# Certificate

This certificate is presented to

## Suyitno

as

## Presenter

in International Conference on Statistics and Mathematics 2014
Surabaya Indonesia, November 27th-28th 2014

Conference organizer,

Dr. rer. pol. Heri Kuswanto, M.Si

# PROCEEDINGS

## International Conference on Statistics and Mathematics

Majapahit Hotel, Surabaya - Indonesia
November 27-28, 2014

ICSM2014

Challenge of Statistical and Mathematical Applications for Interdisciplinary Research

ISSN : 2460 - 0237

# Contents

2014 International Conference on Statistics and Mathematics (ICSM 2014)

# Geographically Weighted Multivariate Weibull Regression Model

**Suyitno[a], Purhadi[b], Sutikno[b], and Irhamah[b]**

*[a]Ph.D. Student of Department of Statistics, Faculty of Mathematics and Natural Sciences, Institut Teknologi Sepuluh Nopember Surabaya, and Program Studi Statistika FMIPA Universitas Mulawarman, Jl. Kuaro, Kampus Gunung Kelua, Samarinda, 75123, Indonesia*

*[b]Department of Statistics, Faculty of Mathematics and Natural Sciences Institut Teknologi Sepuluh Nopember, Kampus ITS Sukolilo, Surabaya, 60111, Indonesia.*

*Email: suyitno.stat.unmul@gmail.com*

**Abstract**

In this study, we propose Geographically Weighted Multivariate Weibull Regression (GWMWR) model, including parameter estimation and hypothesis testing for regression parameters. The GWMWR model is the local model of the multivariate Weibull regression, which the parameter estimation is done on every location (point) where the data is collected, with the weighting of geographical location. The model is constructed from the multivariate survival function of the multivariate Weibull distribution of Lee and Wen, which the scale parameters are expressed in the regression parameters with identical covariates and non-identical regression parameters. The goals of this study are to estimate the parameters of GWMWR model, and to determine the test statistics on hypothesis testing for regression parameters. The analytical method for parameter estimation is maximum likelihood estimation (MLE), with the weighting of geographical location. The maximum likelihood estimator (MLEs) is obtained by Newton-Raphson iterative method. The test statistics for test of goodness of fit and the simultaneous test are Wilk's likelihood ratio statistic, and the test statistic for partial test is Wald statistic.



*Keywords: GWMWR model, MLE, the weighting of geographical location, Wilk's likelihood ratio statistic, Wald statistic.*

## 1. Introduction

The GWMWR model is an extension of multivariate Weibull regression model with the unknown location information. The GWMWR is a multivariate Weibull regression modelling for spatial data, where the spatial data contains both attribute and location information. In the spatial data, the value of the observation in one location depends on the values of the observation in nearby locations [1]. The spatial data is dependent variable, which the data collected from different location will indicate interdependent between the data and location. So, the classical regression modelling for the spatial data does not fit, and a method which can be used for spatial data modelling is the geographically weighted regression (GWR). Therefore, the multivariate Weibull regression modelling for the spatial data is able to use GWMWR model. In the GWMWR model, the parameter estimation is done on every location where the data is collected, and it is based on the weighting of geographical location. When the estimated parameter of the GWMWR model on every location is equal, then we have the global model of multivariate Weibull regression, and it means that the weighting of geographical location does not influence the model.

There are many literatures about the geographically weighted linear regression, but few which discussed about GWMWR model, so in this study, we propose the GWMWR model. The model is used for continuous nonnegative data (not the censored life time data), where many applications on the several fields often provide continuous nonnegative data. The GWMWR model is the local model of the joint probability density function (PDF) of the multivariate Weibull distribution, which is derived from the multivariate survival function of the multivariate distribution of Lee and Wen (2009), [3]. This study is focused on estimating of parameters, and determining the test statistic on the hypothesis testing for regression parameters. The parameter estimation method is MLE, the test statistics for goodness of fit test and simultaneous test are determined by LRT method, and the test statistic for partial test is Wald statistic, which is derived from the asymptotic normality property.

