Fisher Scoring Method for Parameter Estimation of Geographically Weighted Ordinal Logistic Regression (GWOLR) Model

Purnami Widyaningsih¹, Dewi Retno Sari Saputro¹, Aulia Nugrahani Putri¹

¹Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Sebelas Maret, Indonesia.
Email: purnami_w@staff.uns.ac.id, dewiretnoss@gmail.com ,aulianugrahaniputri@gmail.com

Abstract. GWOLR model combines geographically weighted regression (GWR) and (ordinal logistic regression) OLR models. Its parameter estimation employs maximum likelihood estimation. Such parameter estimation, however, yields difficult-to-solve system of nonlinear equations, and therefore numerical approximation approach is required. The iterative approximation approach, in general, uses Newton-Raphson (NR) method. The NR method has a disadvantage—its Hessian matrix is always the second derivatives of each iteration so it does not always produce converging results. With regard to this matter, NR model is modified by substituting its Hessian matrix into Fisher information matrix, which is termed Fisher scoring (FS). The present research seeks to determine GWOLR model parameter estimation using Fisher scoring method and apply the estimation on data of the level of vulnerability to Dengue Hemorrhagic Fever (DHF) in Semarang. The research concludes that health facilities give the greatest contribution to the probability of the number of DHF sufferers in both villages. Based on the number of the sufferers, IR category of DHF in both villages can be determined.

1. Introduction

In nondeterministic models, in order to find out the relationship between two or more variables, regression model is used. The variables involve predictor and response variables. The response variable can take the form of either quantitative or qualitative. A qualitative, or so-called categorical, variable is the measurement result of a variable involving two or more possible values (categories). Reference [1] points out that a categorical variable with two categories is termed a binary or dichotomous variable, while that with more than two categories is called polychotomous variable.

A regression model which can accommodate both categorical response variables and categorical (and/or quantitative) predictor variables is defined as a logistic regression model. It represents the probability of an event occurring resulted from predictor variables. There are two types of polychotomous logistic regression model: ordinal logistic regression and nominal logistic regression. The former takes ordered response variable into account, while the latter does not [2].

Meanwhile, Tobler’s first law of Geography (Tobler in Anselin [1]), “Everything is related to everything else, but near things are more related than distant things”, has laid foundation for regional science, specifically on spatial dimension. Spatial data refer to geographically-oriented data and possess certain coordinate system as a reference framework. This implies that spatial data contain information about geographic location of an area. Spatial analysis, in general, requires location-based data containing the location’s characteristics. It includes modeling which indicates a concept of cause and effect relationship by using a method sourced from spatial and non-spatial data sources to predict the presence of a spatial pattern.

One of spatial models associated with regression model is geographically weighted regression (GWR). Logistic regression model has been developed to represent the relationship between response variable and predictor variables regarding a geographical location where data are observed. According to [3], the model is called geographically weighted logistic regression (GWLR). The GWLR model which involves response variables with ordinal scales is geographically weighted ordinal logistic...
regression (GWOLR). It combines geographically weighted regression (GWR) and (ordinal logistic regression) OLR models.

The GWOLR model includes response variables, predictor variables, and parameters. The parameters, however, are unknown, and therefore a sample-based estimation to obtain parameter values is required. Reference [2] assumes that one method to estimate parameters is maximum likelihood estimation (MLE). Such method, however, results in difficult-to-solve system of nonlinear equations, and therefore a numerical method is required.

Fisher scoring (FS) is a numerical method modified from Newton-Raphson (NR) method using score vectors and Fisher information matrix. The Fisher information plays a key role in statistical inference ([8], [9]). NR iterations employ Hessian matrix of which elements comprise the second derivatives of a likelihood function. The method, however, does not always produce converging results (Farbod [7]). In 2004, [5] compared NR’s and FS’ strengths in term of the computation of maximum likelihood estimation and concluded that FS was better than NR algorithm since FS algorithm converged when NR algorithm did not. In 2013, [6] applied Fisher scoring algorithm for estimating parameters in binary logistic regression model and found out FS’ advantage—its guaranteed convergence. With regard to the gap, the present research seeks to examine parameter estimation using FS method in GWOLR model and provides an example of its application on data of the level of vulnerability to Dengue Hemorrhagic Fever (DHF) in Semarang city of Central Java.

