RESEARCH ARTICLE • OPEN ACCESS

Parameter Estimation of Mixed Geographically Weighted Bivariate Zero-Inflated Negative Binomial Regression Model

Mawadah Putri Islamiati, Purhadi, and Wibawati

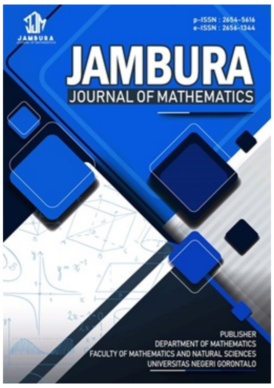
JAMBURA JOURNAL OF MATHEMATICS

Volume 7, Issue 2, Pages 112–119, August 2025

Received 16 May 2025, Revised 1 July 2025, Accepted 3 July 2025, Published 7 July 2025

To Cite this Article : M. P. Islamiati, P. Purhadi, and W. Wibawati,“Parameter Estimation of Mixed Geographically Weighted Bivariate Zero-Inflated Negative Binomial Regression Model”, *Jambura J. Math*, vol. 7, no. 2, pp. 112–119, 2025, *https://doi.org/10.37905/jjom.v7i2.32711*

JOURNAL INFO • JAMBURA JOURNAL OF MATHEMATICS



Homepage	:	http://ejurnal.ung.ac.id/index.php/jjom/index
Journal Abbreviation	:	Jambura J. Math.
Frequency	:	Biannual (February and August)
Publication Language	:	English (preferable), Indonesia
DOI	:	https://doi.org/10.37905/jjom
Online ISSN	:	2656-1344
Editor-in-Chief	:	Hasan S. Panigoro
Publisher	:	Department of Mathematics, Universitas Negeri Gorontalo
Country	:	Indonesia
OAI Address	:	http://ejurnal.ung.ac.id/index.php/jjom/oai
Google Scholar ID	:	iWLjgaUAAAAJ
Email	:	info.jjom@ung.ac.id

Jambura Journal of Mathematics, Volume 7, Issue 2, Pages 112–119, August 2025 https://doi.org/10.37905/jjom.v7i2.32711

Research Article

Parameter Estimation of Mixed Geographically Weighted Bivariate Zero-Inflated Negative Binomial Regression Model

Mawadah Putri Islamiati[1,*], **Purhadi**[1], **Wibawati**[1]

[1]*Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia*

ARTICLE HISTORY
Received 16 May 2025
Revised 1 July 2025
Accepted 3 July 2025
Published 7 July 2025

KEYWORDS
Bivariate Zero-Inflated Negative Binomial
Mixed Geographically Weighted Regression
Spatial Heterogeneity
Maximum Likelihood Estimation
BHHH Algorithm
Excess Zeros

ABSTRACT. *The Bivariate Zero-Inflated Negative Binomial (BZINBR) regression model is commonly used to analyze two correlated count response variables characterized by overdispersion and excess zeros. To account for spatial heterogeneity in predictor effects, the BZINBR model has been extended into the Geographically Weighted BZINBR (GWBZINBR) model. However, predictor effects are not always entirely local; certain global effects may persist across regions. This study proposes the Mixed Geographically Weighted BZINBR (MGWBZINBR) model, which integrates both global and local parameter structures for modeling spatially correlated bivariate count data. The theoretical framework of the MGWBZINBR model is developed, including the derivation of the log-likelihood function, parameter estimation procedures, and hypothesis testing. Parameter estimation is conducted using the Maximum Likelihood Estimation (MLE) method via the iterative Berndt–Hall–Hall–Hausman (BHHH) algorithm. Given the complexity of the likelihood equations and the absence of closed-form solutions, numerical optimization is employed to ensure convergence and stability. The MGWBZINBR model offers a flexible and robust framework for analyzing spatial count data with excess zeros and complex dependency structures. It can be applied in various fields, including public health, ecology, and transportation analysis, to understand the influence of both local and global predictors on spatial phenomena. As the focus of this paper is methodological, empirical and simulation-based applications are intentionally excluded.*

1. Introduction

Poisson regression is widely applied in the analysis of count data, particularly in fields such as public health, demography, and environmental research. This regression method relies on the assumption of equidispersion, meaning that the variance of the response variable equals its mean [1]. However, this assumption is frequently violated in practice due to overdispersion and the presence of a substantial number of excess zeros, rendering the standard Poisson regression model inadequate [2]. To address these limitations, several alternative models have been proposed.

One such model is the Zero-Inflated Poisson (ZIP), designed to handle datasets with a high proportion of zeros. It combines a Poisson distribution for non-zero counts with a separate process modeling the probability of structural zeros [3]. Another model, the Negative Binomial, extends Poisson regression by introducing a dispersion parameter, providing greater flexibility in handling overdispersion [4]. A combination of these two models leads to the Zero-Inflated Negative Binomial (ZINB), which effectively addresses both overdispersion and excess zeros [5].

The effectiveness of these models has been confirmed in prior studies. For instance, Saputro and Qudratullah [6] applied the Maximum Likelihood Estimation (MLE) method and the Expectation Maximization (EM) algorithm to model ZINBR. In another study, Azwarini [7] estimated parameters and performed hypothesis testing in the BZINBR model, showing that all predictors significantly influenced the outcomes. To incorporate spatial heterogeneity in predictor effects, the Geographically Weighted Bivariate Zero-Inflated Negative Binomial Regression (GWBZINBR) model was introduced, allowing parameter estimates to vary locally across spatial units [8]. However, in many cases, not all predictor effects are strictly local. Some predictors may exhibit consistent global effects.

Enforcing complete locality in all predictor effects may reduce model efficiency and interpretability [9]. To overcome this limitation, the Mixed Geographically Weighted Regression (MGWR) model was proposed, combining both global and local coefficients within a single framework [10]. This mixed approach has also been extended to count models with overdispersion and excess zeros, such as the Geographically Weighted Negative Binomial Regression (GWNBR) [11] and Geographically Weighted Bivariate Zero-Inflated Poisson Inverse Gaussian Regression (GWBZIPIGR) [12].

This study aims to develop the theoretical foundation of the Mixed Geographically Weighted Bivariate Zero-Inflated Negative Binomial Regression (MGWBZINBR) model. It includes the derivation of its log-likelihood function, parameter estimation using the MLE method, and hypothesis testing for both local and global components. Due to the absence of closed-form solutions, parameter estimation is implemented using the iterative Berndt–Hall–Hall–Hausman (BHHH) algorithm, known for its computa-

*Corresponding Author.

Email : *mawadahputrii@gmail.com* (M. P. Islamiati)
Homepage : http://ejurnal.ung.ac.id/index.php/jjom/index / E-ISSN : 2656-1344

tional efficiency and numerical stability in complex spatial data analysis.

This paper is presented as a methodological contribution, emphasizing the mathematical formulation, estimation procedures, and structural properties of the MGWBZINBR model. Empirical validation and simulation studies are acknowledged as valuable future directions but are deliberately excluded from the current study.

2. Model

2.1. Bivariate Zero-Inflated Negative Binomial Regression

The Bivariate Zero-Inflated Negative Binomial (BZINB) distribution is an extension of the ZINB distribution for two dependent random variables. BZINB is a mixed distribution combining Zero-Inflated and Negative Binomial distributions, capable of addressing overdispersion when there are many zero values in observed bivariate data [13].

Based on the response variable, BZINB distributions can be classified into two main types, namely Type I and Type II. The Type I BZINB model has a response variable that consists of a single set of values $(Y_1 = 0, Y_2 = 0)$ dan $(Y_1 \neq 0, Y_2 \neq 0)$, while the Type II BZINB regression model has a response variable comprising multiple combinations of values [14].

Assume that there are correlated random variables Y_1 and Y_2, which follow a bivariate zero-inflated negative binomial distribution. For each observation $i = 1, 2, \cdots, n$ the pair:

$$(Y_1, Y_2) \sim BZINB\left(\lambda_1, \lambda_2, p_1, p_2, \tau\right) \tag{1}$$

where:

λ_{ki} : mean of the negative binomial distribution for response k,

p_{ki} : zero-inflation probability for response k,

τ : dispersion parameter.