## 2. Materials and Methods

The joint survival function for the multivariate Weibull distribution of the continuous nonnegative random variable $Y_1, Y_2, \cdots, Y_m$ was proposed by Lee and Wen (2009), [3], which for $0 < \alpha \le 1$; $0 < y_1, \cdots, y_m < \infty$ ; $0 < \lambda_1, \lambda_2, \cdots, \lambda_m < \infty$ ; $0 < \gamma_1, \gamma_2, \cdots, \gamma_m < \infty$, has an expression given by

$$S(y_1, y_2, \cdots, y_m) = \exp[-A^{\alpha}], \text{ with } A = \sum_{k=1}^{m} \left( \frac{y_k}{\lambda_k} \right)^{\frac{\gamma_k}{\alpha}}, \tag{1}$$

where $\alpha$ measures the association among the variables, $\lambda_k$ is the scale parameter and $\gamma_k$ is the shape parameter for variable $Y_k$. The PDF $f(y_1, y_2, \cdots, y_m)$ of the multivariate Weibull distribution can be obtained by differentiating the multivariate survival function (1) with respect to each variable, that is

$$f(y_1, y_2, \cdots, y_m) = \frac{(-1)^m \partial^m S(y_1, y_2, \cdots, y_m)}{\partial y_1 \partial y_2 \cdots \partial y_m}. \tag{2}$$

It follows directly from (1) and (2), then the joint PDF of the multivariate Weibull distribution can be obtained, and it has the expression given by Equation (3)

$$f_1(y_1, y_2, \cdots, y_m) = \left( \prod_{k=1}^{m} \left( \frac{1}{\alpha} \right) \left( \frac{\gamma_k}{\lambda_k} \right) \left( \frac{y_k}{\lambda_k} \right)^{\frac{\gamma_k}{\alpha} - 1} \right) \left( \sum_{l=1}^{m} (-1)^{l+m} C(m, l, \alpha) A^{l\alpha - m} \right) \exp(-A^{\alpha}), \tag{3}$$

where $C(m, l, \alpha)$ is the generalized factorial coefficient (Wahyudi, Purhadi, Irhamah and Sutikno, 2011) [5], which for all positive integer $m, l$ with $l \le m$, and for real number $\alpha$, it is defined by

$$C(m, l, \alpha) = \sum_{\substack{l_1 + l_2 + \cdots + l_m = l \\ l_1 + 2l_2 + \cdots + m l_m = m \\ l_i \text{ is nonnegative integer}}} \binom{m}{l_1, l_2, \cdots, l_m} \prod_{k=1}^{m} \binom{\alpha}{k}^{l_k} \text{ and } \binom{m}{l_1, l_2, \cdots, l_m} = \frac{m!}{l_1! \, l_2! \cdots l_m!}.$$

The PDF (3) can be stated in term of the regression parameters by taking a transformation, that is the scale parameter $\lambda_k$, for $k = 1, 2, \cdots, m$ can be expressed in the regression parameters (Hanagal, 2005) [2], by the following way

$$\lambda_k(\mathbf{x}) = \exp[\boldsymbol{\beta}_k^T \mathbf{x}], \; k = 1, 2, \cdots, m, \tag{4}$$

where $\boldsymbol{\beta}_k^T = [\beta_{k0} \; \beta_{k1} \cdots \beta_{kp}]$ is a vector of the regression parameters with $-\infty < \beta_{kv} < \infty$, $v = 0, 1, \cdots, p$ and $\mathbf{x}^T = [X_0 \; X_1 \cdots X_p]$ is a vector of covariates with $X_0 = 1$. By using a transformation defined by (4), the PDF (3) which is expressed in the regression parameters, is given by Equation (5)

$$f_2(y_1, y_2, \cdots, y_m) = \left( \prod_{k=1}^{m} \frac{\gamma_k}{\alpha} (y_k)^{(\gamma_k/\alpha)-1} \exp[-\frac{\gamma_k}{\alpha} \boldsymbol{\beta}_k^T \mathbf{x} \right) \left( A_1^{\alpha-m} \exp[-A_1^{\alpha}] \right) \left( \sum_{l=1}^{m} (-1)^{l+m} C(m,l,\alpha) A_1^{(l-1)\alpha} \right), \quad (5)$$

with $A_1 = \sum_{k=1}^{m} (y_k)^{\gamma_k/\alpha} \exp[-\frac{\gamma_k}{\alpha} \boldsymbol{\beta}_k^T \mathbf{x}]$. Furthermore, the model (5) is known as the multivariate Weibull regression model.