2. Research Method

The present research presents a theoretical review examining FS method for GWOLR model parameter estimation. Several deepened theories included the modification of NR and FS, the correlation between Fisher Information and FS, and convergence matter of iterative methods. The results of the review will then be applied on data of the level of vulnerability to DHF in Semarang. The data were obtained from Central Agency on Statistics (Badan Pusat Statistik—BPS) and the Department of Health (Dinas Kesehatan) of Semarang.

Some research procedures were conducted to achieve the objective of the research. Firstly, we determined the likelihood function of joint probability density function and formulated its ln-likelihood. Afterwards, we determined a solution which enabled its maximization. In this stage, a constraint was found—a solution to system of nonlinear equations was difficult to find, and therefore numerical approximation using Fisher scoring was required. Next, we applied the results of its parameter estimation on data of the level of vulnerability to DHF in Semarang by categorizing DHF sufferers. According to the Ministry of Health (Kementerian Kesehatan—Kemenkes), DHF sufferers fall into three categories with regard to incidence rate (IR): mild, moderate, and severe. After that, we determined coordinate of each village (kelurahan) in Semarang, distance between villages, weight of each village, and parameter estimation. This way, parameter estimation values and GWOLR model were yielded.

3. Results and Discussion

3.1 GWOLR Model. GWOLR model is the combination of geographically weighted regression (GWR) and ordinal logistic regression (OLR) models of which each parameter depends on location (\(u_i, v_i\)). According to reference [4] GWOLR model with categorical response variable \(K\) is expressed as

\[
\begin{align*}
\text{Logit} \left( P(Y_i < s|x_i) \right) &= a_s(u_i, v_i) + x_i^T \beta(u_i, v_i),
\end{align*}
\]

where \(s = 1, 2, ..., K - 1\) and \(i = 1, 2, ..., n\). Parameters \(a_s(u_i, v_i)\) represent intercepts, \(\beta(u_i, v_i)\) is the vector of regression coefficient for the \(i^{th}\) location sized \(p \times 1\), \(x_i^T\) is the vector of predictor variables for the \(i^{th}\) location sized \(1 \times p\), and \((u_i, v_i)\) represent coordinates (latitude, longitude) of the \(i^{th}\) location.

3.2 Parameter Estimation. The research examined samples of observation (response variables) \(Y_1, Y_2, ..., Y_n\) having \(K\) category and probability of category towards \(x\) \((P_k(x))\), where
$k = 1, 2, ..., K, \sum_{k=1}^{K} P_k(x) = 1$. Since the response variables have $K$ category (multinomial distribution), the likelihood function is expressed as joint probability density function of the multinomial distribution. The joint probability density function is denoted as

$$f(y_k = y_k) = P_k(x) \quad k = 1, 2, ... K$$

Therefore, the likelihood function of $n$ samples of observation is formulated by

$$L(u_i, v_j) = \prod_{i=1}^{n} f(y_{i1})f(y_{i2}) ... f(y_{ik})$$

$$= \prod_{i=1}^{n} \left( P_1(x_i)^{y_{i1}} P_2(x_i)^{y_{i2}} ... P_K(x_i)^{y_{ik}} \right)$$

$$= \prod_{i=1}^{n} \left( \frac{\exp(\alpha_{i1}(u_i, v_j) + x_i^T \beta(u_i, v_j))}{1 + \exp(\alpha_{i1}(u_i, v_j) + x_i^T \beta(u_i, v_j))} \right)^{y_{i1}} \frac{\exp(\alpha_{i2}(u_i, v_j) + x_i^T \beta(u_i, v_j))}{1 + \exp(\alpha_{i2}(u_i, v_j) + x_i^T \beta(u_i, v_j))} - \frac{\exp(\alpha_{i1}(u_i, v_j) + x_i^T \beta(u_i, v_j))}{1 + \exp(\alpha_{i1}(u_i, v_j) + x_i^T \beta(u_i, v_j))} \right)^{y_{i2}} ...$$

$$= \prod_{i=1}^{n} \left( 1 - \frac{\exp(\alpha_{iK-1}(u_i, v_j) + x_i^T \beta(u_i, v_j))}{1 + \exp(\alpha_{iK-1}(u_i, v_j) + x_i^T \beta(u_i, v_j))} \right)^{y_{ik}} \right).$$

