Parameter estimation of mixed geographically weighted Weibull regression model

Suyitno¹ and N W W Sari²

¹Department of Statistics, Mulawarman University Samarinda, Indonesia
²STMIK SPB Airlangga Samarinda, Indonesia

*Corresponding Author: suyitno.stat.unmul@gmail.com

Abstract. This study discusses a Mixed Geographically Weighted Weibull Regression (MGWWR) Model. MGWWR is a regression model developed from a Geographically Weighted Weibull Regression (GWWR) model. Parameter estimation of GWWR model is done locally at every observation location using geographical location weighting. Based on parameter identification result in GWWR model, the certain covariates influencing the GWWR model may be global (the same value) in nature, whilst others are different. Based on consideration this situation, a MGWWR model is proposed, in which some parameters are assumed to be constant and the others are different for every local model in the study area. The aim of this study is to identify the constant and local parameters in GWWR model, and to estimate the MGWWR model parameters using maximum likelihood estimation (MLE) method. Identification of constant and local parameters in GWWR model is initial step to construct the MGWWR model. The results show that test statistic for hypothesis testing on the constant parameter identification is Wilk’s statistic derived from likelihood ratio test (LRT) method, and the maximum likelihood estimator of MGWWR model can be obtained by using the Newton-Raphson iterative method based on the back-fitting procedure.

1. Introduction

A univariate Weibull distribution depends on three parameters, namely location, scale and shape parameters. One special univariate Weibull distribution is scale-shape version, where it has two parameters, namely scale and shape parameters. As Weibull distribution development, it can depend on covariates or supplementary variables namely, scale or shape parameter can be stated as a function of regression parameters [4, 6]. Furthermore, a Weibull distribution in which scale parameter depend on covariates or regression parameters is called Weibull regression model. The regression models can be constructed from the interrelationship functions of Weibull distribution, such as probability density function (PDF), cumulative distribution function (CDF), survival function and hazard function, by transforming scale parameter to function of regression parameters. In many applications, the Weibull regression model may Weibull survival regression or Weibull hazard regression [9].

Several researchers have discussed the Weibull regression models, those are O’Quigley et al. [5] proposed a regression model for survival time studies. The study discussed parameter estimation of the univariate Weibull regression model and the estimator of parameters was calculated by Fortran program. Hanagal [2] proposed bivariate Weibull regression models for the censored life time samples with identical covariates and non-identical regression parameters. The models were derived from a
Marshal-Olkın’s bivariate exponential and a Freund’s bivariate exponential distribution. Suyitno et al. [7, 8, 9] proposed the Weibull regression models for continuous samples (not life time) for univariate, bivariate and multivariate cases. The general parameter estimation method of Weibull regression models which have been discussed before was maximum likelihood estimation (MLE) method.

Motivated by the idea of Geographically Weighted Regression (GWR) model, the Weibull regression model was developed to Geographically Weighted Weibull Regression (GWWR) model, namely it was applied to the spatial data. Parameter estimation of GWWR model is done locally at every location in the studied area by using geographical location weighting, which produces the local models. Considering on the results of the previous studies, the GWR model was outperform than global model [1, 2, 7, 10].

Based on the result of partial testing, some parameters in GWWR model may the same value (constant) for all location, and the others are different, namely several covariates influence globally and the others influence locally to the response. Considering on the situations where certain covariates influencing the GWWR model may be global in nature, whilst others are local, and motivated by Mixed Geographically Weighted Regression (MGWR) method [2], then the GWWR model development to Mixed Geographically Weighted Weibull Regression (MGWWR) model is needed. Statistical method enrichment is needed, especially parameter estimation and test statistics of parameter testing for MGWWR model. Meanwhile, many real problems in several fields can be solved by using MGWWR model. Statistical identifying of global or local factors, and the solution finding, when some factors may influence globally, whilst others influence locally to a problem can use MGWWR modeling.

The initial step of MGWWR model constructing is identifying of the constant parameters in GWWR model. Test statistic for the constant parameters identification is derived by using the maximum likelihood ratio test (LRT) method. The proposed parameter estimation method in this study is maximum likelihood estimation (MLE) based on the back-fitting procedure, and maximum likelihood (ML) estimator is obtained by using the Newton-Raphson iterative method. MLE method is more suitable to hand this problem based on the ML estimator properties, namely consistent, asymptotic normality, asymptotic efficiency and invariance.

The remaining part of this paper is organized as follows. Introducing of univariate Weibull regression (WR) model is discussed in section 2, general model of GWWR model is introduced in section 3. Identification of constant parameters in GWWR model is discussed in the section 4. Parameter estimation of MGWWR model is discussed in section 5, and finally the conclusion of the study will be stated in the section 6.

