Thailand Statistician
April 2021; 19(2): 294-307
http://statassoc.or.th
Contributed paper

Geographically Weighted Regression Model with Covariate Measurement Errors

Ida Mariati Hutabarat* and Yacob Ruru
Department of Mathematics, Faculty of Mathematics and Natural Sciences, Cenderawasih University, Jayapura, Indonesia.
*Corresponding author; e-mail: ida_mariati@yahoo.com

Received: 15 October 2019
Revised: 31 January 2020
Accepted: 6 April 2020

Abstract

In this paper, we determine the parameter estimators of the Geographically Weighted Regression (GWR) model with measurement errors through the mixed linear model approach. In contrast to model that does not pay attention to geographic location factors, GWR is very concerned about geographic location factors. This model will produce the parameter estimators of the local model for each point or location where data is collected. The mixed linear model approach in the GWR model is form a weighting matrix for each observation location. The estimation method used to estimate parameters in the mixed linear model is Restricted Maximum Likelihood (REML) method. Asymptotic normality properties of the estimators are obtained. The estimators are shown to be consistent. From the estimation results obtained, we illustrate the data of malnutrition sufferers in East Java province.

Keywords: Mixed linear model, restricted maximum likelihood method, asymptotic normality properties, weighting matrix.

1. Introduction

Measurement error is an error that arises when a recorded value is not exactly the same as the actual value in relation to a measurement process. The actual value of the covariate is represented by a value obtained through a measurement process that is not necessarily in accordance with the actual value. For example in economic applications related to income issues, if the interviewee could not accurately state their income, the results of the research records would be higher or lower than the actual value. Another example in medicine field was systolic blood pressure (SBP) measurement. During SBP measurement, various sources of errors could occur such as recording equipment errors and administrative errors (Carroll et al. 2006). The presence of measurement errors causes parameter estimators to be biased and inconsistent, which leads to incorrect conclusions (Chen et al. 2011). In overcoming this problem, measurement error models were used.

There were several important sources that may cause measurement errors. Biemer et al. (1991) identified four main sources of measurement errors, namely questionnaire design, data, interviewer and respondent. Based on the estimation of regression curve parameters, Fuller (1987) stated that measurement errors in the slope of the regression curve (Fuller 1987). Moreover, Caroll et al. (2006)

affirmed that errors may cause bias in regression estimators and also lead to not exactly models (Carroll et al. 2006). The presence of measurement errors causes the parameters to be biased and inconsistent (Chen et al. 2011). In solving this problem is a model of measurement error.

The researches on parametric regression models with errors are Carroll et al. (1996), Fuller and Hidiroglou (1978). The results obtained namely measurement error will affect the regression coefficient. In simple regression, the magnitude of the regression coefficient is greater than if calculated without using a measurement error model. This is due to a correction factor of various errors. Nonparametric regression model with measurement error has been developed by Fan and Truong (1993) which predicts the parameter of measurement error model with kernel deconvolution method. In addition, Carroll et al. (1999) had conducted a study to estimate parameters in the measurement error model using the modified spline method. Several studies that have been conducted on nonlinear regression models with measurement errors include: Stefanski and Carroll (1985), in the logistic regression model, namely the process of cardiac development. Stefanski (1987) conducted a study on Generalized Linear Models (GLM). Nakamura (1992) suspected the parameters in the proportional Hazard model.

Spatial effects are common among regions. In some cases, the independent variables observed were related to observations in different regions, especially adjacent areas. The existence of a spatial relationship in the dependent variable will cause the estimation to be inaccurate because the error randomness assumption is violated. To overcome the above problems, we need a regression model that incorporates spatial relationships between regions into the model. The existence of spatial relationship information between regions causes the need to accommodate spatial diversity in the model, so the model used is a spatial regression model.

The according to Li et al. (2009) spatial data are susceptible to measurement errors in covariates. Research for spatial regression models with measurement errors began to develop, because in its application there are variables that cannot be measured precisely which have a spatial effect. In the mixed linear spatial model, Li et al. (2009) used the conditional auto-regressive model (CAR). In addition to CAR, one method for analyzing spatial data is a geographically weighted regression model (GWR) (Fotheringham et al. 2002). GWR is the development of a classical linear regression model. In the linear regression model produces estimator parameters that apply globally, but in the GWR model, a parameter estimator of the model is produced which is local for each observation location.

Differences from previous research; Zare et al. (2012) studied a linear mixed model with measurement errors in effects, they did not included location factors so that the resulting parameter estimators are global. In this paper we will discuss the linear approach to the GWR model with measurement errors. In contrast to models that do not pay attention to location factors, GWR model is very concerned about the location so as to produce estimators of local model parameters for each point or location where the data is collected.

The basic and necessary component of the spatial regression model is the weighting matrix. The weighting matrix is basically a matrix that describes relationships between regions. This matrix reflects the relationship between one location and another (Arbia 2006). The weighting matrix depends on the proximity between the observation locations. The closer a location is, the greater the weight will be. Some literature can be used to determine the amount of weighting for each different location. According to Fotheringham et al. (2002), spatial weighting functions that can be used in GWR are adaptive spatial kernel a bi-square and biweight.

Our study concentrates on determine the parameter estimators of the Geographically Weighted Regression (GWR) model with measurement errors through the mixed linear model approach. This paper is organized as follows: In Section 2, we present the model definition. In Sections 3 and 4, determine the estimated parameters and the algorithm. In Section 5, we state their asymptotic properties for the estimates. A real data example is presented in Section 6. Finally, in Section 7, we present our summary.

2. Model Definition

Mixed linear model with measurement error in covariate with fixed effect can be written as follows:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Ub} + \varepsilon$$

$$\boldsymbol{Z} = \boldsymbol{X} + \boldsymbol{t} \tag{1}$$

In this model, $\boldsymbol{y}$ is an $n \times 1$ vector of random variables whose observed; $\boldsymbol{X}$ is a matrix, which is the fixed effect with dimension $n \times p$ and $\boldsymbol{U}$ is a matrix which is the random effect with dimension $n \times q$; $\boldsymbol{\beta}$ is an $p \times 1$ vector of parameter, which is the fixed effect; $\boldsymbol{b}$ is an $q \times 1$ vector of unobservable random effect with $\boldsymbol{b} \sim N(\boldsymbol{0}, \sigma^2 \sum_b)$; ε is an $n \times 1$ vector of random errors with $\boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{R})$. Variable $\boldsymbol{Z}$ is the observed value of $\boldsymbol{X}$ with the measurement error $\boldsymbol{t}$, where $\boldsymbol{t}$ is a matrix of size $n \times p$ of the distribution of $N\left(0, I \otimes \Lambda\right)$ where $\otimes$ is a croneker and Λ is a known block matrix with dimensions $np \times np$. We assume that $\boldsymbol{b}$, $\boldsymbol{\varepsilon}$ and $\boldsymbol{t}$ are mutually independent.

Let in model (1), covariate $\boldsymbol{X}$ is measured with error and correlation structure derived from random effect. If we replaced $\boldsymbol{X}$ with $\boldsymbol{Z}$, then the estimation derived from score function is generally inconsistent. Some methods are proposed in the measurement error model. In this article, using score corrected method is adopted from Nakamura (1990) as a general approach in the measurement error model. In this method, it is define corrected score function where its expected value related to the distribution of measurement error with score function based on covariate $\boldsymbol{t}$ which is already known.

The first step from mixed linier model approach in GWR model is to form weighted matrix for each location observed. Let $\left(\boldsymbol{W}(u_i, v_i)\right)$ is a spatial weighted matrix of location i which its diagonal element values defined by location i with other location (location j).

The model (1) becomes

$$\boldsymbol{y}^* = \boldsymbol{W}(u_i, v_i)\boldsymbol{X\beta} + \boldsymbol{W}(u_i, v_i)\boldsymbol{Ub} + \boldsymbol{\varepsilon}, \tag{2a}$$

$$\boldsymbol{Z} = \boldsymbol{X} + \boldsymbol{t}, \tag{2b}$$

$\boldsymbol{W}\left(u_i, v_i\right)$ is an $n \times n$ matrix whose off-diagonal elements are zero and whose diagonal elements denote the geographical weighting of each of the n observed data for location i of the form

$$\boldsymbol{W}(u_i, v_i) = \begin{bmatrix} w(u_1, v_1) & 0 & \cdots & 0 \\ 0 & w(u_2, v_2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & w(u_n, v_n) \end{bmatrix}.$$

Hence the weighting matrix has to be computed for each point i and the weights depict the proximity of each data point to the location of i with points in closer proximity carrying more weight in the estimation of the parameters for location i.