Likelihood function serves as exponential function, and therefore to simply the computation, the likelihood is altered to ln-likelihood function. In spatial analysis, to find out closeness between locations, a weight is required. The weight provided, therefore, is in the form of ln-likelihood. If the weight of each location $(u_i, v_i)$ is denoted $w_{ij}(u_i, v_i)$, then the weighted ln-likelihood function is expressed as

$$\ln L(u_i, v_j) = \sum_{i=1}^{n} \left( y_{i1} \ln \left( \frac{\exp(\alpha_{i1}(u_i, v_j) + x_i^T \beta(u_i, v_j))}{1 + \exp(\alpha_{i1}(u_i, v_j) + x_i^T \beta(u_i, v_j))} \right) + y_{i2} \ln \left( \frac{\exp(\alpha_{i2}(u_i, v_j) + x_i^T \beta(u_i, v_j))}{1 + \exp(\alpha_{i2}(u_i, v_j) + x_i^T \beta(u_i, v_j))} \right) - \right.$$

$$\left. \cdots + y_{ik} \ln \left( \frac{\exp(\alpha_{iK-1}(u_i, v_j) + x_i^T \beta(u_i, v_j))}{1 + \exp(\alpha_{iK-1}(u_i, v_j) + x_i^T \beta(u_i, v_j))} \right) \right) w_{ij}(u_i, v_i).$$

Weight $w_{ij}(u_i, v_i)$ is the weight of fixed Gaussian kernel represented as $w_{ij}(u_i, v_i) = \exp\left(-\frac{1}{2}(\frac{d_{ij}}{h})^2\right)$ where $d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$ is the distance between location $(u_i, v_i)$ and location $(u_j, v_j)$, $h$ is bandwidth (which measures the size of neighborhood). A method used to obtain optimum $h$ is cross validation (CV), represented as $CV(h) = \sum_{i=1}^{n} \sum_{k=1}^{K} (y_{i,k} - P_{x_{i,k}}(h))^2$, where $y_{x_{i,k}}$ is
response variable in which observation at location \((u_i, v_i)\) is neglected from estimation, and \(P_{s=1}(h)\) is the value of observation probability estimation at location \((u_i, v_i)\) having \(K\) category. Optimum \(h\) value is obtained from \(h\) which produces minimum CV.

The parameter value which maximizes In-likelihood function can be determined by computing the first derivatives of In-likelihood function on each parameter:

\[
\frac{\partial \ln L(u_i, v_i)}{\partial \alpha_1(u_i, v_i)} = \sum_{i=1}^{n} \left( y_{i1} \frac{1}{1 + e_1} - y_{i2} \left( \frac{e_1(1 + e_2)}{(1 + e_1)(e_2 - e_1)} \right) \right) w_{ij}(u_i, v_i)
\]

\[
\frac{\partial \ln L(u_i, v_i)}{\partial \alpha_2(u_i, v_i)} = \sum_{i=1}^{n} \left( y_{i2} \left( \frac{e_2(1 + e_1)}{(1 + e_2)(e_2 - e_1)} \right) \right) w_{ij}(u_i, v_i)
\]

\[
\vdots
\]

\[
\frac{\partial \ln L(u_i, v_i)}{\partial \alpha_{K-1}(u_i, v_i)} = \sum_{i=1}^{n} \left( y_{iK} \left( \frac{e_{K-1}}{(1 + e_{K-1})} \right) \right) w_{ij}(u_i, v_i)
\]

\[
\frac{\partial \ln L(u_i, v_i)}{\partial \beta(u_i, v_i)} = \sum_{i=1}^{n} \left( y_{i1} \frac{x_i^T}{1 + e_1} - y_{i2} \left( \frac{x_i^T(-1 + e_1 e_2)}{(1 + e_1)(e_2 - e_1)} \right) \right)
\]

\[
= -y_{iK} \frac{x_i^T(e_{K-1})}{1 + e_{K-1}} w_{ij}(u_i, v_i)
\]

Where \(e_1 = \exp(\alpha_1(u_i, v_i) + x_i^T \beta(u_i, v_i))\), \(e_2 = \exp(\alpha_2(u_i, v_i) + x_i^T \beta(u_i, v_i))\), and \(e_{K-1} = \exp(\alpha_{K-1}(u_i, v_i) + x_i^T \beta(u_i, v_i))\).

Moreover, maximum In-likelihood function can be determined by its negative second derivatives. System (1) is a system of nonlinear equations. The value of GWOLR model parameter is the (exact) solution of system (1). The solution of system (1) is difficult to be found so that it is later determined numerically using FS method.