Suppose, coordinate of location for each observation is known, and all of the parameters of local model for every location are different, where they are the continuous function of the location $(u_i, v_i)$ in the geographical study area. Then, the local model of the multivariate Weibull regression (5) or the GWMWR model for the location $\mathbf{u}_i = (u_i, v_i)$ in given by Equation (6)

$$\begin{aligned} f(y_{1i}, y_{2i}, \cdots, y_{mi}) = & \left( \prod_{k=1}^{m} \frac{\gamma_k(\mathbf{u}_i)}{\alpha(\mathbf{u}_i)} (y_{ki})^{(\gamma_k(\mathbf{u}_i)/\alpha(\mathbf{u}_i))-1} \exp[-\frac{\gamma_k(\mathbf{u}_i)}{\alpha(\mathbf{u}_i)} \boldsymbol{\beta}_k^T(\mathbf{u}_i)\mathbf{x}_i \right) \times \\ & \left( A_i^{\alpha(\mathbf{u}_i)-m} \exp[-A_i^{\alpha(\mathbf{u}_i)}] \right) \left( \sum_{l=1}^{m} (-1)^{l+m} C(m,l,\alpha(\mathbf{u}_i)) A_i^{(l-1)\alpha(\mathbf{u}_i)} \right), \end{aligned} \quad (5)$$

with $A_i = \sum_{k=1}^{m} (y_{ki})^{\gamma_k(\mathbf{u}_i)/\alpha(\mathbf{u}_i)} \exp[-\frac{\gamma_k(\mathbf{u}_i)}{\alpha(\mathbf{u}_i)} \boldsymbol{\beta}_k^T(\mathbf{u}_i)\mathbf{x}_i]$ for $i = 1, 2, \cdots, n$. Note that, the GWMWR model consists of $n$ sets of local parameters which for each local model has $1+m(p+2)$ unknown parameters, which should be estimated. The local parameters at location $\mathbf{u}_i$ can be written in the expression of vector as $\boldsymbol{\theta}(\mathbf{u}_i) = [\alpha(\mathbf{u}_i) \ \boldsymbol{\gamma}^T(\mathbf{u}_i) \ \boldsymbol{\beta}_1^T(\mathbf{u}_i) \ \boldsymbol{\beta}_2^T(\mathbf{u}_i) \cdots \boldsymbol{\beta}_m^T(\mathbf{u}_i)]^T$ with $\boldsymbol{\gamma}(\mathbf{u}_i) = [\gamma_1(\mathbf{u}_i) \ \gamma_2(\mathbf{u}_i) \cdots \gamma_m(\mathbf{u}_i)]^T$ and $\boldsymbol{\beta}_k^T(\mathbf{u}_i) = [\beta_{k0}(\mathbf{u}_i) \ \beta_{k1}(\mathbf{u}_i) \cdots \beta_{kp}(\mathbf{u}_i)]^T$, for $k = 1, 2, \cdots, m$ and $i = 1, 2, \cdots, n$.

## 3. Results and Discussion

### *3. 1. Parameter Estimation*

Let $(y_{1i}, y_{2i}, \cdots, y_{mi})$ for $i = 1, 2, \cdots, n$ is a random sample from the multivariate Weibull distribution, $(X_{1i}, X_{2i}, \cdots, X_{pi})$ is the value of the covariate and $\mathbf{u}_i = (u_i, v_i)$ is the coordinate of the location of the $i^{th}$ observation. Based on the sample of size $n$, the likelihood function is defined by

$$\begin{aligned} \mathscr{L}(\boldsymbol{\theta}(\mathbf{u}_i) \mid \mathbf{y}) = \prod_{j=1}^{n} f(\boldsymbol{\theta}(\mathbf{u}_i) \mid y_j) = & \left( \prod_{j=1}^{n} \left( \prod_{k=1}^{m} \frac{\gamma_k(\mathbf{u}_i)}{\alpha(\mathbf{u}_i)} (y_{kj})^{\gamma_k(\mathbf{u}_i)/\alpha(\mathbf{u}_i)} \exp[-\frac{\gamma_k(\mathbf{u}_i)}{\alpha(\mathbf{u}_i)} \boldsymbol{\beta}_k^T(\mathbf{u}_i)\mathbf{x}_j \right) \right) \times \\ & \left( \prod_{j=1}^{n} A_j^{\alpha(\mathbf{u}_i)-m} \right) \left( \prod_{j=1}^{n} \exp[-A_j^{\alpha(\mathbf{u}_i)}] \right) \left( \prod_{j=1}^{n} \left( \sum_{l=1}^{m} (-1)^{l+m} C(m,l,\alpha(\mathbf{u}_i)) A_j^{(l-1)\alpha(\mathbf{u}_i)} \right) \right), \end{aligned} \quad (6)$$