3. Parameter Estimation

One of the estimation methods to estimate the parameter in the mixed linier model is restricted maximum likelihood (REML). If $\boldsymbol{\beta}$ is a fixed effect parameter, expected value and matrix covariance for $\boldsymbol{y}^*$ from (2a) is $E(\boldsymbol{y}^*) = \boldsymbol{WX\beta}$ and $\operatorname{var}(\boldsymbol{y}^*) = \sigma^2\boldsymbol{V}$, where

$$\boldsymbol{V} = \boldsymbol{R} + \boldsymbol{WU}\sum\nolimits_{\boldsymbol{b}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W} = \boldsymbol{R} + \sum_{i=1}^{m}\gamma_i\boldsymbol{WU}_i\boldsymbol{U}_i^T\boldsymbol{W}.$$

Variable $\boldsymbol{y}^*$ has normally distributed, i.e. $\boldsymbol{y}^* \sim N\left(\boldsymbol{WX\beta}, \sigma^2\boldsymbol{V}\right)$, so that log-likelihood function from $\boldsymbol{y}^*$ based on above distribution is

$$l\left(\boldsymbol{\beta},\boldsymbol{\theta}\middle|\boldsymbol{X},\boldsymbol{y}^*\right) = -\frac{n}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2}\log\left(|\boldsymbol{V}|\right) - \frac{1}{2\sigma^2}\left[\left(\boldsymbol{y}^* - \boldsymbol{WX\beta}\right)^{\mathrm{T}}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^* - \boldsymbol{WX\beta}\right)\right],$$

where $\boldsymbol{\theta} = \left(\sigma^2, \boldsymbol{\gamma}\right) = \left(\sigma^2, \gamma_1, \ldots, \gamma_m\right)$ is an element in $\Omega = \left\{\boldsymbol{\theta}; \sigma^2 > 0, \gamma_i \geq 0; (i = 1, \ldots, m)\right\}$.

Conditional distribution of $\boldsymbol{b}|\boldsymbol{y}^*$ is $\boldsymbol{b}|\boldsymbol{y}^* \sim N\left(\sum\nolimits_{\boldsymbol{b}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{WV}^{-1}\left(\boldsymbol{y}^* - \boldsymbol{WX\beta}\right), \sigma^2\sum\nolimits_{\boldsymbol{b}}\boldsymbol{T}\right)$ with $\boldsymbol{T} = \left(\boldsymbol{R} + \boldsymbol{WU}\sum\nolimits_{\boldsymbol{b}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\right)^{-1}$. The probability log function from $\boldsymbol{b}|\boldsymbol{y}^*$ is

$$l_{\boldsymbol{b}}\left(\boldsymbol{\beta},\boldsymbol{\theta}\middle|\boldsymbol{X},\boldsymbol{y}^*\right) = -\frac{q}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2}\log\left(\left|\sum\nolimits_{\boldsymbol{b}}\boldsymbol{T}\right|\right)$$
$$-\frac{1}{2\sigma^2}\left[\left(\boldsymbol{b} - \sum\nolimits_{\boldsymbol{b}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{WV}^{-1}\left(\boldsymbol{y}^* - \boldsymbol{WX\beta}\right)\right)^{\mathrm{T}}\left(\sum\nolimits_{\boldsymbol{b}}\boldsymbol{T}\right)^{-1}\left(\boldsymbol{b} - \sum\nolimits_{\boldsymbol{b}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{WV}^{-1}\left(\boldsymbol{y}^* - \boldsymbol{WX\beta}\right)\right)\right].$$

i) Predictor of fixed effect and random effect

Suppose E represents the conditional expectation value $\boldsymbol{Z}$ if $\boldsymbol{y}^*$ is known. The probability corrected function $l^*\left(\boldsymbol{\beta},\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^*\right)$ must satisfy

$$E\left[\partial l^*\left(\boldsymbol{\beta},\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^*\right)/\partial\boldsymbol{\beta}\right] = \partial l\left(\boldsymbol{\beta},\boldsymbol{\theta}\middle|\boldsymbol{X},\boldsymbol{y}^*\right)/\partial\boldsymbol{\beta}$$
$$E\left[\partial l_1^*\left(\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^*\right)/\partial\sigma^2\right] = \partial l_1\left(\boldsymbol{\theta}\middle|\boldsymbol{X},\boldsymbol{y}^*\right)/\partial\sigma^2$$

and

$$E\left[\partial l_1^*\left(\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^*\right)/\partial\gamma_i\right] = \partial l_1\left(\boldsymbol{\theta}\middle|\boldsymbol{X},\boldsymbol{y}^*\right)/\partial\gamma_i,\ i = 1,\ldots,m,$$

with $l_1\left(\boldsymbol{\theta}\middle|\boldsymbol{X},\boldsymbol{y}^*\right) = l\left(\hat{\boldsymbol{\beta}}(\boldsymbol{\gamma}),\boldsymbol{\theta}\middle|\boldsymbol{X},\boldsymbol{y}^*\right)$, where $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}(\boldsymbol{\gamma})$ is a maximum probability estimation of $\boldsymbol{\beta}$ and $l_1^*\left(\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^*\right) = l^*\left(\hat{\boldsymbol{\beta}}(\boldsymbol{\gamma}),\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^*\right)$ with $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}(\boldsymbol{\gamma})$ is the solution of the equation of $\partial l^*\left(\boldsymbol{\beta},\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^*\right)/\partial\boldsymbol{\beta} = 0.$ The probability corrected function $l_{\boldsymbol{b}}^*\left(\boldsymbol{\beta},\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^*\right)$ is $E\left[\partial l_{\boldsymbol{b}}^*\left(\boldsymbol{\beta},\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^*\right)/\partial\boldsymbol{b}\right] = \partial l_{\boldsymbol{b}}\left(\boldsymbol{\beta},\boldsymbol{\theta}\middle|\boldsymbol{X},\boldsymbol{y}^*\right)/\partial\boldsymbol{b}.$

Theorem 1 *If* $\boldsymbol{t}$ *is a random vector* $n \times p$, *and* $\boldsymbol{V} = \boldsymbol{R} + \boldsymbol{WU}\sum\nolimits_{\boldsymbol{b}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}$, *then*

$$E\left(\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{WV}^{-1}\boldsymbol{WZ}\right) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{WV}^{-1}\boldsymbol{WX} + \operatorname{tr}\left(\boldsymbol{V}^{-1}\right)\boldsymbol{\Lambda}. \qquad (3)$$

Proof: By the substitution of the $\boldsymbol{Z}=\boldsymbol{X}+\boldsymbol{t}$ from (1), then

$$\begin{aligned} E\left(\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{Z}\right) &= E^{*}\left((\boldsymbol{X}+\boldsymbol{t})^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}(\boldsymbol{X}+\boldsymbol{t})\right) \\ &= E^{*}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}+\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{t}+\boldsymbol{t}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}+\boldsymbol{t}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{t}\right) \\ &= E^{*}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}\right)+E^{*}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{t}\right)+E^{*}\left(\boldsymbol{t}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}\right)+E^{*}\left(\boldsymbol{t}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{t}\right) \\ &= \boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}+\mathrm{tr}\left(\boldsymbol{V}^{-1}\right)\boldsymbol{\Lambda}. \end{aligned}$$