In principle, FS method is similar to NR, which is a modification of NR due to convergence matter in NR (Farbod [7]). GWOLR model parameter estimation with FS requires score vectors and Fisher information matrix. A score vector is a vector of which elements are the first derivatives of In-likelihood function towards each parameter. It is expressed:

\[
S = \begin{bmatrix}
\sum_{i=1}^{n} \left( y_{i1} \frac{1}{1 + e_1} - y_{i2} \left( \frac{e_1(1 + e_2)}{(1 + e_1)(e_2 - e_1)} \right) \right) w_{ij}(u_i, v_i) \\
\sum_{i=1}^{n} \left( y_{i2} \left( \frac{e_2(1 + e_1)}{(1 + e_2)(e_2 - e_1)} \right) \right) w_{ij}(u_i, v_i) \\
\vdots \\
\sum_{i=1}^{n} \left( y_{iK} \left( \frac{e_{K-1}}{(1 + e_{K-1})} \right) \right) w_{ij}(u_i, v_i) \\
\sum_{i=1}^{n} \left( y_{i1} \frac{x_i^T}{1 + e_1} - y_{i2} \left( \frac{x_i^T(-1 + e_1 e_2)}{(1 + e_1)(e_2 - e_1)} \right) \right) w_{ij}(u_i, v_i) \\
\sum_{i=1}^{n} \left( -y_{iK} \frac{x_i^T(e_{K-1})}{1 + e_{K-1}} \right) w_{ij}(u_i, v_i)
\end{bmatrix}
\]
Fisher information matrix is a modification of NR algorithm with substitution of its Hessian matrix. Hessian matrix is a matrix whose elements consist of the second derivatives of In-likelihood function towards each parameter. It is denoted:

\[
H = \begin{bmatrix}
\frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_1^2 (u_i, v_i)} & \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_1 (u_i, v_i) \partial \alpha_2 (u_i, v_i)} & \cdots & \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_1 (u_i, v_i) \partial \beta (u_i, v_i)} \\
\frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_2 (u_i, v_i) \partial \alpha_1 (u_i, v_i)} & \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_2^2 (u_i, v_i)} & \cdots & \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_2 (u_i, v_i) \partial \beta (u_i, v_i)} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial^2 \ln L(u_i, v_i)}{\partial \beta (u_i, v_i) \partial \alpha_1 (u_i, v_i)} & \frac{\partial^2 \ln L(u_i, v_i)}{\partial \beta (u_i, v_i) \partial \alpha_2 (u_i, v_i)} & \cdots & \frac{\partial^2 \ln L(u_i, v_i)}{\partial \beta^2 (u_i, v_i)} 
\end{bmatrix}
\]

The use of Hessian matrix in NR algorithm presents a disadvantage: it does not converge, and therefore it has to be modified with Fisher information matrix. The Fisher information matrix is a matrix of which elements consist of the expectation value of the second derivatives of In-likelihood function towards each parameter. It is expressed as

\[
\text{Inf} = \begin{bmatrix}
E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_1^2 (u_i, v_i)} \right) & E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_1 (u_i, v_i) \partial \alpha_2 (u_i, v_i)} \right) & \cdots & E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_1 (u_i, v_i) \partial \beta (u_i, v_i)} \right) \\
E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_2 (u_i, v_i) \partial \alpha_1 (u_i, v_i)} \right) & E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_2^2 (u_i, v_i)} \right) & \cdots & E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_2 (u_i, v_i) \partial \beta (u_i, v_i)} \right) \\
\vdots & \vdots & \ddots & \vdots \\
E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \beta (u_i, v_i) \partial \alpha_1 (u_i, v_i)} \right) & E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \beta (u_i, v_i) \partial \alpha_2 (u_i, v_i)} \right) & \cdots & E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \beta^2 (u_i, v_i)} \right)
\end{bmatrix}
\]

The concept of the modification is called Fisher information which is later so-called FS. The followings are FS algorithm.

1. Choose an initial value \((m=0)\) of parameter vector \(V_0\) which is obtained from parameter estimation value of ordinal logistic regression model.
2. Calculate parameters
   \[V_{m+1} = V_m + \text{Inf}^{-1} S_m,\]  where \(m = 0, 1, 2, \ldots\)
3. Calculating the norm \(V_{m+1} - V_m = \|V_{m+1} - V_m\|\), where
   \[
   \|V_{m+1} - V_m\| = \sqrt{(\alpha_1(m+1)(u_i, v_i) - \alpha_1(m)(u_i, v_i))^2 + (\beta_1(m+1)(u_i, v_i) - \beta_1(m)(u_i, v_i))^2}.
   \]
   \(\|V_{m+1} - V_m\|\) is used to stop iteration with the criterion of \(\|V_{m+1} - V_m\| < \text{error tolerance}\).
   If the criterion is met, the iteration process stops and the parameter estimation value is \(\tilde{V} = V_{m+1}\). On the other hand, if \(\|V_{m+1} - V_m\| > \text{error tolerance}\), the iteration process is repeated from step (2) to step (3).
   After the parameter estimation value is obtained, GWOLR model is formulated as
   \[
   \text{Logit}(P(Y_i < s|x_i)) = \tilde{\alpha}(u_i, v_i) + x_i^T \tilde{\beta}(u_i, v_i).
   \]

