) are the latitude and longitude coordinates of the ith observation, and \(\phi\) is the dispersion parameter with the max value \(-\lambda^{-1}, -m^{-1}\) < \(\phi < \lambda^{-1}\), \(m\) is the largest positive integer on 1 + \(\phi m > 0\) when \(\phi < 0\). If \(\phi = 0\), then there is no dispersion. If \(\phi > 0\), then there is an overdispersion, and if \(\phi < 0\), then there is underdispersion.

The relationship between the response variables \(Y_{i1}\) and \(Y_{i2}\) and predictor variables \(X_{i1}, X_{i2}, \ldots, X_{pi}\) stated as follows:

\[
\log(\hat{\lambda}_i (u, v_i)) = x_i^T \beta_i (u, v_i)
\]

(2)

where:

\[
x_i = [1 \ \ x_{i1} \ \ x_{i2} \ \ \ldots \ \ x_{pi}]^T : i = 1, 2, \ldots, n
\]

\[
\beta_i (u, v_i) = [\beta_{i0} (u, v_i) \ \ \beta_{i1} (u, v_i) \ \ \beta_{i2} (u, v_i) \ \ \ldots \ \ \beta_{ip} (u, v_i)]^T : c = 1, 2, \ \ \beta_i (u, v_i) \ is \ the \ parameter \ vector \ of \ the \ response \ variable \ Y_{ic} \ on \ the \ GWBGR \ model.
\]

2.2 Geographically and Temporally Weighted Bivariate Generalized Poisson Regression (GTWBPR)

The probability density function on equation (1) does not pay attention to temporal heterogeneity. As a result, the resulting parameter estimates are the same for the exact location despite different periods. This result is, of course unrealistic, because there is temporal heterogeneity in the data. Regarding equation (1), then the probability density function of \(Y_{i1}\) and \(Y_{i2}\) on location \((u, v_i)\) and period \(l\), where \(i = 1, 2, \ldots, n; l = 1, 2, \ldots, L\) is:

\[
f \left( y_{i1l}, y_{i2l} \right) = \lambda_{i0} (u, v_i, t_l) \lambda_{i1} (u, v_i, t_l) \lambda_{i2} (u, v_i, t_l)
\]

\[
\exp \left[-\left( \lambda_{i0} (u, v_i, t_l) + \lambda_{i1} (u, v_i, t_l) + \lambda_{i2} (u, v_i, t_l) \right) - y_{i1l} \phi_{i0} (u, v_i, t_l) - y_{i2l} \phi_{i0} (u, v_i, t_l) \right]
\]

\[
\left(\sum_{k=0}^{\min(\eta_{i1l}, \eta_{i2l})} \left( \hat{\lambda}_{i0} (u, v_i, t_l) + k \phi_{i0} (u, v_i, t_l) \right)^{k-1} \frac{1}{k!} \left( \hat{\lambda}_{i1} (u, v_i, t_l) + (\eta_{i1l} - k) \phi_{i1} (u, v_i, t_l) \right)^{\eta_{i1l}-k-1} \right)
\]

\[
\left(\frac{(\lambda_{i2l} (u, v_i, t_l) + (\eta_{i2l} - k) \phi_{i2} (u, v_i, t_l))^{\eta_{i2l}-k-1}}{(\eta_{i2l}-k)!} \right) \exp \left( k \phi_{i0} (u, v_i, t_l) + k \phi_{i1} (u, v_i, t_l) - k \phi_{i2} (u, v_i, t_l) \right) \right)
\]

(3)

To make writing more manageable, then \((u, v_i, t_l)\) is written as a vector \((v_i)\). In modeling the GTWBPR, it is carried out in stages for each \(L\) period where \(L = 1, 2, \ldots, L^*\). By considering equation (2), the model in the \(L\) period is:

\[
\log \left( \hat{\lambda}_l (v_i) \right) = x_{il}^T \beta_{il} (v_i)
\]

(4)

where:

\[
x_{il} = [1 \ \ x_{i1l} \ \ x_{i2l} \ \ \ldots \ \ x_{pll}]^T : i = 1, 2, \ldots, n
\]

\[
\beta_{il} (v_i) = [\beta_{i0l} (v_i) \ \ \beta_{i1l} (v_i) \ \ \beta_{i2l} (v_i) \ \ \ldots \ \ \beta_{ipl} (v_i)]^T : c = 1, 2
\]
2.3 Spatiotemporal Weight
Spatial and temporal effects have been accommodated in the GTWBGPR model by assigning spatiotemporal weight \( w_{it} \) to each observation point (area). As a result, the parameter estimates at each observation point (area) will be different, which shows the local nature of the GTWBGPR model. Weight has an essential role in the spatiotemporal analysis. Spatiotemporal weight describes the relationship between observations, observations close to the \( i \)th regression point have a greater weight than observations that are further away. When forming a weighting matrix, a kernel function is needed. Commonly used kernel functions are as follows:

