Parameter estimation and hypothesis testing of geographically and temporally weighted bivariate Poisson inverse Gaussian regression model

Meylita Sari1*, Sutikno1 and Purhadi1

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

Corresponding E-mail: meylitasari27@gmail.com

Abstract. One of the appropriate methods used to model count data response and its corresponding predictors is Poisson regression. Poisson regression strictly assumes that the mean and variance of response variables should be equal (equidispersion). Nonetheless, some cases of the count data unsatisfied this assumption because variance can be larger than mean (over-dispersion). If overdispersion is violated, causing the underestimate standard error. Furthermore, this will lead to incorrect conclusions in the statistical test. Thus, a suitable method for modelling this kind of data needs to develop. One alternative model to outcome the overdispersion issue in bivariate response variable is the Bivariate Poisson Inverse Gaussian Regression (BPIGR) model. The BPIGR model can produce a global model for all locations. On the other hand, each location and time have different geographic conditions, social, cultural, and economical so that Geographically and Temporally Bivariate Poisson Inverse Gaussian Regression (GTWBPIGR) is needed. The weighting function spatial-temporal in GTWBPIGR generates a different local model for each period. GTWBPIGR model solves the overdispersion case and generates global models for each period and location. The parameter estimation of the GTWBPIGR model uses the Maximum Likelihood Estimation (MLE) method, followed by Newton Raphson iteration. Meanwhile, the test statistics on the hypothesis testing is simultaneously testing of the GTWBPIGR model is obtained with the Maximum Likelihood Ratio Test (MLRT) approach, using n large samples of the statistical test is chi-square distribution. Moreover, the test statistics for partially testing used the Z-test statistic.

Keywords: Overdispersion, BPIGR, GTWBPIGR, MLE, MLRT

Track Name: Advanced Technology and Renewable Energy

1. Introduction

The Poisson distribution is a discrete distribution with a random variable value in positive integers, so it is a good choice for modelling count data. The Poisson distribution is only determined by one parameter that defines the distribution's mean and variance. Poisson regression has assumptions when the mean and variance of the response variable must be the same value (equidispersion). However, the mean and variance of count data are often not the same [1]; often, the variance is greater than the mean.
(overdispersion). If the overdispersion case is ignored, it can lead to an underestimate of the standard error estimation. In overcoming overdispersion cases, several models are formed, which are a combination of the Poisson distribution with several distributions, both discrete and continuous (mixed Poisson distribution). The mixed Poisson distribution is an alternative solution for the overdispersion case; one of the distributions used is the Poisson Inverse Gaussian (PIG). PIG is a combination of the Poisson distribution and Gaussian inverse [2].

The combined modelling of two or more data counts has received much attention in recent years. The bivariate count model is a case with a pair of correlated response variables and needs to be estimated together. The Poisson bivariate is the most widely used model for bivariate computations. However, it is not possible for overdispersion or under dispersion cases [3]. Thus, to overcome overdispersion, some bivariate mixing of the Poisson distribution is required. Poisson bivariate mixed distribution is an alternative solution to negotiating over-dispersion cases.

BPIGR will produce a global model that is considered valid at all data collection locations [4]. However, every location and time have different geographical, social, cultural, and economic conditions. Therefore, a regression model will be developed that considers the spatial and temporal effects, such as Geographically and Temporally Weighted Bivariate Poisson Inverse Gaussian Regression (GTWBPiGR). The GTWR parameter may be estimated anywhere in the study area given the dependent variable and where one or more independent variables have been measured in places whose location is known at any given time [5]. The spatial-temporal weighting function in the GTWBPiGR model produces a different local model for each period. The GTWBPiGR model is used to overcome over-dispersion and to generate localized models for each period. In this paper, the case of over-dispersion and developing localized models for each period can be solved using the proposed GTWBPiGR model. The parameter estimation of the GTWBPiGR model uses the Maximum Likelihood Estimation (MLE) method with Newton Raphson numerical iteration. At the same time, the test statistics on hypothesis testing are simultaneously obtained by the Maximum Likelihood Ratio Test (MLRT) method approach, using n large samples of the statistical test is chi-square distribution. Moreover, the test statistics for partially testing used the Z-test statistic approach using n large samples of the statistical test is normal (0,1) distribution.

