

Geographically weighted Poisson regression models with different kernels: application to road traffic accident data

Ghanim Al-Hasani, Md Asaduzzaman^a and Abdel-Hamid Soliman

Department of Engineering, School of Digital, Technologies and Arts, Staffordshire University, UK

ABSTRACT

Geographically weighted Poisson regression models (GWPR) are the class of spatial count regression models that capture the localisation effect on various influencing factors on the dependent variable. The main challenge with the GWPR models is to set appropriate kernel function to give weights for each neighbouring point during the model calibration. In this paper, we consider GWPR models for many different kernel functions, including box-car, bi-square, tri-cube, exponential and Gaussian function. Likelihood function, parameter estimation and model selection criteria have been shown in details. We applied the model formulation to the road traffic accident data in Oman as the country is one of the largest road traffic accident-prone countries in the Gulf region. Akaike information criterion (AIC), corrected Akaike information criterion (AICc) and geographically weighted deviance (GWD) have been used to assess the model fitting. The model with the exponential kernel weighted function provides the best fit for the data and captures the spatial heterogeneity and factors better with the exponential kernel weighting function.

KEYWORDS

Geographically weighted Poisson regression; kernel weighting function; model diagnostics; road traffic accidents; spatial modelling.

1. Introduction

1.1. Background

Spatial models are found to be an elegant technique to capture the localisation effect on various influencing factors in many different areas including health, energy, transport, business, earth and environment sciences. As the effect of factors may vary with locations, spatial modelling techniques provides significantly improved estimates and better prediction in comparison to non-spatial models [21]. Ignoring the spatial effect in regression modelling may cause bias and a higher standard error in the estimation of model parameters and a lower amount of regression variation explained by the model [4]. Many studies have been carried out to date, in which spatial models capture spatial effects in a variety of ways based on the different form of spatial distributions. Some of those spatial models are: spatial autoregressive (SAR) models [3, 32], spatial error models (SEM) [1, 31], multiple memberships models (MMM) [12], extended multiple membership models (EMMM) [7]. Although these models are spatial models, the

CONTACT Md Asaduzzaman, Department Engineering, Staffordshire University, Stoke-on-Trent ST4 2DE, United Kingdom. Email: md.asaduzzaman@staffs.ac.uk

parameter estimates remain fixed for all locations. The spatial variations in these models are taken into account only through the spatial error structure [29]. However, there is another type of spatial modelling approach that provides a set of local models obtained by the calibration of multiple geographical entities. Two widely used forms of such models are– geographically weighted regression (GWR) models when the response variable is measured in continuous scale [8, 15, 36] and geographically weighted Poisson regression (GWPR) models when the response variable is a count [21, 27, 29]. These geographically weighted models allow the parameters to vary through the spatial units in a study region to capture the local factors. Therefore, these models focus on the geographic difference of factors affecting the outcome variable [17]. Despite these promising features of a geographically weighted model, there are two main challenges involved in fitting the geographically weighted regression models, which are the selection of bandwidth and kernel weighting function.

During the calibration of a geographically weighted model, the bandwidth, also referred to as window size, controls the fit of a localised model. Small bandwidths result in high spatial variation while the large bandwidth leads to the estimates close to the global models [24]. Generally, there are two suggested methods in the literature for the selection of bandwidth– classical fixed bandwidth approach or adaptive bandwidth approach [11]. Several studies found that adaptive (optimal) bandwidth performs better than fixed bandwidth [11, 20, 36]. Recently, [11] compared the fixed and adaptive (optimal) bandwidth and recommended that adaptive bandwidth is optimal due to its performance, which is estimated through the integrated squared error between a given density estimate and the true density under a given scenario.

The main challenge of a GWR and GWPR model is to find the most suitable kernel weighting function which gives weights for the neighbouring observations during model calibration. If the generated weights through a kernel function are not accurate then the parameter estimates have large standard errors and the fitted model gives less reliable predictions [21]. There are five different kernel functions used for GWR or GWPR modelling: box-car, bi-square, tri-cube, exponential and Gaussian weighted function [21]. These weighting functions generate weights for each observation and use them to calibrate the parameter estimates according to the spatial position of data points [21]. However, the potential problem for a modeller is to select the best possible kernel weighting function to generate the weights [8].

Several studies applied different kernel functions for the geographically weighted models in many areas including road traffic accidents. [21] chose an adaptive bi-square kernel for weighting schemes in a GWPR model. [29] found that Gaussian weighted functions with adaptive bandwidth are the most suitable in GWRP models while comparing the bi-square and the Gaussian kernel weighting functions. Although some studies [17, 27, 37] compared the bi-square and Gaussian kernel functions with both fixed and adaptive bandwidths. Broadly speaking, these empirical studies concentrated only on two kernel functions, bi-square and Gaussian weighted functions, to inspect spatial factors in GWPR models. Despite these efforts, it remains unclear which kernel function would provide the best fit for a geographically weighted Poisson regression (GWPR) model.