with $A_j = \sum_{k=1}^{m} (y_{kj})^{\gamma_k(\mathbf{u}_i)/\alpha(\mathbf{u}_i)} \exp[-\frac{\gamma_k(\mathbf{u}_i)}{\alpha(\mathbf{u}_i)} \boldsymbol{\beta}_k^T(\mathbf{u}_i)\mathbf{x}_j]$ and $\mathbf{x}_j = [X_{0j} \ X_{1j} \cdots X_{pj}]^T$, $X_{0j} = 1$.

In the GWMWR model, the influence of the geographical location factor is given by the weighting of geographical location on the natural logarithm of the likelihood function. The weighting of geographical location factor can be constructed by the weighting function, which one of the approaches is the Gauss function. By using Gauss function, the weighting ( $w_{ij}$ ) of the $j^{th}$ observation which has the coordinate of location $\mathbf{u}_j$, for the model at the location $\mathbf{u}_i$, is calculated by

$$w_{ij} = \exp(-\frac{1}{2}(d_{ij}/b)^2),\ j = 1,2,\cdots,n, \tag{7}$$

where $d_{ij}$ is the Euclidian distance between the location $\mathbf{u}_i$ and $\mathbf{u}_j$, and $b$ is a bandwidth with $b > 0$. If the coordinate of location for each observation is given, then the natural logarithm of the likelihood function with the weighting of geographical location, can be formulated as

$$L(\boldsymbol{\theta}(\mathbf{u}_i)) \mid \mathbf{y}, \mathbf{w}) = \ln \mathscr{L}(\boldsymbol{\theta}(\mathbf{u}_i)) \mid \mathbf{y}) = \sum_{q=1}^{4} L_q(\boldsymbol{\theta}(\mathbf{u}_i) \mid \mathbf{y}, \mathbf{w}), \tag{8}$$

where: $L_1(\boldsymbol{\theta}(\mathbf{u}_i)) \mid \mathbf{y}, \mathbf{w}) = \sum_{j=1}^{n} w_{ij} \sum_{k=1}^{m} \{\ln \gamma_k(\mathbf{u}_i) - \ln a(\mathbf{u}_i) + (\frac{\gamma_k(\mathbf{u}_i)}{a(\mathbf{u}_i)} - 1)\ln y_{kj} - \frac{\gamma_k(\mathbf{u}_i)}{a(\mathbf{u}_i)}\boldsymbol{\beta}_k^T(\mathbf{u}_i)\mathbf{x}_j\}$;

$$L_2(\boldsymbol{\theta}(\mathbf{u}_i) \mid \mathbf{y}, \mathbf{w}) = \sum_{j=1}^{n} w_{ij}(a(\mathbf{u}_i) - m)\ln \lambda_j;\ L_3(\boldsymbol{\theta}(\mathbf{u}_i) \mid \mathbf{y}, \mathbf{w}) = -\sum_{j=1}^{n} w_{ij}\lambda_j^{a(\mathbf{u}_i)} \text{ and}$$

$$L_4(\boldsymbol{\theta}(\mathbf{u}_i) \mid \mathbf{y}, \mathbf{w}) = \sum_{j=1}^{n} w_{ij} \ln Q_j(\mathbf{u}_i) \text{ with } Q_j(\mathbf{u}_i) = \sum_{l=1}^{m} (-1)^{l+m} C(m, l, a(\mathbf{u}_i))\lambda_j^{(l-1)a(\mathbf{u}_i)}.$$