Fixed Gaussian Kernel Functions

\[
\hat{w}_{it} = \exp\left(-\frac{1}{2} \frac{d_{it}}{q} \right)
\]

Fixed Bisquare Kernel Functions

\[
\hat{w}_{it} = \begin{cases} 
1 - \left( \frac{d_{it}}{q} \right)^2, & \text{if } d_{it} \leq q \\
0, & \text{if } d_{it} > q 
\end{cases}
\]

Adaptive Gaussian Kernel Functions

\[
\hat{w}_{it} = \exp\left(-\frac{1}{2} \frac{d_{it}}{q_d} \right)
\]

Adaptive Bisquare Kernel Functions

\[
\hat{w}_{it} = \begin{cases} 
1 - \left( \frac{d_{it}}{q_d} \right)^2, & \text{if } d_{it} \leq q_d \\
0, & \text{if } d_{it} > q_d 
\end{cases}
\]

\( d_{it} \) is the euclidean distance, that is \( d_{it} = \sqrt{(u_i - u_r)^2 + (v_i - v_r)^2 + \tau_t (t_i - t_r)^2} \) [12], and \( q_d \) is spatiotemporal bandwidth. The selection of the optimum bandwidth affects the accuracy of the regression results. Bandwidth values that are too large will cause significant parameter bias. Bandwidth values that are too small will cause the model to be too coarse (under smoothing). GCV, which is one of the easy bandwidth choices in its application, is as follows [7]:

\[
GCV = \frac{nL \sum_{i=1}^{L} \sum_{r=1}^{R} \left[ y_{it} - \hat{y}_{it} (Q_{it}) \right]^T \left[ y_{it} - \hat{y}_{it} (Q_{it}) \right]}{(nL - v_i)^2}
\]

with \( (nL - v_i) \) is the number of observations minus the number of the parameters.

Calculation of \( \hat{w}_{it} \) at \( L \) period is obtained from:

\[
\hat{w}_{it} = A_i (d_{it}) ; d_{it} = \sqrt{(u_i - u_r)^2 + (v_i - v_r)^2 + \tau_t (t_i - t_r)^2} ; i = 1, 2, \ldots, L
\]

So that for the random samples \( Y_{i1}, Y_{i2}; i = 1, 2, \ldots, n \) using weight \( \hat{w}_{i1} \). Random samples \( Y_{i2}, Y_{i2}; i = 1, 2, \ldots, n \) using weight \( \hat{w}_{i2} \). And so on until the random samples \( Y_{iL}, Y_{iL}; i = 1, 2, \ldots, n \) using weight \( \hat{w}_{iL} \).
3. Results and Discussion

3.1 Parameter Estimation

GTWBGP model parameter estimation using MLE. In estimating the parameters in the GTWBGP, it is carried out gradually in each period by including observations of the previous period. When estimating L period parameter using n number of data in period 1, 2, up to L or require random samples \( Y_{1t}, Y_{2t}, Y_{3t}, \ldots, Y_{Lt} \), where \( L=1,2,\ldots,L^n \). The first step in estimating the L period parameter is to form the likelihood function of the probability density function on equation (3), namely:

\[
L\left( \lambda_{il} (v_i), \beta_{il} (v_i), \beta_{2l} (v_i), \phi_{il} (v_i), \phi_{2l} (v_i), \phi_{2l} (v_i); i=1,2,\ldots,n \right)
= \prod_{i=1}^{L} \prod_{t=1}^{n} \left( \lambda_{il} (v_i) \left[ e^{x^T \beta_{il} (v_i)} - \lambda_{il} (v_i) \right] \right) \left( e^{x^T \beta_{2l} (v_i)} - \lambda_{il} (v_i) \right) \exp \left( -e^{x^T \beta_{il} (v_i)} - e^{x^T \beta_{2l} (v_i)} + \lambda_{il} (v_i) - y_{it} \phi_{2l} (v_i) \right)
- y_{2t} \phi_{2l} (v_i) \left[ \sum_{k=0}^{\min(y_{it},y_{2t})} \left( \frac{\lambda_{il} (v_i) + k \phi_{2l} (v_i)}{k!} \right) \right]^{y_{2t}-k+1} \exp \left[ k \phi_{2l} (v_i) + k \phi_{2l} (v_i) - k \phi_{2l} (v_i) \right] \right]
\]