Each observation $i$is associated with a covariate vector $\boldsymbol{x_i} = \begin{bmatrix} 1 & x_{1i} & x_{2i} & \cdots & x_{hi} \end{bmatrix}^T_{1\times(1+q)}$. Where h is the number of predictor variables, and the first element represents the intercept term. For each response variable $k = 1, 2$ the regression coefficients are denoted by β_k for the count component and δ_k for the zero-inflation component. Let q_i denote the exposure variable for observation i.

The BZINB regression model is divided into the Bivariate Negative Binomial model and the Zero-Inflated model, which are formulated as follows:

- Bivariate Negative Binomial Model (λ_{ki}) :

$$\lambda_{ki} = q_{ki}\text{exp}\left(\boldsymbol{x}_i^T{}_k\right);\ k = 1, 2. \tag{2}$$

- Zero-Inflated Model :

$$\begin{aligned} \log\ it\left(p_{ki}\right) &= \ln\left(\frac{p_{ki}}{1 - p_{ki}}\right) = \boldsymbol{x}_i^T d_k\ ;\ k{=}\ 1, 2, \\ p_{ki} &= \frac{exp\left(\boldsymbol{x}_i^T d_k\right)}{1 + exp\left(\boldsymbol{x}_i^T d_k\right),} \quad \text{and} \\ 1{-}p_{ki} &= \frac{1}{1 + exp\left(\boldsymbol{x}_i^T d_k\right)}. \end{aligned} \tag{3}$$

Parameter estimation for the BZINBR model is obtained using the Maximum Likelihood Estimation (MLE) approach.

2.2. Geographically Weighted Bivariate Zero-Inflated Negative Binomial Regression

The Geographically Weighted Bivariate Zero-Inflated Negative Binomial Regression (GWBZINBR) model is an extension of the BZINBR model that incorporates spatial or geographical characteristics to produce locally varying parameter estimates. This enhancement is achieved by assigning each observation a location coordinate, represented as a spatial weight matrix (W_{ii*}), based on latitude and longitude values.

The parameter estimation for the GWBZINBR model is conducted using the Maximum Likelihood Estimation (MLE) method, implemented via the Berndt–Hall–Hall–Hausman (BHHH) iterative algorithm [8]. Suppose we have a vector of observations

$$\begin{aligned} &(Y_{1i}, Y_{2i}) \sim BZINB\left({}_1\left(\boldsymbol{u}_i\right), {}_2\left(\boldsymbol{u}_i\right), d_1\left(\boldsymbol{u}_i\right), d_2\left(\boldsymbol{u}_i\right), t\right) \\ &\quad \text{for}\ \ i = 1, 2, \cdots, n. \end{aligned} \tag{4}$$

Then the GWBZINBR model is:

$$\begin{aligned} \lambda_{ki} &= q_{ki}\text{exp}\left(\boldsymbol{x}_i^T{}_k\left(\boldsymbol{u}_i\right)\right);\ k = 1, 2;\ i = 1, 2, \cdots, n, \\ p_{ki} &= \frac{\text{exp}\left(\boldsymbol{x}_i^T d_k\left(\boldsymbol{u}_i\right)\right)}{1 + \text{exp}\left(\boldsymbol{x}_i^T d_k\left(\boldsymbol{u}_i\right)\right)}, \\ 1 - p_{ki} &= \frac{1}{1 + \text{exp}\left(\boldsymbol{x}_i^T d_k\left(\boldsymbol{u}_i\right)\right)}. \end{aligned} \tag{5}$$

The local log-likelihood function is then optimized using numerical routines, and spatial weights (W_{ii*}) are used to ensure local estimation at each location (u_i, v_i).

2.3. Mixed Weighted Bivariate Zero-Inflated Negative Binomial Regression (MGWBZINBR)

The MGWBZINBR framework is an advancement of the GW-BZINBR model, achieved by incorporating spatial heterogeneity and employing a mixed coefficient approach. This modeling approach integrates local and global parameter estimation to capture spatially heterogeneous relationships in bivariate count data, as outlined by [15]. It is particularly suitable for analyzing bivariate count data characterized by overdispersion, excess zeros, and spatially varying associations between covariates and outcomes. The typical representation of the Mixed Bivariate Geographically Weighted Regression framework which incorporates an exposure variable (q_i) is employed to model the relationship among covariates and response variables, is given as follows:

$$\begin{aligned} E(Y_{ki}) = \mu_{ki} = q_i\, \text{exp}\big(\mathbf{x}_i^{*T}\boldsymbol{\beta}_k^*(\mathbf{u}_i) + \mathbf{x}_i^{**T}\boldsymbol{\gamma}_k\big), \\ k = 1, 2;\ i = 1, 2, \ldots, n. \end{aligned} \tag{6}$$

For local predictors, the coefficient and covariate matrices for response $k = 1, 2$ are:

$$\begin{aligned} \boldsymbol{\beta}_k^*(\mathbf{u}_i) &= \left[\beta_{0,k}^*(\mathbf{u}_i),\ \beta_{1,k}^*(\mathbf{u}_i),\ \ldots,\ \beta_{q,k}^*(\mathbf{u}_i)\right]^T, \\ \mathbf{x}_i^* &= \left[1,\ X_{1i},\ X_{2i},\ \ldots,\ X_{qi}\right]^T. \end{aligned} \tag{7}$$

For global predictors:

$$\begin{aligned} \boldsymbol{\gamma}_k &= \left[\gamma_{(q+1),k},\ \gamma_{(q+2),k},\ \ldots,\ \gamma_{p,k}\right]^T, \\ \mathbf{x}_i^{**} &= \left[X_{(q+1),i},\ X_{(q+2),i},\ \ldots,\ X_{p,i}\right]^T. \end{aligned} \tag{8}$$

Assume that (Y_{1i}, Y_{2i}) follows a BZINB distribution parameterized by $\boldsymbol{\beta_1}\left(\boldsymbol{u}_i\right), \boldsymbol{\beta_2}\left(\boldsymbol{u}_i\right), \boldsymbol{\delta_1}\left(\boldsymbol{u}_i\right), \boldsymbol{\delta_2}\left(\boldsymbol{u}_i\right), \tau$ Under this assumption, the associated joint probability function can be written as:

- Bivariate Negative Binomial Model (λ_{ki}) :

$$\lambda_{ki} = q_{ki} \exp\left(\mathbf{x}_i^T \boldsymbol{\beta}_k(\mathbf{u}_i) + \mathbf{x}_i^{**T} \gamma_k\right); \quad k = 1, 2;\ i = 1, 2, \ldots, n. \tag{9}$$

- Zero-Inflated Model :

$$\begin{aligned} \log it(p_{ki}) &= \ln\left(\frac{p_{ki}}{1 - p_{ki}}\right), \\ p_{ki} &= \frac{\exp\left(\mathbf{x}_i^T \boldsymbol{\delta}_k(\mathbf{u}_i) + \mathbf{x}_i^{**T} \gamma_k\right)}{1 + \exp\left(\mathbf{x}_i^T \boldsymbol{\delta}_k(\mathbf{u}_i) + \mathbf{x}_i^{**T} \gamma_k\right)}, \\ 1 - p_{ki} &= \frac{1}{1 + \exp\left(\mathbf{x}_i^T \boldsymbol{\delta}_k(\mathbf{u}_i) + \mathbf{x}_i^{**T} \gamma_k\right)}; \\ & k = 1, 2;\ i = 1, 2, \ldots, n. \end{aligned} \tag{10}$$

The corresponding joint probability distribution can be expressed as:

$$P\left(Y_{1i} = y_{1i}, Y_{2i} = y_{2i}\right) = \begin{cases} A_1, (y_{1i} = 0, y_{2i} = 0) \\ B_1, (y_{1i} = 0, y_{2i} > 0) \\ C_1, (y_{1i} > 0, y_{2i} = 0) \\ D_1, (y_{1i} > 0, y_{2i} > 0) \end{cases}. \tag{11}$$