3.3 Application. Data used in the application include data of the level of vulnerability to DHF in Semarang in 2014. The level of vulnerability to DHF comprises three categories (K=3) based on IR, namely mild, moderate, and severe, which are used as response variables, while the predictor variables are population density \((X_1)\), the number of population aged 0-14 years \((X_2)\), the number of semi-
permanent houses ($X_3$), the availability of health facilities ($X_4$), and larvae-free rate ($X_5$), and therefore the p value is 5 ($p=5$). Thus, there are seven parameters, consisting of $V = [\alpha_1(u; v_1); \alpha_2(u; v_1); \beta_1(u; v_1); \beta_2(u; v_1); \beta_3(u; v_1); \beta_4(u; v_1); \beta_5(u; v_1)]$

Semarang city covers 16 sub-districts and 177 villages, and therefore $n=177$. Based on equation (3.1), GWOLR has 177 models for intercept parameter 1 and 177 models for intercept parameter 2. Each intercept parameter was provided with 2 samples of villages. The procedures included categorizing DHF sufferers, determining coordinate (latitude, longitude) of each village, calculating the distance between villages, calculating weight of each village and calculating parameter estimation. The initial value used is $\alpha_1 = -4.843250, \alpha_2 = -2.804115, \beta_1 = 2.623848\times10^{-6}, \beta_2 = -4.601760\times10^{-5}, \beta_3 = 3.561353\times10^{-4}, \beta_4 = -7.387467\times10^{-1}, \beta_5 = -1.737042\times10^{-2}$ and therefore $S_0=[-0.1517446, 1.1334551, \ldots, -74.6099972]^T$ is obtained.

The result of parameter estimation with error tolerance of 0.0001 is obtained in 15th iteration. The results of parameter estimation for Kuningan and Tinjomoyo villages are presented in Table 1.

| Villa | $\alpha_1(u; v_1)$ | $\alpha_2(u; v_1)$ | $\beta_1(u; v_1)$ | $\beta_2(u; v_1)$ | $\beta_3(u; v_1)$ | $\beta_4(u; v_1)$ | $\beta_5(u; v_1)$ |
|-------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| Ku    | -4.704688        | -2.829467        | 0.000042         | -0.000048        | 0.000533        | -0.638496        | -0.017453        |
| Ti    | -4.835065        | -2.821396        | -0.000004        | -0.000047        | 0.000376        | -0.746907        | -0.018490        |

In reference to Table 1, the following GWOLR models are obtained.

(1) GWOLR model of data of DHF for Kuningan village is

$$Logit(P(Y_1 \leq 1|x_i)) = -4.704688 + 0.000042x_{i1} - 0.000048x_{i2} + 0.000533x_{i3} - 0.638496x_{i4} - 0.017453x_{i5}$$

$$Logit(P(Y_1 \leq 2|x_i)) = -2.829467 + 0.000042x_{i1} - 0.000048x_{i2} + 0.000533x_{i3} - 0.638496x_{i4} - 0.017453x_{i5}$$

(2) GWOLR model of data of DHF for Tinjomoyo village is

$$Logit(P(Y_2 \leq 1|x_i)) = -4.835065 - 0.000044x_{i1} - 0.000047x_{i2} + 0.00376x_{i3} - 0.746907x_{i4} - 0.018490x_{i5}$$

$$Logit(P(Y_2 \leq 2|x_i)) = -2.821396 + 0.000044x_{i1} - 0.000047x_{i2} + 0.00376x_{i3} - 0.746907x_{i4} - 0.018490x_{i5}$$

On the basis of the models obtained, for Kuningan village, additions of one population ($X_j$) and one unit of semi-permanent house ($X_3$) lead to 0.0042% and 0.0533% increase in the number of DHF sufferers, so the village belongs to moderate or severe DHF IR category. Additions of one population aged 0-14 ($X_2$), one unit of health facilities ($X_4$), and 1% larvae-free rate ($X_5$) cause the probability of the number of DHF sufferers to reduce 0.0048%, 63.8496%, and 1.7453%, so the village is classified mild or moderate DHF IR category.

