

# Review on Geographically Weighted Regression (GWR) approach in spatial analysis

Ayuna Sulekan, Shariffah Suhaila Syed Jamaludin\*

Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, Skudai 81310, Johor, Malaysia

\* Corresponding author: [suhailasj@utm.my](mailto:suhailasj@utm.my)

## Article history

Received 10 January 2019  
 Revised 1 April 2019  
 Accepted 30 November 2019  
 Published Online 15 April 2020

## Abstract

In spatial analysis, it is important to identify the nature of the relationship that exists between variables. Normally, it is done by estimating parameters with observations which taken from different spatial units that across a study area where parameters are assumed to be constant across space. However, this is not so as the spatial non-stationarity is a condition in which a simple model cannot explain the relationship between some sets of variables. The nature of the model must alter over space to reflect the structure within the data. Non-stationarity means that the relationship between variables under study varies from one location to another depending on physical factors of the environment that are spatially autocorrelated. Geographically Weighted Regression (GWR) is a technique in which it applied to capture the variation by calibrating a multiple regression model, which allows different relationships to exist at different points in space. A robust algorithm has been successfully used in spatial analysis. GWR can theoretically integrate geographical location, altitude, and other factors for spatial analysis estimations, and reflects the non-stationary spatial relationship between these variables. The main goal of this study is to review the potential of the GWR in modelling the spatial relationship between variables either dependent or independent and its used as the spatial prediction models. Based on the application of GWR such as house property indicates that GWR is the best model in estimating the parameters. Hence, from the GWR model, the significance of the variation can also be tested.

**Keywords:** Geographically Weighted Regression (GWR), Spatial Modelling, Spatial Non-stationarity.

© 2020 Penerbit UTM Press. All rights reserved

## INTRODUCTION

The technique of linear regression has long resided in the analytical toolbox of the quantitative geographer. As a general technique for investigating the linkage between geographical variables, the method has featured in virtually countless publications. However, this is surprising in some respects, as the technique itself takes no account of location in its analysis of relationships between variables. The well-known components of a regression model are  $X$ , a matrix containing a set of independent or predictor variables and  $y$ , a vector of dependent or response variables. The understanding of this method is when it is applied to geographical data in which each case corresponds to a geographical location. However, there may be situations when the nature of such models is not fixed over space. This is referred to here as spatial non-stationarity. Several more explicit ways of incorporating space have been considered. For example, Casetti (1972) proposed the expansion method where coefficients in the regression model were expressed as explicit functions of the spatial locations of the cases. The expansion method has been introduced to overcome the weaknesses of linear regression as the technique of regression plays no role in modelling processes when it is applied to geographical data in which each case corresponds to a geographical location. The nature of the models which is not fixed over spaces is referred as spatial non-stationarity. The advantage of a technique of this sort is that, once the model has been calibrated, it is possible to map the variation in the original regression parameters, and to gain some understanding of the

spatial patterns in the association between the predictor or response variables.

In Malaysia, only few studies have employed the method of GWR in spatial variation and modelling (Eboy and Samat, 2015; Jamhuri *et al.*, 2016). GWR is used as spatial regression tools in modelling property rating valuation over Kota Kinabalu, Sabah (Eboy and Samat, 2015) while Jamhuri *et al.* (2016) developed a toolbox's extension of GWR with the application to the forestry sector. However, the best of our knowledge, the application of GWR in spatial modelling of rainfall data is still new in Malaysia. Our challenge now is to investigate on how GWR can be applied as a new geostatistical tool in the context of climate data in Malaysia and will the proposed GWR in spatial modelling technique be more efficient model compared to the traditional modelling such as Kriging method.

Spatial modelling is an essential analytical process of spatial analysis which is conducted with geographical information system (GIS) due to describe the properties and processes of a given spatial features. According to McKenzie and Ryan (1999), Robinson and Metternicht (2006), a number of prediction methods have been suggested to interpolate data from sparse sampling points into continuous surfaces, varying from regression methods to geostatistical methods such as ordinary kriging, ordinary cokriging, regression kriging and other hybrid techniques. Regression methods involved such as simple linear regressions, nonlinear regressions, inverse distance weighting, generalized linear models and regression trees. Spatial non-stationarity is a condition where a simple model cannot explain the relationship between some sets of variables. The nature of the model

must alter over space to reflect the structure within the data. Non-stationarity means that the relationship between variable under study differs from one location to another depending on physical factors of the environment which are spatially autocorrelated. Non-stationarity is a conceptual data in which it is difficult to model due to the changing of mean and its variance. Since non-stationarity is unpredictable, it can not be modelled and forecasted.