(i) For $Y_{1i} = 0, Y_{2i} = 0$,

$$\begin{aligned} P\left(Y_{1i} = 0, Y_{2i} = 0\right) &= p_{1i} \cdot p_{2i} + p_{1i}\left(1 - p_{2i}\right)\left(\frac{1}{1 + \tau\lambda_{2i}}\right)^{\frac{1}{\tau}} \\ &+ p_{2i}\left(1 - p_{1i}\right)\left(\frac{1}{1 + \tau\lambda_{1i}}\right)^{\frac{1}{\tau}} \\ &+ \left(1 - p_{1i}\right)\left(1 - p_{2i}\right)\left(\frac{1}{1 + \tau\left(\lambda_{1i} + \lambda_{2i}\right)}\right)^{\frac{1}{\tau}}. \end{aligned} \tag{12}$$

(ii) For $Y_{1i} > 0, Y_{2i} = 0$ or $Y_{1i} = 1, 2, \cdots; Y_{2i} = 0$,

$$\begin{aligned} P(Y_{1i} > 0, Y_{2i} = 0) &= p_{2i}(1 - p_{1i})\frac{\Gamma\left(y_{1i} + \frac{1}{\tau}\right)}{\Gamma\left(\frac{1}{\tau}\right) y_{1i}!}\left(\frac{1}{1 + \tau\lambda_{1i}}\right)^{\frac{1}{\tau}} \\ &\left(\frac{\tau\lambda_{1i}}{1 + \tau\lambda_{1i}}\right)^{y_{1i}} + (1 - p_{1i})(1 - p_{2i})\frac{\Gamma\left(y_{1i} + \frac{1}{\tau}\right)}{\Gamma\left(\frac{1}{\tau}\right) y_{1i}!} \\ &\times \left(\frac{1}{1 + \tau(\lambda_{1i} + \lambda_{2i})}\right)^{\frac{1}{\tau}}\left(\frac{\tau(\lambda_{1i} + \lambda_{2i})}{1 + \tau(\lambda_{1i} + \lambda_{2i})}\right)^{y_{1i}}. \end{aligned} \tag{13}$$

(iii) For $Y_{1i} = 1, 2, \cdots; Y_{2i} = 0$ or $Y_{1i} > 0, Y_{2i} = 0$,

$$\begin{aligned} P(Y_{1i} = 0, Y_{2i} > 0) &= p_{1i}(1 - p_{2i})\frac{\Gamma\left(y_{2i} + \frac{1}{\tau}\right)}{\Gamma\left(\frac{1}{\tau}\right) y_{2i}!}\left(\frac{1}{1 + \tau\lambda_{2i}}\right)^{\frac{1}{\tau}} \\ &\left(\frac{\tau\lambda_{2i}}{1 + \tau\lambda_{2i}}\right)^{y_{2i}} + (1 - p_{1i})(1 - p_{2i})\frac{\Gamma\left(y_{2i} + \frac{1}{\tau}\right)}{\Gamma\left(\frac{1}{\tau}\right) y_{2i}!} \\ &\times \left(\frac{1}{1 + \tau(\lambda_{1i} + \lambda_{2i})}\right)^{\frac{1}{\tau}}\left(\frac{\tau(\lambda_{1i} + \lambda_{2i})}{1 + \tau(\lambda_{1i} + \lambda_{2i})}\right)^{y_{2i}}. \end{aligned} \tag{14}$$

(iv) For $Y_{1i} > 0, Y_{2i} > 0$ or $Y_{1i} = 1, 2, \cdots; Y_{2i} = 1, 2, \cdots$,

$$\begin{aligned} P(Y_{1i} > 0, Y_{2i} > 0) &= (1 - p_{1i})(1 - p_{2i})\frac{\Gamma\left(y_{1i} + y_{2i} + \frac{1}{\tau}\right)}{\Gamma\left(\frac{1}{\tau}\right) y_{1i}!\, y_{2i}!} \\ &\times \left(\frac{1}{1 + \tau(\lambda_{1i} + \lambda_{2i})}\right)^{\frac{1}{\tau}}\left(\frac{\tau(\lambda_{1i} + \lambda_{2i})}{1 + \tau(\lambda_{1i} + \lambda_{2i})}\right)^{y_{1i} + y_{2i}}. \end{aligned} \tag{15}$$

3. Results and Discussion

The parameter estimation of the MGWBZINBR model produces both global and local parameter estimates. This estimation is carried out using MLE method. Let

$$\left(Y_{1i}, Y_{2i}\right) \sim BZINB\left(\boldsymbol{\beta}_1\left(\boldsymbol{u}_i\right), \boldsymbol{\beta}_2\left(\boldsymbol{u}_i\right), \boldsymbol{\delta}_1\left(\boldsymbol{u}_i\right), \boldsymbol{\delta}_2\left(\boldsymbol{u}_i\right), \tau\right); \quad i = 1, 2, \cdots, n.$$

Then the MGWBZINBR model is defined as follows:

- Model for the Bivariate Negative Binomial Component:

$$\lambda_{ki} = q_{ki} \exp\left(\mathbf{x}_i^T \boldsymbol{\beta}_k(\mathbf{u}_i) + \mathbf{x}_i^{**T} \gamma_k\right); \quad k = 1, 2;\ i = 1, 2, \ldots, n.$$

- Model for the Zero-Inflated Component:

$$\begin{aligned} \log it(p_{ki}) &= \ln\left(\frac{p_{ki}}{1 - p_{ki}}\right), \\ p_{ki} &= \frac{\exp\left(\mathbf{x}_i^T \boldsymbol{\delta}_k(\mathbf{u}_i) + \mathbf{x}_i^{**T} \gamma_k\right)}{1 + \exp\left(\mathbf{x}_i^T \boldsymbol{\delta}_k(\mathbf{u}_i) + \mathbf{x}_i^{**T} \gamma_k\right)}, \\ 1 - p_{ki} &= \frac{1}{1 + \exp\left(\mathbf{x}_i^T \boldsymbol{\delta}_k(\mathbf{u}_i) + \mathbf{x}_i^{**T} \gamma_k\right)}; \\ & k = 1, 2;\ i = 1, 2, \ldots, n. \end{aligned}$$

The MGWBZINBR model's joint probability structure can be expressed as follows:

$$P(Y_{1i} = y_{1i}, Y_{2i} = y_{2i}) = \begin{cases} A_1, & (y_{1i} = 0, y_{2i} = 0) \\ B_1, & (y_{1i} = 0, y_{2i} > 0) \\ C_1, & (y_{1i} > 0, y_{2i} = 0) \\ D_1, & (y_{1i} > 0, y_{2i} > 0) \end{cases}.$$

The forms of A_i, B_i, C_i and D_i are explained in detail in eq. (12), (13), (14), and (15). Meanwhile, the parameters to be estimated in the MGWBZINBR model can be represented in matrix form as follows:

$$\boldsymbol{\theta}_{i^*, MGWBZINBR} = \left[\boldsymbol{\beta}_1^*(\mathbf{u}_{i^*})\ \ \boldsymbol{\beta}_2^*(\mathbf{u}_{i^*})\ \ \boldsymbol{\delta}_1^*(\mathbf{u}_{i^*})\ \ \boldsymbol{\delta}_2^*(\mathbf{u}_{i^*})\ \ \gamma_1\ \ \gamma_2\ \ \tau\right]^T. \tag{16}$$

The probability function employed in this model is formulated as follows:

$$\begin{aligned} P(y_{1i}, y_{2i} \mid \boldsymbol{\theta}_{i^*, MGWBZINBR}) &= P\left(y_{1i}, y_{2i}; \boldsymbol{\beta}_1^*(\mathbf{u}_i), \boldsymbol{\beta}_2^*(\mathbf{u}_i), \boldsymbol{\delta}_1^*(\mathbf{u}_i), \right. \\ &\left. \boldsymbol{\delta}_2^*(\mathbf{u}_i), \gamma_1, \gamma_2, \tau\right) \\ &= \left((A_i)^{1 - b_i - c_i - d_i}(B_i)^{b_i}(C_i)^{c_i}(D_i)^{d_i}\right) \end{aligned}$$

The joint probability functions are transformed into the form of a natural logarithm (ln) probability function to estimate the parameters for the *i*-th observation, as formulated below:

$$\ln P\left(y_{1i}, y_{2i} \mid \boldsymbol{\theta}_{i^*, MGWBZINBR}\right) = \left(1 - b_i - c_i - d_i\right) \ln A_i$$

$$+ (b_i)\ln B_i + (c_i)\ln C_i + (d_i)\ln D_i.$$

These derivatives, with respect to each parameter, are determined under four distinct data scenarios, namely:

$(Y_{1i}=0, Y_{2i}=0)$, $(Y_{1i}>0, Y_{2i}=0)$ or $Y_{1i}=1,2,\ldots; Y_{2i}=0$, $(Y_{1i}=0, Y_{2i}>0)$ or $Y_{1i}=0; Y_{2i}=1,2,\ldots$ and $(Y_{1i}>0, Y_{2i}>0)$ or $Y_{1i}=1,2,\ldots; Y_{2i}=1,2,\ldots$.