(7)

where \( x_{ij} = [1, x_{it}, x_{2t}, \ldots, x_{pj}]^T ; i=1,2,\ldots,n; j=1,2,\ldots,L \)

Based on equation (7), the log-likelihood of GTWBGP is obtained as follows:

\[
Q_L = \log L\left( \lambda_{il} (v_i), \beta_{il} (v_i), \beta_{2l} (v_i), \phi_{il} (v_i), \phi_{2l} (v_i), \phi_{2l} (v_i); i=1,2,\ldots,n \right)
\]

(8)

where:

\[
B_{ij} = \sum_{k=0}^{\min(y_{it},y_{2t})} \left[ (k-1) \log \left( \frac{\lambda_{il} (v_i) + k \phi_{2l} (v_i)}{k!} \right) - \log (y_{it} - k - 1) \right]
\]

Then the log-likelihood function is weighted by spatiotemporal weight \( (w_{ij}) \) for each point of observation are:

\[
Q_L = \sum_{l=1}^{L} \sum_{r=1}^{n} w_{ir} \log \lambda_{il} (v_i) + \sum_{l=1}^{L} \sum_{r=1}^{n} w_{ir} \log \left( e^{x^T \beta_{il} (v_i)} - \lambda_{il} (v_i) \right) + \sum_{l=1}^{L} \sum_{r=1}^{n} w_{ir} \log \left( e^{x^T \beta_{2l} (v_i)} - \lambda_{il} (v_i) \right)
- \sum_{l=1}^{L} \sum_{r=1}^{n} w_{ir} e^{x^T \beta_{2l} (v_i)} + \sum_{l=1}^{L} \sum_{r=1}^{n} w_{ir} \lambda_{il} (v_i) - \sum_{l=1}^{L} \sum_{r=1}^{n} w_{ir} y_{ir} \phi_{2l} (v_i)
- \sum_{l=1}^{L} \sum_{r=1}^{n} w_{ir} \phi_{2l} (v_i) + \sum_{l=1}^{L} \sum_{r=1}^{n} w_{ir} B_{ir}
\]

(9)
Furthermore, equation (9) is derived for each parameter. Suppose that:

\[ s_0 = \frac{k-1}{\lambda_{0L}(v_i) + k\bar{\phi}_{0L}(v_i)} \]

\[ s_1 = \left( y_{1ri} - k - 1 \right) \left( e^{v_i\hat{\beta}_{2L}(v_i)} - \lambda_{0L}(v_i) \right) + \left( y_{1ri} - k \right) \bar{\phi}_{1L}(v_i) \]

\[ s_2 = \left( y_{2ri} - k - 1 \right) \left( e^{v_i\hat{\beta}_{2L}(v_i)} - \lambda_{0L}(v_i) \right) + \left( y_{2ri} - k \right) \bar{\phi}_{2L}(v_i) \]

So that the following results are obtained:

\[
\frac{\partial Q^*_L(v_i)}{\partial \lambda_{0L}(v_i)} = \sum_{l=1}^{L} \sum_{l'=1}^{L} W_{l'l'} \hat{m}_{0L}(v_{ii}) - \sum_{l=1}^{L} \sum_{l'=1}^{L} W_{l'l'} \left( e^{v_i\hat{\beta}_{1L}(v_i)} - \lambda_{0L}(v_i) \right) - \sum_{l=1}^{L} \sum_{l'=1}^{L} W_{l'l'} \left( e^{v_i\hat{\beta}_{2L}(v_i)} - \lambda_{0L}(v_i) \right) \\
+ \sum_{l=1}^{L} \sum_{l'=1}^{L} W_{l'l'} + \sum_{l=1}^{L} \sum_{l'=1}^{L} \min\left(y_{1ri},y_{2ri}\right) \sum_{k=0}^{n} \left[ s_0 - s_1 - s_2 \right] 
\]