Using (3), it is obtained l^{*} and $l_{\boldsymbol{b}}^{*}$ as follows

$$\begin{aligned} l^{*}\left(\boldsymbol{\beta},\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^{*}\right) = &-\frac{n}{2}\log\left(2\pi\sigma^{2}\right)-\frac{1}{2}\log\left(\left|\boldsymbol{V}\right|\right) \\ &-\frac{1}{2\sigma^{2}}\left\{\left(\boldsymbol{y}^{*}-\boldsymbol{W}\boldsymbol{Z}\boldsymbol{\beta}\right)^{\mathrm{T}}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^{*}-\boldsymbol{W}\boldsymbol{Z}\boldsymbol{\beta}\right)-\mathrm{tr}\left(\boldsymbol{V}^{-1}\right)\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{\beta}\right\} \end{aligned} \tag{4}$$

and

$$\begin{aligned} l_{\boldsymbol{b}}^{*}\left(\boldsymbol{\beta},\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^{*}\right) = &-\frac{q}{2}\log\left(2\pi\sigma^{2}\right)-\frac{1}{2}\log\left(\left|\textstyle\sum_{\boldsymbol{b}}\boldsymbol{T}\right|\right)-\frac{1}{2\sigma^{2}}\left\{\left(\boldsymbol{b}-\textstyle\sum_{\boldsymbol{b}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^{*}-\boldsymbol{W}\boldsymbol{Z}\boldsymbol{\beta}\right)\right)^{\mathrm{T}}\left(\textstyle\sum_{\boldsymbol{b}}\boldsymbol{T}\right)^{-1}\right. \\ &\left.\left(\boldsymbol{b}-\textstyle\sum_{\boldsymbol{b}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^{*}-\boldsymbol{W}\boldsymbol{Z}\boldsymbol{\beta}\right)\right)-\mathrm{tr}(\boldsymbol{I}-\boldsymbol{V}^{-1})\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{\beta}\right\}. \end{aligned} \tag{5}$$

If in (4) is derived to $\boldsymbol{\beta}$, then the results are equated with zero, then we obtained

$$\begin{aligned} &\frac{\partial l^{*}\left(\boldsymbol{\beta},\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^{*}\right)}{\partial\boldsymbol{\beta}}=\boldsymbol{0} \\ &-\frac{1}{2\sigma^{2}}\left[-2\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^{*}+2\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{Z}\boldsymbol{\beta}-2\mathrm{tr}(\boldsymbol{V}^{-1})\boldsymbol{\Lambda}\boldsymbol{\beta}\right]=\boldsymbol{0} \\ &\left[\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{Z}-\mathrm{tr}(\boldsymbol{V}^{-1})\boldsymbol{\Lambda}\right]\boldsymbol{\beta}=\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^{*} \\ &\boldsymbol{\beta}=\left(\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{Z}-\mathrm{tr}(\boldsymbol{V}^{-1})\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^{*}. \end{aligned}$$

So, estimating the corrected score for $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}}_{c}=\left(\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{Z}-\mathrm{tr}(\boldsymbol{V}^{-1})\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^{*}. \tag{6}$$

If in (5) is derived to $\boldsymbol{b}$, then the results is equated with zero, then we obtained

$$\begin{aligned} &\frac{\partial l_{\boldsymbol{b}}^{*}\left(\boldsymbol{\beta},\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^{*}\right)}{\partial\boldsymbol{b}}=\boldsymbol{0} \\ &-\frac{1}{2\sigma^{2}}\left[2\left(\textstyle\sum_{\boldsymbol{b}}\boldsymbol{T}\right)^{-1}\boldsymbol{b}-2\left(\textstyle\sum_{\boldsymbol{b}}\boldsymbol{T}\right)^{-1}\textstyle\sum_{\boldsymbol{b}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^{*}-\boldsymbol{W}\boldsymbol{Z}\hat{\boldsymbol{\beta}}\right)\right]=\boldsymbol{0} \\ &\left[\left(\textstyle\sum_{\boldsymbol{b}}\boldsymbol{T}\right)^{-1}\boldsymbol{b}=\left(\textstyle\sum_{\boldsymbol{b}}\boldsymbol{T}\right)^{-1}\textstyle\sum_{\boldsymbol{b}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^{*}-\boldsymbol{W}\boldsymbol{Z}\boldsymbol{\beta}\right)\right] \\ &\qquad\boldsymbol{b}=\textstyle\sum_{\boldsymbol{b}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^{*}-\boldsymbol{W}\boldsymbol{Z}\boldsymbol{\beta}\right). \end{aligned}$$

So, estimating the corrected score for $\boldsymbol{b}$ is:

$$\hat{\boldsymbol{b}}_{c}=\textstyle\sum_{\boldsymbol{b}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^{*}-\boldsymbol{W}\boldsymbol{Z}\hat{\boldsymbol{\beta}}_{c}\right)=\textstyle\sum_{\boldsymbol{b}}\tilde{\boldsymbol{\upsilon}}_{c}\ \text{ with }\ \boldsymbol{\upsilon}_{c}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^{*}-\boldsymbol{W}\boldsymbol{Z}\hat{\boldsymbol{\beta}}_{c}\right). \tag{7}$$

ii) Variant component estimator

The probability corrected log function for estimation of $\boldsymbol{\theta}$ is

$$l_1^*\left(\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^*\right)=-\frac{n}{2}\log\left(2\pi\sigma^2\right)-\frac{1}{2}\log\left(\left|\boldsymbol{V}\right|\right) \\ -\frac{1}{2\sigma^2}\left[\left(\boldsymbol{y}^*-\boldsymbol{WZ}\hat{\boldsymbol{\beta}}_c\right)^{\mathrm{T}}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^*-\boldsymbol{WZ}\hat{\boldsymbol{\beta}}_c\right)-\mathrm{tr}(\boldsymbol{V}^{-1})\hat{\boldsymbol{\beta}}_{\mathrm{c}}^{\mathrm{T}}\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}_c\right]. \quad (8)$$

Estimating the corrected score is a value of vector $\boldsymbol{\theta}\in\boldsymbol{\Omega}$ which maximises l_1^*. If in (8) is derived to σ^2, then the results is equated to zero, then we obtained

$$\frac{\partial l_1^*\left(\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^*\right)}{\partial\sigma^2}=0$$

$$-\frac{n}{2\sigma^2}+\frac{1}{2(\sigma^2)^2}\left[\left(\boldsymbol{y}^*-\boldsymbol{WZ}\hat{\boldsymbol{\beta}}_c\right)^{\mathrm{T}}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^*-\boldsymbol{WZ}\hat{\boldsymbol{\beta}}_c\right)-\mathrm{tr}(\boldsymbol{V}^{-1})\hat{\boldsymbol{\beta}}_{\mathrm{c}}^{\mathrm{T}}\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}_c\right]=0$$

$$n\sigma^2=\left[\left(\boldsymbol{y}^*-\boldsymbol{WZ}\hat{\boldsymbol{\beta}}_c\right)^{\mathrm{T}}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^*-\boldsymbol{WZ}\hat{\boldsymbol{\beta}}_c\right)-\mathrm{tr}(\boldsymbol{V}^{-1})\hat{\boldsymbol{\beta}}_{\mathrm{c}}^{\mathrm{T}}\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}_c\right]$$

$$\sigma^2=\frac{1}{n}\left[\left(\boldsymbol{y}^*-\boldsymbol{WZ}\hat{\boldsymbol{\beta}}_c\right)^{\mathrm{T}}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^*-\boldsymbol{WZ}\hat{\boldsymbol{\beta}}_c\right)-tr(\boldsymbol{V}^{-1})\hat{\boldsymbol{\beta}}_{\mathrm{c}}^{\mathrm{T}}\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}_c\right].$$

So, the estimating corrected score for σ^2 is

$$\hat{\sigma}_c^2=\frac{1}{n}\left[\left(\boldsymbol{y}^*-\boldsymbol{WZ}\hat{\boldsymbol{\beta}}_c\right)^{\mathrm{T}}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^*-\boldsymbol{WZ}\hat{\boldsymbol{\beta}}_c\right)-\mathrm{tr}\left(\boldsymbol{V}^{-1}\right)\hat{\boldsymbol{\beta}}_{\mathrm{c}}^{\mathrm{T}}\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}_c\right]. \quad (9)$$

By using relationship $\left|\boldsymbol{V}\right|=\left|\boldsymbol{R}\right|\left|\boldsymbol{I}+\boldsymbol{U}^{\mathrm{T}}\boldsymbol{WR}^{-1}\boldsymbol{WU}\sum_{\boldsymbol{b}}\right|$ from (8) it is obtained,

$$l_1^*\left(\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^*\right)=-\frac{n}{2}\log\left(2\pi\sigma^2\right)-\frac{1}{2}\log\left(\left|\boldsymbol{R}\right|\right)-\frac{1}{2}\log\left(\left|\boldsymbol{I}+\boldsymbol{U}^{\mathrm{T}}\boldsymbol{WR}^{-1}\boldsymbol{WU}\sum_{\boldsymbol{b}}\right|\right) \\ -\frac{1}{2\sigma^2}\left[\left(\boldsymbol{y}^*-\boldsymbol{WZ}\hat{\boldsymbol{\beta}}_c\right)^{\mathrm{T}}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^*-\boldsymbol{WZ}\hat{\boldsymbol{\beta}}_c\right)-\mathrm{tr}(\boldsymbol{V}^{-1})\hat{\boldsymbol{\beta}}_{\mathrm{c}}^{\mathrm{T}}\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}_c\right]. \quad (10)$$