The first-order derivative with respect to parameter $(\boldsymbol{\beta}_1^*(\boldsymbol{u}_i), \boldsymbol{\beta}_2^*(\boldsymbol{u}_i), \boldsymbol{\delta}_1^*(\boldsymbol{u}_i), \boldsymbol{\delta}_2^*(\boldsymbol{u}_i), \gamma_1, \gamma_2, \tau)$ for one of the conditions, namely the condition $(Y_{1i}=0, Y_{2i}=0)$ is given by:

$$\begin{aligned} ZL_{1i} &= \exp\left(\mathbf{x}_i^{*T}\boldsymbol{\delta}_1^*(\mathbf{u}_i^*) + \mathbf{x}_i^{**T}\gamma_1\right), \\ ZL_{2i} &= \exp\left(\mathbf{x}_i^{*T}\boldsymbol{\delta}_2^*(\mathbf{u}_i^*) + \mathbf{x}_i^{**T}\gamma_2\right). \end{aligned}$$

- The derivative with respect to $(\boldsymbol{\beta}_1^*(\boldsymbol{u}_i))$ under the condition $(Y_{1i}=0, Y_{2i}=0)$:

$$\frac{\partial \ln P(\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \boldsymbol{\beta_1}^*(\mathbf{u}_{i_*})} = \frac{N_{\boldsymbol{P_1\beta_1}}}{D_{\boldsymbol{P_1\beta_1}}},$$

where,

$$\begin{aligned} N_{\boldsymbol{P_1\beta_1}} &= \mathbf{x}_{i_*}^T \cdot \tau^2 \cdot ZI_{2i} \cdot \lambda_{1i}(1+\tau\lambda_{1i})^{\tau-1} \\ &\quad + \mathbf{x}_{i_*}^T \cdot \tau^2 \cdot \lambda_{1i}\left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau-1}, \\ D_{\boldsymbol{P_1\beta_1}} &= ZI_{1i}\cdot ZI_{2i} + ZI_{1i}(1+\tau\lambda_{2i})^{\tau} + ZI_{2i}(1+\tau\lambda_{1i})^{\tau} \\ &\quad + \left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau}. \end{aligned}$$

- The derivative with respect to $(\boldsymbol{\beta}_2^*(\boldsymbol{u}_i))$ under the condition $(Y_{1i}=0, Y_{2i}=0)$:

$$\frac{\partial \ln P(\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \boldsymbol{\beta}_2^*(\boldsymbol{u}_{i_*})} = \frac{N_{\boldsymbol{P_1\beta_2}}}{D_{\boldsymbol{P_1\beta_2}}},$$

where,

$$\begin{aligned} N_{\boldsymbol{P_1\beta_2}} &= \boldsymbol{x}_{i_*}^T \cdot \tau^2 \cdot ZI_{1i} \cdot \lambda_{2i}(1+\tau\lambda_{2i})^{\tau-1} \\ &\quad + \boldsymbol{x}_{i_*}^T \cdot \tau^2 \cdot \lambda_{2i}\left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau-1}, \\ D_{\boldsymbol{P_1\beta_2}} &= ZI_{1i}\cdot ZI_{2i} + ZI_{1i}(1+\tau\lambda_{2i})^{\tau} + ZI_{2i}(1+\tau\lambda_{1i})^{\tau} \\ &\quad + \left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau}. \end{aligned}$$

- The derivative with respect to $(\boldsymbol{\delta}_1^*(\boldsymbol{u}_i))$ under the condition $(Y_{1i}=0, Y_{2i}=0)$:

$$\frac{\partial \ln P(\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \boldsymbol{\delta}_1^*(\boldsymbol{u}_{i_*})} = \frac{N_{\boldsymbol{P_1\delta_1}}}{D_{\boldsymbol{P_1\delta_1}}},$$

where,

$$\begin{aligned} N_{\boldsymbol{P_1\delta_1}} &= -\boldsymbol{x}_{i_*}^T \cdot ZI_{1i}\left(1+\tau(ZI_{2i})\right) + \boldsymbol{x}_{i_*}^T \cdot ZI_{1i}\cdot ZI_{2i} \\ &\quad + \boldsymbol{x}_{i_*}^T \cdot ZI_{1i}\left(1+\tau\lambda_{2i}\right)^{\tau}, \\ D_{\boldsymbol{P_1\delta_1}} &= \left(1+\tau(ZI_{1i})\right)\left(1+\tau(ZI_{2i})\right) + ZI_{1i}\cdot ZI_{2i} \\ &\quad + ZI_{1i}\left(1+\tau\lambda_{2i}\right)^{\tau} + ZI_{2i}\left(1+\tau\lambda_{1i}\right)^{\tau} \\ &\quad + \left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau}. \end{aligned}$$

- The derivative with respect to $(\boldsymbol{\delta}_2^*(\boldsymbol{u}_i))$ under the condition $(Y_{1i}=0, Y_{2i}=0)$:

$$\frac{\partial \ln P(\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \boldsymbol{\delta}_2^*(\boldsymbol{u}_{i_*})} = \frac{N_{\boldsymbol{P_1\delta_2}}}{D_{\boldsymbol{P_1\delta_2}}},$$

where,

$$\begin{aligned} N_{\boldsymbol{P_1\delta_2}} &= -\boldsymbol{x}_{i_*}^T \cdot ZI_{2i}\left(1+\tau(ZI_{1i})\right) + \boldsymbol{x}_{i_*}^T \cdot ZI_{1i}\cdot ZI_{2i} \\ &\quad + \boldsymbol{x}_{i_*}^T \cdot ZI_{2i}\left(1+\tau\lambda_{2i}\right)^{\tau}, \\ D_{\boldsymbol{P_1\delta_2}} &= \left(1+\tau(ZI_{1i})\right)\left(1+\tau(ZI_{2i})\right) + ZI_{1i}\cdot ZI_{2i} \\ &\quad + ZI_{1i}\left(1+\tau\lambda_{2i}\right)^{\tau} + ZI_{2i}\left(1+\tau\lambda_{1i}\right)^{\tau} \\ &\quad + \left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau}. \end{aligned}$$

- The derivative with respect to (γ_1) under the condition $(Y_{1i}=0, Y_{2i}=0)$:

$$\frac{\partial \ln P(\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \gamma_1} = \frac{N_{\boldsymbol{P_1\gamma_1}}}{D_{\boldsymbol{P_1\gamma_1}}},$$

where,

$$\begin{aligned} N_{\boldsymbol{P_1\gamma_1}} &= -\boldsymbol{x}_{i_*}^T \cdot ZI_{1i}\left(1+\tau(ZI_{2i})\right) + \boldsymbol{x}_{i_*}^T \cdot ZI_{1i}\cdot ZI_{2i} \\ &\quad + \boldsymbol{x}_{i_*}^T \cdot ZI_{1i}\left(1+\tau\lambda_{2i}\right)^{\tau} + \boldsymbol{x}_{i_*}^T \cdot ZI_{2i}\left(1+\tau\lambda_{1i}\right)^{\tau} \\ &\quad + \boldsymbol{x}_{i_*}^T\left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau}, \\ D_{\boldsymbol{P_1\gamma_1}} &= \left(1+\tau(ZI_{1i})\right)\left(1+\tau(ZI_{2i})\right) + ZI_{1i}\cdot ZI_{2i} \\ &\quad + ZI_{1i}\left(1+\tau\lambda_{2i}\right)^{\tau} + ZI_{2i}\left(1+\tau\lambda_{1i}\right)^{\tau} \\ &\quad + \left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau}. \end{aligned}$$