For Tinjomoyo village, additions of one unit of semi-permanent house ($X_3$) contribute to 0.0376% increase in probability of the number of DHF sufferers, therefore the village is included in
moderate or severe DHF IR category. Meanwhile, additions of one population \((X_1)\), one population aged 0-14 \((X_2)\), one unit of health facilities \((X_4)\), and 1% larvae-free rate \((X_5)\) cause the probability of the number of DHF sufferers to reduce 0.0004\%, 0.0047\%, 74.6907\%, and 1.849\%, so the village belongs to mild or moderate DHF fever IR category.

4. Conclusions and Suggestions

(1) The result of GWOLR model parameter estimation using Fisher scoring method is: \( \hat{\varphi} = \hat{V}_{m+1} \), where \( \hat{V}_{m+1} = V_m + 1n_{m^{-1}}S_m \). The initial value obtained from the value of ordinal logistic regression model parameter estimation is provided, and therefore the GWOLR model is notated:

\[
\text{Logit} (P(Y_i < s | x_i)) = \alpha_i(u, v_i) + x_i^T \beta(u, v_i)
\]

where \(s = 1, 2, ..., K - 1\).

(2) The Fisher information is expanded through NR algorithmic modification. It is notated in the form of matrix which is so-called Fisher information matrix. Fisher-information matrix is the modification of NR algorithm which substitutes its Hessian matrix. The fisher information matrix is a matrix of which elements consist of expectation values of the second derivatives of likelihood function towards each parameter, stated as

\[
\text{Inf} = \begin{bmatrix}
E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_1^2(u_i, v_i)} \right) & E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_1 \partial \alpha_2(u_i, v_i)} \right) & \cdots & E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_1 \partial \beta(u_i, v_i)} \right) \\
E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_2^2(u_i, v_i)} \right) & E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_2 \partial \alpha_3(u_i, v_i)} \right) & \cdots & E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \alpha_2 \partial \beta(u_i, v_i)} \right) \\
E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \beta^2(u_i, v_i)} \right) & E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \beta \partial \alpha_1(u_i, v_i)} \right) & \cdots & E \left( \frac{\partial^2 \ln L(u_i, v_i)}{\partial \beta \partial \alpha_2(u_i, v_i)} \right)
\end{bmatrix}
\]

The present research employed FS for GWOLR model parameter estimation, while methods for parameter estimation do not only employ FS so that developing other methods is highly possible.

References

[1] McCullagh, P. and J. A. Nelder, “Generalized Linear Models,” second ed., Chapman and Hall, 1983.
[2] Hosmer, D. W. and S. Lemeshow, “Applied Logistic Regression,” John Wiley and Sons, Inc., USA, 2000.
[3] Atkinson, P. M., S. E. German, D. A. sear, and M. J. Clark, “Exploring the Relations Between Riverbank Erosion and Geomorphological Control Using Geographically Weighted Logistic Regression”, Ohio: Ohio State University, vol. 35, pp. 58-82, 2003.
[4] Purhadi, M. Rifada, and P. Wulandari, “Geographically Weighted Ordinal Logistic Regression Model,” International Journal of Mathematics and Computation, vol. 16, pp. 116-216, 2012.
[5] Schworer, A. and P. Hovey, “Newton Raphson versus Fisher Scoring Algorithms in Calculating Maximum Likelihood Estimates,” Dayton, 2004.
[6] Marius, O. U. and O. I. C. Anaene, “Estimating the Fisher’s Scoring Matrix Formula from Logistic Model,” American Journal of Theoretical and Applied Statistics,” vol. 2, pp. 221-227, 2013.
[7] Farbod, D., Ebrahimpour, M., Ghayournorad, Z., (2010), Maximum Likelihood Estimation for Distribution Generated by Cauchy Stable Law, International Journal of Mathematics and Computation, Vol. 7, No. J10, pp. 23-28.
[8] Lehman, E. L., & Casella, G. (1998). *Theory of Point Estimation 2nd edition*. New York, NY: Springer.

[9] Spanos, A. (1999). *Probability Theory and Statistical Inference: Econometric Modeling with Observational Data*. Cambridge, UK: Cambridge University Press.

[10] Tobler, W.R., 1970. A *Computer Movie Simulating Urban Growth in the Detroit Region*, Economic Geography, 46: 234-240.

[11] Anselin, L., *Spatial econometrics : methods and models*, Kluwer Academic Publishers, Dordrecht, 1998.