\[
\frac{\partial Q^*_L(v_i)}{\partial \beta_{VL}(v_i)} = \sum_{l=1}^{L} \sum_{l'=1}^{L} W_{l'l'} e^{v_i\hat{\beta}_{1L}(v_i)} - \sum_{l=1}^{L} \sum_{l'=1}^{L} W_{l'l'} e^{v_i\hat{\beta}_{2L}(v_i)} \hat{X}_{ril} + \sum_{l=1}^{L} \sum_{l'=1}^{L} W_{l'l'} \sum_{k=0}^{n} \left[ s_i e^{v_i\hat{\beta}_{2L}(v_i)} \hat{X}_{ril} \right] 
\]

\[
\frac{\partial Q^*_L(v_i)}{\partial \beta_{VL}(v_i)} = \sum_{l=1}^{L} \sum_{l'=1}^{L} W_{l'l'} - \sum_{l=1}^{L} \sum_{l'=1}^{L} e^{v_i\hat{\beta}_{2L}(v_i)} \hat{X}_{ril} - \sum_{l=1}^{L} \sum_{l'=1}^{L} W_{l'l'} \sum_{k=0}^{n} \left[ s_i e^{v_i\hat{\beta}_{2L}(v_i)} \hat{X}_{ril} \right] 
\]

\[
\frac{\partial Q^*_L(v_i)}{\partial \phi_{0L}(v_i)} = \sum_{l=1}^{L} \sum_{l'=1}^{L} W_{l'l'} \sum_{k=0}^{n} \left[ s_0 \right] - \sum_{l=1}^{L} \sum_{l'=1}^{L} W_{l'l'} \sum_{k=0}^{n} \left[ s_1 \right] - \sum_{l=1}^{L} \sum_{l'=1}^{L} W_{l'l'} \sum_{k=0}^{n} \left[ s_2 \right] 
\]

The results of the first derivative above are not closed-form, so the Newton-Raphson iteration solves it. The Newton-Raphson iteration equation is:

\[
\hat{\theta}_L(v_i)_{(m+1)} = \hat{\theta}_L(v_i)_{(m)} - H^{-1} \left( \hat{\theta}_L(v_i)_{(m)} \right) g \left( \hat{\theta}_L(v_i)_{(m)} \right) 
\]

\[
\hat{\theta}_L(v_i)_{(m)} \text{ is the parameter estimator vector in the } m \text{th iteration, while } \left( \hat{\theta}_L(v_i)_{(m)} \right) \text{ is the gradient vector with the parameters } \hat{\theta}_L(v_i)_{(m)}, \text{ and } H^{-1} \left( \hat{\theta}_L(v_i)_{(m)} \right) \text{ is the Hessian matrix. The iteration above will stop if } \left\| \hat{\theta}_L(v_i)_{(m+1)} - \hat{\theta}_L(v_i)_{(m)} \right\| \leq \varepsilon, \text{ and } \varepsilon > 0 \text{ is very small. So that the parameter estimator vector is obtained, namely:} \\
\hat{\theta}_L(v_i) = \left[ \hat{\lambda}_{0L}(v_i) \hat{\beta}_{VL}(v_i) \hat{\beta}_{VL}(v_i) \hat{\phi}_{0L}(v_i) \hat{\phi}_{1L}(v_i) \hat{\phi}_{2L}(v_i) \right]_i, i=1,2,\ldots,n 
\]
3.2 Hypothesis Testing

After getting the parameter estimator for the GTWBGPR model, the next step is testing the parameters. There is simultaneous and partial parameter testing. Hypothesis testing is carried out in stages in each period \( L \) where \( L = 1, 2, \ldots, L^* \). Simultaneous testing using the MLRT method. Simultaneous testing is carried out to measure the significance of the predictor variable to response variables simultaneously in the GTWBGPR model. Simultaneous hypothesis testing is

\[
H_0: \beta_{t,kl}(v_i) = \beta_{t,2l}(v_i) = \ldots = \beta_{t,ql}(v_i) = 0; c = 1, 2; i = 1, 2, \ldots, n \\
H_1: \text{There is at least one } \beta_{t,jl}(v_i) \neq 0; c = 1, 2; j = 1, 2, \ldots, p
\]

The test statistic in the MLRT method requires the maximum value of the log-likelihood function under \( H_0 \) and under the population. Therefore, the first step is to determine the set of parameters under \( H_0 (\omega_k) \) and under the population.