Although Brunson *et al.* (1996) was a pioneer of technique GWR, Fotheringham *et al.* (2002) were fully described the GWR by saying that Geographically Weighted Regression (GWR) technique was the first selecting a bandwidth that being involved at that time for an isotropic spatial weights kernel. A fixed bandwidth of Gaussian kernel was chosen through leave-one-out cross-validation ways. Since all the  $n$  regressions were well fitted in each step, so the demanding of the choices bandwidth could be happened. According to Brunson *et al.* (1998), he expands the idea of Casetti (1972) which is the expansion method one is required to provide an explicit function to describe the variation of each element of the vector of regression coefficient ( $\beta$ ) over space. The method proposed here was Geographically Weighted Regression (GWR) which attempts to overcome this problem by providing a nonparametric estimate of  $B_i(p_i)$ .

The method of GWR was specifically designed to deal with issues of non-stationarity (Fotheringham *et al.*, 2002). GWR handles the issues by measuring local relationships between the target variable and explanatory variables at various locations. The assumption of the stationarity structural stability over space might not be realistic due to the variations in relationships between the dependent and independent variables over space. Hence, spatial non-stationarity should be considered in analysing spatial data. Yu *et al.* (2009) proposed a paper in investigating the potential of spatial non-stationarity of the relationship between dependent and independent variables by applying the GWR technique. The calibration of GWR will then employs the geographically weighted local least squares regression approach. Basically, their studies attempt to highlight the issues of non-normality on spatial non-stationarity in GWR analysis for the first time. Ehlkes *et al.* (2014) stated that GWR was specifically designed as an extension of traditional regression in describing the relationship between variables. GWR can incorporate, detect and account for spatial non-stationarity.

GWR is a technique or method in which it is applied to capture the variation through a multiple regression model calibration which allows different relationships to exist at different points in space. Besides, the method of GWR also allows us to establish the relationship between dependent variables and independent variables and its used as the spatial prediction models. The purpose of this study is to review the effectiveness of GWR model in dealing with the issues of spatial non-stationarity in estimating the parameters. In this study, we can also see that GWR model not only can explore the variation of parameters but also can test the significance of the parameters.

**METHODS**

GWR is an exploratory technique which mainly intended to indicate where the non-stationarity takes place (Bivand, 2017). Simple linear regression is frequently used as modeling tools in geographical analysis in which the dependent variable is modelled as a linear function of a set of  $n$  independent variable which known as predictor variables (Dobson, 1990).

A global regression model can be written as :

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \varepsilon_i \tag{1}$$

where  $y_i$  is the  $i^{th}$  observation of the dependent variable,  $x_{ik}$  is the  $i^{th}$  observation of the  $k^{th}$  independent variable, the  $\varepsilon_i$  are independent normally distributed error terms with zero means and each  $\beta_k$  must be determined from a sample of  $n$  observation. GWR is a relatively simple technique which extends the traditional regression framework of the equation (1) by allowing the local parameter rather than global parameter to be estimated so that the model is rewritten as:

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \tag{2}$$

where  $(u_i, v_i)$  is the coordinates of  $i^{th}$  point in space and  $\beta_k(u_i, v_i)$  is the realization of continuous function  $\beta_k(u, v)$  at point  $i$  which means that allows to be a continuous surface of parameter values and measurements of this surface at certain point denotes as the spatial variability of the surface. Note that (1) is a special case to (2) in which the parameter surface is assumed to be constant over space. Thus, the expression of GWR in (2) is a recognition of spatial variation in relationship might exist and will then provide a way which can be measured. The calibrating of (2) is assumed implicitly that the observed data near to location  $i$ , could have more influence on the estimation of the  $\beta_k(u_i, v_i)$  rather than locate farther from  $i$ . Basically, (2) measures the relationships which exist in the model around each point  $i$ . Weighted least square will then provide a basic to understand on how the GWR operates. In weighted least square, a weighting factor is applied to each square difference before minimizing, so that the inaccuracy of some predictions will carry more penalty than others. In GWR, an observation is weighted proximately location  $i$ , so that the weighting of an observation is varies with  $i$ . Data obtained from observations which close to  $i$  are weighted more than data obtained from observation far away, that is:

$$\hat{\beta}(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) y \tag{3}$$

where  $\hat{\beta}$  is an estimate of  $\beta$  and  $W(u_i, v_i)$  is a matrix of  $n$  by  $n$  whose off-diagonal elements are zero and whose diagonal elements denote the geographically weighted of observed data for point  $i$ . Note that, there is no reasons that  $i$  must be the location of a data point. Local estimates of the parameters can in fact be derive for any point in space regardless of whether that point is one at which data have been observed.

**Choice of spatial weighting function**

It has been stated in GWR that  $W(u_i, v_i)$  or more convenient terms,  $W(i)$  is a weighting scheme based on the proximity of point  $i$  to the sampling location around  $i$  without an explicit relationship being stated. The choice of a relationship will be considered here. Firstly, consider the implicit weighting scheme of the Ordinary Least Square (OLS) framework in (1) as

$$w_{ij} = 1 \quad \forall i, j \tag{4}$$

where  $j$  represents a specific point in space at which data are observed and  $i$  represents any point in space for which parameter are estimated. In global model, each observation has a weight of unity. Based on locality, initial step towards weighting may be excluded from model's calibration observation which further than some distance  $d$  from locality. This can be done by setting their weights to zero. The weighting function are:

$$w_{ij} = \begin{cases} 1 & \text{if } d_{ij} < d, \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where  $d_{ij}$  is distance between the location of observation  $i$  and  $j$ . Equation in (5) will simplify the procedure of calibration since only a subset of data will be used in calibration for each regression point. However, equation (5) will suffer discontinuity problems as  $i$  varies around the study area, the regression coefficient would drastically change as one sample point moves into or out of the circular buffer around  $i$ . It is possible to relate  $d_{ij}$  to  $w_{ij}$  with a continuous function as one way to combat the problem of discontinuous of weights. One obvious choice is:

$$w_{ij} = \exp \left[ -\frac{1}{2} \left( \frac{d_{ij}}{b} \right)^2 \right] \tag{6}$$

where  $b$  is referred to as the bandwidth. If  $i$  and  $j$  coincide in which  $i$  also happens to be a point in space where the data are observed, the weighting of data at that point will be unity and the weighting of the

other data will decrease according to a Gaussian curve as the distance between  $i$  and  $j$  increases. In latter case, the inclusion of data in calibration procedure become fractional. Here the value of the weight would decay gradually with distance to the extend when  $d_{ij} = b$  the weighting would be 0.05. Alternative weighting function could be

$$w_{ij} = \begin{cases} \left[ 1 - \left( \frac{d_{ij}}{b} \right)^2 \right]^2 & \text{if } d_{ij} < b, \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

The equation (7) can be referred to the kernel function and denoted with letter  $K$  as in  $w_{ij} = K(d_{ij})$ . In each case, as  $d$  does in basic step function, the constant  $b$  will provide some control range of the geographical data. When a certain distance is reached, the degree of weighting would decay with distance and then will suddenly drop to zero.

**Calibrating the weighting function**

One difficulty in GWR is that the estimated parameters in which it depends on the weighting function. For example, in Equation (5), as  $d$  become larger, the nearest will be the solution model to OLS and if  $d$  is equal to the maximum distance between points in the system, then two models; OLS and GWR will be equal. Equivalently, the weights tend to one for all pairs of the points as  $b$  tends to infinity so that the estimated parameters become uniform and GWR becomes equivalent to OLS. Several numbers of criteria have been proposed in selecting a suitable bandwidth. Consider the selection of  $b$  in (6). One possibility is to choose  $b$  using a ‘least square’ criterion. One way to proceed would be to minimize the quantity,