2. Materials and Methods

2.1 Bivariate Poisson Inverse Gaussian Regression Models

Bivariate Poisson Inverse Gaussian distribution (BPIG) has two correlated count variables. Suppose two random variables independent of each other and have a Poisson distribution, $Y_1$ and $Y_2$ have mean $\mu_1$ and $\mu_2$. The variable $\nu$ is a random variable with an Inverse Gaussian distribution with the joint probability density function as follows:

$$g(\nu; \tau) = \left(2\pi \tau^3\right)^{-\frac{1}{2}} e^{-\frac{(\nu-\tau)^2}{2\tau}}, \nu > 0$$  \hspace{1cm} (1)$$

The BPIG distribution is based on the Inverse Gaussian and Poisson mixture distribution which has the following joint density function:

$$f(y_1, y_2; \mu_1, \mu_2, \tau) = \left(\frac{2\pi}{\tau}\right)^{\frac{1}{2}} e^{-2\left(\mu_1^2 + \mu_2^2\right) / 2\tau} K_0(\sqrt{2\tau (\mu_1 + \mu_2)}), y \geq 0$$  \hspace{1cm} (2)$$

with $s = y_1 + y_2 - \frac{1}{2}$ and $z = \sqrt{\frac{1}{\tau^2} + \frac{2(\mu_1 + \mu_2)}{\tau}}$. so, $K_0(z) = K_{y_1+y_2-\frac{1}{2}} \left(\frac{1}{\sqrt{2\tau (\mu_1 + \mu_2)}}\right)$ is the third type of modification of the Bessel function [6].
A bivariate regression model is a regression model with two correlated response variables and one or more predictor variables. BPIGR is used when the two response variables \((Y_1, Y_2)\) are experiencing overdispersion. The joint probability density function for \(Y_1\) and \(Y_2\) is:

\[
f(y_j; \beta_j, \tau, j = 1, 2) = e^{\gamma K(z)} \left( \frac{2}{\pi \tau} \right)^{\frac{3}{2}} \left( 1 + 2 \tau \sum_{j=1}^{\infty} \mu_j \right) \left( \frac{2^2 \sum_{j=1}^{\infty} y_j - 1}{4} \right) \prod_{j=1}^{\infty} \frac{\mu_j}{y_j!}
\]

(3)

Modelling bivariate Poisson inverse Gaussian regression (BPIGR) requires a logarithm (log) function. Parameters \(\mu_{ij}\) are associated with predictor variables using the In link function so that the BPIG regression model can be written as follows:

\[
\log \mu_{ij} = x_i^T \beta_j
\]

with \(E(Y_{ij}) = \mu_{ij} = \exp(x_i^T \beta_j)\); \(x_i = [1 \ x_{i1} \ x_{i2} \ ... \ x_{ip}]\) is the vector of the predictor variable in the \(i\)-th observation and the \(j\)-th response variable \((i = 1, 2, ..., n)\). \(\beta_j = [\beta_{j0} \ \beta_{j1} \ \beta_{j2} \ ... \ \beta_{jp}]^T; j = 1, 2\) is the vector of the \((p + 1) \times 1\) dimensionless regression coefficient in the \(j\)-th response variable.

The parameter estimation of the BPIGR model was obtained by the MLE method with Newton Raphson numerical iteration. While the test statistics on hypothesis testing are simultaneously obtained by the MLRT method.

2.2 Geographically and Temporally Weighted Bivariate Poisson Inverse Gaussian Regression Models

The GTWBPIGR model is a development of the GWBPIGR model by considering the temporal aspect. Based on the BPIG equation (3), the joint probability function \(y_{1lt}\) and \(y_{2lt}\) is

\[
(y_{1lt}, y_{2lt}) \sim GTWBPIGR(\beta_{1lt}(u_l, v_l, t_l), \beta_{2lt}(u_l, v_l, t_l), \tau_l)
\]

\[
f(y_{lt} \mid \beta_{lt}(u_l, v_l, t_l), \tau_l) = e^{\gamma K_d(z(u_l, v_l, t_l))} \left( \frac{2}{\pi \tau_l} \right)^{\frac{3}{2}} \left( 1 + 2 \tau_l \sum_{j=1}^{\infty} e^{x_{ij}^T \beta_{ij}(u_l, v_l, t_l)} \right) \left( \frac{2^2 \sum_{j=1}^{\infty} y_{jlt} - 1}{4} \right) \prod_{j=1}^{\infty} \frac{e^{x_{ij}^T \beta_{ij}(u_l, v_l, t_l)}}{y_{jlt}!}
\]

(4)

with \(l\)-th is the period, \(u_{jl} = \exp(x_{ijl}^T \beta_{jl}(u_l, v_l, t_l))\); \(x_{ijl} = [1 \ x_{i1l} \ x_{i2l} \ ... \ x_{ipl}]\) is the vector of the predictor variable in the \(i\)-th observation and the \(j\)-th response variable; \((i = 1, 2, ..., n; \ l = 1, 2, ..., L\ dan j=1,2)\). \(\beta_{jl}(u_l, v_l, t_l) = [\beta_{j1l}(u_l, v_l, t_l) \ \beta_{j2l}(u_l, v_l, t_l) \ ... \ \beta_{jpl}(u_l, v_l, t_l)]^T\) are vectors of regression coefficients that are \((K + 1) \times 1\) dimensioned on the response variable in the \(j\)-th in the \(l\)-period with a temporal-spatial weighting matrix. \(u_l, v_l\) are location latitude and longitude.