The set of parameters under \( H_0 (\omega_k) \):

\[
\omega_k = \{ \lambda_{s0L}(v_i), \beta_{s0L}(v_i), \beta_{s20L}(v_i), \phi_{s0L}(v_i), \phi_{s02L}(v_i), \phi_{s20L}(v_i), \phi_{s22L}(v_i); i = 1, 2, \ldots, n \}
\]

The set of parameters under the population:

\[
\Omega_k = \{ \lambda_{qL}(v_i), \beta_{qL}(v_i), \beta_{q2L}(v_i), \phi_{qL}(v_i), \phi_{q2L}(v_i), \phi_{q22L}(v_i); i = 1, 2, \ldots, n \}
\]

The likelihood function under the population \( L(\Omega_k) \) is the same as equation (7), while the likelihood function under \( H_0 (L(\omega_k)) \) is as follows:

\[
L(\omega_k) = L_1(\lambda_{s0L}(v_i), \beta_{s0L}(v_i), \beta_{s20L}(v_i), \phi_{s0L}(v_i), \phi_{s02L}(v_i), \phi_{s20L}(v_i), \phi_{s22L}(v_i); i = 1, 2, \ldots, n)
\]

\[
= \prod_{i=1}^{L} \prod_{j=1}^{n} \left[ e^{\beta_{s0L}(v_i)} - \lambda_{s0L}(v_i) \right] \left[ e^{\beta_{s20L}(v_i)} - \lambda_{s0L}(v_i) \right] \exp\left\{ -e^{\beta_{s0L}(v_i)} - e^{\beta_{s20L}(v_i)} \right\}
\]

\[
+ \lambda_{s0L}(v_i) - y_{id} \phi_{s0L}(v_i) - y_{2id} \phi_{s20L}(v_i) \left[ \sum_{k=0}^{\min(n, y_{id})} \left( \lambda_{s0L}(v_i) + k \phi_{s0L}(v_i) \right)^k \left( \lambda_{s20L}(v_i) + (y_{id} - k) \phi_{s20L}(v_i) \right)^{y_{id} - k} \right]^{y_{2id} - 1} \left( y_{2id} - k ! \right)
\]

\[
(18)
\]

\[
Q_{s0L} = \log L(\lambda_{s0L}(v_i), \beta_{s0L}(v_i), \beta_{s20L}(v_i), \phi_{s0L}(v_i), \phi_{s02L}(v_i), \phi_{s20L}(v_i), \phi_{s22L}(v_i); i = 1, 2, \ldots, n)
\]

\[
= \sum_{i=1}^{L} \sum_{j=1}^{n} \log \lambda_{s0L}(v_i) + \sum_{i=1}^{L} \sum_{j=1}^{n} \log \left( e^{\beta_{s0L}(v_i)} - \lambda_{s0L}(v_i) \right) + \sum_{i=1}^{L} \sum_{j=1}^{n} \log \left( e^{\beta_{s20L}(v_i)} - \lambda_{s0L}(v_i) \right) - \sum_{i=1}^{L} \sum_{j=1}^{n} e^{\beta_{s20L}(v_i)} - \sum_{i=1}^{L} \sum_{j=1}^{n} y_{id} \phi_{s0L}(v_i) - \sum_{i=1}^{L} \sum_{j=1}^{n} y_{2id} \phi_{s20L}(v_i) + \sum_{i=1}^{L} \sum_{j=1}^{n} B_{ij}
\]

(19)
where:

\[
B_{it} = \sum_{k=0}^{\min(y_{it}, y_{2it})} \left[ (k-1) \log \left( \lambda_{o0L}(v_i) + k \varphi_{o0L}(v_i) \right) - \log k + (y_{it} - k - 1) \right] \\
\log \left( e^{\beta_{o0L}(v_i) - \lambda_{o0L}(v_i)} + (y_{it} - k) \varphi_{o0L}(v_i) \right) - \log (y_{it} - k + 1) + (y_{2it} - k - 1) \\
\log \left( e^{\beta_{o2L}(v_i) - \lambda_{o0L}(v_i)} + (y_{2it} - k) \varphi_{o2L}(v_i) \right) - \log (y_{2it} - k + 1) + k \varphi_{o0L}(v_i) + k \varphi_{o2L}(v_i) \\
- k \varphi_{o0L}(v_i) 
\]

Next is to determine \( Q_{oL}^* \), namely, the log-likelihood under \( H_0 \) in equation (19) multiplied by the spatiotemporal weight \( w_{ir} \) to obtain the parameters estimator under \( H_0 \) at Lth period and ith location.