2. Weibull Regression Model

Let $Y$ be continuous nonnegative random variable which follows scale-shape version of Weibull distribution. The probability density function (PDF) is defined as

$$f(y) = \frac{y}{\lambda} \left(\frac{y}{\lambda}\right)^{\gamma-1} \exp\left[-\left(\frac{y}{\lambda}\right)^\gamma\right], y > 0; \gamma > 0; \lambda > 0,$$

with $\gamma$ is shape parameter and $\lambda$ is scale parameters [6]. The survival function of scale-shape version of Weibull distribution is given by

$$S(y) = P(Y > y) = \exp\left[-\left(\frac{y}{\lambda}\right)^\gamma\right], y > 0; \gamma > 0; \lambda > 0,$$
and the hazard function is given by
\[ h(y) = \frac{y}{\lambda \left( \frac{y}{\lambda} \right)^{\gamma - 1}}, \quad y > 0; \gamma > 0; \lambda > 0. \]  
(3)

A general method to estimate parameters of Weibull distribution is MLE. Considering on PDF given by equation (1), the general moment of univariate Weibull distribution can be expressed as
\[ E(Y^r) = \lambda^r \Gamma \left( \frac{r}{\gamma} + 1 \right); \quad E(Y) = \lambda \Gamma \left( \frac{1}{\gamma} + 1 \right), \]
(4)
with $\Gamma(.)$ is Gamma function.

As a Weibull distribution development, scale parameter ($\lambda$) on the equations (1), (2), (3) and (4) can depend on covariates, namely it can be stated in the function of covariates or regression parameters [3, 4, 6]. Furthermore, a Weibull distribution in which scale parameter is stated in the term of regression parameters is called a univariate Weibull regression (WR) model [7, 8, 9]. Let scale parameter $\lambda$ be positive real valued, then it can be expressed in the terms of the regression parameters as follows
\[ \lambda = \lambda(\beta, x) = \exp[\beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p] = \exp[\beta^T x], \]
(5)
with $\beta = [\beta_0, \beta_1, \cdots, \beta_p]^T$ is a $(p+1)$ dimensional vector of regression parameters and $x = [x_0, x_1, \cdots, x_p]^T$ is a $(p+1)$ dimensional vector of covariates, with $x_0 = 1$. The WR models can be obtained by substituting scale parameter defined in equation (5) to the survival function (2), hazard function (3) or mean of Y (4). Based on the equation (2), (3), (4) and (5), a Weibull survival regression model has form
\[ S(y, \theta) = \exp[-y^\gamma \exp[-\gamma \beta^T x]], \]
(6)
a Weibull hazard regression model can be expressed
\[ h(y, \theta) = \gamma y^{\gamma - 1} \exp[-\gamma \beta^T x], \]
(7)
and a Weibull regression model for mean given by
\[ \mu_i(\theta, x) = \Gamma \left( \frac{1}{\gamma} + 1 \right) \exp[\beta^T x], \]
(8)
with $\theta = [\gamma, \beta_0, \cdots, \beta_p]^T$ is a $(p + 2)$ dimensional vector. Parameter estimation of WR models given by equation (6), (7) and (8) can use MLE method, and the maximum likelihood (ML) estimator can be obtained by using iterative Newton-Raphson method [7, 8, 9].

3. Geographically Weighted Weibull Regression (GWWR) Model

GWWR model is WR model applied to spatial data. Spatial data is data containing both attribute and location information. There is a spatial heterogeneity or spatial effect on the spatial data, which indicates interdependence between geographical location and the data. Spatial heterogeneity in response data causes covariates data will give different effect to response data at different location.
Spatial heterogeneity in response data will cause the values of regression parameters at every location are different, therefore applying of global model (WR) to spatial data will produce an invalid model. Based on the GWR idea, an appropriate model to apply the WR model to spatial data is GWWR. Parameter estimating of GWWR model is done locally at every location using the spatial weighting (geographical location weighting). The spatial weighting can be calculated by using a selected weighting function.

The most common weighting function applied in practice is adaptive Gaussian [1]. Let, \( w_j \) be a weight placed to an observation at location \( \mathbf{u}_j = [u_{ij}, u_{2j}]^T \) for a model at location \( \mathbf{u}_i = [u_{ij}, u_{2i}]^T \), with \( u_{ij} \) and \( u_{2j} \) are latitude and longitude position respectively. The spatial weight \( (w_j) \) can be calculated by using an adaptive Gaussian weighting function as follows:

\[
 w_j = \exp \left( -\frac{1}{2} \left( \frac{d_{ij}}{b_j} \right)^2 \right), \quad j = 1, 2, \cdots, n, \tag{9}
\]

with \( b_j \) is an adaptive bandwidth. Based on the equation (9), the weighting function depends on the bandwidth selection, and therefore, determination of optimum bandwidth is a crucial problem. One of methods to obtain optimum bandwidth is Akaike Information Criterion (AIC). The AIC score for a model with parameter \( \theta \) is calculated as follows

\[
 AIC = -2L(\hat{\theta}) + 2v_0, \tag{10}
\]

with \( \hat{\theta} \) is ML estimator, \( L(.) \) is log-likelihood function and \( v_0 \) is the dimension of \( \theta \) [11]. An alternative method to obtain optimum bandwidth is Bayesian Information Criterion (BIC). The value of BIC is calculated as

\[
 BIC = -2L(\hat{\theta}) + \log n + 2v_0, \tag{11}
\]

with \( n \) is the sample size [1]. Considering on the equations (10) and (11), the optimum bandwidth is \( b_0 = \{b_j, i = 1, 2, \cdots, n\} \) selected as desirable value of \( b_j \) so that \( AIC(b_0) \) or \( BIC(b_0) \) is minimum.