- The derivative with respect to (γ_2) under the condition $(Y_{1i}=0, Y_{2i}=0)$:

$$\frac{\partial \ln P(\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \gamma_2} = \frac{N_{\boldsymbol{P_1\gamma_2}}}{D_{\boldsymbol{P_1\gamma_2}}},$$

where,

$$\begin{aligned} N_{\boldsymbol{P_1\gamma_2}} &= -\boldsymbol{x}_{i_*}^T \cdot ZI_{2i}\left(1+\tau(ZI_{1i})\right) + \boldsymbol{x}_{i_*}^T \cdot ZI_{1i}\cdot ZI_{2i} \\ &\quad + \boldsymbol{x}_{i_*}^T \cdot \tau^2 \cdot ZI_{1i}\cdot\lambda_{2i}\left(1+\tau\lambda_{2i}\right)^{\tau-1} \\ &\quad + \boldsymbol{x}_{i_*}^T \cdot \tau^2 \cdot \lambda_{2i}\left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau-1}, \\ D_{\boldsymbol{P_1\gamma_2}} &= \left(1+\tau(ZI_{1i})\right)\left(1+\tau(ZI_{2i})\right) + ZI_{1i}\cdot ZI_{2i} \\ &\quad + ZI_{1i}\left(1+\tau\lambda_{2i}\right)^{\tau} + \left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau}. \end{aligned}$$

In the MGWBZINBR model, the dispersion parameter t is set based on the estimate derived from the BZINBR model. To efficiently estimate the population parameters, this study applies a likelihood-based approach that seeks to optimize the probability of observing the given data under a known distribution:

$$\begin{aligned} L(\boldsymbol{\theta}_{i_*,MGWBZINBR}) &= \prod_{i=1}^{n} P\left(y_{1i}, y_{2i} \,|\, \boldsymbol{\theta}_{i_*,MGWBZINBR}\right) \\ &= \prod_{i=1}^{n} P\left(y_{1i}, y_{2i}; \boldsymbol{\beta}_1^*(\boldsymbol{u}_i), \boldsymbol{\beta}_2^*(\boldsymbol{u}_i), \boldsymbol{\delta}_1^*(\boldsymbol{u}_i),\right. \\ &\qquad \left.\boldsymbol{\delta}_2^*(\boldsymbol{u}_i), \gamma_1, \gamma_2, \tau\right) \\ &= \prod_{i=1}^{n} \left((A_i)^{1-b_i-c_i-d_i}(B_i)^{b_i}(C_i)^{c_i}(D_i)^{d_i}\right)^{w_{i_*}} \end{aligned}$$

$$= \left(L_1\left(\boldsymbol{\theta}_{MGWBZINBR}\right)\right)\left(L_2\left(\boldsymbol{\theta}_{MGWBZINBR}\right)\right) \left(L_3\left(\boldsymbol{\theta}_{MGWBZINBR}\right)\right)\left(L_4\left(\boldsymbol{\theta}_{MGWBZINBR}\right)\right).$$

To estimate the likelihood for the *i*-th observation, the function is first transformed into its natural logarithmic (ln) form, considering the spatial weight w_{ii*} as follows:

$$\begin{aligned}\ln L\left(\boldsymbol{\theta}_{i_*,MGWBZINBR}\right) &= \sum_{i=1}^{n} w_{ii_*} \ln\left(P\left(y_{1i}, y_{2i}|\boldsymbol{\theta}_{i_*,MGWBZINBR}\right)\right)\\ &= \sum_{i=1}^{n} w_{ii_*}(1-b_i-c_i-d_i)\ln A_i + \sum_{i=1}^{n} w_{ii_*}(b_i)\ln B_i\\ &\quad + \sum_{i=1}^{n} w_{ii_*}(c_i)\ln C_i + \sum_{i=1}^{n} w_{ii_*}(d_i)\ln D_i\\ &= \ln\left(L_1\left(\boldsymbol{\theta}_{MGWBZINBR}\right)\right) + \ln\left(L_2\left(\boldsymbol{\theta}_{MGWBZINBR}\right)\right)\\ &\quad + \ln\left(L_3\left(\boldsymbol{\theta}_{MGWBZINBR}\right)\right) + \ln\left(L_4\left(\boldsymbol{\theta}_{MGWBZINBR}\right)\right).\end{aligned}$$

Subsequently, the partial derivatives of the log-likelihood function are computed with respect to each parameter:

$$\ln L\left(\boldsymbol{\theta}_{MGWBZINBR}\right) = \ln L\left(\boldsymbol{\beta}_1^*(\boldsymbol{u}_i), \boldsymbol{\beta}_2^*(\boldsymbol{u}_i), \boldsymbol{\delta}_1^*(\boldsymbol{u}_i), \boldsymbol{\delta}_2^*(\boldsymbol{u}_i), \gamma_1, \gamma_2, \tau\right).$$

with respect to each parameter and then equate them to zero in order to obtain the parameter estimates:

$$\begin{aligned}\frac{\partial \ln L\left(\boldsymbol{\theta}_{i_*,MGWBZINBR}\right)}{\partial \boldsymbol{\theta}_{i_*,MGWBZINBR}} &= \frac{\partial \ln\left(L_1\left(\boldsymbol{\theta}_{i_*,MGWBZINBR}\right)\right)}{\partial \boldsymbol{\theta}_{i_*,MGWBZINBR}} + \frac{\partial \ln\left(L_2\left(\boldsymbol{\theta}_{i_*,MGWBZINBR}\right)\right)}{\partial \boldsymbol{\theta}_{i_*,MGWBZINBR}}\\ &\quad + \frac{\partial \ln\left(L_3\left(\boldsymbol{\theta}_{i_*,MGWBZINBR}\right)\right)}{\partial \boldsymbol{\theta}_{i_*,MGWBZINBR}} + \frac{\partial \ln\left(L_4\left(\boldsymbol{\theta}_{i_*,MGWBZINBR}\right)\right)}{\partial \boldsymbol{\theta}_{i_*,MGWBZINBR}}.\end{aligned}$$

The first-order derivatives of the log-likelihood function for the MGWBZINBR model are obtained using the following formulation, the following expression is employed:

$$\ln L\left(\boldsymbol{\theta}_{MGWBZINBR}\right) = \ln L\left(\boldsymbol{\beta}_1^*(\boldsymbol{u}_i), \boldsymbol{\beta}_2^*(\boldsymbol{u}_i), \boldsymbol{\delta}_1^*(\boldsymbol{u}_i), \boldsymbol{\delta}_2^*(\boldsymbol{u}_i), \gamma_1, \gamma_2, \tau\right).$$