The parameter estimation GTWBPIGR model using the MLE method is carried out in stages in each period by including observations of the previous period. The parameter estimation in period first uses \(n\) data in period 1. In period 2 uses \(n\) data in period 1 and \(n\) data in period 2. The parameter estimation in period \(L\) uses total \(n\) data in periods 1, 2 to \(L\). The spatial-temporal weighting function in the GTWBPIGR model uses \(w_{ijl}\) \([5]\).
For the weighting function for the period \( l = 1 \), the temporal-spatial weighting is obtained with the same Euclidean distance value as the GWR weight. During one period, weighting calculations are only performed using the data for period \( l \) total \( n \) observations to obtain the weights for the \( i \)-location. When period \( l = 2 \), the weighted calculation uses data for periods 1 and 2 with a total of 2\( n \) observations. So that the weights \( w_{l12} \ldots w_{l2n} \) for the \( i \)-location are obtained. And so on for three periods to \( L \) periods.

For temporal-spatial weighting in determining the Euclidean distance between the location and time of observation with the formula in the following equation.

\[
d_{i\ast} = \sqrt{(u_i - u_{i'})^2 + (v_i - v_{i'})^2 + \gamma(t_i - t_{i'})^2}
\]

So we get a temporal-spatial weighting with the Adaptive Bisquare Kernel Function, follows as

\[
w_{i\ast\ast} = \left(1 - \left(\frac{d_{i\ast\ast}}{b_{ij}}\right)^2\right) ; \quad \text{for } d_{i\ast\ast} \leq b_{ij}
\]

\[
w_{i\ast\ast} = 0 ; \quad \text{for } d_{i\ast\ast} > b_{ij}
\]

with, \( d_{l\ast\ast} \) is the Euclidean distance and \( b_{ij} \) is bandwidth.

The selection of the optimum bandwidth can be made by using the Generalized Cross-Validation (GCV) method. This GCV method is defined as follows:

\[
GCV = \min_h \left\{ \sum_{i=1}^{n} \frac{\left(\mathbf{y}_i - \mathbf{y}_{si}(b_{ij})\right)^T \left(\mathbf{y}_i - \mathbf{y}_{si}(b_{ij})\right)}{(n-v_i)^2} \right\}
\]

with:

\( \mathbf{y}_{si}(h) \): estimator value \( \mathbf{y}_i \) when site observations \((u_i, v_i)\) are not included in the estimate

\( \mathbf{y}_i \): \( i \)-th response variable

\( b_{ij} \): bandwidth

\( n \): number of observations

\( v_i \): \( Tr(S) \)

\( S \): a matrix that represents the estimated value of GTWBPIGR

\[
S = \begin{pmatrix}
x_1 \left( X^T W(u_i, v_i) X \right)^{-1} X^T W(u_i, v_i) \\
x_2 \left( X^T W(u_2, v_2) X \right)^{-1} X^T W(u_2, v_2) \\
\vdots \\
x_n \left( X^T W(u_n, v_n) X \right)^{-1} X^T W(u_n, v_n)
\end{pmatrix}
\]

The process of obtaining bandwidth that minimizes GCV value can be done using the golden section search technique [7].

Furthermore, for test statistics on the hypothesis testing simultaneously, the GTWBPIGR model is obtained with the Maximum Likelihood Ratio Test (MLRT) approach. Using \( n \) large samples of the statistical test is chi-square distribution. Moreover, the test statistics for partially testing used the Z-test statistic.

3. Results and Discussion

3.1 Parameter Estimation of GTWBPIGR Model
The parameter estimator of the GTWBPIGR model uses MLE. The presumed parameters $\beta_{y1}(u_i, v_i, t_i)$, $\beta_{y2}(u_i, v_i, t_i)$ and $\tau_j$. The MLE method is carried out by taking the previous n random samples $(Y_{1tt}, Y_{2tt}, X_{1tt}, X_{2tt}, ..., X_{ptt})$ with $k = 1, 2, ..., p$ and $i = 1, 2, ..., n$. The form of the model has been mentioned in equation (4). The joint probability density function of $Y_{1tt}$ and $Y_{2tt}$ as follows:

$$f(y_{1l}, y_{2l} \mid \beta_{y1}(u_i, v_i, t_i), \beta_{y2}(u_i, v_i, t_i), \tau_j)$$

$$= e^{\tau_j} K_{il}(z(u_i, v_i, t_i)) \left( \frac{2}{\pi \tau_j} \right)^{\frac{1}{2}} \left( 1 + 2 \tau_j \sum_{j=1}^{2} e^{x_{jy}^T \beta_{yj}(u_i, v_i, t_i)} \right)^{\frac{2 \sum_{j=1}^{y_{jl}} - 1}{4}} \prod_{j=1}^{2} e^{y_{jl} \beta_{yj}(u_i, v_i, t_i)}$$