Suppose, every point coordinate of location in the study area is known, then from equation (5), the scale parameter stated in terms of regression parameters at location \( \mathbf{u}_i \) has a form

\[
 \lambda(\mathbf{p}(\mathbf{u}_i)) = \exp[\mathbf{p}^T(\mathbf{u}_i)\mathbf{x}_i], \tag{12}
\]

with \( \mathbf{p}(\mathbf{u}_i) = [\beta_0(\mathbf{u}_i), \beta_1(\mathbf{u}_i), \cdots, \beta_p(\mathbf{u}_i)]^T \) depend on the geographical location \( \mathbf{u}_i = [u_{ij}, u_{2j}]^T \). Considering on the equation (6), (7) and (12), a GWWR of survival at location \( \mathbf{u}_i \) has a form

\[
 S(y_i, \mathbf{p}(\mathbf{u}_i)) = \exp \left[ -y_i^{\gamma} \exp \left[ -\gamma \mathbf{y}_i^{\gamma} \mathbf{p}^T(\mathbf{u}_i)\mathbf{x}_i \right] \right], \tag{13}
\]

and a GWWR of hazard at location \( \mathbf{u}_i \) has a form

\[
 h(y_i, \mathbf{p}(\mathbf{u}_i)) = \gamma y_i^{\gamma-1} \exp \left[ -\gamma \mathbf{y}_i^{\gamma} \mathbf{p}^T(\mathbf{u}_i)\mathbf{x}_i \right],
\]

with \( \mathbf{p}(\mathbf{u}_i) = [\gamma(\mathbf{u}_i), \beta_1(\mathbf{u}_i), \beta_2(\mathbf{u}_i), \cdots, \beta_p(\mathbf{u}_i)]^T \) and \( \mathbf{x}_i = [1, X_1, X_2, \cdots, X_p]^T \).
Parameter estimation method of GWWR model was MLE. The ML estimator is obtained by maximizing the likelihood function or log-likelihood function, because their maxima are attained at the same points. Suppose, there are spatial data of response sample \( y_i, i = 1, 2, \cdots, n \) which are taken from Weibull distribution \( W(\gamma(u), \exp[\beta^T X_j]) \), and \( X_{i1}, X_{i2}, \cdots, X_{ip}, i = 1, 2, \cdots, n \) are covariates data. Suppose, the point coordinates of all locations in the study area are known, and there is a spatial heterogeneity on the responses data \( y_i, i = 1, 2, \cdots, n \). Based on the equation (1) and (12), the likelihood function placed spatial weighting \( w_j \) for GWWR model at location \( u_j \) is defined as

\[
L(\theta(u_j)) = \prod_{j=1}^{n} \left[ f(\theta(u_j)|y_j,x_j) \right]^{w_j} = \prod_{j=1}^{n} \left[ h(\theta(u_j))S(\theta(u_j)|y_j,x_j) \right]^{w_j} \\
= \prod_{j=1}^{n} \left[ \gamma(u_j)y_j^{(w)} \exp\left[ -\gamma(u_j)\exp[\beta^T(u_j)x_j] \right] \exp\left[ -y_j^{(w)} \exp\left[ -\gamma(u_j)\beta^T(u_j)x_j \right] \right] \right]^{w_j} .
\] (14)

Taking natural logarithm on both sides of likelihood function (14), a log-likelihood function placed weighting \( w_j \) is obtained, which has a form as follows

\[
\ell(\theta(u_j)) = \ln L(\theta(u_j)) = \sum_{j=1}^{n} w_j \left[ \ln \gamma(u_j) + (\gamma(u_j) - 1) \ln y_j - \gamma(u_j)\beta^T(u_j)x_j - y_j^{(w)} \exp\left[ -\gamma(u_j)\beta^T(u_j)x_j \right] \right]
\] (15)

The ML estimator of GWWR model (\( \hat{\theta} \)) is obtained by solving a likelihood equation

\[
\frac{\partial \ell(\theta)}{\partial \theta} = 0,
\] (16)

with \( \theta \) is vector of zeros with dimension \( p + 2 \) and \( \partial \ell(\theta)/\partial \theta \) is a gradient vector. Let a gradient vector in the left hand side of equation (16) is a scalar which has a general form

\[
g(\theta(u_j)) = \frac{\partial \ell(\theta(u_j))}{\partial \theta(u_j)} = \left[ \frac{\partial \ell(\theta(u_j))}{\partial \gamma(u_j)} \frac{\partial \ell(\theta(u_j))}{\partial \beta(u_j)} \frac{\partial \ell(\theta(u_j))}{\partial \beta(u_j)} \cdots \frac{\partial \ell(\theta(u_j))}{\partial \beta_p(u_j)} \right]^T.
\] (17)