For one of the conditions, namely the condition $(Y_{1i}=0, Y_{2i}=0)$ is given by:

$$\boldsymbol{g}\left(\boldsymbol{\theta}_{i_*,MGWBZINBR}\right) = \begin{bmatrix} \frac{\partial \ln L(\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \boldsymbol{\beta}_1^*(\boldsymbol{u}_{i_*})} \\ \frac{\partial \ln L(\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \boldsymbol{\beta}_2^*(\boldsymbol{u}_{i_*})} \\ \frac{\partial \ln L(\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \boldsymbol{\delta}_1^*(\boldsymbol{u}_{i_*})} \\ \frac{\partial \ln L(\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \boldsymbol{\delta}_2^*(\boldsymbol{u}_{i_*})} \\ \frac{\partial \ln L(\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \gamma_1} \\ \frac{\partial \ln L(\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \gamma_2} \\ \frac{\partial \ln L(\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \tau} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} w_{ii_*} \frac{\partial \ln P(y_{1i},y_{2i}\,|\,\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \boldsymbol{\beta}_1^*(\boldsymbol{u}_{i_*})} \\ \sum_{i=1}^{n} w_{ii_*} \frac{\partial \ln P(y_{1i},y_{2i}\,|\,\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \boldsymbol{\beta}_2^*(\boldsymbol{u}_{i_*})} \\ \sum_{i=1}^{n} w_{ii_*} \frac{\partial \ln P(y_{1i},y_{2i}\,|\,\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \boldsymbol{\delta}_1^*(\boldsymbol{u}_{i_*})} \\ \sum_{i=1}^{n} w_{ii_*} \frac{\partial \ln P(y_{1i},y_{2i}\,|\,\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \boldsymbol{\delta}_2^*(\boldsymbol{u}_{i_*})} \\ \sum_{i=1}^{n} w_{ii_*} \frac{\partial \ln P(y_{1i},y_{2i}\,|\,\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \gamma_1} \\ \sum_{i=1}^{n} w_{ii_*} \frac{\partial \ln P(y_{1i},y_{2i}\,|\,\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \gamma_2} \\ \sum_{i=1}^{n} w_{ii_*} \frac{\partial \ln P(y_{1i},y_{2i}\,|\,\boldsymbol{\theta}_{i_*,MGWBZINBR})}{\partial \tau} \end{bmatrix}. \tag{17}$$

- The derivative $L\left(\boldsymbol{\theta}_{i*,MGWBZINBR}\right)$ with respect to $\left(\boldsymbol{\beta}_1^*\left(\boldsymbol{u}_i\right)\right)$ under the condition $(Y_{1i}=0, Y_{2i}=0)$:

$$\frac{\partial L\left(\boldsymbol{\theta}_{i_*,MGWBZINBR}\right)}{\partial \boldsymbol{\beta}_2^*(\boldsymbol{u}_{i_*})} = \sum_{i=1}^{n}\left(\frac{N_{\boldsymbol{L_1\beta_1}}}{D_{\boldsymbol{L_1\beta_1}}}\right) w_{ii_*},$$

where,

$$\begin{aligned}N_{\boldsymbol{L_1\beta_1}} &= \boldsymbol{x}_{i_*}^T \cdot \tau^2 \cdot ZI_{2i} \cdot \lambda_{1i}(1+\tau\lambda_{1i})^{\tau-1}\\ &\quad + \boldsymbol{x}_{i_*}^T \cdot \tau^2 \cdot \lambda_{1i}\left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau-1},\\ D_{\boldsymbol{L_1\beta_1}} &= ZI_{1i}\cdot ZI_{2i} + ZI_{1i}(1+\tau\lambda_{2i})^{\tau} + ZI_{2i}(1+\tau\lambda_{1i})^{\tau}\\ &\quad + \left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau}.\end{aligned}$$

- The derivative $L\left(\boldsymbol{\theta}_{i*,MGWBZINBR}\right)$ with respect to $\left(\boldsymbol{\beta}_2^*\left(\boldsymbol{u}_i\right)\right)$ under the condition $(Y_{1i}=0, Y_{2i}=0)$:

$$\frac{\partial L\left(\boldsymbol{\theta}_{i_*,MGWBZINBR}\right)}{\partial \boldsymbol{\beta}_2^*(\boldsymbol{u}_{i_*})} = \sum_{i=1}^{n}\left(\frac{N_{\boldsymbol{L_1\beta_2}}}{D_{\boldsymbol{L_1\beta_2}}}\right) w_{ii_*},$$

where,

$$\begin{aligned}N_{\boldsymbol{L_1\beta_2}} &= \boldsymbol{x}_{i_*}^T \cdot \tau^2 \cdot ZI_{1i} \cdot \lambda_{2i}(1+\tau\lambda_{2i})^{\tau-1}\\ &\quad + \boldsymbol{x}_{i_*}^T \cdot \tau^2 \cdot \lambda_{2i}\left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau-1},\\ D_{\boldsymbol{L_1\beta_2}} &= ZI_{1i}\cdot ZI_{2i} + ZI_{1i}(1+\tau\lambda_{2i})^{\tau} + ZI_{2i}(1+\tau\lambda_{1i})^{\tau}\\ &\quad + \left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau}.\end{aligned}$$

- The derivative $L\left(\boldsymbol{\theta}_{i*,MGWBZINBR}\right)$ with respect to $\left(\boldsymbol{\delta}_1^*\left(\boldsymbol{u}_i\right)\right)$ under the condition $(Y_{1i}=0, Y_{2i}=0)$:

$$\frac{\partial L\left(\boldsymbol{\theta}_{i_*,MGWBZINBR}\right)}{\partial \boldsymbol{\delta}_1^*(\boldsymbol{u}_{i_*})} = \sum_{i=1}^{n}\left(\frac{N_{\boldsymbol{L_1\delta_1}}}{D_{\boldsymbol{L_1\delta_1}}}\right) w_{ii_*},$$

where,

$$\begin{aligned}N_{\boldsymbol{L_1\delta_1}} &= -\boldsymbol{x}_{i_*}^T \cdot ZI_{1i}\left(1+\tau(ZI_{2i})\right) + \boldsymbol{x}_{i_*}^T \cdot ZI_{1i}\cdot ZI_{2i}\\ &\quad + \boldsymbol{x}_{i_*}^T \cdot ZI_{1i}\left(1+\tau\lambda_{2i}\right)^{\tau},\\ D_{\boldsymbol{L_1\delta_1}} &= \left(1+\tau(ZI_{1i})\right)\left(1+\tau(ZI_{2i})\right) + ZI_{1i}\cdot ZI_{2i}\\ &\quad + ZI_{1i}\left(1+\tau\lambda_{2i}\right)^{\tau} + ZI_{2i}\left(1+\tau\lambda_{1i}\right)^{\tau}\\ &\quad + \left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau}.\end{aligned}$$

- The derivative $L\left(\boldsymbol{\theta}_{i*,MGWBZINBR}\right)$ with respect to $\left(\boldsymbol{\delta}_2^*\left(\boldsymbol{u}_i\right)\right)$ under the condition $(Y_{1i}=0, Y_{2i}=0)$:

$$\frac{\partial L\left(\boldsymbol{\theta}_{i_*,MGWBZINBR}\right)}{\partial \boldsymbol{\delta}_2^*(\boldsymbol{u}_{i_*})} = \sum_{i=1}^{n}\left(\frac{N_{\boldsymbol{L_1\delta_2}}}{D_{\boldsymbol{L_1\delta_2}}}\right) w_{ii_*},$$

where,

$$\begin{aligned} N_{\boldsymbol{L_1\delta_2}} &= -\boldsymbol{x}_{i_*}^T \cdot ZI_{2i}\left(1+\tau(ZI_{1i})\right) + \boldsymbol{x}_{i_*}^T \cdot ZI_{1i} \cdot ZI_{2i} \\ &\quad + \boldsymbol{x}_{i_*}^T \cdot ZI_{2i}\left(1+\tau\lambda_{2i}\right)^{\tau}, \\ D_{\boldsymbol{L_1\delta_2}} &= \left(1+\tau(ZI_{1i})\right)\left(1+\tau(ZI_{2i})\right) + ZI_{1i} \cdot ZI_{2i} \\ &\quad + ZI_{1i}\left(1+\tau\lambda_{2i}\right)^{\tau} + ZI_{2i}\left(1+\tau\lambda_{1i}\right)^{\tau} \\ &\quad + \left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau}. \end{aligned}$$

