A statistical analysis for geographical weighted regression

Nur Edayu Z\textsuperscript{1} and Syerrina Z\textsuperscript{2}

\textsuperscript{1,2}School of Informatics and Applied Mathematics, University Malaysia Terengganu, 21300 Kuala Terengganu, Malaysia

nuredayu23@gmail.com, syerrina@umt.edu.my

Abstract. Geographically Weighted Regression (GWR) model provides a spatial nonstationarity analysis for regression that gives different relationships to exist at different points in space. Moreover, GWR model provides a better specification to model rather than Ordinary Least Square Regression (OLS). It can be easily applied to spatial data by using the R packages that are available in R programming or GWR4 software. Though, the statistical test studies for GWR infrequently explored due to its complexity problem. Currently, the GWR studies are only limited for exploration on the relationship between the variables in study field. While the validation process for hypothesis testing of GWR model is difficult to conduct due to multiple parameter estimations for each location. The main purpose of this paper is to summaries a statistical inferences analysis for GWR model. The hypothesis testing will be derived by testing the variation of the parameter in the model. This study gives a better understanding of the different type of statistical inferences procedures that can be applied by researchers and this encourage more application for statistical inferences study from different fields.

1. Introduction

Many studies on the application of GWR have been explored in many different fields such as geography, environmental study, healthy, socioeconomic, violent, and most of the data consist of the geographical location. The research tends to compare between OLS and GWR for their study. For example, [1] perform research about the relationship between agricultural landscape and urbanization for ecological planning and management. They analyze the spatial varying relationship between urbanization indicators and changes in metrics explaining the agricultural landscape pattern by comparing OLS and GWR. Their result established that GWR model provide a better result and it is potential to provide references for ecological purposes. In addition, the study on the relationship between coastal community incomes in Setiu Wetlands in Terengganu, Malaysia has been conducted by [2]. The result shows that the GWR model exposes a relationship between the coastal community incomes with several independent variables and each location provides different significant variable that affecting the income.

In view of the fact that, the GWR model also offer a better detail for spatial data which the researcher can easily apply it. Nowadays, several statistical software such as GWR4 software, R programming, and ArcGIS system are applicable for beginners and researchers. This gives a new experience to the user for better spatial data analysis. Recent studies showed that, in term of statistical testing, the GWR model provides a complex theoretical framework and the statistical test was only applicable for simulation data [3][4][5]. The initial study on the statistical test of GWR was conducted by [3], where they start to build a theoretical framework for goodness-of-fit of the GWR.
model and it is compared with OLS model. Moreover, they also focus on determining the variation of parameters that exist in GWR model. A detailed procedure for selecting the significant independent variable also showed in this study. Meanwhile, [4] extend the studies by introducing mixed GWR model that recognized as Mixed Geographical Weighted Regression (MGWR) model. The MGWR model assumed that a coefficient need to fixed and other coefficients is varying. Hence, some procedure should be conducted for determining the type of coefficient before performing the hypothesis testing.

Perhaps, this theoretical framework should be simpler for better application to real-world data. Currently, only a single study that applied statistical testing to the real-world data [6]. Their study performs hypothesis testing on Dublin voter data. They also mapping spatial analysis and they include some adjustments test. A recent study by [7], they practice computer simulation analysis to test the significance of each parameter. Their study explores the application of the hypothesis testing analysis. However, this study does not discuss the hypothesis testing for each parameter at each location. Somehow, the issue is still new and it is complex multiple inference problems. So, the main objective of this paper is to review some of the statistical testing method for this model. The hypothesis testing will be derived by testing the variation of the parameter in the model. In addition, the problem of classical hypothesis testing will be solved by using approximated null distributions from the chosen statistical hypothesis. Hence, this study only discusses a different theoretical part that can be used to GWR model while the application of the model is not covered. The first section is an introduction to GWR model, a second section for the methodology of GWR model and the third section is a summary of hypothesis testing of GWR.