$$z = \sum_{i=1}^n [y_i - \hat{y}_i(b)]^2 \quad (8)$$

where  $\hat{y}_i(b)$  is the fitted value of  $y_i$  using bandwidth of  $b$ . To find the fitted value of  $y_i$ , it is necessary to estimate  $\beta_k(u_i, v_i)$  at each points of data and then combine these with the  $x$ -values at these points. Suppose  $b$  is made to be very small so that the weighting of all points will become negligible except for  $i$  itself. The value of equation (8) will become zero if the fitted values at the sampled points are more likely to the actual values. The idea of optimizing the value of  $b$  become zero is not a good way to do so due to two reasons. First, the parameter of such models will be undefined in the limiting case and second, the estimating will move wildly throughout space to give locally a good fitted value at each point  $i$ . Cross-validation (CV) approach is suggested as the solution for local regression and kernel density. Here, a score of the form

$$CV = \sum_{i=1}^n [y_i - \hat{y}_{\neq i}(b)]^2 \quad (9)$$

is used where  $\hat{y}_{\neq i}(b)$  is the fitted value of  $y_i$  with the observation for point  $i$  omitted from the calibrating process. Since  $b$  becomes very small, the model is calibrated only on samples which are closed to  $i$  and not at  $i$  itself. This is because the approach has desirable property in countering the ‘wrap around’ effect. A similar method of deriving the bandwidth in which provides a trade-off between goodness-of-fit and degrees of freedom is to minimize the Akaike Information Criterion (AIC) which is defined in GWR as:

$$AIC = 2n \log_e(\hat{\sigma}) + n \log_e(2\pi) + n \left\{ \frac{n + \text{tr}(\mathbf{S})}{n - 2 - \text{tr}(\mathbf{S})} \right\} \quad (10)$$

where  $n$  is the sample size,  $\hat{\sigma}$  is the estimate standard deviation of the error term and  $\text{tr}(\mathbf{S})$  denotes as the trace of the hat matrix which is a function of the bandwidth. Since AIC can be used in Poisson and logistic GWR as well as in linear models, AIC has an advantage to be more general in application rather than CV statistics. Besides that, AIC

can also be used in assessing whether the GWR provides better fit than global model by considering the varying degrees of freedom in two models.

**Testing the significance of spatial non-stationarity**

In this section, the development of a hypothesis test will be considered to justify the use of spatially weighting regression model.

$$\begin{aligned} H_0: \beta_{ij} &= \beta & \forall i, \\ H_1: \beta_{ij} &\neq \beta & \text{not all the same } \forall i. \end{aligned} \quad (11)$$

Variance of  $\beta_{ij}$  across  $i$  is one of the useful statistical test to measure the variability of  $\beta_{ij}$  as  $i$  varies for a fixed  $j$ .

$$v_j = \frac{\sum_i (\beta_{ij} - \beta_j)^2}{N} \quad (12)$$

Under the null hypothesis,  $H_0$ , we assume that  $\beta_{ij}$  do not vary with  $i$  for variable  $j$ . This suggests that, if the GWR model were calibrated with locations of the observations which is randomly assigned to the predictor and response variables, a little difference will then occur in patterns of  $\beta_{ij}$ . This happen when  $\beta_{ij}$  are fixed over space and spatial location should not affect the calibration. Therefore, Monte Carlo is used to test and to compare the distribution of the  $v_j$  under the randomization hypothesis. The procedure for a given  $j$  is as follows:

- a. Make a note of  $v_j$  for the correctly located observation.
- b. Randomly ‘scramble’ the location of  $p_j$  among the observation.
- c. Repeat the previous step  $P - 1$  times, note that  $v_j$  each time.
- d. Compute the rank of  $v_j$  for the correctly located case,  $R$ .
- e. The  $p$ -value for the randomization hypothesis is  $\frac{R}{P}$ .

**RESULTS AND DISCUSSION**

In this section, we will discuss the application of GWR in two different cases that highlighted different approach in the analysis. The results are presented in the forms of tables which were adapted from the articles.

First case study focused on spatial varying relationship between house price and floor area done by Lu *et al.* (2011). In their study, the performance of GWR method was compared to OLS.

Second case study discussed the issue of non-stationarity and how the GWR method can be used to handle these issues. We reviewed the work done by Yu *et al.* (2009) which focused on the relationship between tobacco outlet density and demographic factors.