- The derivative $L\left(\boldsymbol{\theta}_{i*,MGWBZINBR}\right)$ with respect to $(\boldsymbol{\gamma}_1)$ under the condition $(Y_{1i}=0, Y_{2i}=0)$:

$$\frac{\partial L\left(\boldsymbol{\theta}_{i_*,MGWBZINBR}\right)}{\partial \boldsymbol{\gamma}_1} = \sum_{i=1}^{n}\left(\frac{N_{\boldsymbol{L_1\gamma_1}}}{D_{\boldsymbol{L_1\gamma_1}}}\right) w_{ii_*},$$

where,

$$\begin{aligned} N_{\boldsymbol{L_1\gamma_1}} &= -\boldsymbol{x}_{i_*}^T \cdot ZI_{1i}\left(1+\tau(ZI_{2i})\right) + \boldsymbol{x}_{i_*}^T \cdot ZI_{1i} \cdot ZI_{2i} \\ &\quad + \boldsymbol{x}_{i_*}^T \cdot ZI_{1i}\left(1+\tau\lambda_{2i}\right)^{\tau} + \boldsymbol{x}_{i_*}^T\left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau}, \\ D_{\boldsymbol{L_1\gamma_1}} &= \left(1+\tau(ZI_{1i})\right)\left(1+\tau(ZI_{2i})\right) + ZI_{1i} \cdot ZI_{2i} \\ &\quad + ZI_{1i}\left(1+\tau\lambda_{2i}\right)^{\tau} + ZI_{2i}\left(1+\tau\lambda_{1i}\right)^{\tau} \\ &\quad + \left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau}. \end{aligned}$$

- The derivative $L\left(\boldsymbol{\theta}_{i*,MGWBZINBR}\right)$ with respect to $(\boldsymbol{\gamma}_2)$ under the condition $(Y_{1i}=0, Y_{2i}=0)$:

$$\frac{\partial L\left(\boldsymbol{\theta}_{i_*,MGWBZINBR}\right)}{\partial \boldsymbol{\gamma}_2} = \sum_{i=1}^{n}\left(\frac{N_{\boldsymbol{L_1\gamma_2}}}{D_{\boldsymbol{L_1\gamma_2}}}\right) w_{ii_*},$$

where,

$$\begin{aligned} N_{\boldsymbol{L_1\gamma_2}} &= -\boldsymbol{x}_{i_*}^T \cdot ZI_{2i}\left(1+\tau(ZI_{1i})\right) + \boldsymbol{x}_{i_*}^T \cdot ZI_{1i} \cdot ZI_{2i} \\ &\quad + \boldsymbol{x}_{i_*}^T \cdot \tau^2 \cdot ZI_{1i} \cdot \lambda_{2i}\left(1+\tau\lambda_{2i}\right)^{\tau-1} \\ &\quad + \boldsymbol{x}_{i_*}^T \cdot \tau^2 \cdot \lambda_{2i}\left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau-1}, \\ D_{\boldsymbol{L_1\gamma_2}} &= \left(1+\tau(ZI_{1i})\right)\left(1+\tau(ZI_{2i})\right) + ZI_{1i} \cdot ZI_{2i} \\ &\quad + ZI_{1i}\left(1+\tau\lambda_{2i}\right)^{\tau} + ZI_{2i}\left(1+\tau\lambda_{1i}\right)^{\tau} \\ &\quad + \left(1+\tau(\lambda_{1i}+\lambda_{2i})\right)^{\tau}. \end{aligned}$$

Summary of Derivation Structure. The derivation of the score function is divided into two levels. First, partial derivatives of the log-probability function are calculated for four data scenarios:

1. $(Y_{1i}=0, Y_{2i}=0)$
2. $(Y_{1i}>0, Y_{2i}=0)$ or $Y_{1i}=1,2,...; Y_{2i}=0$
3. $(Y_{1i}=0, Y_{2i}>0)$or $Y_{1i}=0\,; Y_{2i}=1,2,...$
4. $(Y_{1i}>0, Y_{2i}>0)$ or $Y_{1i}=1,2,...; Y_{2i}=1,2,....$

Each case contributes a specific component to the score function. The complete gradient vector is then assembled and used in the iterative BHHH algorithm to estimate the localized parameter vector $\theta_{i*,MGWBZINBR}$.

The local estimation procedure of the MGWBZINBR model is derived based on the partial derivatives of the log-likelihood function for each spatial location. Figure 1 summarizes the iterative steps performed using the BHHH algorithm.

As shown in Figure 1, the parameter vector is updated iteratively using the score information derived from the log-likelihood function and the approximated Hessian matrix until the convergence criterion is satisfied. The estimation procedure of the MGWBZINBR model is summarized in the following diagram to illustrate the overall structure, from case-based derivations to iterative optimization.

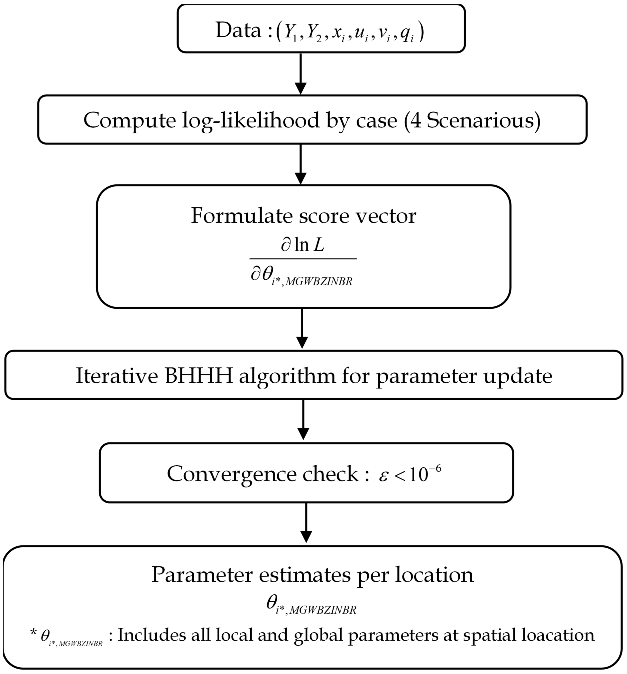


Figure 1. Estimation procedure of the MGWBZINBR model using the BHHH algorithm

Analyzing of the first-order derivatives of the MGWBZINBR Type II model parameters under the four conditions, it is found that the resulting expressions do not have closed-form solutions. Consequently, the parameter estimates cannot be obtained analytically. Hence, numerical optimization is performed through the iterative procedure developed by BHHH.

The steps involved in the BHHH iteration process are described as follows : The first step begins by specifying the initial values of the parameter estimates at iteration $m=0$. The initial parameter vector for location (u_{i*}, v_{i*}) in the MGWBZINBR model follows the same structure as defined previously in eq. (16), consisting of both local and global parameter blocks.

In the second step, the gradient vector of the log-likelihood function for the MGWBZINBR model is computed and follows the structure outlined in the previous derivations.

The structure of the gradient vector $\boldsymbol{g}\left(\boldsymbol{\theta}_{i*,MGWBZINBR}\right)$ follows the arrangement of partial derivatives with respect to each model parameter, as previously defined in eq. (17). Each component is calculated as a weighted sum of individual gradient contributions from each observation, denoted by $\boldsymbol{g_i}\left(\boldsymbol{\theta}_{i*,MGWBZINBR}\right)$, multiplied by the spatial kernel weight w_{ii*}. These individual contributions are based on the partial derivatives of the log-probability function $\ln P\left(y_{1i}, y_{2i}\,|\boldsymbol{\theta}_{i*,MGWBZINBR}\right)$, and reflect the sensitivity of the likelihood to changes in each parameter at observation iii. This formulation ensures that the gradient computation accounts for spatial heterogeneity across locations.

The third step is to construct the Hessian matrix as the negative definite product of the gradient vectors, given by:

$$\boldsymbol{H}\left(\widehat{\boldsymbol{\theta}}_{i_*,MGWBZINBR}\right) = -\sum_{i=1}^{n} w_{ii_*}\boldsymbol{g}_i\left(\widehat{\boldsymbol{\theta}}_{i_*,MGWBZINBR}\right) \boldsymbol{g}_i\left(\widehat{\boldsymbol{\theta}}_{i_*,MGWBZINBR}\right)^T.$$

The next step, involves substituting the current parameter values $\widehat{\boldsymbol{\theta}}^{(0)}_{i^*,MGWBZINBR}$ into both the gradient vector $\boldsymbol{g}\left(\widehat{\boldsymbol{\theta}}_{i*,MGWBZINBR}\right)$ and the Hessian matrix $\boldsymbol{H}\left(\widehat{\boldsymbol{\theta}}_{i*,MGWBZINBR}\right)$. These updated values serve as the basis for computing the direction and magnitude of the parameter adjustments.