Based on the sample, the maximum likelihood estimators of $\boldsymbol{\theta}(\mathbf{u}_i)$ denoted by $\hat{\boldsymbol{\theta}}(\mathbf{u}_i)$ is the value of $\boldsymbol{\theta}(\mathbf{u}_i)$ that maximize likelihood function $\mathscr{L}(\boldsymbol{\theta}(\mathbf{u}_i) \mid \mathbf{y})$. Since, the likelihood function (7) is differentiable with respect to all component of vector $\boldsymbol{\theta}(\mathbf{u}_i)$, then $\hat{\boldsymbol{\theta}}(\mathbf{u}_i)$ that maximize likelihood function is often obtained by solving the equations $\partial\mathscr{L}(\boldsymbol{\theta}(\mathbf{u}_i))/\partial\boldsymbol{\theta}(\mathbf{u}_i) = \mathbf{0}$ with $\mathbf{0}$ is a vector of zeros. In most situations, it is more convenient to work with the natural logarithm of the likelihood function, with the weigthing of geographical location, which the maxima is attained at the same points as those of $\mathscr{L}(\boldsymbol{\theta}(\mathbf{u}_i) \mid \mathbf{y})$. So, the MLEs $\hat{\boldsymbol{\theta}}(\mathbf{u}_i)$ are the roots of the system all likelihood equations

$$\frac{\partial L(\boldsymbol{\theta}(\mathbf{u}_i))}{\partial\boldsymbol{\theta}(\mathbf{u}_i)} = \left[\frac{\partial L(\boldsymbol{\theta}(\mathbf{u}_i))}{\partial a(\mathbf{u}_i)}\ \frac{\partial L(\boldsymbol{\theta}(\mathbf{u}_i))}{\partial\boldsymbol{\gamma}^T(\mathbf{u}_i)}\ \frac{\partial L(\boldsymbol{\theta}(\mathbf{u}_i))}{\partial\boldsymbol{\beta}^T(\mathbf{u}_i)}\right]^T = \mathbf{0}, \tag{9}$$

where the $1 + m(p+2)$ dimensional vector $\mathbf{g}(\boldsymbol{\theta}(\mathbf{u}_i)) = L(\boldsymbol{\theta}(\mathbf{u}_i))/\partial\boldsymbol{\theta}(\mathbf{u}_i)$ on the left hand side of (10) is called the score vector. Based on the expression of log likelihood function (9), the likelihood equation (10) does not have closed form solutions, because it is a system of interdependent non linear equations. Hence, the MLEs can be obtained by Newton-Raphson iterative method. Furthermore, to obtain the MLEs $\hat{\boldsymbol{\theta}}(\mathbf{u}_i)$ using Newton-Raphson algorithm can use the formula

$$\hat{\boldsymbol{\theta}}(\mathbf{u}_i)^{(q+1)} = \hat{\boldsymbol{\theta}}(\mathbf{u}_i)^{(q)} - \mathbf{H}^{-1}(\hat{\boldsymbol{\theta}}(\mathbf{u}_i)^{(q)})\mathbf{g}(\hat{\boldsymbol{\theta}}(\mathbf{u}_i)^{(q)}),\ q = 0,1,2,\cdots, \tag{10}$$

where $\mathbf{H}(\boldsymbol{\theta}(\mathbf{u}_i))$ is the Hessian matrix, which has size $(1+m(p+2)) \times (1+m(p+2))$. Based on the Equation (9), to calculate the score vector $\mathbf{g}(\hat{\boldsymbol{\theta}}(\mathbf{u}_i)$ and the Hessian matrix $\mathbf{H}(\boldsymbol{\theta}(\mathbf{u}_i))$ can be formulated as follows

$$\mathbf{g}(\boldsymbol{\theta}(\mathbf{u}_i)) = \frac{\partial L(\boldsymbol{\theta}(\mathbf{u}_i))}{\partial \boldsymbol{\theta}(\mathbf{u}_i)} = \sum_{q=1}^{4} \mathbf{g}_q(\boldsymbol{\theta}(\mathbf{u}_i)) \text{ and } \mathbf{H}(\boldsymbol{\theta}(\mathbf{u}_i)) = \frac{\partial^2 L(\boldsymbol{\theta}(\mathbf{u}_i))}{\partial \boldsymbol{\theta}(\mathbf{u}_i)\partial \boldsymbol{\theta}^T(\mathbf{u}_i)} = \sum_{q=1}^{4} \mathbf{H}_q(\boldsymbol{\theta}(\mathbf{u}_i)). \quad (11)$$

*3.2. Hypothesis Testing For Regression Parameter*

Hypothesis testing for regression parameters in the GWMWR model involves goodness of fit, simultaneous and partial test. The goal of the testing of goodness of fit is to test whether the geographical location factor influences the model. Hypothesis testing for goodness of fit test takes the form:

$H_0 : \beta_{kv}(\mathbf{u}_i) = \beta_{kv}$ for $k = 1, 2, \cdots, m$; $v = 1, 2, \cdots, p$ (no influence of the geographical factor on the model)

$H_1$ : At least one $\beta_{kv}(\mathbf{u}_i) \neq \beta_{kv}$ for $i = 1, 2, \cdots, n$ (there is influence of the geographical factor on the model).