2. Methodology

2.1. The model and the estimation of the parameters

In the spatial field, an ordinary least square model has been widely used as the powerful tools for statistical analysis purposes. Moreover, OLS is able to discover the nature of associations among variables [8]. Ordinary Least Square (OLS) model is the traditional regression framework or also known as or global model. In order to estimate the parameter, it is common to apply for OLS. It can be written as:

\[ y = \beta_0 + \sum_{k=1}^{M} \beta_k x_{ki} + u_i \]  \hspace{1cm} (1)

where \( y_i \) is the observation of the dependent variable \( y \), \( \beta_k \) (k= 1,2,...M) represents the regression coefficients, \( x_{ki} \) is the \( i \)th value of \( x_k \) and \( u_i \) are the independent normally distributed error terms with zero mean and constant variance. Even though the OLS is widely used but it is still lack in term of spatial effect. The spatial effect might exist in regression and this severely might give an impact to the significance level. Thus, some spatial regressions have been introduced in order to cover the weakness of OLS.

One of the new spatial regressions frameworks that has been introduced as GWR model [9]. GWR model provides a spatial nonstationarity analysis for regression that gives different relationships to occur at different points in space. In other words, GWR also takes into account the aspect of geography and coordinate as a weighting in calculating the model parameter. The main purpose of GWR is to produce predictions and parameter estimates for at a set of locations. Each location is referring to regression points. Moreover, the location has been included in regression as the important factor and it is written as:

\[ y = \beta_0(u_i, v_i) + \sum_{k=1}^{M} \beta_k(u_i, v_i)x_{ik} + \epsilon_i \]  \hspace{1cm} (2)
where $\beta_0(u_i,v_i)$ and $\beta_k(u_i,v_i)$ are GWR coefficients in subdistrict i; location point of subdistrict i is defined by latitude and longitude coordinates $(u_i,v_i)$ and $\epsilon$ is a random error term that is independently normally distributed with zero mean and common variance $\sigma^2$.

2.2. The weighting system applicable for GWR
In the GWR modeling, there is some distance function that can be used such as Euclidean distance, Minkowski distance, Manhattan distance and the great circle distance [6]. Firstly, the latitude and longitude coordinates $(u_i,v_i)$ for each subdistrict is identity. Next, these geographical coordinates were used to indicate a Euclidean distance between observed data in subdistrict i in a village and observed data in subdistrict j in a village:

$$d = \sqrt{(\mu_i - \mu_j)^2 + (v_i - v_j)^2}$$

The distance in equation (3) is a basic measurement in weighting the data for parameter estimation. The more weight of the data is obtained if the distance between subdistrict is near. There are four different type of weighting function that can be shown in Table 1.

| Table 1 The different type of weighted functions |
|-----------------------------------------------|
| Type of Weighted | Functions |
| Fixed Gaussian | $\psi_{ij} = \exp(-d_{ij}^2 / b^2)$ |
| Fixed bi-square | $\psi_{ij} = \begin{cases} 1 - d_{ij}^2 / b^2 & d_{ij} < b \\ 0 & d_{ij} > b \end{cases}$ |
| Adaptive bi-square | $\psi_{ij} = \begin{cases} 1 - d_{ij}^2 / b_i(k) & d_{ij} < b_i(k) \\ 0 & d_{ij} > b_i(k) \end{cases}$ |
| Adaptive Gaussian | $\psi_{ij} = \exp(-d_{ij}^2 / b_i^2(k))$ |

where $b > 0$ is bandwidth constant in which its designating was done by cross-validation method and $\psi_{ij}$ is the weight value of observation at a location $j$ for estimating the coefficient at a location $i$ [9]. Meanwhile, $d_{ij}$ is the Euclidean distance between $i$ and $j$ and $b_{i(k)}$ is an adaptive bandwidth size defined as the $k$th nearest neighborhood distance [10]. Besides, the Gaussian function provide the weight as one for the estimated parameter of subdistrict and the weight with uniform decrease value for other subdistrict data, the distance between subdistrict increase as the weight continue to reduce. A Gaussian function take on in constructing weighted matrix:

$$W(\mu, v) = \text{diag}(\psi_1, \psi_2, \ldots, \psi_n)$$

where $0 \leq \psi_i \leq 1$ is weight data for subdistrict to estimate the parameter. Every observed data has one weighted matrix $W(\mu, v)$ in estimating the parameter. By using the algebraic matrix approach, the estimation of the parameter $\hat{\beta}(\mu, v) = (\beta_0(\mu, v), \beta_1(\mu, v))^T$ in subdistrict i by means of weighted least squares (WLS) method is express in [5]:

$$\hat{\beta}(\mu, v) = \left[ X^T W(\mu, v) X \right]^{-1} X^T W(\mu, v) Y$$

For all observed data in a particular subdistrict usually share the common estimated parameter. This is because the reality that the data weighting only connect the distances among the subdistricts.
3. Result and Discussion
This section discusses on a summary of the statistical inference test for GWR model that has recently studied.

3.1. Statistical Testing for Spatial non-stationary based on the GWR model [3]
A first statistical study of GWR has been conducted by [6], they found that model in the equation (1) assume that the slopes of the parameter to be general over the observed area. Unlike, the OLS in equation (1), the model in equation (2) gives the parameter to vary in space. Although, GWR model consists of a free form that is not execute because the number of parameter rise as the number of observation increases. Thus, new strategies should be constructed to represent the variation of the parameter. This strategy will control the number of degree of freedom used for each regression.

They conduct a further statistical test on GWR by extracting for testing the goodness of fit, the variation of the parameters, and their approximated distribution. In addition, they succeeded to present that it is possible for testing spatial non-stationary in a conventional way by running some simulation cases.

3.1.1. Testing the goodness-of-fit
For the assumption: The error term $\varepsilon_{1, \varepsilon_{2, \ldots, \varepsilon_{n}}}$ is independently identically distributed as a normal distribution with zero mean and constant variance $\sigma^2$. Let $\hat{\mathbf{y}} = (\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n)^T$ be the vector of the fitted values and let $\hat{\mathbf{e}} = (\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_n)^T$ be the vector of the residual.

$$\hat{\mathbf{y}} = \mathbf{L} \mathbf{Y}$$

(10)

where

$$
\mathbf{L} = \begin{bmatrix}
X^T_1 \left[ X^T_1 W(1) X \right]^{-1} X^T W(1) \\
X^T_2 \left[ X^T_1 W(2) X \right]^{-1} X^T W(2) \\
\vdots \\
X^T_n \left[ X^T_1 W(n) X \right]^{-1} X^T W(n)
\end{bmatrix}
$$

(11)

is a $n \times n$ matrix and $\mathbf{I}$ is an identity matrix of order $n$. Then, a residual sum of square (RSS) is computed as shown in equation (10). The $RSS_\varepsilon$ is describe as a quadratic form in normal variables with a symmetric and positive semidefinite matrix $(1-L)^T (1-L) Y$.

$$RSS_\varepsilon = \mathbf{Y}^T (1-L)^T (1-L) \mathbf{Y}$$

(12)

where

$$\delta_1 = tr[(1-L)^T (1-L)]$$

(13)

Next, if the null hypothesis - $H_0$: there is no significant difference between OLS model and GWR model for the given data – is true, the quantity $RSS_\varepsilon / RSS_0$ is close to one. If not, it likely to be small. Then, the $F_1$ is computed to check the hypothesis. If it is small, it indicates that $F_1$ support alternative hypothesis GWR model has improved goodness of fit. Let.

$$F_1 = \frac{RSS_\varepsilon}{RSS_0} \frac{\delta_1}{n - p - 1}$$

(14)
where \( RSS \) is a residual sum of square

\[ \therefore \text{Reject } H_0 \text{ if } F_1 < F_{1-p}(\delta_1^2 / \delta_2, n-p-1). \]