If $\partial\boldsymbol{V}/\partial\gamma_i=\boldsymbol{WU}_i\boldsymbol{U}_i^T\boldsymbol{W}$, $\partial\boldsymbol{V}^{-1}/\partial\gamma_i=-\boldsymbol{V}^{-1}\boldsymbol{WU}_i\boldsymbol{U}_i^T\boldsymbol{WV}^{-1}$ and $\partial\sum/\partial\gamma_i=\mathrm{diag}\left(0,\ldots,0,I_{qi},0,\ldots,0\right)$ and in (10) is derived to $\sigma^2\gamma_i$, then the result is equated to zero, then we obtained

$$\frac{\partial l_1^*\left(\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^*\right)}{\partial\gamma_i}=0,\quad i=1,\ldots,m$$

$$\frac{\partial l_1^*\left(\boldsymbol{\theta}\middle|\boldsymbol{Z},\boldsymbol{y}^*\right)}{\partial\gamma_i}=-\frac{1}{2}\mathrm{tr}\left[\left(\boldsymbol{I}+\boldsymbol{U}^{\mathrm{T}}\boldsymbol{WR}^{-1}\boldsymbol{WU}\sum_{\boldsymbol{b}}\right)\boldsymbol{U}^{\mathrm{T}}\boldsymbol{WR}^{-1}\boldsymbol{WU}\left(\partial\sum_{\boldsymbol{b}}/\partial\gamma_i\right)\right] \\ +\frac{1}{2\sigma^2}\left[\left(\boldsymbol{y}^*-\boldsymbol{WZ}\hat{\boldsymbol{\beta}}_c\right)^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{WU}_i\boldsymbol{U}_i^T\boldsymbol{WV}^{-1}\left(\boldsymbol{y}^*-\boldsymbol{WZ}\hat{\boldsymbol{\beta}}_c\right)-\mathrm{tr}\left(\boldsymbol{V}^{-1}\boldsymbol{WU}_i\boldsymbol{U}_i^T\boldsymbol{WV}^{-1}\right)\hat{\boldsymbol{\beta}}_{\mathrm{c}}^{\mathrm{T}}\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}_c\right] \\ =\frac{1}{2}\gamma_i^{-1}\left(q_i-\mathrm{tr}(T_{ii})\right)+\frac{1}{2\sigma^2}\left[\hat{\boldsymbol{b}}_{ic}^T\hat{\boldsymbol{b}}_{ic}-\gamma_i^{-2}\mathrm{tr}(\boldsymbol{D}_i^T\boldsymbol{D}_i)\hat{\boldsymbol{\beta}}_c^T\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}_c\right]=0. \quad (11)$$

From (11), the estimating corrected score of $\sigma_1^2,\ldots,\sigma_m^2$ is

$$\hat{\sigma}_{ic}^2=\frac{1}{q_i-\mathrm{tr}(T_{ii})}\left[\hat{\boldsymbol{b}}_{ic}^T\hat{\boldsymbol{b}}_c-\mathrm{tr}\left(\hat{\boldsymbol{D}}_i^T\hat{\boldsymbol{D}}_i\right)\hat{\boldsymbol{\beta}}_c^T\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}_c\right],\quad i=1,\ldots,m,$$

where $\boldsymbol{D}_i=\hat{\gamma}_{ic}\boldsymbol{U}_i^T\boldsymbol{WV}^{-1}=\left(\hat{\sigma}_{ic}^2/\hat{\sigma}_c^2\right)\boldsymbol{U}_i^T\boldsymbol{WV}^{-1}$.

4. Algorithm of Parameter Estimation

The following theorem is an extension adopted from Harvill (1977) and Fellner (1986) for linear mixed models with measurement errors.

Theorem 2 *If $\hat{\boldsymbol{\beta}}_c$ and $\tilde{\boldsymbol{\upsilon}}_c$ are the $p\times 1$ and $q\times 1$ components of any solution to the linear system*

$$\begin{bmatrix} \boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\boldsymbol{W}\boldsymbol{Z}-\mathrm{tr}(\boldsymbol{V}^{-1})\boldsymbol{\Lambda} & \boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\boldsymbol{W}\boldsymbol{U}\sum_{\boldsymbol{b}} \\ \boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\boldsymbol{W}\boldsymbol{Z} & \boldsymbol{I}+\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\boldsymbol{W}\boldsymbol{U}\sum_{\boldsymbol{b}} \end{bmatrix}\begin{bmatrix}\tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\upsilon}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\boldsymbol{y}^* \\ \boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\boldsymbol{y}^*\end{bmatrix}, \tag{12}$$

is $\tilde{\boldsymbol{\beta}}=\hat{\boldsymbol{\beta}}_c$ and $\tilde{\boldsymbol{\upsilon}}=\tilde{\boldsymbol{\upsilon}}_c$, which $\hat{\boldsymbol{\beta}}_c$ and $\tilde{\boldsymbol{\upsilon}}_c$ each is given in (6) and (7).

Proof:

$$\left(\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\boldsymbol{W}\boldsymbol{Z}-\mathrm{tr}(\boldsymbol{V}^{-1})\boldsymbol{\Lambda}\right)\tilde{\boldsymbol{\beta}}+\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\boldsymbol{W}\boldsymbol{U}\sum_{\boldsymbol{b}}\tilde{\boldsymbol{\upsilon}}=\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\boldsymbol{y}^* \tag{13}$$

$$\begin{aligned}&\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\boldsymbol{W}\boldsymbol{Z}\tilde{\boldsymbol{\beta}}+\left(\boldsymbol{I}+\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\boldsymbol{W}\boldsymbol{U}\sum_{\boldsymbol{b}}\right)\tilde{\boldsymbol{\upsilon}}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\boldsymbol{y}^* \\ &\tilde{\boldsymbol{\upsilon}}=\left(\boldsymbol{I}+\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\boldsymbol{W}\boldsymbol{U}\sum_{\boldsymbol{b}}\right)^{-1}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\left(\boldsymbol{y}^*-\boldsymbol{W}\boldsymbol{Z}\tilde{\boldsymbol{\beta}}\right).\end{aligned} \tag{14}$$

By substituting $\tilde{\boldsymbol{\upsilon}}_c$ to (13) we obtained

$$\left(\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{Z}-\mathrm{tr}(\boldsymbol{V}^{-1})\boldsymbol{\Lambda}\right)\tilde{\boldsymbol{\beta}}=\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^*$$

$$\tilde{\boldsymbol{\beta}}=\left(\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{Z}-\mathrm{tr}(\boldsymbol{V}^{-1})\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^*.$$

Corollary 1 *Based on Theorem 2, that is $\tilde{\boldsymbol{\upsilon}}=\tilde{\boldsymbol{\upsilon}}_c$, then we obtained*

$$\tilde{\boldsymbol{\upsilon}}_c=\left(\boldsymbol{I}+\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\boldsymbol{W}\boldsymbol{U}\sum_{\boldsymbol{b}}\right)\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}^{-1}\left(\boldsymbol{y}^*-\boldsymbol{W}\boldsymbol{Z}\hat{\boldsymbol{\beta}}_c\right).$$

An iterative algorithm is needed to calculate the corrected score estimation θ. Estimation steps are as follows (Fellner 1986):

Step 1 Perform an iteration from $t=0$ in $\sigma^{2(0)}$ and $\sigma_i^{2(0)}, i=1,\ldots,m$.

Step 2 Calculate estimator $\hat{\boldsymbol{\beta}}_c^{(t)}$ and $\hat{\boldsymbol{b}}_{1c}^{(t)},\ldots,\hat{\boldsymbol{b}}_{mc}^{(t)}$ as a linear equation (12).