**Case study I**

Lu *et al.* (2011) studied the analysis of a non-Euclidean distance metric on London house price data using GWR. Basically, the idea of their study focused on GWR model in which it was used to explore the spatial varying relationship between house price and floor area in London. Sample of 372 properties sold within London area during 2001 were used in their studies. Their study area were divided by the river Thames and this makes the network distance (ND) significantly different from Euclidean distance (ED) which affected by the density of the bridges along river. A regression model between price and floor area has been proposed and its GWR expression was written as:

$$P_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)FLRAREA_i. \quad (13)$$

To facilitate the analysis, the sampled houses location has been used as regression points. Their study focused on comparing between two types of distance; the network distance and euclidean distance, which are more significant in measuring the house price. Lu *et al.* (2011) proposed five different ways in calibrating the model. First model refers to an ordinary least square (OLS) while the second model is a GWR model. Then, the model was calibrated using ED and ND respectively with fixed and adaptive spatial kernels. Cross-validation (CV) approach was chosen to employ the selection for bandwidth of each calibration

in GWR. Table 1 showed that all the performance of GWR approaches was better than OLS.

**Table 1** Comparison of Akaike Information Criterion (AIC) and adjusted  $R^2$  values for different model calibration.

	OLS Model	GWR Model (Fixed & ED)	GWR Model (Adaptive & ED)	GWR Model (Fixed & ND)	GWR Model (Adaptive & ND)
AIC	9529.11	9065.08	9078.04	9057.17	9078.18
Adjusted $R^2$	0.46719	0.82418	0.81805	0.82546	0.81556
	16	92	45	11	02

After the model calibration, results show a significant improvement for fixed kernel and ND based on the smallest AIC values and the largest adjusted  $R^2$ . Clearly, results from calibration of GWR model with fixed kernels and ND confirmed that the performance of GWR had achieved its feasible improvement.

**Case study II**

Regression analysis is a statistical technique that is most frequently used in investigating the relationship between given observation and a set of factors. Yu *et al.* (2009) examined the relationships between three demographic factors at census tract level such as percentage of African American residents, percentage of Hispanic residents and median household income and the tobacco outlets density in the tracts. These studies covered all 15037 retail outlets licensed selling tobacco in 2004. The number of tobacco outlets per 10 km of roadway in a census tract, one of the primary linkage between tobacco and accessibility of outlet population was used to calculate the tobacco outlets density. They conducted the preliminary analyses at both municipality and census block group levels. Although there was about 1950 census tract in New Jersey, this studies only focused on 1938 residential tracts which were recorded the demographic data.

Generally, Yu *et al.* (2009) examined and estimated the relationship expression written as:

$$outlet.dn = f(pct.aa, pct.hisp, med.inc) \tag{14}$$

where *outlet.dn* was representing the outlet density in a particular census tract, *pct.aa* would represents the percentage of African American, *pct.hisp* would represents the percentage of Hispanics, *med.inc* would represents the median household income and linear functional form (*f*) would preferred be the estimation.

According to Leung *et al.* (2000) and Fotheringham *et al.* (2002), GWR models were fitted the data better than OLS models due to the decrease of degrees of freedom. ANOVA test which has been developed by Brunson *et al.* (1999) would then be used to statistically test whether GWR model was improved significantly from OLS model or not by considering the decrease of degrees of freedom. Table 2(a) and Table 2(b) were reported as the results of OLS estimation of both transformed and non-transformed data. Based on the results, the used of usual diagnostic statistics had indicated both models are significant and selected demographic factors explained that the variation of the outlet densities were around 45% which included either dependent variable was transformed or not. Clearly, two estimations with transformed model and non-transformed model were varied in GWR analysis.

**Table 2(a)** OLS results for non-transformed data.

	Estimate	Std. error	t value	Pr(<  t )
(Intercept)	2.849	0.570	4.998	$6.32 \times 10^{-7}$
Pct.AA	5.261	0.705	7.465	$1.26 \times 10^{-13}$
Pct. Hisp.	28.53	0.949	30.065	$< 2 \times 10^{-16}$
Median Inc.	$-3.468 \times 10^{-5}$	$7.394 \times 10^{-6}$	-4.691	$2.91 \times 10^{-6}$
Adjusted $R^2 = 0.4406$ ; F-statistic= 509.5 on 3 and 1934 DF				
p-value:< $2.2 \times 10^{-16}$ ; AIC= 12771.28				