### 3.1.2. Testing the variation of each set of parameter

After the last model selection, the testing process of each set of the parameter in the model is performed across the study region. Firstly, the hypothesis testing will be as follow

\[ H_0 := \beta_{1k} = \beta_{2k} = \ldots = \beta_{nk} \text{ for a given } k \]

\[ H_1 : \text{not all } \beta_{ik} (i=1,2,\ldots,n) \text{ are equal}. \]

Secondly, the spatial variation of the parameter is estimated using \( V^2 \) as the state in equation (15), where \( J \) and \( B \) is a \( n \times n \) matrix

\[
V_k^2 = e'[\frac{1}{n}B^T(1 - \frac{1}{n}J)B]e
\]  

(15)

The distribution of \( \frac{V^2}{\gamma^2} \) can be approximated by a \( \chi^2 \) distribution with \( \frac{1}{\gamma^2} \) degree of freedom, where

\[ \gamma_i = n\left[\frac{1}{n}B^T(1 - \frac{1}{n}J)B\right]^{1/2} \]

(16)

Next, \( \hat{\sigma}^2 \) is an unbiased estimator for \( \delta_2^2 \), it is written as

\[ \hat{\sigma}^2 = \frac{RSS}{\delta_2} \]

(17)

Thus, for the statistical testing \( F_3 \),

\[ F_3 = \frac{V_k^2/\gamma_k}{\hat{\sigma}^2} \]

(18)

\[ \therefore \text{Reject } H_0 \text{ if } F_3 \geq F_{3,\alpha}(\gamma_1^2 / \gamma_2, \hat{\sigma}_k^2 / \delta_2) \]

### 3.2. A note on mixed GWR model [4]

They introduce mixed GWR where it gives additional detail notes about the spatial relationship of a data. It is a combination of global estimate and local estimate. It can write as:

\[
y = \beta_0 + \sum_{k=1}^{R} B_k x_{ik} + \sum_{k=R+1}^{M} \beta_k (u_i, v_i) x_{ik} + \epsilon_i
\]

(19)

A mixed GWR is more user-friendly for applying real data sets. However, this method required an analyst to decide the coefficient status either fixed or varying.

#### 3.2.1. Test statistics in [3] was applied for deriving p-value.

The p-value is same as \( F_3 \) in testing the variation of each set of a parameter in the equation (18).

The value of \( F_k \) for testing \( H_0 \) vs \( H_1 \). So,

\[ p(k) = P_{H_0}(F_3(k) > f(k)) \]

(20)

where \( f(k) \) is the observed value of \( F(k) \).

\[ \therefore \text{Reject } H_0 \text{ if } p(k) < \alpha \]

#### 3.2.2. The three-moment \( \chi^2 \) approximation approach

The purpose of \( \chi^2 \) approximation is to estimated the distribution of a quadratic form in normal variables. For example, the linear function \( \chi^2 \) variable with a degree of freedom \( a + b \chi^2_\alpha \). The P-value can be summaries as equation (21)
\[ p(k) = \begin{cases} \left( P[\chi_2^2 > d - \frac{1}{b} \text{tr}[M_1 - f(k)M_2]] \right), & \text{if } \text{tr}[M_1 - f(k)M_2] > 0 \\ \left( P[\chi_2^2 < d - \frac{1}{b} \text{tr}[M_1 - f(k)M_2]] \right), & \text{if } \text{tr}[M_1 - f(k)M_2] < 0 \end{cases} \]

(21)

where

\[
\begin{align*}
   b &= \frac{\text{tr}[M_1 - f(k)M_2]^3}{\text{tr}[M_1 - f(k)M_2]^2} \\
   d &= \frac{(\text{tr}[M_1 - f(k)M_2])^3}{(\text{tr}[M_1 - f(k)M_2])^2}
\end{align*}
\]

(22)

where \(\text{tr}(.)\) is a trace matrix and \(M\) is a positive semidefinite \(n \times n\) matrices. This formula is illogical as \(\text{tr}[M_1 - f(k)M_2]^3 = 0\). The writers mention that invalid situation rarely appears in practice.