The test statistic for testing of goodness of fit is determined by likelihood ratio test (LRT) method, which is formed by comparing the maximum of the likelihood function under population ( $H_1$ ) and under $H_0$ . Under $H_0$ , the set of the parameters is given by $\Omega_0 = \{\alpha, \boldsymbol{\gamma}^T, \boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T, \cdots, \boldsymbol{\beta}_m^T\}$ , that is the set of parameters on the multivariate Weibull regression or the global model given by equation (5). The set of parameters under $H_1$ is given by $\Omega_1 = \{\alpha(\mathbf{u}_i), \boldsymbol{\gamma}^T(\mathbf{u}_i), \boldsymbol{\beta}_1^T(\mathbf{u}_i), \boldsymbol{\beta}_2^T(\mathbf{u}_i), \cdots, \boldsymbol{\beta}_m^T(\mathbf{u}_i), i = 1, 2, \cdots, n\}$ , that is the set of parameters on the GWMWR model. Based on the LRT method, the test statistic for goodness of fit test is given by Wilk's likelihood ratio statistic, which has the form

$$G_1 = 2(L(\hat{\Omega}_1) - L(\hat{\Omega}_0)), \quad (12)$$

where $L(\hat{\Omega}_1)$ and $L(\hat{\Omega}_0)$ are the maximum of the natural logarithm of the likelihood function under $H_1$ and under $H_0$ respectively, where they are given by:

$L(\hat{\Omega}_1) = \sum_{q=1}^{4} L_q(\hat{\Omega}_1)$ with :

$$L_1(\hat{\Omega}_1) = \sum_{j=1}^{n}\sum_{k=1}^{m} \{\ln \hat{\gamma}_k(\mathbf{u}_i) - \ln \hat{\alpha}(\mathbf{u}_i) + (\frac{\hat{\gamma}_k(\mathbf{u}_i)}{\hat{\alpha}(\mathbf{u}_i)} - 1)\ln y_{kj} - \frac{\hat{\gamma}_k(\mathbf{u}_i)}{\hat{\alpha}(\mathbf{u}_i)} \hat{\boldsymbol{\beta}}_k^T(\mathbf{u}_i)\mathbf{x}_j\} ;$$

$$L_2(\hat{\Omega}_1) = \sum_{j=1}^{n} (\hat{\alpha}(\mathbf{u}_i) - m)\ln \hat{A}_j ; \; L_3(\hat{\Omega}_1) = -\sum_{j=1}^{n} \hat{A}_j^{\hat{\alpha}(\mathbf{u}_i)} ; \; L_4(\hat{\Omega}_1) = \sum_{j=1}^{n} \ln \hat{Q}_j(\mathbf{u}_i) ;$$

$$\hat{Q}_j(\mathbf{u}_i) = \sum_{l=1}^{m} (-1)^{l-m} C(m, l, \hat{\alpha}(\mathbf{u}_i)) \hat{A}_j^{(l-1)\hat{\alpha}(\mathbf{u}_i)} ; \; \hat{A}_j = \sum_{k=1}^{m} (y_{kj})^{\hat{\gamma}_k(\mathbf{u}_i)/\hat{\alpha}(\mathbf{u}_i)} \exp[-\frac{\hat{\gamma}_k(\mathbf{u}_i)}{\hat{\alpha}(\mathbf{u}_i)} \hat{\boldsymbol{\beta}}_k^T(\mathbf{u}_i)\mathbf{x}_j],$$

and $L(\hat{\Omega}_0) = \sum_{q=1}^{4} L_q(\hat{\Omega}_0)$ with :

$$L_1(\hat{\Omega}_0) = \sum_{i=1}^{n}\sum_{k=1}^{m} \{\ln \hat{\gamma}_k - \ln \hat{\alpha} + (\frac{\hat{\gamma}_k}{\hat{\alpha}} - 1)\ln y_{ki} - \frac{\hat{\gamma}_k}{\hat{\alpha}} \hat{\boldsymbol{\beta}}_k^T \mathbf{x}_i\} ; \; L_2(\hat{\Omega}_0) = \sum_{i=1}^{n} (\hat{\alpha} - m)\ln \hat{A}_i ; \; L_3(\hat{\Omega}_0) = -\sum_{i=1}^{n} \hat{A}_i^{\hat{\alpha}} ;$$