The components of gradient vector (17) can be expressed in the general form as follows

\[
\frac{\partial L(\theta(u_j))}{\partial \gamma(u_j)} = \sum_{j=1}^{n} w_j \left[ \frac{1}{\gamma(u_j)} + \ln y_j - \beta^T(u_j)x_j - y_j^{(w)} \ln y_j \exp\left[ -\gamma(u_j)\beta^T(u_j)x_j \right] \right] + \\
\sum_{j=1}^{n} w_j \left[ y_j^{(w)} \beta^T(u_j)x_j \ln y_j \exp\left[ -\gamma(u_j)\beta^T(u_j)x_j \right] \right] ;
\] (18)

\[
\frac{\partial \ell(\theta(u_j))}{\partial \beta_k(u_j)} = \sum_{j=1}^{n} w_j \left[ -\beta_k(u_j)X_{kj} + \gamma(u_j)y_j^{(w)}X_{kj} \exp\left[ -\gamma(u_j)\beta^T(u_j)x_j \right] \right], k = 0, 1, 2, \cdots, p .
\] (19)
Considering on the component expressions of gradient vector given by equation (18) and (19), the likelihood equation (16) is a system of nonlinear equations which has not a closed form solution to obtain ML estimator. The alternative method to solve likelihood equation (16) is the iterative Newton-Raphson method, and the ML estimator \( \hat{\theta}(u_j) \) is estimated by the roots of likelihood equation (16). Computation of gradient vector and Hessian matrix to apply the Newton-Raphson algorithm is needed. The Hessian matrix for GWWR model at location \( u_j \) has an expression

\[
H(\theta(u_j))_{(p+2)\times(p+2)} = \frac{\partial^2 \ell(\theta(u_j))}{\partial \theta(u_j) \theta^T(u_j)},
\]

and it elements can be expressed in the general form as follows

\[
\frac{\partial^2 \ell(\theta(u_j))}{\partial \gamma(u_j) \beta(u_j)} = \sum_{j=1}^{n} w_j \left( \frac{1}{\gamma^2(u_j)} - \gamma_j \ln(\gamma_j) \right) \exp \left[ -\gamma(u_j) \beta^T(u_j) x_j \right] + 2 \sum_{j=1}^{n} w_j \left( 2 \gamma_j \ln(\gamma_j) \beta^T(u_j) x_j \right) + \sum_{j=1}^{n} w_j \left( \gamma_j \beta^T(u_j) x_j \right)^2 \exp \left[ -\gamma(u_j) \beta^T(u_j) x_j \right];
\]

\[
\frac{\partial^2 \ell(\theta(u_j))}{\partial \gamma(u_j) \gamma(u_j)} = \sum_{j=1}^{n} w_j \left( \gamma(u_j) y_j x_j \beta^T(u_j) x_j \right) \exp \left[ -\gamma(u_j) \beta^T(u_j) x_j \right] - \sum_{j=1}^{n} w_j \left( \gamma(u_j) y_j x_j \beta^T(u_j) x_j \right) \exp \left[ -\gamma(u_j) \beta^T(u_j) x_j \right] k=1,2,\cdots,p
\]

\[
\frac{\partial^2 \ell(\theta(u_j))}{\partial \beta(u_j) \beta(u_j)} = \sum_{j=1}^{n} w_j \left( -\gamma^2(u_j) y_j x_j \gamma(u_j) \beta^T(u_j) x_j \right) \exp \left[ -\gamma(u_j) \beta^T(u_j) x_j \right],
\]

for \( k,t=0,1,2,\cdots,p \). After the gradient vector and the Hessian matrix are calculated, now the Newton-Raphson algorithm can be applied to obtain \( \hat{\theta}(u_j) \) maximizing log-likelihood function (15). The Newton-Raphson algorithm is given by

\[
\hat{\theta}(u_j)^{(m+1)} = \hat{\theta}(u_j)^{(m)} - \left[ H(\hat{\theta}(u_j)) \right]^{-1} g(\hat{\theta}(u_j)), \text{ for } m = 0,1,2,\cdots
\]

As note, the parameter estimator of global model (WR) \( \hat{\theta} = [\hat{\gamma} \hat{\beta}_0 \hat{\beta}_1 \cdots \hat{\beta}_p]^T \) is a special solution of likelihood equation (16) when the spatial weighting \( w_j \) is one for every location \( (i,j=1,2,\cdots,n) \).

4. Identification of Constant Parameters in GWWR Model

Based on the partial testing result, some parameters in GWWR model may constant for all location and the others are different. Considering of the situations where certain covariates influencing the GWWR model may be global in nature, whist others are local, and motivated by a MGWR method,
then the GWWR model development to Mixed Geographically Weighted univariate Weibull Regression (MGWWR) model is needed.