\[
Q_{oL}^* = \sum_{i=1}^{n} \sum_{r=1}^{L} w_{ir} \log \lambda_{o0L}(v_i) + \sum_{i=1}^{n} \sum_{r=1}^{L} w_{ir} \log \left( e^{\beta_{o1L}(v_i) - \lambda_{o0L}(v_i)} \right) + \sum_{i=1}^{n} \sum_{r=1}^{L} w_{ir} \log \left( e^{\beta_{o2L}(v_i) - \lambda_{o0L}(v_i)} \right) \\
- \sum_{i=1}^{n} \sum_{r=1}^{L} w_{ir} e^{\beta_{o1L}(v_i)} - \sum_{i=1}^{n} \sum_{r=1}^{L} w_{ir} \lambda_{o0L}(v_i) - \sum_{i=1}^{n} \sum_{r=1}^{L} w_{ir} \varphi_{o0L}(v_i) \\
- \sum_{i=1}^{n} \sum_{r=1}^{L} w_{ir} \varphi_{o2L}(v_i) \\ + \sum_{i=1}^{n} \sum_{r=1}^{L} w_{ir} \left( y_{1ir} - k - 1 \right) \\
+ \sum_{i=1}^{n} \sum_{r=1}^{L} w_{ir} \left( y_{2ir} - k - 1 \right) \\
- \sum_{i=1}^{n} \sum_{r=1}^{L} w_{ir} \left( y_{1ir} - k \right) \varphi_{o0L}(v_i) \\
- \sum_{i=1}^{n} \sum_{r=1}^{L} w_{ir} \left( y_{2ir} - k \right) \varphi_{o2L}(v_i) \\
\sum_{k=0}^{\min(y_{1ir}, y_{2ir})} \left[ s_{o0} - s_{o1} - s_{o2} \right] 
\]

Then look for the first derivative of \( Q_{oL}^* \) for each parameter under \( H_0 \) and equalize it to zero.

Suppose that:

\[
s_{o0} = \left( \lambda_{o0L}(v_i) + k \varphi_{o0L}(v_i) \right) \\
\left( y_{1ir} - k - 1 \right) \\
\left( y_{2ir} - k - 1 \right) \\
\left( y_{1ir} - k \right) \varphi_{o0L}(v_i) \\
\left( y_{2ir} - k \right) \varphi_{o2L}(v_i) \\
\sum_{k=0}^{\min(y_{1ir}, y_{2ir})} \left[ s_{o0} - s_{o1} - s_{o2} \right] 
\]

Then the following results are obtained:

\[
\frac{\partial Q_{oL}^*}{\partial \lambda_{o0L}(v_i)} = \sum_{i=1}^{n} \sum_{r=1}^{L} w_{ir} \lambda_{o0L}(v_i) - \sum_{i=1}^{n} \sum_{r=1}^{L} \left( e^{\beta_{o1L}(v_i) - \lambda_{o0L}(v_i)} \right) - \sum_{i=1}^{n} \sum_{r=1}^{L} \left( e^{\beta_{o2L}(v_i) - \lambda_{o0L}(v_i)} \right) \\
+ \sum_{i=1}^{n} \sum_{r=1}^{L} w_{ir} \lambda_{o0L}(v_i) + \sum_{i=1}^{n} \sum_{r=1}^{L} w_{ir} \varphi_{o0L}(v_i) \\
\sum_{k=0}^{\min(y_{1ir}, y_{2ir})} \left[ s_{o0} - s_{o1} - s_{o2} \right] 
\]

\[
\frac{\partial Q_{oL}^*}{\partial \beta_{o0L}(v_i)} = \sum_{i=1}^{n} \sum_{r=1}^{L} \left( e^{\beta_{o1L}(v_i)} \right) - \sum_{i=1}^{n} \sum_{r=1}^{L} \left( e^{\beta_{o2L}(v_i)} \right) - \sum_{i=1}^{n} \sum_{r=1}^{L} \sum_{k=0}^{\min(y_{1ir}, y_{2ir})} \left[ s_{o0} e^{\beta_{o1L}(v_i)} \right] 
\]

\[
\frac{\partial Q_{oL}^*}{\partial \beta_{o2L}(v_i)} = \sum_{i=1}^{n} \sum_{r=1}^{L} \left( e^{\beta_{o2L}(v_i)} \right) - \sum_{i=1}^{n} \sum_{r=1}^{L} \sum_{k=0}^{\min(y_{1ir}, y_{2ir})} \left[ s_{o2} e^{\beta_{o2L}(v_i)} \right] 
\]