$$L_3(\hat{\Omega}_0)=\sum_{j=1}^{n}\ln\hat{Q}_{1j}\ ;\ \hat{Q}_{1j}=\sum_{l=1}^{m}(-1)^{l+m}C(m,l,\hat{\alpha})\hat{\lambda}_{1j}^{(l-1)\hat{\alpha}}\ \text{ and }\ \hat{\lambda}_{1r}=\sum_{k=1}^{m}(y_{kj})^{\hat{\gamma}_k/\hat{\alpha}}\exp[-\frac{\hat{\gamma}_k}{\hat{\alpha}}\hat{\beta}_k^T\mathbf{x}_j].$$

The $G_1$ statistic follows $\chi^2$ distribution with $(n-1)mp$ degree of freedom.

The goal of the simultaneous test is to confirm the certain covariates that are relevant. The simultaneous test uses the form:

$H_0:\beta_{k1}(\mathbf{u}_i)=\beta_{k2}(\mathbf{u}_i)=\cdots=\beta_{kp}(\mathbf{u}_i)=0$ for $k=1,2,\cdots,m$,

$H_1$ : At least one $\beta_{kv}(\mathbf{u}_i)\neq 0$ for $k=1,2,\cdots,m$, $v=1,2,\cdots,p$, $i=1,2,\cdots,n$.

By the same way, the test statistic for simultaneous test is determined by LRT method. The set of the parameters under $H_0$ is given by $\omega=\{\alpha(\mathbf{u}_i),\gamma_1(\mathbf{u}_i),\cdots,\gamma_m(\mathbf{u}_i),\beta_{01}(\mathbf{u}_i),\cdots,\beta_{0m}(\mathbf{u}_i),i=1,2,\cdots,n\}$ and the set of the parameters under $H_1$ is $\Omega_1=\{\alpha(\mathbf{u}_i),\boldsymbol{\gamma}^T(\mathbf{u}_i),\boldsymbol{\beta}_1^T(\mathbf{u}_i),\boldsymbol{\beta}_2^T(\mathbf{u}_i),\cdots,\boldsymbol{\beta}_m^T(\mathbf{u}_i),i=1,2,\cdots,n\}$, that is the set of the parameters on the GWMWR model. Based on the LRT method, the test statistic for simultaneous test is Wilk's likelihood ratio statistic, which has the form:

$$G_2=2(L(\hat{\Omega}_1)-L(\hat{\omega})) \tag{13}$$

where $L(\hat{\Omega}_1)$ and $L(\hat{\omega})$ are the maximum of the natural logarithm of the likelihood function under $H_1$ and under $H_0$ respectively, where $L(\hat{\omega})$ is given by

$$L(\hat{\omega})=\sum_{q=1}^{4}L_q(\hat{\omega})\ \text{with}:\ \hat{\lambda}_{1j}=\sum_{k=1}^{m}(y_{kj})^{\hat{\gamma}_k(\mathbf{u}_i)/\hat{\alpha}(\mathbf{u}_i)}\exp[-\frac{\hat{\gamma}_k(\mathbf{u}_i)}{\hat{\alpha}(\mathbf{u}_i)}\hat{\beta}_{k0}(\mathbf{u}_i)]$$

$$;\ L_1(\hat{\omega})=\sum_{j=1}^{n}\sum_{k=1}^{m}\{\ln\hat{\gamma}_k(\mathbf{u}_i)-\ln\hat{\alpha}(\mathbf{u}_i)+(\frac{\hat{\gamma}_k(\mathbf{u}_i)}{\hat{\alpha}(\mathbf{u}_i)}-1)\ln y_{kj}-\frac{\hat{\gamma}_k(\mathbf{u}_i)}{\hat{\alpha}(\mathbf{u}_i)}\hat{\beta}_{k0}(\mathbf{u}_i)\}\ ;$$

$$L_2(\hat{\omega})=\sum_{j=1}^{n}(\hat{\alpha}(\mathbf{u}_i)-m)\ln\hat{\lambda}_{1j}\ ;$$

$$L_3(\hat{\omega})=-\sum_{j=1}^{n}\hat{\lambda}_{1j}^{\hat{\alpha}(\mathbf{u}_i)}\ \text{ and }\ L_4(\hat{\omega})=\sum_{j=1}^{n}\ln\hat{Q}_{2j}(\mathbf{u}_i)\ \text{ for }\ \hat{Q}_{2j}(\mathbf{u}_i)=\sum_{l=1}^{m}(-1)^{l+m}C(m,l,\hat{\alpha}(\mathbf{u}_i))\hat{\lambda}_{2j}^{(l-1)\hat{\alpha}(\mathbf{u}_i)}.$$