The initial step of MGWWR model constructing is determining of constant parameters (the values are same for each location) and and local parameters (the values are different for each location). The hypothesis form for testing whether \( \hat{\beta}_k(u_j) \) (the coefficient of the \( k \)-th explanatory variable \( X_k \); \( k = 1, 2, \cdots, p \) is constant across the studied area is given by

\[
H_0 : \hat{\beta}_k(u_1) = \hat{\beta}_k(u_2) = \cdots = \hat{\beta}_k(u_n) = \hat{\beta}_k,
\]

\[
H_1 : \text{not all } \hat{\beta}_k(u_j) \text{ for } i = 1, 2, \cdots, n \text{ are equal}.
\]

**Theorem:**
A test statistic for null hypothesis testing stated (23) is

\[
G_k = 2[\ell(\hat{\Omega}) - \ell(\hat{\omega})] \approx \sum_{i=1}^{k} \frac{(\hat{\beta}_i(u_j) - \hat{\beta}_k)^2}{\text{var}[\hat{\beta}_i(u_j)]},
\]

with \( G_k \sim \chi^2_n 

**Proof:**
The estimators of the \( k \)-th regression parameter of GWWR model at \( n \) location are \( \hat{\beta}_{k_e} = [\hat{\beta}_1(u_1), \hat{\beta}_k(u_2), \cdots, \hat{\beta}_n(u_n)]^T \). The test statistic for hypothesis testing stated in (23) is derived by using likelihood ratio test (LRT) method. Let \( \Omega = \{\theta(u_1), \theta(u_2), \cdots, \theta(u_n)\} \) be a parameter set under population and its maximum likelihood (ML) estimator is \( \hat{\Omega} = \{\hat{\theta}(u_1), \hat{\theta}(u_2), \cdots, \hat{\theta}(u_n)\} \), with \( \hat{\theta}(u_j) = [\hat{\theta}_1(u_j), \hat{\theta}_k(u_j), \cdots, \hat{\theta}_n(u_j)]^T \) for \( i = 1, 2, \cdots, n \). Considering on the ML estimator properties, it was known that \( \theta(u_j) \sim N(\hat{\theta}(u_j), [I(\hat{\theta}(u_j))]^{-1}) \), \( n \to \infty \), with \( [I(\hat{\theta}(u_j))] = -H(\hat{\theta}(u_j)) \) is Fisher information matrix. Let \( \omega = \{\gamma \beta_1, \beta_2, \cdots, \beta_n\} \) be a parameter set under null hypothesis, and its ML estimator is \( \hat{\omega} = [\hat{\gamma}, \hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_n] \). The test statistic for hypothesis testing (23) is Wilk’s likelihood ratio derived from LRT method, that is

\[
G_k = 2[\ell(\hat{\Omega}) - \ell(\hat{\omega})] = \sum_{i=1}^{k} 2[\ell(\hat{\theta}(u_j) | y, x_j) - \ell(\hat{\omega} | y, x_j)].
\]

Let, under null hypothesis expressed by equation (23), it is known that \( E[\hat{\beta}_i(u_j)] = \hat{\beta}_i(u_j) = \hat{\beta}_i \), then from equation (24a) and considering a theorem stated in [11], it is obtained

\[
2[\ell(\hat{\theta}(u_j) | y, x_j) - \ell(\hat{\omega} | y, x_j)] \approx \frac{[\hat{\beta}_i(u_j) - \hat{\beta}_i]^2}{\text{var}[\hat{\beta}_i(u_j)]}, i = 1, 2, \cdots, n \cdot
\]

Based on the asymptotic properties of ML estimator and null hypothesis stated in (23), it is obtained \( \hat{\beta}_i(u_j) - \hat{\beta}_i = \hat{\beta}_i(u_j) - \hat{\beta}_i(u_j) - \mu \to Z \), \( n \to \infty \), with \( Z \sim N(0, \text{var}[\hat{\beta}_i(u_j)]) \), and \( \frac{\hat{\beta}_i(u_j) - \hat{\beta}_i}{\text{SE}[\hat{\beta}_i(u_j)]} \to Z_0 \), \( n \to \infty \), with \( Z_0 \sim N(0, 1) \) or

\[
\frac{(\hat{\beta}_i(u_j) - \hat{\beta}_i)^2}{\text{var}[\hat{\beta}_i(u_j)]} \sim \chi^2_i, i = 1, 2, \cdots, n.
\]
Considering on the equation (25) and (26), a test statistic \(G_k\) given by equation (24) can be approximated by

\[ G_k \approx \sum_{i=1}^{n} \left( \frac{\hat{\beta}_i(u_i) - \beta_i}{\text{var}[\hat{\beta}_i(u_i)]} \right)^2, \quad k = 1, 2, \ldots, p, \]

with \(\text{var}[\hat{\beta}_i(u_i)]\) is estimated by the \((k + 2)\)-th diagonal element of matrix \([I(\hat{\theta}(u_i))]^{-1}\). Based on equation (26), it is obtained the test statistics \(G_k \sim X_a^2\) Null hypothesis stated in (23) will be rejected at significant level \(\alpha\), if the value of \(G_k > X_{a,\alpha}^2\).