\[
\frac{\partial Q_{oL}^*}{\partial \varphi_{o0L}(v_i)} = \sum_{i=1}^{n} \sum_{r=1}^{L} \sum_{k=0}^{\min(y_{1ir}, y_{2ir})} \left[ s_{o0} k - k \right] 
\]
The first derivative results above are not closed-form, so the Newton-Raphson iteration solves it to obtain the parameter estimator under $H_0$, that is:

$$\hat{\theta}_{\text{OLS}} (v_i) = \begin{bmatrix} \hat{\lambda}_{\text{OLS}} (v_i), \hat{\beta}_{\text{OLS}} (v_i) \end{bmatrix}^\top : i = 1, 2, \ldots, n.$$  (27)

These parameter estimators maximize the likelihood function under $H_0$ of the period $L$, namely:

$$L(\hat{\theta}_L) = \max_{\theta_L} L(\theta_L)$$  (28)

These parameter estimators at (17) maximize the likelihood function under the population of the period $L$, namely:

$$L(\hat{\Omega}_L) = \max_{\Omega_L} L(\Omega_L)$$  (29)

Based on equations (28) and (29) then the odds ratio is determined as follows:

$$\Lambda_L = \frac{L(\hat{\theta}_L)}{L(\hat{\Omega}_L)} < K_\alpha$$  (30)

where $K_\alpha$ is a constant that depends on $\alpha$ and $0 \leq K_\alpha \leq 1$. Furthermore, equation (30) can be written as follows:

$$\log \left( \frac{L(\hat{\theta}_L)}{L(\hat{\Omega}_L)} \right) < \log K_\alpha$$

$$-2 \log \left( \frac{L(\hat{\theta}_L)}{L(\hat{\Omega}_L)} \right) > -2 \log K_\alpha$$

$$G^2_L(\hat{\theta}_L) = -2 \log \frac{L(\hat{\theta}_L)}{L(\hat{\Omega}_L)} = 2 \left( \log L(\hat{\Omega}_L) - \log L(\hat{\theta}_L) \right)$$

Based on the steps carried out by Berliana et al. [14], it was found that $G^2_L(\hat{\theta}_L) \sim \chi^2_{(a-b)}$. $G^2_L(\hat{\theta}_L)$ is the $\chi^2$ distribution approach with degrees of freedom $a-b$, namely the number of parameters under the population minus the number of parameters under $H_0$. So that $H_0$ is rejected if $G^2_L(\hat{\theta}_L) > \chi^2_{(a-2, pL)}$, which means that there are predictor variables that affect the response variable.

If the result of the simultaneous test rejects $H_0$, then it is followed by partial tests. According to Triyanto et al. [15], parameter estimators obtained by MLE were asymptotically normally distributed. So that it is obtained:
\[ \hat{\theta}_L(v_i) \xrightarrow{d_{\alpha \to 0}} N\left(\theta_L(v_i), \left[I\left(\hat{\theta}_L(v_i)\right)\right]^{-1}\right) \]  

(31)

then it is obtained:

\[ Z = \left[I\left(\hat{\theta}_L(v_i)\right)\right]^{-1/2} \left(\hat{\theta}_L(v_i) - \theta_L(v_i)\right) \xrightarrow{d_{\alpha \to 0}} N(0, 1) \]  

(32)

So that for partial testing using the Z test. The testing hypothesis for parameter \( \beta \) is:

\[ H_0 : \beta_{i \ell} (v_i) = 0 \]
\[ H_1 : \beta_{i \ell} (v_i) \neq 0 \]

This test is done for \( c = 1, 2; j = 1, 2, \ldots, p; i = 1, 2, \ldots, n \).

By paying attention to equation (32), the test statistic used is:

\[ Z = \frac{\hat{\beta}_{i \ell} (v_i)}{SE(\hat{\beta}_{i \ell} (v_i))} \]  

(33)

where \( SE(\hat{\beta}_{i \ell} (v_i)) = \sqrt{\text{var}(\hat{\beta}_{i \ell} (v_i))} \). \( \text{var}(\hat{\beta}_{i \ell} (v_i)) \) is obtained from the main diagonal of the covariance variance matrix \( \left[I\left(\hat{\theta}_L(v_i)\right)\right]^{-1} \).

Reject \( H_0 \) if \( |Z| > Z_{\alpha/2} \).