**Table 2(b)** OLS results for transformed data.

	Estimate	Std. error	t value	Pr(<  t )
(Intercept)	1.020	0.116	8.785	$< 2 \times 10^{-16}$
Pct.AA	1.201	0.144	8.369	$< 2 \times 10^{-16}$
Pct. Hisp.	4.741	0.193	24.527	$< 2 \times 10^{-16}$
Median Inc.	$-1.747 \times 10^{-5}$	$1.506 \times 10^{-6}$	-11.599	$< 2 \times 10^{-16}$
Adjusted $R^2 = 0.451$ ; F-statistic= 531.5 on 3 and 1934 DF				
p-value:< $2.2 \times 10^{-16}$ ; AIC= 6604.26				

AIC values of both OLS estimation has been used as an alternative measure of goodness-of-fit and then will be compared with the AIC values in GWR. In terms of statistic, application of GWR in the models was justified as the significantly improve exists for both models in their corresponding OLS counterparts even though degrees of freedom was reduced. Table 3(a) and Table 3(b) recorded the test of ANOVA for GWR models improvements over the OLS models for both transformed and non-transformed data by considering added computation complexity and the decrease of degrees of freedom.

**Table 3(a)** ANOVA test for GWR models' improvement over OLS of non-transformed data.

	SS <sup>a</sup>	DF <sup>a</sup>	MS <sup>a</sup>	F	Pr (> F)
OLS residuals	82150.12	1934.00	42.477	n.a.	n.a.
GWR improvement	27220.63	136.437	199.51	n.a.	n.a.
GWR residuals	54929.48	1797.56	30.558	6.528	$< 2.2 \times 10^{-16}$
With 1011.269 and 1836.223 degrees of freedom for the F test; AIC= 12211.29					

**Table 3(b)** ANOVA test for GWR models' improvement over OLS of transformed data.

	SS <sup>a</sup>	DF <sup>a</sup>	MS <sup>a</sup>	F	Pr (> F)
OLS residuals	3408.9	1934.000	1.7626	n.a.	n.a.
GWR improvement	1252.9	180.9464	6.9242	n.a.	n.a.
GWR residuals	2155.9	1753.053	1.2298	5.630	$< 2.2 \times 10^{-16}$
With 1062.870 and 1804.046 degrees of freedom for the F test; AIC= 6014.77					

where SS<sup>a</sup> represents sum of squares, DF represents degrees of freedom and MS represents mean of squares.

Spatial non-stationarity test for the coefficients of independent variables in GWR models was one way in applying the GWR techniques. Fotheringham *et al.* (2002) used Monte Carlo simulation to test a non-parametric on a spatial non-stationarity. While Brunson *et al.* (1996) preferred natural choice for a formal statistical test of spatial non-stationarity by using the sample variance of the estimated coefficients. The null hypothesis of all estimated coefficients which is same with particular independent variables will then be tested by a constructed F test has been fully explained in Leung *et al.* (2000).

In this study, Yu *et al.* were using a practical yet natural choice of testing spatial non-stationarity to provide sufficient insights for understanding the potential impact of non-normality on non-stationarity. The coefficients of two GWR models were estimated with non-transformed and transformed tobacco outlet density and these two models will then be tested the null hypothesis of each independent variables to see their spatial non-stationarity. These stationarity test of two models will be compared to determining the impact of non-normality on spatial non-stationarity.

Results of stationarity test for both GWR models of transformed model and non-transformed model were listed in Table 4(a) and Table 4(b) respectively. Clearly, Table 4(a) and Table 4(b) illustrated the non-normality distribution of the regression through the relationships of the spatial non-stationarity. Based on Table 4(a), there were no significant

variation of the relationships across spaces by looking at the  $F$  statistical test for percentages of African Americans (Pct.AA) and percentages of Hispanics (Pct.Hisp.) in the census tracts. While median household income was experienced a significant variation over space. If the basic assumption was invalid, these particular results of justification could be wrongly as in classical regression analysis, the normal distribution can be stationary without effect the final analytical results.