Step 3 Calculate

$$\hat{\sigma}_c^{2(t+1)}=\frac{1}{n}\left[\left(\boldsymbol{y}^*-\boldsymbol{W}\boldsymbol{Z}\hat{\boldsymbol{\beta}}_c^{(t)}\right)^{\mathrm{T}}\boldsymbol{V}^{-1(t)}\left(\boldsymbol{y}^*-\boldsymbol{W}\boldsymbol{Z}\hat{\boldsymbol{\beta}}_c^{(t)}\right)-\mathrm{tr}(\boldsymbol{V}^{-1(t)})\hat{\boldsymbol{\beta}}_c^{(t)T}\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}_c^{(t)}\right],$$

and

$$\hat{\sigma}_{ic}^{2(t+1)}=\frac{1}{q_i-tr\left(T_{ii}\right)}\left[\hat{\boldsymbol{b}}_{ic}^{(t)T}\hat{\boldsymbol{b}}_{ic}^{(t)}-tr\left(\hat{\boldsymbol{D}}_i^{(t)T}\hat{\boldsymbol{D}}_i^{(t)}\right)\hat{\boldsymbol{\beta}}_c^{(t)T}\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}_c^{(t)}\right].$$

Step 4 If it is convergent, specify $\hat{\sigma}_c^2=\hat{\sigma}_c^{2(t+1)}$ and $\hat{\sigma}_{ic}^2=\hat{\sigma}_{ic}^{2(t+1)}$.

Step 5 If it is not reach a converging parameter estimator, the Step 2 is done again until it reaches convergence.

5. Asymptotic Properties

In this section, the asymptotic properties of estimation will be examined. Please note that y^* components are not independent each other. All derivatives related to a function are assumed to exist and parameters can be identified. In this case it is assumed that as $n\to\infty$, the limit of the following functions exist: $n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}$, $n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{U}, n^{-1}\left(\boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{U}+\boldsymbol{\Sigma}_{\boldsymbol{b}}^{-1}\right), n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-2}\boldsymbol{W}\boldsymbol{X}, n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{A}_i\boldsymbol{W}\boldsymbol{X}$, $n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{A}_i\boldsymbol{V}\boldsymbol{A}_i\boldsymbol{W}\boldsymbol{X}$, $n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{A}_i^2\boldsymbol{W}\boldsymbol{X}$, $n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{A}_i\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}$, $n^{-1}\mathrm{tr}(\boldsymbol{A}_i)$, $n^{-1}\mathrm{tr}(\boldsymbol{V}^{-1})$, $n^{-1}\mathrm{tr}(\boldsymbol{V}^{-2})$, $n^{-1}\mathrm{tr}(\boldsymbol{A}_i^2)$, and $n^{-1}\mathrm{tr}(\boldsymbol{A}_i\boldsymbol{V}\boldsymbol{A}_i)$, where $\boldsymbol{A}_i=\boldsymbol{D}_i^{\mathrm{T}}\boldsymbol{D}_i$; $\boldsymbol{D}_i=\gamma_i\boldsymbol{U}_i^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}$.

Lemma 1 *Under the above conditions, we have*

$$\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{Z}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}+\mathrm{tr}(\boldsymbol{V}^{-1})\boldsymbol{\Lambda}+\boldsymbol{O}_p(n^{-1/2}). \tag{15}$$

Proof: Using $\boldsymbol{Z}=\boldsymbol{X}+\boldsymbol{t}$ from (2b), we have

$$n^{-1}\left\{\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{Z}-\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}-\mathrm{tr}(\boldsymbol{V}^{-1})\boldsymbol{\Lambda}\right\}=n^{-1}\left\{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{t}+\boldsymbol{t}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}+\mathbf{C}\right\},$$

where $\mathbf{C}=\boldsymbol{t}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{t}-\mathrm{tr}(\boldsymbol{V}^{-1})\boldsymbol{\Lambda}$. Since $\boldsymbol{t}\sim N(0,\boldsymbol{I}\otimes\boldsymbol{\Lambda})$, then

$$n^{-1/2}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{t}\sim N(0,n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-2}\boldsymbol{W}\boldsymbol{X}\otimes\boldsymbol{\Lambda}.$$

By assumption, the limit $n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-2}\boldsymbol{W}\boldsymbol{X}$ exists as $n\to\infty$, obtained $n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{t}=\boldsymbol{O}_p(n^{-1/2})$ and $n^{-1}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}=\boldsymbol{O}_p(n^{-1/2})$. Suppose the element $\mathbf{C}$ at the (a,b) is $\boldsymbol{C}_{ab}$. Then

$$\boldsymbol{C}_{ab}=\sum_{i=1}^{n}\sum_{i=1}^{n}\boldsymbol{t}_{ia}\boldsymbol{V}^{ij}\boldsymbol{t}_{jb}-\sum_{i=1}^{n}\boldsymbol{V}^{ii}\boldsymbol{\Lambda}_{ab},$$

where $\boldsymbol{t}=\left(\boldsymbol{t}_{ia}\right)$, $\boldsymbol{V}^{-1}=\boldsymbol{V}^{ij}$, $\boldsymbol{\Lambda}=\left(\boldsymbol{\Lambda}_{ab}\right)$, $i,j=1,2,\ldots,n,$ and $a,b=1,2,\ldots,p.$ $E\left(\boldsymbol{t}_{ia}\boldsymbol{t}_{ib}\right)=\boldsymbol{\Lambda}_{ab}$ and $E\left(\boldsymbol{C}_{ab}\right)=0.$ Furthermore,

$$\begin{aligned}E\left(\boldsymbol{C}_{ab}\right)^2&=\sum_{i,j}\sum_{k,l}E\left(\boldsymbol{t}_{ia}\boldsymbol{t}_{jb}\boldsymbol{t}_{ka}\boldsymbol{t}_{lb}\right)\boldsymbol{V}^{ij}\boldsymbol{V}^{kl}-\boldsymbol{\Lambda}_{ab}^2\left\{\mathrm{tr}(\boldsymbol{V}^{-1})\right\}^2\\&=\left(\boldsymbol{\Lambda}_{aa}\boldsymbol{\Lambda}_{bb}+\boldsymbol{\Lambda}_{ab}^2\right)\mathrm{tr}(\boldsymbol{V}^{-2}).\end{aligned}$$

By assumption the limit $n^{-1}\mathrm{tr}(\boldsymbol{V}^{-2})$ exists, $E\left(n^{-1/2}\boldsymbol{C}_{ab}\right)^2=\boldsymbol{O}_p(1)$ for $n\to\infty$ and $n^{(-1)}C=O_p(n^{(-1/2)})$. By combining all the results above, we get (15).

Theorem 3 $\hat{\boldsymbol{\beta}}_c$ *is asymptotically normally distributed. The asymptotic mean and variance of* $\hat{\boldsymbol{\beta}}_c$ *are respectively given as* $\boldsymbol{\beta}$ *at the (2a) and*

$$\mathrm{avar}(\hat{\boldsymbol{\beta}}_c)=\sigma^2\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}\right)^{-1}+\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}\right)^{-1}\boldsymbol{B}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}\right)^{-1}, \tag{16}$$

where $\boldsymbol{B}=\left\{\sigma^2\mathrm{tr}(\boldsymbol{V}^{-1})\boldsymbol{\beta}^T\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-2}\boldsymbol{W}\boldsymbol{X}\right)\boldsymbol{\beta}\right\}\boldsymbol{\Lambda}.$

Proof: Let $\boldsymbol{V}^{-1/2}=\boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Gamma}^{\mathrm{T}}$ denote the spectral decomposition of $\boldsymbol{V}^{-1/2}$, where $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}}=\boldsymbol{I}_n$, $\Phi=\mathrm{diag}\left(\lambda_1^{-1/2},\ldots,\lambda_n^{-1/2}\right)$ and λ_i are the eigen values of $\boldsymbol{V}$ and $\boldsymbol{\xi}=\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^*/\sqrt{n}$, then we have

$$\boldsymbol{\xi}=\frac{1}{\sqrt{n}}\boldsymbol{Z}^T\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^*=\frac{1}{\sqrt{n}}\boldsymbol{Z}^T\boldsymbol{W}\boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Gamma}^T\boldsymbol{V}^{-1/2}\boldsymbol{y}^*=\frac{1}{\sqrt{n}}\tilde{\boldsymbol{Z}}^T\boldsymbol{W}\boldsymbol{\Phi}\tilde{\mathbf{y}}^*,$$

where

$$\tilde{\mathbf{Z}}=\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{W}\mathbf{Z}\sim\mathrm{N}\left(\boldsymbol{\Gamma}^{\mathrm{T}}\mathbf{W}\mathbf{X},\boldsymbol{I}_n\otimes\boldsymbol{\Lambda}\right),\ \tilde{\mathbf{y}}^*=\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{V}^{-1/2}\boldsymbol{y}^*\sim\mathrm{N}\left(\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{V}^{-1/2}\boldsymbol{W}\boldsymbol{X}\boldsymbol{\beta},\sigma^2\boldsymbol{I}_n\right).$$