5. Parameter estimation of MGWWR Model

Suppose, considering on the result of parameter identification in GWWR model, there are \(q\) constant parameters and \(p - q + 2\) local parameters included a shape parameter \((\gamma(u_i))\) and an intercept \(\beta_0(u_i)\). After re-ordering the covariates and considering on the GWWR model (13), a MGWWR model for survival function at location \(u_i\) has an expression

\[ S(y_i, \theta, \theta(u_i)) = \exp\left[-y_i^{-\gamma(u_i)} \exp\left[-\gamma(u_i) \left( \beta^T u_i + \beta(u_i) x_i \right) \right] \right], \quad (27) \]

with \(\theta = \beta = [\beta_0, \beta_1, \ldots, \beta_k]^T\); \(x_i = [X_1, X_2, \ldots, X_p]^T\); \(\beta(u_i) = [\beta_0(u_i), \beta_{q+1}(u_i), \beta_{q+2}(u_i), \ldots, \beta_q(u_i)]^T\); \(\theta(u_i) = [\gamma(u_i), \beta^T(u_i)]^T\) and \(x_i = [1, X_{q+1}, X_{q+2}, \ldots, X_p]^T\). A MGWWR hazard at location \(u_i\) has a form

\[ h(y_i, \theta, \theta(u_i)) = \gamma(u_i) y_i^{\gamma(u_i)-1} \exp\left[-\gamma(u_i) \left( \beta^T u_i + \beta(u_i) x_i \right) \right]. \quad (28) \]

The proposed parameter estimation method of MGWWR model in this study is MLE based on the back-fitting procedure. MLE method is more suitable to hand this problem based on good asymptotic and invariance properties of ML estimator. Parameter estimation is consisted of two steps, which firstly is local parameters estimation, and secondly is constant (global) parameters estimation. When local parameter is estimated, the value of global parameters are equal to their estimator obtained from global parameter estimating before, and vice versa, local parameters are equal to their estimator obtained from local parameter estimating before, when global parameter is estimated. The ML estimator will be obtained easier by maximizing the log-likelihood (the natural logarithm of likelihood) function, because the maximum of both likelihood and log-likelihood function will be attained at the same point.

The local parameter estimation of MGWWR model is done locally at every location. Log-likelihood function at location \(u_i\) is obtained based on the equation (14) and (15), that is

\[ \ell(\theta, \theta(u_i)) = \sum_{j=1}^{n} w_j \left( \ln \gamma(u_i) + (\gamma(u_i) - 1) \ln y_j - \gamma(u_i) \left[ \hat{\beta}^T x_{ij} + \beta^T(u_i) x_{ij} \right] \right) \]

\[ - \sum_{j=1}^{n} w_j \left( y_j^{-\gamma(u_i)} \exp\left[-\gamma(u_i) \left( \hat{\beta}^T u_i + \beta(u_i) x_i \right) \right] \right) \]

with \(\hat{\beta}_j = [\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k]^T\) is ML estimator of WR (global model) given by equation (6), (7) or (8), and \(\theta(u_i) = [\gamma(u_i), \beta_0(u_i), \beta_{q+1}(u_i), \beta_{q+2}(u_i), \ldots, \beta_q(u_i)]^T\). The ML estimator of local parameter at location
\(\mathbf{u}\) is the solution of likelihood equation \(\ell(\mathbf{0}_i(\mathbf{u})) / \partial \mathbf{0}_i(\mathbf{u}) = 0\). However, based on equation (29), the closed form of ML estimator cannot be found analytically, since likelihood equation is consisted of interdependence non linear equations which do not have exact solution. The ML estimator of local parameters can be approximated by roots of likelihood equation obtained numerically by using the iterative Newton-Raphson (N-R) method. To estimate the ML estimator by using the N-R algorithm, computation of gradient vector and Hessian matrix is needed. The gradient vector components and the elements of Hessian matrix to estimate the local parameters using N-R algorithm can be obtained with similar way on GWWR model given by equation (18), (19) and (20)-(22) respectively. The general form of gradient vector for the local parameter estimation is given by

\[
g_{\ell}(\mathbf{0}(\mathbf{u}))_{(p,q+2)p+1} = \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))}{\partial \mathbf{0}_i(\mathbf{u})} = \left[ \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))}{\partial \gamma(\mathbf{u})} \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))}{\partial \beta_0(\mathbf{u})} \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))}{\partial \beta_{q+1}(\mathbf{u})} \ldots \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))}{\partial \beta_p(\mathbf{u})} \right]^T, \tag{30}
\]