(5)

where:

$$s_l = y_{1l} + y_{2l} - \frac{1}{2} ; \quad \mu_{yj}(u_i, v_i, t_i) = e^{x_{jy}^T \beta_{yj}(u_i, v_i, t_i)} ; \quad z(u_i, v_i, t_i) = \sum_{j=1}^{\tau_j} \left( 1 + 2 \tau_j \sum_{j=1}^{\mu_{yj}(u_i, v_i, t_i)} \right)^{\frac{1}{2}}$$

$$K_{il}(z(u_i, v_i, t_i)) = K_{y_{il}} \left( \sum_{j=1}^{\tau_j} \frac{1}{2} \left( 1 + 2 \tau_j \sum_{j=1}^{\mu_{yj}(u_i, v_i, t_i)} \right)^{\frac{1}{2}} \right)$$

Parameter estimation using data from the previous period. Suppose that period $L$ uses data information from $n$ samples of period 1, $n$ samples of period two until $n$ samples of period $L$ [8]. So that the likelihood function for the population in period $L$ of equation (5) is

$$L(\beta_{yL}(u_i, v_i, t_i); \tau_j; j = 1, 2; i = 1, 2, ..., n; l = 1, 2, ..., L)$$

$$= \prod_{i=1}^{L} \prod_{l=1}^{n} e^{\tau_j} K_{il}(z(u_i, v_i, t_i)) \left( \frac{2}{\pi \tau_j} \right)^{\frac{1}{2}} \left( 1 + 2 \tau_j \sum_{j=1}^{2} e^{x_{jy}^T \beta_{yj}(u_i, v_i, t_i)} \right)^{\frac{2 \sum_{j=1}^{y_{jl}} - 1}{4}} \prod_{j=1}^{2} e^{y_{jl} \beta_{yj}(u_i, v_i, t_i)}$$

(6)

Next, the log-likelihood function for equation (6) is as follows:

$$Q_{L} = \log L(\beta_{yL}(u_i, v_i, t_i); \beta_{y2}(u_i, v_i, t_i); \tau_j; i = 1, 2, ..., n)$$

$$= \sum_{i=1}^{L} \sum_{l=1}^{n} \log e^{\tau_j} K_{il}(z(u_i, v_i, t_i)) \left( \frac{2}{\pi \tau_j} \right)^{\frac{1}{2}} \left( 1 + 2 \tau_j \sum_{j=1}^{2} e^{x_{jy}^T \beta_{yj}(u_i, v_i, t_i)} \right)^{\frac{2 \sum_{j=1}^{y_{jl}} - 1}{4}} \prod_{j=1}^{2} e^{y_{jl} \beta_{yj}(u_i, v_i, t_i)}$$

(7)

So we get the log-likelihood function to estimate the $\tau^*$ location with $w_{\tau_j}$ is spatial-temporal weight as follows:


\(Q^{*}_L = \log L(\beta_{1L}(u_i, v_j, t_j); \beta_{2L}(u_i, v_j, t_j); \tau_L); t' = 1, 2, ..., n)\)

\[
\frac{1}{\tau_L} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \log K_{z}(z(u_i, v_j, t_j)) + \log \left( \frac{2}{\pi \tau_L} \right)^{\frac{3}{2}} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \log \left( \frac{2 \tau_L - 1}{4} \right)
\]

\[
\log \left( 1 + 2 \tau_L \sum_{i=1}^{n} \sum_{j=1}^{n} e^{\beta_{z}(u_i, v_j, t_j)} \right) + \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \log(y_{ij})
\]

To obtain the parameter estimation of the GTWBPIGR model, the functions in the \(Q^{*}_L\) equation are derived respectively to \(\beta_{1L}(u_i, v_j, t_j), \beta_{2L}(u_i, v_j, t_j)\) and \(\tau_L\), as follows:

\[
\frac{\partial Q^{*}_L}{\partial \beta_{1L}(u_i, v_j, t_j)} = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \frac{\partial \log K_{z}(z(u_i, v_j, t_j))}{\partial \beta_{1L}(u_i, v_j, t_j)} + \frac{\tau_L}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left( y_{ij} + y_{2ij} - \frac{1}{2} \right) x_{ij}^T
\]

\[
\frac{\partial Q^{*}_L}{\partial \beta_{2L}(u_i, v_j, t_j)} = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left[ y_{ij} + y_{2ij} - \frac{1}{2} \right] \frac{\mu_{z}(u_i, v_j, t_j) x_{ij}^T}{\sqrt{1 + 2 \tau_L \sum_{i=1}^{n} \sum_{j=1}^{n} \mu_{z}(u_i, v_j, t_j)}}
\]

\[
\frac{\partial Q^{*}_L}{\partial \tau_L} = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left[ y_{ij} - M(y_{i1}, y_{2i}) \mu_{z}(u_i, v_j, t_j) \right] x_{ij}^T
\]