The element α^{th} of $\boldsymbol{\xi}$ is

$$\boldsymbol{\xi}_\alpha = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\boldsymbol{Z}}_{i\alpha}\lambda_i^{-1/2}\tilde{\boldsymbol{y}}_i^* = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\alpha_i.$$

Since α_i are independent and the limit of $\text{var}(\xi_\alpha)$ exists as $n \to \infty,$ by the central limit theorem, ξ_α is asymptotically normal. It follows from (6) and (15) that,

$$\begin{aligned}\hat{\boldsymbol{\beta}}_c &= \left\{n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X} + \boldsymbol{O}_p(n^{-1/2})\right\}^{-1} n^{-1}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^* \\ &= \left\{\boldsymbol{I}_p + \boldsymbol{O}_p(n^{-1/2})\right\}^{-1}\left(n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}\right)^{-1} n^{-1}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^* \\ &= \left\{\boldsymbol{I}_p + \boldsymbol{O}_p(n^{-1/2})\right\}\left(n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}\right)^{-1} n^{-1}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^*,\end{aligned}$$

where $I_p + O_p(n^{(-1/2)})^{(-1)} = I_p + O_p(n^{(-1/2)})$ is obtained from Taylor series expansion. So

$$\sqrt{n}\hat{\boldsymbol{\beta}}_c = \left\{\boldsymbol{I}_p + \boldsymbol{O}_p(n^{-1/2})\right\}\left(n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}\right)^{-1}\frac{1}{\sqrt{n}}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^*. \tag{17}$$

Moreover, since the limit of $n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}$ exists, and let $\boldsymbol{M} = n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X},$ then (17) can be written as

$$\begin{aligned}\sqrt{n}\hat{\boldsymbol{\beta}}_c &= \left(n^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}\right)^{-1}\frac{1}{\sqrt{n}}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^* + \boldsymbol{O}_p(n^{-1/2}) \\ &= \boldsymbol{M}^{-1}\boldsymbol{\xi} + \boldsymbol{O}_p(n^{-1/2}).\end{aligned} \tag{18}$$

From $E\left(\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^*\right) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}\boldsymbol{\beta}$ or $E(\boldsymbol{\xi}) = \sqrt{n}\boldsymbol{M}\boldsymbol{\beta}$ that $\sqrt{n}(\hat{\boldsymbol{\beta}}_c - \boldsymbol{\beta})$ is asymptotically normal with mean 0. To find the asymptotic variance of $\hat{\boldsymbol{\beta}}_c,$ from (18) is rewritten as follows

$$\begin{aligned}\sqrt{n}(\hat{\boldsymbol{\beta}}_c - \boldsymbol{\beta}) &= \boldsymbol{M}^{-1}\boldsymbol{\xi} - \boldsymbol{M}^{-1}\boldsymbol{M}\sqrt{n}\boldsymbol{\beta}_c + \boldsymbol{O}_p(n^{-1/2}) \\ &= \boldsymbol{M}^{-1}\left(\boldsymbol{\xi} - E(\boldsymbol{\xi})\right) + \boldsymbol{O}_p(n^{-1/2}).\end{aligned}$$

So we have $\text{var}(\sqrt{n}\hat{\boldsymbol{\beta}}_c) = \boldsymbol{M}^{-1}\,\text{var}(\boldsymbol{\xi})\boldsymbol{M}^{-1}.$ The variance of $\boldsymbol{\xi}$ is

$$\begin{aligned}\text{var}(\boldsymbol{\xi}) &= E^+\left\{\text{var}^*(\boldsymbol{\xi})\right\} + \text{var}^+\left\{E^*(\boldsymbol{\xi})\right\} \\ &= n^{-1}E^+\left\{\text{var}^*(\boldsymbol{\xi})\right\} + n^{-1}\,\text{var}^+\left\{E^*(\boldsymbol{\xi})\right\} \\ &= n^{-1}E^+\left(\boldsymbol{y}^{*T}\boldsymbol{V}^{-2}\boldsymbol{y}^*\boldsymbol{\Lambda}\right) + n^{-1}\sigma^2\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}\right).\end{aligned}$$

Since $E^+\left(\boldsymbol{y}^{*T}\boldsymbol{V}^{-2}\boldsymbol{y}^*\right) = \sigma^2\text{tr}(\boldsymbol{V}^{-1}) + \boldsymbol{\beta}^T\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-2}\boldsymbol{W}\boldsymbol{X}\right)\boldsymbol{\beta},$ so $\text{var}(\boldsymbol{\xi}) = n^{-1}\left\{\boldsymbol{B} + \sigma^2\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X}\right)\right\}$ whose limit exists as $n \to \infty.$ This completes the proof.

Theorem 3 is an extension (3) and (4) of Nakamura (1990) to the linear mixed models, and show that $\hat{\boldsymbol{\beta}}_c$ is consistent (see (18)).

Corollary 2 *Let* $\boldsymbol{\beta}$ *be given in (2a) the true value, then* $\hat{\boldsymbol{\beta}}_c$ *is consistent in probability and* $\hat{\boldsymbol{\beta}}_c - \boldsymbol{\beta} = \boldsymbol{O}_p(n^{-1/2}).$

Theorem 4 *Let* $\boldsymbol{b}$ *given in (2a), then* $\hat{\boldsymbol{b}}_c - \boldsymbol{b} = \boldsymbol{O}_p(n^{-1/2})$ *and it's asymptotically normally distributed with the asymptotic variance*

$$\text{avar}(\hat{\boldsymbol{b}}_c - \boldsymbol{b}) = M_1^{-1}M_2\text{avar}(\hat{\boldsymbol{\beta}})M_2^TM_1^{-1},$$

where $M_1 = n^{-1}\left(U^T U + \Sigma^{-1}\right)$ *and* $M_2 = n^{-1}U^T X,$ *whose limits exists by assumption.*

Proof: From (7),

$$\begin{aligned}\hat{\boldsymbol{b}}_c - \boldsymbol{b} &= -\left(U^T U + \Sigma^{-1}\right)^{-1} U^T Z(\hat{\boldsymbol{\beta}}_c - \boldsymbol{\beta}) \\ &= -\left\{n^{-1}\left(U^T U + \Sigma^{-1}\right)\right\}^{-1}\left\{n^{-1}U^T X + \boldsymbol{O}_p(n^{-1/2})\right\}(\hat{\boldsymbol{\beta}}_c - \boldsymbol{\beta}) \\ &= -M_1^{-1}M_2(\hat{\boldsymbol{\beta}}_c - \boldsymbol{\beta}) + \boldsymbol{O}_p(n^{-1})\end{aligned}$$

We use the result $U^T Z = U^T X + O_p(n^{-1/2})$ whose proof is similar to (15). Then from Theorem 3, we get the desired results.

Lemma 2 *Under conditions,* $n\hat{\sigma}_c^2$ *has asymptotic representation*

$$\boldsymbol{y}^{*\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y}^* - \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{M}^{-1}\boldsymbol{\xi} + \boldsymbol{O}_p(n^{1/2}),$$

where $\boldsymbol{\xi} = n^{-1/2}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^*$ *is asymptotically normal.*

Proof: Using (9), we have

$$\begin{aligned}\hat{\sigma}_c^2 &= \frac{1}{n}\left[\left(\boldsymbol{y}^* - \boldsymbol{W}\boldsymbol{Z}\hat{\boldsymbol{\beta}}_c\right)^{\mathrm{T}}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^* - \boldsymbol{W}\boldsymbol{Z}\hat{\boldsymbol{\beta}}_c\right) - \mathrm{tr}(\boldsymbol{V}^{-1})\hat{\boldsymbol{\beta}}_c^{\mathrm{T}}\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}_c\right] \\ &= \frac{1}{n}\left[\boldsymbol{y}^{*\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y}^* - \hat{\boldsymbol{\beta}}_c^T\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^*\right].\end{aligned}$$

Since $\left(\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{Z}\right) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{X} + \mathrm{tr}(\boldsymbol{V}^{-1})\boldsymbol{\Lambda} + \boldsymbol{O}_p(n^{-1/2}),$ then $n^{-1/2}\hat{\boldsymbol{\beta}}_c = \boldsymbol{M}^{-1}\boldsymbol{\xi} + \boldsymbol{O}_p(n^{-1/2})$ and $\hat{\boldsymbol{\beta}}_c\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^* = \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{M}^{-1}\boldsymbol{\xi} + \boldsymbol{O}_p(n^{1/2})$ and Lemma 2 will be proved.