3.3 Statistical software for multiple GWR model using the method of Maximum Likelihood Ratio Test(MGWR)[5]

MGWR model is an improvement of GWR model with the model parameter that is local to each point. The vector error in MGWR model is a random and the distributed normal multivariate with zero mean and variance covariance.

3.3.1. Testing the variation of each set of parameter

In order to conduct hypothesis testing for spatial analysis, the writers perform it by differentiate the fitness of the parameters coefficient between multivariate linear regression model and MGWR model. To test either the weighting \(W(\mu_i, v_i)\) that used in the process is equal to 1. Thus, the hypothesis testing is as follow.

\[ H_0 : \beta_{ib}(u_i, v_i) = \beta_{ib} \text{ for a given k and h (no influence of geographical factor)} \]
\[ H_1 : \text{at least one } \beta_{ib}(u_i, v_i) \neq \beta_{ib} \]

Ratio likelihood test is used to compare the multivariate linear model and the MGWR model based on the F test. The test will reject \(H_0\) if \(\land < \land_0 < 1\) with the value \(0 < \land_0 < 1\) is expressed as:

\[
\land_0^2 = \frac{\sum_{i=1}^{n} \frac{1}{\hat{\sigma}_i^2} \left( Y^T (1-S)^T (1-S) Y \right)}{n}
\]

(23)

where \(S\) and \(M\) is a \(n \times n\) matrices, with a significance level \((\alpha)\) then the decision taken will reject \(H_0\) if the value of \(F_{hitung}^* < F_{1-\alpha}(p, n - p - 1)\) or \(P_{value} < \alpha\).

3.3.2. Testing simulation of MGWR model

By using MLRT, the statistical testing was performed as follows

\[ H_0 : \beta_{ib}(u_i, v_i) = \beta_{2b}(u_i, v_i) = \ldots = \beta_{ib}(u_i, v_i) = 0 \]
\[ H_1 : \text{at least one } \beta_{ib}(u_i, v_i) \neq 0 \]

Ratio likelihood tests the MGWR model based on the F test. The test will reject \(H_0\) if \(\land < \land_0 < 1\) with the value \(0 < \land_0 < 1\) is expressed as:
The equation can be simplified as
\[ S_{\text{RSS}} = \sum (y_i - \hat{y}_i)^2 = \sum (y_i - \hat{y}_i)^2 / (n - 2) \]

where \( y_i \) and \( \hat{y}_i \) are the observed and predicted values for the \( i \)th data point, respectively. The term for the variation of data can be solved by computing the local standard errors. According to [10], the weight of the data point might be small if the data points lie far from regression point. Thus, the standard errors of GWR parameter are deduced following steps:

1. First, let the estimators of the local parameter estimates as a “D”, \( \hat{\beta}(u_i, v_i) = Dy \). Now, the equation can be simplified as:
\[ D = (X^TW(u_i, v_i)X)^{-1}X^TW(u_i, v_i) \]

2. The variance of parameter estimates given by
\[ \text{Var}(\hat{\beta}(u_i, v_i)) = DD^T\sigma^2 \]
where \( \sigma^2 \) is the normalized residual sum of square (RSS) from the local regression. It is defined as
\[ \sigma^2 = \sum (y_i - \hat{y}_i)^2 / (n - 2) \]