With the same procedure, we get:

\[
\frac{\partial Q^{*}_L}{\partial \beta_{2L}(u_i, v_j, t_j)} = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left[ y_{ij} - \mu_{z}(u_i, v_j, t_j) M(y_{i1}, y_{2i}) \right] x_{ij}^T
\]

Furthermore, the first derivative of the \(Q^{*}_L\) to parameter \(\tau_L\) is formulated as below:

\[
\frac{\partial Q^{*}_L}{\partial \tau_L} = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left[ y_{ij} - \mu_{z}(u_i, v_j, t_j) M(y_{i1}, y_{2i}) \right] x_{ij}^T
\]
\[
\frac{\partial Q_j}{\partial \tau_L} = \frac{\partial}{\partial \tau_L} \left[ \frac{1}{\tau_L} \sum_{i=1}^{L} \sum_{t=1}^{n} w_{i,t} \log K_p(z(u_{i,t}, v_{i,t}, L_{i,t})) + \log \left( \frac{2}{\pi \tau_L} \right)^{\frac{1}{2}} \sum_{i=1}^{L} \sum_{t=1}^{n} w_{i,t} \sum_{j=1}^{2} \frac{y_{ij} - 1}{4} \times \log \left( 1 + \frac{2\tau_L}{2} \sum_{j=1}^{2} (y_{ij} - 1)^2 \right) \right] 
\]

\[
= -\frac{1}{\tau_L^2} \sum_{i=1}^{L} \sum_{t=1}^{n} w_{i,t} - \frac{1}{2\tau_L} \sum_{i=1}^{L} \sum_{t=1}^{n} w_{i,t} + \sum_{i=1}^{L} \sum_{t=1}^{n} \frac{\partial \log K_p(z(u_{i,t}, v_{i,t}, L_{i,t}))}{\partial \tau_L} + \frac{1}{\tau_L} \left( 1 + 2\tau_L \sum_{j=1}^{2} \mu_{jl}(u_{i,t}, v_{i,t}, L_{i,t}) \right) \]

\[
\frac{\partial \log K_p(z(u_{i,t}, v_{i,t}, L_{i,t}))}{\partial \tau_L} = \frac{1 + \tau_L \sum_{j=1}^{2} \mu_{jl}(u_{i,t}, v_{i,t}, t_{i,j})}{\tau_L} \left( 1 + 2\tau_L \sum_{j=1}^{2} \mu_{jl}(u_{i,t}, v_{i,t}, L_{i,t}) \right) \]

\[
\frac{\partial \log K_p(z(u_{i,t}, v_{i,t}, L_{i,t}))}{\partial \tau_L} = \frac{1}{\tau_L} \left( 1 + 2\tau_L \sum_{j=1}^{2} \mu_{jl}(u_{i,t}, v_{i,t}, L_{i,t}) \right) \]

(13)

where: \( M(y_{i1}, y_{i2}) = \frac{1}{1 + 2\tau_L \sum_{j=1}^{2} \mu_{jl}(u_{i,t}, v_{i,t}, L_{i,t})} \)

Equations (11)–(13) equate to zero and produce a non-explicit form. Hence, an iterative method needs to be applied for estimating parameters, so Newton Raphson's iteration will solve it. So that the Newton-Raphson iteration algorithm is as follows:

Step 1. Determine the initial value for the parameters

\[
\theta_L^{(0)} = [\beta_{1L}^{(0)}(u_{i,t}, v_{i,t}, t_{i,j}) \quad \beta_{2L}^{(0)}(u_{i,t}, v_{i,t}, L_{i,t}) \quad \tau_L^{(0)}]^T.
\]

The initial value of parameters \( \theta_L^{(0)} \) is using univariate Poisson regression. The initial value \( \tau_L \) used the average of the observed overdispersion based on the variance of PIGD.
Step 2. Determine the gradient vector $g(\theta_L^{(m)}(u_{i}, v_{i}, t_{i}))$, which is the elements that consist of the first derivative of the log-likelihood

$$g(\theta_L^{(m)}(u_{i}, v_{i}, t_{i})) = \left[ \frac{\partial Q_L^*}{\partial \beta_{1L}(u_{i}, v_{i}, t_{i})}, \frac{\partial Q_L^*}{\partial \beta_{2L}(u_{i}, v_{i}, t_{i})}, \frac{\partial Q_L^*}{\partial \tau_L^*} \right]$$

function.