and general form of Hessian matrix is

\[
H_{\ell}(\mathbf{u}) = \begin{bmatrix}
\frac{\partial^2 \ell(\mathbf{0}_i(\mathbf{u}))}{\partial \gamma(\mathbf{u})^2} & \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))/\partial \gamma(\mathbf{u})}{\partial \beta_0(\mathbf{u})} & \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))/\partial \beta_{q+1}(\mathbf{u})}{\partial \beta_0(\mathbf{u})} & \ldots & \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))/\partial \beta_p(\mathbf{u})}{\partial \beta_0(\mathbf{u})} \\
\frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))/\partial \gamma(\mathbf{u})}{\partial \beta_0(\mathbf{u})} & \frac{\partial^2 \ell(\mathbf{0}_i(\mathbf{u}))}{\partial \gamma(\mathbf{u})^2} & \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))/\partial \beta_{q+1}(\mathbf{u})}{\partial \beta_0(\mathbf{u})} & \ldots & \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))/\partial \beta_p(\mathbf{u})}{\partial \beta_0(\mathbf{u})} \\
\frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))/\partial \gamma(\mathbf{u})}{\partial \beta_{q+1}(\mathbf{u})} & \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))/\partial \beta_{q+1}(\mathbf{u})}{\partial \beta_0(\mathbf{u})} & \frac{\partial^2 \ell(\mathbf{0}_i(\mathbf{u}))}{\partial \beta_{q+1}(\mathbf{u})^2} & \ldots & \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))/\partial \beta_p(\mathbf{u})}{\partial \beta_{q+1}(\mathbf{u})} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))/\partial \gamma(\mathbf{u})}{\partial \beta_p(\mathbf{u})} & \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))/\partial \beta_{q+1}(\mathbf{u})}{\partial \beta_0(\mathbf{u})} & \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))/\partial \beta_{q+1}(\mathbf{u})}{\partial \beta_0(\mathbf{u})} & \frac{\partial \ell(\mathbf{0}_i(\mathbf{u}))/\partial \beta_p(\mathbf{u})}{\partial \beta_p(\mathbf{u})} \\
\end{bmatrix}. \tag{31}
\]

After the gradient vector (30) and the Hessian matrix (31) are calculated, now the Newton-Raphson algorithm can be applied to obtain estimator of local parameter (\(\hat{\mathbf{0}}_i(\mathbf{u})\)). The next step is parameter estimation of global parameter.

Based on the equation (15) and (29), log-likelihood function to estimate the global parameter part is

\[
\ell(\mathbf{0}) = \sum_{j=1}^{n} \left[ \ln \hat{\gamma}(\mathbf{u}) + (\hat{\gamma}(\mathbf{u}) - 1) \ln y_j - \hat{\gamma}(\mathbf{u}) \left[ \beta_j x_{j\gamma} + \hat{\gamma}(\mathbf{u}) x_{j\gamma} \right] \right] - \sum_{j=1}^{n} \left( y_j^{\ell(\mathbf{u})} \exp \left[ -\hat{\gamma}(\mathbf{u}) (\beta_j x_{j\gamma} + \hat{\gamma}(\mathbf{u}) x_{j\gamma}) \right] \right), \tag{32}
\]

with \(\mathbf{0} = [\beta_1, \beta_2, \ldots, \beta_p]^T\), and \(\hat{\mathbf{0}}_i(\mathbf{u}) = [\hat{\gamma}(\mathbf{u}), \hat{\beta}_0(\mathbf{u}), \hat{\beta}_{q+1}(\mathbf{u}), \hat{\beta}_{q+2}(\mathbf{u}), \ldots, \hat{\beta}_p(\mathbf{u})]^T\) is obtained from the local parameter estimation before (the first step). Based on the expression of log-likelihood (32), the exact ML estimator of global parameter cannot be found analytically, and it can be approximated by using iterative Newton-Raphson method. The gradient vector \(g_{\ell}(\mathbf{0})\) has a form

\[
g_{\ell}(\mathbf{0})_{p+1} = \frac{\partial \ell(\mathbf{0})}{\partial \mathbf{0}} = \left[ \frac{\partial \ell(\mathbf{0})}{\partial \beta_1} \frac{\partial \ell(\mathbf{0})}{\partial \beta_2} \ldots \frac{\partial \ell(\mathbf{0})}{\partial \beta_p} \right]^T, \tag{33}
\]
and Hessian matrix $H_{c}(\theta_{c})$ for estimating of constant parameters can be expressed as