Lemma 3 *For* $\hat{\sigma}_c^2,$ $E\left(\hat{\sigma}_c^2\right) = \sigma^2 + \boldsymbol{O}\left(\frac{1}{n}\right)$ *and* $\mathrm{var}\left(n\hat{\sigma}_c^2\right) = \boldsymbol{O}\left(\frac{1}{n}\right).$

Proof: Using Lemma 2 and the following relations the result is obtained

$$\begin{aligned}&E\left(\boldsymbol{\xi}\right) = \sqrt{n}\ \boldsymbol{M}\boldsymbol{\beta}_c, \\ &\mathrm{var}\left(\boldsymbol{\xi}\right) = \boldsymbol{B} + \sigma^2\ \boldsymbol{M}, \\ &E\left(\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{M}^{-1}\boldsymbol{\xi}\right) = p\sigma^2 + \mathrm{tr}(\boldsymbol{M}^{-1}\boldsymbol{B}) + n\boldsymbol{\beta}_c^T\boldsymbol{M}\boldsymbol{\beta}_c, \\ &\mathrm{var}\left(\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{M}^{-1}\boldsymbol{\xi}\right) = 2p\sigma^4 + 4\sigma^4\mathrm{tr}(\boldsymbol{M}^{-1}\boldsymbol{B}) + 2tr(\boldsymbol{M}^{-1}\boldsymbol{B})^2 + 4n\boldsymbol{\beta}_c^T\boldsymbol{B}\boldsymbol{\beta}_c + 4n\sigma^2\boldsymbol{\beta}_c^T\boldsymbol{M}\boldsymbol{\beta}_c, \\ &E\left(\boldsymbol{y}^{*\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y}^*\right) = n\sigma^2 + n\boldsymbol{\beta}_c^T\boldsymbol{M}\boldsymbol{\beta}_c, \\ &\mathrm{var}\left(\boldsymbol{y}^{*\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y}^*\right) = 2n\sigma^4 + 4n\sigma^2\boldsymbol{\beta}_c^T\boldsymbol{M}\boldsymbol{\beta}_c, \\ &\mathrm{cov}\left(\boldsymbol{y}^{*\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y}^*, \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{M}^{-1}\boldsymbol{\xi}\right) = 2p\sigma^4 + 4n\sigma^2\boldsymbol{\beta}_c^T\boldsymbol{M}\boldsymbol{\beta}_c.\end{aligned}$$

For Lemma 2,

$$
\begin{aligned}
E\left(n\hat{\sigma}_c^2\right) &= E\left(\boldsymbol{y}^{*\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y}^{*} - \xi^{\mathrm{T}}\boldsymbol{M}^{-1}\xi + \boldsymbol{O}_p\left(n^{1/2}\right)\right) \\
&= E\left(\boldsymbol{y}^{*\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y}^{*}\right) - E\left(\xi^{\mathrm{T}}\boldsymbol{M}^{-1}\xi\right) \\
&= n\sigma^2 + n\boldsymbol{\beta}_c^T\boldsymbol{M}\boldsymbol{\beta}_c - p\sigma^2 + \mathrm{tr}(\boldsymbol{M}^{-1}\boldsymbol{B}) + n\boldsymbol{\beta}_c^T\boldsymbol{M}\boldsymbol{\beta}_c \\
&= n\sigma^2 - p\sigma^2 + \boldsymbol{O}(1)
\end{aligned}
$$

$$
E\left(\hat{\sigma}_c^2\right) = \sigma^2 - \frac{p\sigma^2}{n} + \boldsymbol{O}\left(\frac{1}{n}\right) = \sigma^2 + \boldsymbol{O}\left(\frac{1}{n}\right).
$$

$$
\begin{aligned}
\mathrm{var}\left(n\hat{\sigma}_c^2\right) &= \mathrm{var}\left(\boldsymbol{y}^{*\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y}^{*} - \xi^{\mathrm{T}}\boldsymbol{M}^{-1}\xi + \boldsymbol{O}_p(n^{-1/2})\right) \\
&= \mathrm{var}\left(\boldsymbol{y}^{*\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y}^{*}\right) + \mathrm{var}\left(\xi^{\mathrm{T}}\boldsymbol{M}^{-1}\xi\right) + 2\,\mathrm{cov}\left(\boldsymbol{y}^{*\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y}^{*}, \xi^{\mathrm{T}}\boldsymbol{M}^{-1}\xi\right) \\
&= \boldsymbol{O}(n)
\end{aligned}
$$

$$
\mathrm{var}\left(\hat{\sigma}_c^2\right) = \boldsymbol{O}\left(\frac{1}{n}\right).
$$

Theorem 5 *The mean and variance of* $\sqrt{n}\left(\hat{\sigma}_c^2 - \sigma^2\right)$ *convergen to 0 and* $\boldsymbol{O}(1)$, *respectively, as* $n \to \infty$. *Hence* $\hat{\sigma}_c^2$ *is a consistent estimator of* σ^2.

Proof: It is straightforward using Lemma 3.

6. An Example

We apply this estimation method using research data from Hutabarat et al. (2015). The data collected is the percentage of the poor population and the malnutrition committee in East Java Province consisting of 38 districts. The occurrence of cases of spatial heterogeneity in the percentage of malnourished children under five years of age in East Java province indicates that the parameters of the regression model are influenced by observational location factors, in this case the geographical location of the district. Geographical factors are one of the causes of nutritional status disparities between regions. Covariates of the percentage of poor people who affect nutritional status have spatial effects and experience measurement errors. Therefore, it is necessary to do modeling by accommodating location factors, namely the GWR model with measurement errors.

Estimating the model parameters is obtained by entering a weighting for each observation location. The weighting used is adaptive kernel bi-square function. The weighting value used depends on the distance between the locations of observation. Summary statistics of the estimated parameters of the model parameters for each location (u_i, v_i); for $i = 1, 2, ..., 38$ are presented in Table 1. Real data set is analyzed using Software R version 3.5.3 (R Core Team 2019).

Table 1 Summary statistics of GWR model with measurement error

Parameters	Parameter Coefficients				
	Minimum	First Quartile	Median	Third Quartile	Maximum
β_0	1.552	2.742	3.437	5.038	5.656
β_1	−0.297	−0.126	0.171	0.237	0.357
σ_b^2	1.606	2.147	2.616	2.704	2.849
σ_ε^2	0.602	0.805	0.979	1.014	1.069

The estimation of β_0 and β_1 at each observation location can be seen in Figures 1-2. Based on the two figures, it can be seen that and are quite diverse in each district. From Figure 2, it can be seen that in general the coefficient is positive, meaning that the large percentage of sufferers of malnutrition is increasing with the increasing percentage of the poor population.