Step 3. Determine the Hessian matrix $H(\theta_L^{(m)}(u_{i}, v_{i}, t_{i}))$, which is the elements consist of the second derivative of the log-likelihood function, as follows

$$H(\theta_L^{(m)}(u_{i}, v_{i}, t_{i})) = \left[ \begin{array}{ccc}
\frac{\partial^2 Q_L^*}{\partial \beta_{1L}(u_{i}, v_{i}, t_{i})^2} & \frac{\partial^2 Q_L^*}{\partial \beta_{1L}(u_{i}, v_{i}, t_{i}) \partial \beta_{2L}(u_{i}, v_{i}, t_{i})} & \frac{\partial^2 Q_L^*}{\partial \beta_{1L}(u_{i}, v_{i}, t_{i}) \partial \tau_L^*} \\
\frac{\partial^2 Q_L^*}{\partial \beta_{2L}(u_{i}, v_{i}, t_{i}) \partial \beta_{1L}(u_{i}, v_{i}, t_{i})} & \frac{\partial^2 Q_L^*}{\partial \beta_{2L}(u_{i}, v_{i}, t_{i})^2} & \frac{\partial^2 Q_L^*}{\partial \beta_{2L}(u_{i}, v_{i}, t_{i}) \partial \tau_L^*} \\
\frac{\partial^2 Q_L^*}{\partial \tau_L^* \partial \beta_{1L}(u_{i}, v_{i}, t_{i})} & \frac{\partial^2 Q_L^*}{\partial \tau_L^* \partial \beta_{2L}(u_{i}, v_{i}, t_{i})} & \frac{\partial^2 Q_L^*}{\partial \tau_L^*^2} 
\end{array} \right]$$

Step 4. Start working on the Newton–Raphson iteration using the following equation:

$$\theta_L^{(m+1)}(u_{i}, v_{i}, t_{i}) = \theta_L^{(m)}(u_{i}, v_{i}, t_{i}) - H^{-1}(\theta_L^{(m)}(u_{i}, v_{i}, t_{i})))g(\theta_L^{(m)}(u_{i}, v_{i}, t_{i}))$$

with $\theta_L^{(m)}(u_{i}, v_{i}, t_{i}) = \left[ \beta_{1L}(u_{i}, v_{i}, t_{i}), \beta_{2L}(u_{i}, v_{i}, t_{i}), \tau_L^* \right]$ and $m = 1, 2, \ldots, m^*$

Step 5. The iteration will stop when $\left\| \theta_L^{(m+1)}(u_{i}, v_{i}, t_{i}) - \theta_L^{(m)}(u_{i}, v_{i}, t_{i}) \right\| \leq \varepsilon$ with $\varepsilon$ is a very small value and will produce the estimator value for each parameter.

3.2 Hypothesis Testing of GTWBPIGR Model

The hypothesis testing of the GTWBPIGR model using the Maximum Likelihood Ratio Test (MLRT) method both simultaneously and partially [9]. Hypothesis testing is carried out in stages each period as the parameter estimates, with an L’s period $L = 1, 2, \ldots, L^*$. This method involves two likelihood functions: $L(\Omega_L)$ is the maximum likelihood value for a model involving predictor variables and $L(\omega_L)$ is the maximum likelihood value for a simple model without involving predictor variables.

Simultaneously hypothesis testing is performed to determine the significance of the regression parameters in the model with the following hypothesis:

$H_0 : \beta_{jL}(u_{i}, v_{i}, t_{i}) = \beta_{jL}(u_{i}, v_{i}, t_{i}) = \ldots = \beta_{jL}(u_{i}, v_{i}, t_{i}) = 0$; and the alternative hypothesis is at least one $H_1 : \beta_{jL}(u_{i}, v_{i}, t_{i}) \neq 0; j = 1, 2; i = 1, 2, \ldots, n; k = 1, 2, \ldots, p$. Where $k$ is the number of predictor variables.

Let $\Omega_L^*$ be a set of parameters under population with $\Omega_L = \{\beta_{1L}(u_{i}, v_{i}, t_{i}), \beta_{2L}(u_{i}, v_{i}, t_{i}), \tau_L^*; i = 1, 2, \ldots, n\}$

The likelihood function $L(\Omega_{L^*})$ for each model, as follows:

$$L(\Omega_{L^*}) = \prod_{j=1}^{l} \prod_{t=1}^{T} f(y_{ijt}, y_{ijt}; \beta_{1L}(u_{i}, v_{i}, t_{i}), \beta_{2L}(u_{i}, v_{i}, t_{i}), \tau_L)$$