$$H_{c}(\theta_{c}) = \frac{\partial^2 \ell(\theta_{c})}{\partial \theta_{c} \partial \theta_{c}^T} = \begin{bmatrix} \frac{\partial^2 \ell(\theta_{c})}{\partial \theta_{c}^2} & \ldots & \frac{\partial^2 \ell(\theta_{c})}{\partial \theta_{c} \partial \theta_{c}} & \ldots & \frac{\partial^2 \ell(\theta_{c})}{\partial \theta_{c} \partial \theta_{c}} \\
\frac{\partial \theta_{c}^2}{\partial \theta_{c}^2} & \ldots & \frac{\partial \theta_{c} \partial \theta_{c}}{\partial \theta_{c}^2} & \ldots & \frac{\partial \theta_{c} \partial \theta_{c}}{\partial \theta_{c}^2} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\frac{\partial \theta_{c} \partial \theta_{c}}{\partial \theta_{c}^2} & \frac{\partial \theta_{c} \partial \theta_{c}}{\partial \theta_{c}^2} & \ldots & \frac{\partial \theta_{c} \partial \theta_{c}}{\partial \theta_{c}^2} & \frac{\partial \theta_{c} \partial \theta_{c}}{\partial \theta_{c}^2} \\
\frac{\partial \theta_{c} \partial \theta_{c}}{\partial \theta_{c}^2} & \frac{\partial \theta_{c} \partial \theta_{c}}{\partial \theta_{c}^2} & \ldots & \frac{\partial \theta_{c} \partial \theta_{c}}{\partial \theta_{c}^2} & \frac{\partial \theta_{c} \partial \theta_{c}}{\partial \theta_{c}^2} 
\end{bmatrix}_{j=q}\,$$

Furthermore, the Newton-Raphson algorithm can be applied to obtain estimator of global parameter $(\hat{\theta}_{c})$ after gradient vector (33) and the Hessian matrix (33) are calculated. The parameter estimation procedure based on the back-fitting procedure above is repeated $m$ times until a conditions $\left\| \hat{\theta}_{c}^{(m+1)} - \hat{\theta}_{c}^{(m)} \right\| < \varepsilon$ is fulfilled, with $\varepsilon$ is small positive real number and

$$\hat{\theta}_{c} = [\hat{\theta}_{c} \hat{\theta}_{c}(u_{i})]^T = [\hat{\beta}_{1} \hat{\beta}_{2} \ldots \hat{\beta}_{q} \hat{\gamma}(u_{i}) \hat{\beta}_{q+1}(u_{i}) \hat{\beta}_{q+2}(u_{i}) \ldots \hat{\beta}_{p}(u_{i})]^T.$$

### 6. Conclusion

MGWWR is a regression model constructed from a GWWR model, in which some parameters are assumed to be constant and the others are different for every local model in the study area. The constant and local parameter is obtained based on the identification testing of covariates influencing in GWWR model. A test statistic of constant parameter identifying test is Wilk’s likelihood ratio test from likelihood ratio test (LRT) method. Parameter estimation of MGWWR model is maximum likelihood estimation method based on the back-fitting procedure, and ML estimator is obtained by using the iterative Newton-Raphson method. Parameter estimation is consisted two steps, namely step of local parameter estimation and step of constant (global) parameter estimation. The study of MGWWR model will be continued, that is hypotheses testing on MGWWR model and applying to the data will be discussed in the next paper.

### Acknowledgments

The authors would like to thank to the reviewers and the associate editor for their constructive comments. Author is very thankful for motivation support given by Lecturer of Mathematics Department, Mulawarman University.

### References

[1] Alexander S F, Chris B, Martin C 2002 Geographically Weighted Regression: The Analysis of Spatially Varying Relationships John Wiley & Sons pp 44-61.

[2] Chang L M 2005 Geographically Weighted Regression Technique for Spatial Data Analysis, School of Science Xi’an Jiaotong University.

[3] David D H 2005 A Bivariate Weibull Regression Model, Economic Quality Control 20 pp 143-50.

[4] Jerald F L 2003 Statistical Model and Methods for Lifetime Data, John Wiley & Sons, Hoboken New Jersey, pp 269-70.

[5] John O’Q, Ann R 1980 WEIBULL: A Regression Model for Survival Time Studies, Computer Programs in Biomedicine 12 pp 14-8.

[6] Horst R 2009, The Weibull Distribution, A Handbook (New York: CRC Press, Talyor & Francis Group) pp 30-4, 170-2.
[7] Suyitno, Purhadi, Sutikno, and Irhamah 2016 Parameter Estimation of Geographically Weighted Trivariate Weibull Regression Model, *Applied Mathematical Sciences* **10** pp 861-78.

[8] Suyitno, Purhadi, Sutikno, and Irhamah 2017 Multivariate Weibull Regression Model, Far East *Journal of Mathematical Sciences (FJMS)* **101**(9) pp 1977-92.

[9] Suyitno 2017 Penaksiran Parameter dan Pengujian Hipotesis Model Regresi Weibull Univariat, Jurnal Eksponensial, 8(2) pp 179-184.

[10] Triyanto, Purhadi, Bambang W O, Santi W P 2015 Parameter Estimation of Geographically Weighted Multivariate Poisson Regression, *Applied Mathematical Sciences*, **9** pp 4081-93.

[11] Yudi P 2001 *In All Likelihood, Statistical Modeling and Inference Using Likelihood*, (Oxford:Clarendon Press) pp 256-9.