Furthermore, the $G_2$ statistic follows $\chi^2$ distribution with $mp$ degree of freedom.

Lastly, partial test for regression parameter aims to determine any significant parameters affecting the response variable. Form of the hypothesis as follows:

$H_0:\beta_{kv}(\mathbf{u}_i)=0$,

$H_1:\beta_{kv}(\mathbf{u}_i)\neq 0$ for $k=1,2,\cdots,m$ and $v=1,2,\cdots,p$.

Based on the asymptotic normality properties of the MLEs, the test statistic for partial test is given by Wald statistic, which has the form

$$Z=\frac{\hat{\beta}_{kv}(\mathbf{u}_i)}{SE(\hat{\beta}_{kv}(\mathbf{u}_i))}, \tag{14}$$

and it has approximately the standard normal distribution (Pawitan, 2001), [4]. Let $\hat{\beta}_{kv}(\mathbf{u}_i)$ is the $r^{th}$ component of the vector $\hat{\theta}(\mathbf{u}_i)$, then the standard error of $\hat{\beta}_{kv}(\mathbf{u}_i)$ is given by the estimated standard deviation $SE(\hat{\beta}_{kv}(\mathbf{u}_i)) = \sqrt{h^{rr}}$, which $h^{rr}$ is the $((k-1)(p+1)+v+m+2)^{th}$ diagonal term of $[I(\hat{\theta}(\mathbf{u}_i))]^{-1}$ for $k = 1, 2, \cdots, m$, and $v = 1, 2, \cdots, p$, with $I(\theta((\mathbf{u}_i)) = -E[\mathbf{H}(\theta(\mathbf{u}_i))]$ is the Fisher information matrix.

## 4. Conclusion

In this study is focused on the GWMWR model, which is used for continuous nonnegative data. The GWMWR model is the local model of the multivariate Weibull regression model, which the parameter estimation is done on every location where the data is collected. The analytical method for parameter estimation is MLE, by the weighting of geographical location factor. The maximum likelihood estimator can be obtained by solving the likelihood equation, which is the interdependent non-linear system, by using the Newton-Raphson iterative method. Hypothesis testing for regression parameters in the GWMWR model involves the test of goodness of fit, simultaneous test and partial test. The test statistics for goodness of fit and the simultaneous test are Wilk's likelihood ratio statistic, which are determined by LRT method, and the test statistic for partial test is Wald statistic, which is derived from the asymptotic normality property of the maximum likelihood estimator. Furthermore, for the next study, the convergence of the Newton-Raphson algorithm needs to be discussed for obtaining the initial value and evaluating the efficiency of method.

## References

[1] Fotheringham, A.S., C. Brunsdon, M. Charlton, 2002. Geographically Weighted Regression the analysis of spatially varying relationships. John Wiley and Sons Ltd. ISBN: 0-471-49616-2, 2-10, 56-62.

[2] Hanagal, D.D., 2005. Weibull Extension of a Bivariate Weibull Regression Model. Economic Quality Control, 20: No.2, 149-154.

[3] Lee, C.K. and M.J. Wen, 2009. A Multivariate Weibull Distribution. Pak.J.Stat.Operat. Res., 5, No. 2, 55-66.

[4] Pawitan, Y., 2001. In All Likelihood, Statistical Modelling and Inference Using Likelihood. Oxford University Press Inc., New York, ISBN: 0 19 850765 8, pp: 243-244, 256-259.

[5] Wahyudi, I., Purhadi, Irhamah and Sutikno, 2011. The Development of Parameter Estimation on Hazard Rate of Trivariate WeibullDistribution. American Journal of Biostatistics 2011, 2 (2) 26-35. DOI: 10.3844/ajbssp.2011.26.35.