And then, let $\omega_L$ as a set of parameters under the null hypothesis $\omega_L = \{\beta_{00L}(u_{i}, v_{i}, t_{i}), \beta_{20L}(u_{i}, v_{i}, t_{i}), \tau_{0L}; i = 1, 2, \ldots, n\}$.
The likelihood function $L(\omega_j)$ for each model, as follows:

$$L(\omega_j) = \prod_{i=1}^{n} \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\tau_{\omega_j}}} \exp\left(-\frac{1}{2\tau_{\omega_j}} \sum_{i=1}^{n} (z_{ui} - u_{i,j})^2 \right) + \frac{1}{2\pi\tau_{\omega_j}} \sum_{j=1}^{n} \frac{2}{\pi\tau_{\omega_j}} \sum_{j=1}^{n} e^{eta_{\omega_{ij}}(u_{ij},v_{ij},t_j)} \right) \left(2\sum_{j=1}^{n} \sum_{j=1}^{n} e^{\beta_{\omega_{ij}}(u_{ij},v_{ij},t_j)} \right)$$

So, we get the log-likelihood function under the null hypothesis, $l(\omega_L)$ with weighted spatial-temporal $w_L$ as follows:

$$l(\omega_L) = \frac{1}{\tau_{\omega_L}} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \log K_{\omega_L}(z_{ui},u_{ij},v_{ij},t_j) + \frac{n}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \log \left(2\pi\tau_{\omega_L} \right) - \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} y_{ij} \left(1 + 2\tau_{\omega_L} \sum_{j=1}^{n} e^{eta_{\omega_{ij}}(u_{ij},v_{ij},t_j)} \right)$$

Then, derived $l(\omega_L)$ to the parameters $\beta_{000}(u_{i,j},v_{i,j},t_j)$, $\beta_{001}(u_{i,j},v_{i,j},t_j)$ and $\tau_{\omega_L}$, as below:

$$\frac{\partial l(\omega_L)}{\partial \beta_{000}(u_{i,j},v_{i,j},t_j)} = \frac{1}{\tau_{\omega_L}} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left[ y_{ij} - M(y_{1a},y_{2a}) e^{\beta_{000}(u_{i,j},v_{i,j},t_j)} \right]$$

$$\frac{\partial l(\omega_L)}{\partial \beta_{001}(u_{i,j},v_{i,j},t_j)} = \frac{1}{\tau_{\omega_L}} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left[ y_{2a} - M(y_{1a},y_{2a}) e^{\beta_{001}(u_{i,j},v_{i,j},t_j)} \right]$$

$$\frac{\partial l(\omega_L)}{\partial \tau_{\omega_L}} = \frac{1}{\tau_{\omega_L}} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left[ \frac{M(y_{1a},y_{2a}) \left(1 + 2\tau_{\omega_L} \sum_{j=1}^{n} e^{eta_{\omega_{ij}}(u_{ij},v_{ij},t_j)} \right)}{\tau_{\omega_L}^2} \right] - \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{j=1}^{n} w_{ij} y_{ij} \left(1 + 2\tau_{\omega_L} \sum_{j=1}^{n} e^{eta_{\omega_{ij}}(u_{ij},v_{ij},t_j)} \right)$$

Next, determine the odds ratio $\Lambda_L = \frac{L(\hat{\omega}_L)}{L(\hat{\Omega}_L)}$.

The decision rejects $H_0$ if $\Lambda_L > \Lambda_{L0}$, where $\Lambda_L < \Lambda_{L0} < 1$, and $\alpha = P(\Lambda_L > \Lambda_{L0} \mid H_0$ is true). So for decision making, test statistics are used:

$$G_L^2 = -\log \Lambda^2_L$$

$$= 2 \left( n L \left( \frac{1}{\tau_L} - \frac{1}{\tau_{L0}} - \frac{1}{2} \ln \left( \frac{\tau_{L0}}{\tau_L} \right) \right) + \sum_{i=1}^{n} \sum_{i=1}^{n} \left( \ln K_{\omega_L}(z_{ui},u_{ij},v_{ij},t_j) \right) - \ln K_{\omega_L}(z_{ui},u_{ij},v_{ij},t_j) \right)$$

$$+ \frac{1}{2} \sum_{i=1}^{n} \sum_{i=1}^{n} \left( \frac{2}{\tau_{L0}} \sum_{j=1}^{n} e^{eta_{\omega_{ij}}(u_{ij},v_{ij},t_j)} \right) - \ln \left( 1 + 2\tau_{\omega_L} \sum_{j=1}^{n} e^{\beta_{\omega_{ij}}(u_{ij},v_{ij},t_j)} \right)$$

$$+ \frac{1}{2} \sum_{i=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{n} y_{ij} \left( x_{i,j} \beta_{\omega_{ij}}(u_{ij},v_{ij},t_j) - \beta_{\omega_{ij}}(u_{ij},v_{ij},t_j) \right)$$

Based on steps carried out by Magdalena et al. [10], it was found that test statistic $G_L^2$ approximates the $\chi^2$ distribution with $v$ degrees of freedom where $v$ is the number of parameters under population minus the number of parameters below $H_0$. The $H_0$ rejected criterion is $G_L^2 > \chi^2_{\alpha,v}$. 