The testing hypothesis for parameter \( \lambda_i \) is:

\[ H_0 : \lambda_{iL} (v_i) = 0 \]
\[ H_1 : \lambda_{iL} (v_i) \neq 0 \]

This test is done for \( i = 1, 2, \ldots, n \).

By paying attention to equation (32), the test statistic used is:

\[ Z = \frac{\hat{\lambda}_{iL} (v_i)}{SE(\hat{\lambda}_{iL} (v_i))} \]  

(34)

Reject \( H_0 \) if \( |Z| > Z_{\alpha/2} \).

The testing hypothesis for parameter \( \varphi \) is:

\[ H_0 : \varphi_{i \ell} (v_i) = 0 \]
\[ H_1 : \varphi_{i \ell} (v_i) \neq 0 \]

This test is done for \( c = 1, 2; i = 1, 2, \ldots, n \).

By paying attention to equation (32), the test statistic used is:

\[ Z = \frac{\hat{\varphi}_{i \ell} (v_i)}{SE(\hat{\varphi}_{i \ell} (v_i))} \]  

(35)

Reject \( H_0 \) if \( |Z| > Z_{\alpha/2} \).
4. Conclusions

GTWBGPGR is a development of GWBGPR. GTWBGPGR can accommodate spatial and temporal heterogeneity. Parameter estimation in GTWBGPGR using MLE. The results obtained are not closed-form, so they are solved by Newton-Raphson iteration. Furthermore, the simultaneous hypothesis testing using MLRT and test statistic distribution was a chi-square distribution approximation. The partial testing using Z test.

References

[1] Consul P and Famoye F 1992 Generalized poisson regression model Communications in Statistics-Theory and Methods 21(1) p 89-109
[2] Omzen I and Famoye F 2007 Count regression models with an application to zoological data containing structural zeros Journal of Data Science 5(4) p 491-502
[3] Adams S O, Bamanga M A, Olanrewaju S O, Yahaya H U and Akano R O 2020 Modeling covid-19 cases in nigeria using some selected count data regression models International Journal of Healthcare and Medical Sciences 6(4) p 64-73
[4] Vernic R 1997 On the bivariate generalized poisson distribution ASTIN Bulletin: The Journal of the IAA 27(1) p 23-32
[5] Famoye F 2012 Comparisons of some bivariate regression models Journal of Statistical Computation and Simulation 82(7) p 937-949
[6] Zamani H, Faroughi P and Ismail N 2016 Bivariate generalized poisson regression model: applications on health care data Empirical Economics 51(4) p 1607-1621
[7] Fotheringham A S, Brunsdon C and Charlton M 2003 Geographically weighted regression: the analysis of spatially varying relationships John Wiley & Sons
[8] Purhadi, Sutikno, Berliana, S M and Setiawan D I 2021 Geographically weighted bivariate generalized Poisson regression: application to infant and maternal mortality data Letters in Spatial and Resource Sciences p 1-21
[9] Gujarati D N and Porter D C 2003 Basic econometrics (ed.) Singapore: McGrew Hill Book Co
[10] Fotheringham A S, Crespo R and Yao J 2015 Geographical and temporal weighted regression (GTWR) Geographical Analysis 47(4) p 431-452
[11] Ma X, Zhang J, Ding C and Wang Y 2018 A geographically and temporally weighted regression model to explore the spatiotemporal influence of built environment on transit ridership Computers, Environment and Urban Systems 70 p 113-124.
[12] Huang B, Wu B and Barry M 2010 Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices International Journal of Geographical Information Science 24(3) p 383-401
[13] Allo C B G, Otok B W and Purhadi 2019 Estimation parameter of generalized poisson regression model using generalized method of moments and its application IOP Conference Series: Materials Science and Engineering 546(5) p 052050
[14] Berliana S M, Purhadi, Sutikno and Rahayu S P 2020 Parameter estimation and hypothesis testing of geographically weighted multivariate generalized poisson regression Mathematics 8(9) p 1523
[15] Triyanto, Purhadi, Otok B W and Purnami S W 2015 Parameter estimation of geographically weighed multivariate poisson regression. Appl. Math. Sci, 9 p 4081-4093.

Acknowledgments

The authors are grateful to Statistics Indonesia (Badan Pusat Statistik), which has supported this research financially. The authors also thank lecturers and college friends at the Sepuluh Nopember Institute of Technology (Institut Teknologi Sepuluh Nopember) for their guidance and support.