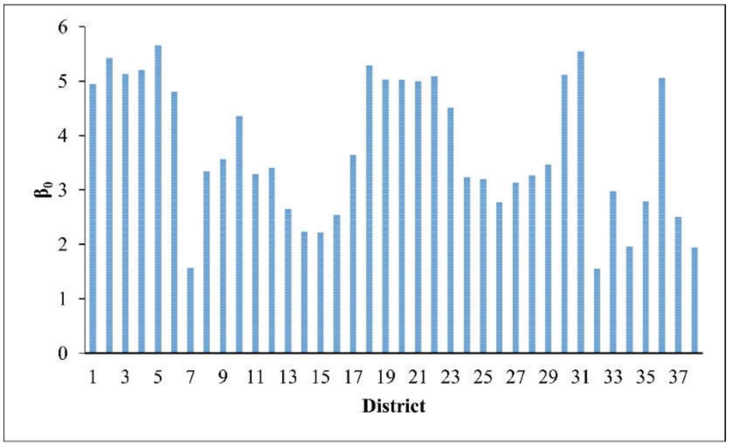

Figure 1 Estimator β_0 of the GWR model with measurement error

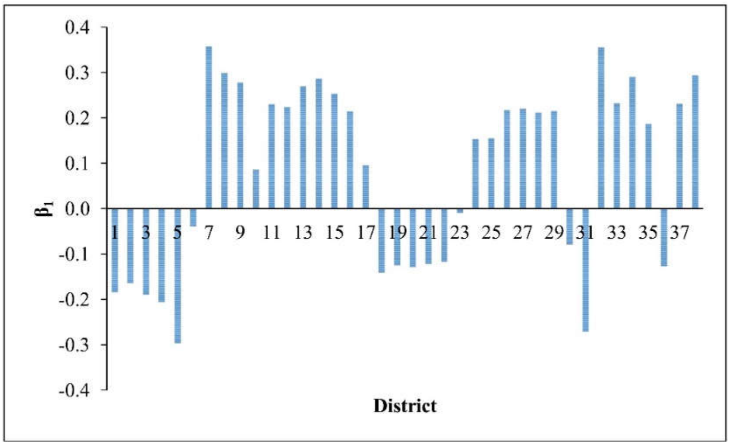

Figure 2 Estimator β_1 of the GWR model with measurement error

Comparison of the GWR model and the GWR model with measurement errors based on the error variance estimate σ_ε^2 and the variance component σ_b^2 for each location are presented in Figures 3 and 4, respectively. From Figures 3 and 4, it can be seen that the error variance estimate σ_ε^2 and the variance component σ_b^2 of the GWR model with measurement errors tend to be smaller than the GWR model.

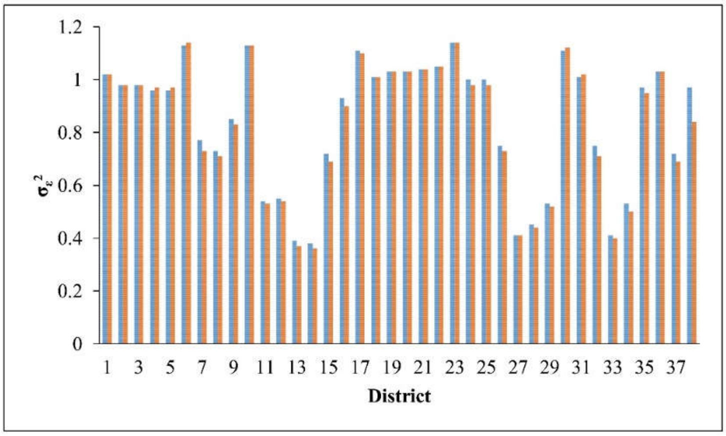

Figure 3 The error variance estimate σ_ε^2 of the GWR model (■), GWR model with measurement error (■) for 38 districts.

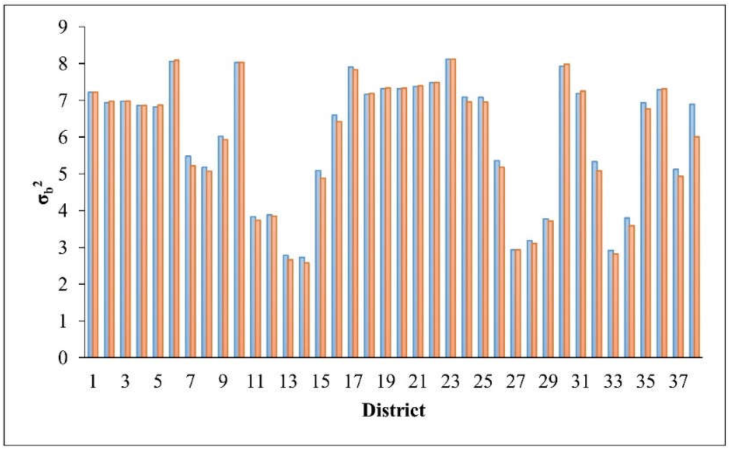

Figure 4 The variance component σ_b^2 of the GWR model (■), GWR model with measurement error (■) for 38 districts.

7. Summary

- The GWR model estimator parameters with measurement errors for the fixed estimator are

$$\hat{\boldsymbol{\beta}}_c = \left(\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{W}\boldsymbol{Z} - \mathrm{tr}(\boldsymbol{V}^{-1})\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\boldsymbol{y}^*.$$

- The estimator for random effect is

$$\hat{\boldsymbol{b}}_c = \sum\nolimits_b \boldsymbol{U}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^* - \boldsymbol{W}\boldsymbol{Z}\hat{\boldsymbol{\beta}}_c\right).$$

- The estimator for the component variance is

$$\hat{\sigma}_c^2 = \frac{1}{n}\left[\left(\boldsymbol{y}^* - \boldsymbol{W}\boldsymbol{Z}\hat{\boldsymbol{\beta}}_c\right)^{\mathrm{T}}\boldsymbol{V}^{-1}\left(\boldsymbol{y}^* - \boldsymbol{W}\boldsymbol{Z}\hat{\boldsymbol{\beta}}_c\right) - \mathrm{tr}(\boldsymbol{V}^{-1})\hat{\boldsymbol{\beta}}_c^{\mathrm{T}}\boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}_c\right].$$

Acknowledgements

This work was supported by Universitas Cenderawasih under Advanced Research Grant (Hibah Penelitian Unggulan) Year 2019, Contract No. 287/UN20.2.2/PP/2019. The authors also acknowledged the two referees for their constructive suggestions that have made this work better.

References

Arbia, G. Spatial econometrics: Statistical foundations and applications to regional convergence. New York: Springer; 2006.

Biemer PP, Groves RM, Lyberg LE, Mathiowetz NA, Sudman S. Measurement errors in surveys. New York: John Wiley & Sons; 1991.

Carroll RJ, Kuchenhoff H, Lombard F, Stefanski LA. Asymptotics for the SIMEX estimator in structural measurement error models. J Am Stat Assoc. 1996; 91(433): 242-250.

Carroll RJ, Maca JD, Ruppert D. Nonparametric regression in the presence of measurement error. Biometrika, 1999; 86(3): 541-554.

Carroll RJ, Ruppert D, Stefanski LA. Measurement error in nonlinear models. New York: Champman and Hall; 2006.

Chen X, Hong H, Nekipelov D. Nonlinear models of measurement errors. J Econ Lit. 2011; 49(4): 901–937.

Fan J, Truong YK. Nonparametric regression with errors in variables. Ann Stat. 1993; 21(4): 1900-1925.

Fellner WH. Robust estimation of variance components. Technometrics. 1986; 28(1): 51-60.

Fotheringham AS, Brunsdon C, Charlton M. Geographically weighted regression, the analysis of spatially varying relationships. New York: John Wiley & Sons; 2002.

Fuller WA, Hidiroglou MA. Regression Estimation after correction for attenuation. J Am Stat Assoc. 1978; 73: 99-104.

Fuller WA. Measurement error models. New York: John Wiley & Sons; 1987.

Harvill DA. Maximum likelihood approaches to variance component estimation and related problems. J Am Stat Assoc. 1977; 72(358): 320-338.

Hutabarat IM, Saefuddin A, Hardinsyah, Djuraidah A. Estimation of percentage on malnutrition occurrences in East Java using geographically weighted regression model. Makara J. Health Res. 2015; 19(3): 92-98.

Li Y, Tang H, Lin X. Spatial linear mixed models with covariate measurement errors. Stat Sinica. 2009; 19: 1077-1093.

Nakamura T. Corrected score function for errors-in-variables models: methodology and application to generalized linear models. Biometrika. 1990; 77(1): 127-137.

Nakamura T. Proportional hazards model with covariates subject to measurement error. Biometrics. 1992; 48(3): 829-838.

R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna: Austria; 2019 [cited 2019 September 2]. Available from: https://www.R-project.org.

Stefanski LA, Carrol RJ. Covariate Measurement Error in Logistic Regression. Ann Stat. 1985; 13(4): 1335-1351.

Stefanski LA, Carroll, RJ. Conditional Scores and Optimal Scores for Generalized Linear Measurement Error Models. Biometrika. 1987; 74: 703-716.

Zare K, Rasekh A, Rasekhi AA. Estimation of variance components in linear mixed measurement error models. Stat Pap. 2012; 53: 849-863.