**Table 4(a)** Stationarity test for GWR models of non-transformed data.

	Intercept	Pct.AA	Pct.Hisp.	Median Income
$F$ statistic	5.027	1.066	1.006	1.793
Numerator d.f.	491.937	75.798	106.078	325.729
Denominator d.f.	1927.644	1927.644	1927.644	1927.644
$p$ values	0.000	0.331	0.466	0.000

**Table 4(b)** Stationarity test for GWR models of transformed data.

	Intercept	Pct.AA	Pct.Hisp.	Median Income
$F$ statistic	3.280	2.375	1.746	2.156
Numerator d.f.	465.845	85.059	112.748	324.023
Denominator d.f.	1920.951	1920.951	1920.951	1920.951
$p$ values	0.000	0.000	0.000	0.000

From examinations, the data sets with and without transformed were acceptable in terms of its statistical inference when it was applied on global ordinary least squares regression. These data sets give a similar conclusion in which both supports the common understanding between tobacco outlet distribution and demographic factors.

## CONCLUSION

The analysis of spatially varying relationships using geographically weighted regression (GWR) has been widely employed in variety of application such as in finding the relationship between house price data and demographic factors, health disease and climate factors and rainfall with its elevation. Based on the results of the above-mentioned studies, GWR is the best method that often used to handle the issues of non-stationarity that exist in data sets at varies location. Continuation of this study, GWR can be used in establishing the relationship of various variables at different locations.

## ACKNOWLEDGEMENT

This work was financially supported by the Universiti Teknologi Malaysia under the Research University Grant (QJ130000.2526.17H95) and Ministry of Higher Education Malaysia.

## REFERENCES

- Bivand, R., 2017. Geographically Weighted Regression. *Applied Spatial Data Analysis with R*, (2008), pp.305–308.
- Brunsdon, C., Fotheringham, A. S., Charlton, M., 1999. Some notes on parametric significance tests for Geographically Weighted Regression. *Journal of Regional Science*, 39(3), pp.497–524.
- Brunsdon, C., Fotheringham, A. S., Charlton, M. E., 1996. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, 28(4), pp.281–298.
- Brunsdon, C., Fotheringham, S., Charlton, M., 1998. Geographically Weighted Regression. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(3), pp.431–443.
- Casetti, E., 1972. Generating models by the expansion method: Applications to geographical research. *Geographical analysis*, 4(1), pp.81–91.
- Dobson, A. J., 1990. *An Introduction to Generalized Linear Models*. London: Chapman & Hall/CRC.
- Ebov, O. V., Samat, N., 2015. Modeling property rating valuation using Geographical Weighted Regression (GWR) and Spatial Regression Model (SRM): The case of Kota Kinabalu, Sabah. *Geografia-Malaysian Journal of Society and Space*, 11(11).
- Ehlikes, L., Krefis, A. C., Kreuls, B., Krumpkamp, R., Adjei, O., Ayim-Akonor, M., Kobbe, R., et al., 2014. Geographically weighted regression of land cover determinants of Plasmodium falciparum transmission in the Ashanti Region of Ghana. *International Journal of Health Geographics*, 13, pp.35.
- Fotheringham, A. S. S., Brunsdon, C., Charlton, M., 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. England: Wiley.
- Jamhuri, J., Azhar, B. M. S., Puan, C. L., Norizah, K., 2016. GWR-PM - Spatial variation relationship analysis with Geographically Weighted Regression (GWR) - An application at Peninsular Malaysia. *IOP Conference Series: Earth and Environmental Science*, 37, p.12032. Available at: <http://stacks.iop.org/1755-1315/37/i=1/a=012032?key=crossref.a228250cfa07597971c8b7bec4acd380>.
- Leung, Y., Mei, C. L., 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.
- Lu, B., Charlton, M., Fotheringham, A. S., 2011. Geographically Weighted Regression using a non-euclidean distance metric with a study on London House Price Data. , 7, pp.92–97.
- McKenzie, N. J., Ryan, P. J., 1999. Spatial prediction of soil properties using environmental correlation. *Geoderma*, 89(1–2), pp.67–94.
- Robinson, T. P., Metternicht, G., 2006. Testing the performance of spatial interpolation techniques for mapping soil properties. *Computers and Electronics in Agriculture*, 50(2), pp.97–108.
- Yu, D., Peterson, N. A., Reid, R. J., 2009. Exploring the Impact of non-normality on spatial non-stationarity in Geographically Weighted Regression Analyses: Tobacco Outlet Density in New Jersey. *GIScience & Remote Sensing*, 46(3), pp.329–346.