3. The matrix S is known as the hat matrix, which maps \( \hat{y} \) on to \( y \) by \( \hat{y} = Sy \). The \( i^{th} \) row of \( S \), \( r_i \) is given by \( r_i = X_iD_i \) (where \( X_i \) is an independent data vector at \( i \)) [10]. In GWR, the term for the effective degrees of freedom of the residual is written as \( n - 2\text{tr}(S) - \text{tr}(S^TS) \). Meanwhile, the term for the effective number of parameter is written as \( 2\text{tr}(S) - \text{tr}(S^TS) \) [3]. After the variance was estimated, the local standard errors can be obtained by:
\[ \text{SE}(\hat{\beta}_i) = \left( \text{Var}(\hat{\beta}_i) \right)^{1/2} \]
\[ \text{SE}(\hat{\beta}_i) = \left( \sum (y_i - \hat{y}_i)^2 / (n - 2) / (n - 2\text{tr}(S) - \text{tr}(S^TS)) \right)^{1/2} \]

where \( \hat{\beta}_i \) is a short-term notation for \( \hat{\beta}_i(u_i, v_i) \).

### 3.4.2. Hypothesis testing procedure on individual GWR coefficient

This procedure is used to determine any significance parameters affecting the response variable. Hence, the hypothesis testing can be shown as follows:

\[ H_0 : \hat{\beta}_{ik}(u_i, v_i) = 0 \]
\[ H_1 : \hat{\beta}_{ik}(u_i, v_i) \neq 0 \]
for \( k = 1, 2, \ldots, p \) and \( i = 1, 2, \ldots, n \).
The partial testing of GWR is solved by obtaining the estimated parameter of \( \hat{\beta}_{ik}(u_i,v_i) \) and standard error is \( SE(\hat{\beta}_{ik}(u_i,v_i)) \). The \( SE(\hat{\beta}_{ik}(u_i,v_i)) \) used to test the level of significance of each location by using t test [5]. Then, t-test statistics can be measured by:

\[
t_{\text{calculate}} = \frac{\hat{\beta}_{ik}(u_i,v_i)}{SE(\hat{\beta}_{ik}(u_i,v_i))}
\]

(30)

\[t_{\text{tabulated}} = n - 2tr(S) - tr(S'^T S)
\]

(31)

\( t_{\text{calculate}} \) will follow t distribution with a significance level \( (\alpha) \), then reject \( H_0 \) if the value of \( t_{\text{calculate}} > t_{\text{tabulated}} \).

3.5. The adjustment test on high order multiple inference problems

The GWR hypothesis testing allows each location to present the significant variable affecting the dependent variable. This gives more precise significant relationship rather than a global model (OLS). Yet, it will be more complex since it involves a large number of simultaneous t-test regarding the number of location \( (i) \) and the number of parameters \( (\beta_i) \). There will be hundreds or thousands of tests that need to be run in order to determine whether parameters are locally significant or not. After conducting the multiple hypotheses testing, the adjustment on significance level should be further investigated. If not, the higher multiple order will occur. In addition, this adjustment test is used to control the significance level and it controls the type I errors.

There are many available approaches that adjust each test such as Benjamini Hochberg, Benjamini Yekutieli, Bonferroni approaches, and Fotheringham-Bryne. For example, these approaches can be applied to GWR model, but the Fotheringham-Bryne is specially designed for GWR model [6]. Moreover, the detailed on the Bonferroni-style adjustment can be explained as ensuing. First, assume the probability of rejecting one or more true null hypothesis (i.e. the family-wise error rate, (FWER)) be expressed as \( \xi \) (with t the number of tests). Followed by the FWER for testing purpose that also controlled as \( \xi \), or less, by selecting

\[
\alpha = \frac{\xi}{1 + p_e - \frac{p_e}{np}}
\]

(32)

where \( \alpha \) is the probability of a type I error in the ith test; where \( p_e \) is the effective number of parameters in GWR model; \( np \) is the number of parameters in each individual local regression; and \( n \) is the sample size. Furthermore, [6] have applied this approach on the Dublin 2004 voter turnout data. They found that the Benjamini Hochberg and Benjamini Yekutieli approach give the same result as un-adjusted p-values. Meanwhile, Bonferroni approaches and Fotheringham-Bryne adjustment give the parallel results.