Suppose the results of hypothesis testing decide to reject $H_0$. In that case, the conclusion obtained is that the predictor variables affect the response variable and a degree of error of 5%.

If the simultaneous testing decision is rejected $H_0$, the next step is to partially test the parameters to determine which parameters significantly affect the model. The test statistic $Z$ approximates the Normal Standard distribution [11]. The hypothesis used is: the null hypothesis is $\beta_{jkl}(u_i,v_i,t_i) = 0$ vs $H_1: \beta_{jkl}(u_i,v_i,t_i) \neq 0$ with $j = 1, 2, L = 1, 2, ..., L^*$, and $k = 1, 2, ..., p$.

The test statistics used is

$$Z = \frac{\hat{\beta}_{jkl}}{se(\hat{\beta}_{jkl})}$$

$$se(\hat{\beta}_{jkl}(u_i,v_i,t_i)) = \sqrt{\hat{\var}{\hat{\beta}_{jkl}(u_i,v_i,t_i)}}$$

where the $\hat{\var}{\hat{\beta}_{jkl}(u_i,v_i,t_i)}$ value is obtained from the principal diagonal element of the covariance variant matrix of the following equation.

$$\hat{\var}{\hat{\beta}_l} = -\mathbf{H}^{-1}(\hat{\beta}_l)$$

The area of rejection $H_0$ is $|Z| > Z_\alpha$ with $\alpha$ is the level of significance.

Furthermore, to get partial hypothesis testing on the following $\tau_L$ parameters, the hypothesis used is:

$H_0: \tau_L = 0$

$H_1: \tau_L \neq 0$

The test statistical used is:

$$Z = \frac{\hat{\tau}_L}{se(\hat{\tau}_L)}$$

$$se(\hat{\tau}_L) = \sqrt{\hat{\var}{\hat{\tau}_L}}$$

where the $\hat{\var}{\hat{\tau}_L}$ value is obtained from the principal diagonal element of the covariance variant matrix of the following equation.

The area of rejection is the level of significance.

$$\hat{\var}{\hat{\tau}_L} = -\mathbf{H}^{-1}(\hat{\tau}_l)$$

The area of rejection $H_0$ is $|Z| > Z_\alpha$ with $\alpha$ is the level of significance.

4. Conclusion

We developed a global regression to accommodate spatial and temporal aspects, called a geographically and temporally weighted bivariate Poisson inverse Gaussian regression (GTWBPIGR). We have shown the step-by-step procedure to estimate the parameters and test statistics for the GTWBPIGR model hypothesis testing. Parameter estimation obtains by using the Maximum Likelihood Estimation (MLE) method. The equation that gets is not closed-form; it continues with the Newton-Raphson algorithm. Meanwhile, hypothesis testing of GTWBPIGR model by using Maximum Likelihood Ratio Test (MLRT) method.

References

[1] A. C. Cameron and P. K. Trivedi 1998 Regression Analysis of Count Data, USA: Cambridge University Press.

[2] D Karlis, & I. Ntzoufras 2005 Bivariate Poisson and diagonal inflated bivariate Poisson regression models in R Journal of statistical software 14 pp. 1-36

[3] J M Hilbe 2014 Modelling count data Cambridge University Press New York

[4] M S Holla 1966 On a Poisson inverse Gaussian distribution Metrik 11 pp 115-121

[5] Huang, B, Wu, B, & Barry, M 2010 Geographically and temporally weighted regression for modelling Spatio-temporal variation in house prices International Journal of Geographical Information Science 24 (3) 383-401
[6] G E Willmot 1987 The Poisson-Inverse Gaussian Distribution as An Alternative to The Negative Binomial,” Scandinavian Actuarial Journal 34 pp. 113-127
[7] Fotheringham A S, Brunsdon C, dan Charlton, M 2002 Geographically Weighted Regression: The Analysis of Spatially Varying Relationship Willey Chichester.
[8] Fotheringham, A S, Crespo, R, & Yao, J 2015 Geographical and temporal weighted regression (GTWR) Geographical Analysis 47 (4) 431-452
[9] N R Draper and R Smith, 1992 Applied Regression Analysis 2nd Editions New York: John Wiley and Sons Inc
[10] Triyanto, Purnadi, Otok, B W, & Purnami, S W 2015 Parameter estimation of geographically weighted multivariate Poisson regression. Appl. Math. Sci 9 4081-4093

Acknowledgements
Authors wishing to acknowledge assistance or encouragement from lecturers and college friends at the Sepuluh Nopember Institute of Technology (Institut Teknologi Sepuluh Nopember) about this research.