The convergence tolerance is set as $\varepsilon = 10^{-6}$, with a maximum iteration limit of $t* = 10.000$ A convergence tolerance of 10^{-6} is commonly recommended to obtain statistically optimal estimates [16]. This value balances numerical precision and computational efficiency: smaller thresholds may yield slightly more accurate results at the cost of longer runtimes and potential instability, while larger values may lead to premature convergence and biased estimates.

The next step is to update the parameter values at iteration m based on the previously calculated gradient and Hessian. The update is performed using the BHHH iterative formula, by subtracting the product of the inverse Hessian and the gradient vector from the previous parameter values. This process is repeated iteratively until convergence is achieved:

$$\widehat{\boldsymbol{\theta}}^{(m+1)}_{i_*,MGWBZINBR} = \widehat{\boldsymbol{\theta}}^{(m)}_{i_*,MGWBZINBR} - \boldsymbol{H}^{-1}\left(\widehat{\boldsymbol{\theta}}^{(m)}_{i_*,MGWBZINBR}\right) \cdot \boldsymbol{g}\left(\widehat{\boldsymbol{\theta}}^{(m)}_{i_*,MGWBZINBR}\right).$$

Next, the updated parameter values are used to recalculate the gradient vector, and the process continues by returning to the initial step. The iteration will be terminated once the convergence criterion is satisfied. Specifically, the stopping condition is defined as:

$$\left\|\widehat{\boldsymbol{\theta}}^{(m+1)}_{i_*,MGWBZINBR} - \widehat{\boldsymbol{\theta}}^{(m)}_{i_*,MGWBZINBR}\right\| \leq \varepsilon,$$

where e is a small positive constant close to zero. This criterion ensures that the parameter estimates have stabilized across successive iterations. After the iteration reaches convergence, the final parameter estimates are then defined as:

$$\widehat{\boldsymbol{\theta}}_{i_*,MGWBZINBR} = \widehat{\boldsymbol{\theta}}^{(M)}_{i_*,MGWBZINBR},$$

where M indicates the final iteration at which convergence is attained. This iterative process is applied individually at each spatial location $i = 1, 2, \cdots, n$ allowing the model to generate location-specific parameter estimates.

At the final stage, a covariance matrix is derived for each location to quantify the variability of the estimated parameters. This matrix is approximated using the inverse of the Hessian matrix, as follows:

$$\text{Cov}_{\widehat{\boldsymbol{\theta}}_{i_*,MGWBZINBR}} \overset{n\to\infty}{\approxeq} -1\left(\widehat{\boldsymbol{\theta}}_{i_*,MGWBZINBR}\right).$$

This covariance matrix provides a measure of the precision of the estimated parameters and can be used for further inference, such as constructing confidence intervals or performing hypothesis testing.

4. Conclusion

The parameter estimation of the Mixed Geographically Weighted Bivariate Zero-Inflated Negative Binomial Regression (MGWBZINBR) model is conducted using the Maximum Likelihood Estimation (MLE) method. Due to the absence of closed-form solutions for the derived likelihood equations, parameter estimation is performed through the Berndt–Hall–Hall–Hausman (BHHH) iterative algorithm, ensuring convergence and numerical stability. This study contributes to the theoretical development of spatial count regression models by introducing a robust and flexible framework capable of modeling bivariate data characterized by overdispersion and excess zeros. The MGWBZINBR model allows for both local and global variations in predictor effects, making it suitable for spatially heterogeneous data contexts. To preserve the focus on methodological formulation, empirical or simulation-based applications are intentionally excluded. Future research is encouraged to implement the MGWBZINBR model in real-world data scenarios across disciplines such as public health, urban planning, and environmental studies, or to explore alternative estimation techniques that may enhance computational efficiency.

Author Contributions. **Mawadah Putri Islamiati**: Formal analysis, writing—original draft preparation, writing—review and editing, visualization. **Purhadi**: Conceptualization, methodology, supervision, validation. **Wibawati**: Conceptualization, methodology, supervision, validation. All authors have read and agreed to the published version of the manuscript.

Acknowledgement. The authors would like to express their appreciation to the editors and reviewers for their careful review, insightful feedback, and constructive suggestions that have contributed significantly to improving the quality of this paper.

Funding. This research received no external funding.

Conflict of interest. The authors declare no conflict of interest related to this article.

Data availability. Not applicable.

References

[1] J. M. Hilbe, *Negative Binomial Regression*. Cambridge, U.K.: Cambridge Univ. Press, 2011, doi: 10.1017/CBO9780511973420.

[2] P. C. Consul and F. Famoye, "Generalized Poisson regression model," *Commun. Stat. Theory Methods*, vol. 21, no. 1, pp. 89–109, Jan. 1992, doi: 10.1080/03610929208830766.

[3] D. Lambert, "Zero-inflated Poisson regression, with an application to defects in manufacturing," *Technometrics*, vol. 34, no. 1, p. 1, Feb. 1992, doi: 10.2307/1269547.

[4] J. F. Lawless, "Negative binomial and mixed Poisson regression," *Can. J. Stat.*, vol. 15, no. 3, pp. 209–225, Sep. 1987, doi: 10.2307/3314912.

[5] M. Ridout, J. Hinde, and C. G. B. Demétrio, "A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives," *Biometrics*, vol. 57, no. 1, pp. 219–223, Mar. 2001, doi: 10.1111/j.0006-341X.2001.00219.x.

[6] M. I. A. Saputro and M. F. Qudratullah, "Estimation of zero-inflated negative binomial regression parameters using the maximum likelihood method (case study: factors affecting infant mortality in Wonogiri in 2015)," in *Proc. Int. Conf. Sci. Eng.*, vol. 4, pp. 240–254, 2021.

[7] R. Azwarini, "Hipotesis pada Model Regresi Bivariate Zero-Inflated Negative Binomial," Undergraduate Thesis, Institut Teknologi Sepuluh Nopember, 2023.

[8] H. J. Sari, "Estimasi Parameter dan Pengujian Hipotesis pada Model Geographically Weighted Bivariate Zero-Inflated Negative Binomial Regression," Undergraduate Thesis, Institut Teknologi Sepuluh Nopember, 2024.

[9] C. Zeng *et al.*, "Mapping soil organic matter concentration at different scales using a mixed geographically weighted regression method," *Geoderma*, vol. 281, pp. 69–82, Nov. 2016, doi: 10.1016/j.geoderma.2016.06.033.

[10] B. Lu, M. Charlton, P. Harris, and A. S. Fotheringham, "Geographically weighted regression with a non-Euclidean distance metric: A case study using hedonic house price data," *Int. J. Geogr. Inf. Sci.*, vol. 28, no. 4, pp. 660–681, 2014, doi: 10.1080/13658816.2013.865739.

[11] A. R. da Silva and M. D. R. de Sousa, "Geographically weighted zero-inflated negative binomial regression: A general case for count data," *Spat. Stat.*, vol. 58, p. 100790, 2023, doi: 10.1016/j.spasta.2023.100790.

[12] M. S. Nur, A. Choiruddin, *et al.*, "Parameter estimation and hypothesis testing of geographically weighted bivariate zero-inflated Poisson inverse Gaussian regression models," in *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 1115, no. 1, p. 012043, Mar. 2021, doi: 10.1088/1757-899X/1115/1/012043.

[13] R. Fitriani and I. G. N. M. Jaya, "Spatial modeling of confirmed COVID-19 pandemic in East Java Province by geographically weighted negative binomial regression," *Commun. Math. Biol. Neurosci.*, vol. 2020, pp. 1–17, 2020, doi: 10.28919/cmbn/4874.

[14] N. Ismail and H. Zamani, "Estimation of claim count data using negative binomial, generalized Poisson, zero-inflated negative binomial and zero-inflated generalized Poisson regression models," *Casualty Actuar. Soc. E-Forum*, no. 1992, pp. 1–28, 2013.

[15] A. S. Fotheringham, C. Brunsdon, and M. Charlton, *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester, U.K.: Wiley, 2002.

[16] J. M. Hilbe, *Modeling Count Data*. Cambridge, U.K.: Cambridge Univ. Press, 2014, doi: 10.1017/CBO9781139236065.