4. Conclusion

This study was summarized about the statistical tests that have been used for GWR model. The details of different hypothesis testing of GWR model which focus on the theoretical part also cover in this study. Nevertheless, the application of the statistical inferences is not covered under this study. The significance of this study is it will help the researcher to differentiate between different hypothesis procedures that are available for GWR model. Currently, the study by [6] has applied the statistical testing on the real world data. While other studies by [3] and [4] have applied on theoretical part and simulation part only for hypothesis testing of GWR model. In addition, more
simulation should be done particularly related to several structures in the independent variables. This will help the researcher to understand the future estimate methods in difficult surroundings [3].

Besides, this study will encourage researcher to apply some statistical analysis on their GWR model and it will be more application study about statistical inferences. Otherwise, this also helps to identify significant factor affecting the spatial study and this will improve the decision-making process. Now, it able carry out the significance test in conventional statistical manner. Lastly, the researcher also may do some adjustments test manually towards their analysis, especially for the larger sample size. According to [4], the techniques to adjust the significance level in GWR continue to be further study because the Bonferroni adjustment is too conventional for this associated multiple tests.

In future, a new programming should be done for this spatial statistical analysis especially GWR. This programming also able to present the amount of error that exists during analysis and it is able to take fast correction and accurate analysis. As a concern, this will encourage more spatial studies by using GWR model and the researcher could explain their studies with more significant.

Acknowledgment

This research was supported by the Niche Research Grant Scheme (NRGS) UMT - P1(R) (Second Phase)-Vote No.: 53131/30, Ministry of Higher Education Malaysia.

References

[1] Su S, Xiao R and Zhang Y 2012 Multi-scale analysis of spatially varying relationships between agricultural landscape patterns and urbanization using geographically weighted regression *Applied Geography* 32(2) pp 360–375

[2] Syerrina Z, Naeim A R, Safiiah L M and Nuredayu Z 2017 Explorative Spatial Analysis of Coastal Community Incomes in Setiu Wetlands: Geographically Weighted Regression *International Journal of Applied Engineering Research ISSN* 12(18) pp 973–4562

[3] Leu Y, Mei C L and Zhang W X 2000 Statistical tests for spatial nonstationarity based on the geographically weighted regression model *Environment and Planning A* 32(1) pp 9–32

[4] Mei C, He S and Fang K 2004 A note on the mixed geographically weighted regression model *Journal of Regional Science* 44(1) pp 143–57

[5] Harini S, Purhadi P, Mashuri M and Sunaryo S 2012 Statistical test for multivariate geographically weighted regression model using the method of maximum likelihood ratio test *International Journal of Applied Mathematics and Statistics* 29(5) pp 110–15

[6] Lu B, Harris P, Charlton M, Brunsdon C, Lu B and Harris P 2014 Geo-spatial Information Science The GWmodel R package : further topics for exploring spatial heterogeneity using geographically weighted models *Geo-Spatial Information Science* 17(2) pp 85–101

[7] Leong Y Y and Yue J C 2017 A modification to geographically weighted regression. *International Journal of Health Geographics* pp 1–18 (Preprint https://doi.org/10.1186/s12942-017-0085-9)

[8] Dobson and Annette J 2002 An introduction to generalized linear models. (Boca Raton: Chapman & Hall/CRC)

[9] Fotheringham A S, Brunsdon C, & Charlton M 2002 Geographically weighted regression: the analysis of spatially varying relationships Wiley (Preprint https://www.wiley.com/enmy/Geographically+Weighted+Regression:+The+Analysis+of+Spatially+Varying+Relationships+-+p-9780471496168)

[10] Waller L A, Zhu L, Gotway C A, Gorman D M and Grunewald P J 2007 Quantifying geographic variations in associations between alcohol distribution and violence: A comparison of geographically weighted regression and spatially varying coefficient models *Stochastic Environmental Research and Risk Assessment* 21(5) pp 573–588