In this study, we compared the suitability of a GWPR model for five different kernel weighting func-

tions: box-car, bi-square, tri-cube, exponential and Gaussian weighted function. The model formulation has been shown in details including the likelihood function, estimation of parameter through calibration of models for different kernel weighting functions. The framework has been applied to the road traffic accident (RTA) data in Oman to explore different factors associated with RTA and give insights into geographical variations of such factors. Several diagnostic tools, namely, Akaike information criterion (AIC), corrected Akaike Information Criterion (AICc) and geographically weighted deviance (GWD) have been used to find the most suitable model with the associated kernel weighting function.

The rest of the paper is structured as: model development, formulation and estimation of the likelihood function in Section 2, data and software in Section 3, results and discussion in Section 4 and some concluding remarks in Section 5.

2. Methods

2.1. Poisson count model and likelihood function

Spatial Poisson model, a special form of spatial count model, allows parameter values to vary with spatial unites \mathbf{u}_i which is a vector describing the location i [21, 27]. A spatial Poisson model can be written as

$$Y_i \sim \text{Poisson} \left[\exp \left(\sum_k \beta_k(\mathbf{u}_i) x_{ik} \right) \right],$$

where x_{ik} is k th explanatory variable in location i , β_k is the parameter for the k th explanatory variable and $\mathbf{u}_i = (u_{li}, u_{hi})$ is the vector describing the latitude and longitude at location i . The model is defined with a geographically varying coefficient $\beta_k(\mathbf{u}_i)$ that is a function of location \mathbf{u}_i and known as geographically weighted Poisson regression (GWPR). The covariate form of the model GWPR with a group of predictors, in which the parameters are allowed to vary over space can be written as

$$\ln(\mathbf{Y}) = \ln(\beta_0(\mathbf{u}_i)) + \beta_1(\mathbf{u}_i)\mathbf{X}_1 + \beta_2(\mathbf{u}_i)\mathbf{X}_2 + \dots + \beta_K(\mathbf{u}_i)\mathbf{X}_K + \varepsilon_i,$$

where β_k is the function of the location $\mathbf{u}_i = (u_{li}, u_{hi})$ denotes the two dimensional coordinates of the i th point in space. The parameters in the model $\beta = (\beta_0, \beta_1, \dots, \beta_K)$ are allowed to be different between locations. Thus, the GWPR modelling framework addresses the spatial heterogeneity.

To estimate the parameters of a geographically weighted Poisson regression, a form of maximum likelihood principal is considered [21, 27]. The method is analogous to a locally maximum likelihood principle and can be referred to as geographically weighted likelihood method. The geographically weighted log-likelihood at location \mathbf{u}_i can be given by

$$L(\mathbf{u}_i) = \sum_{j=1}^N \left(-\hat{Y}_j(\beta(\mathbf{u}_i)) + Y_j \log \hat{Y}_j(\beta(\mathbf{u}_i)) \right) \cdot \omega_{ij}(\|\mathbf{u}_i - \mathbf{u}_j\|),$$

where

$$\hat{Y}_j(\boldsymbol{\beta}(\mathbf{u}_i)) = \exp\left(\sum_k \hat{\beta}_k(\mathbf{u}_i)x_{jk}\right),$$

is the predicted number of events at location $\mathbf{u}_i = (u_i, u_{h_i})$ with estimated parameter vector $\hat{\boldsymbol{\beta}}$ at regression point i , ω_{ij} is the geographical weight of the j th observation at i th regression point. The weights ω_{ij} of the observations decreases gradually as the distance between the regression point i and the observation at location j becomes larger.

The model needs to be calibrated as used in geographically weighted regression and smoothed with a spatial weighting function [21, 27]. The spatial weighting function is, in general, the spatial weighting kernel. In particular, the calibration of the GWPR model is performed to generate the localised effect by weighting each observation in the data set according to a proximity point or centred point. Based on a kernel weighting scheme, observations which are close to the proximity points are highly weighted than further points [14]. However, finding the appropriate kernel is crucial as the weights of the neighbouring observations contribute to the estimates through a distance decay function.

2.2. Kernel weighting functions

There are several kernel functions used for weighting in GWPR model including box-car, bi-square, tri-cube, exponential and Gaussian [5, 16, 23, 35]. The simple form of a kernel function is the box-car kernel function. It is a discontinuous function and uses only the observations that are within a distance, say b , from the GWPR model calibration point. However, the box-car function is computationally efficient as it uses a smaller subset of observations to fit the local model at each GWPR model calibration point. The bi-square is also a discontinuous function and gives null weights to spatial units observation with a distance greater than the bandwidth [16]. Similar to bi-square, the tri-cube kernel gives a cubic weight. The exponential kernel is a continuous function of the distance between two spatial observation points or a calibration and an observation points. The weights can be a maximum of equal to 1 for an inspection at the GWPR model calibration point and will have an exponential decrease with the increase of the distance between calibration points [5, 16]. Gaussian kernel function provides the classical Gaussian weights. Table 1 shows the functional forms of different kernel functions.

2.3. Likelihood estimation and scoring method for GWPR

In this section, we provide the details of model formulation including the derivation of likelihood function, model calibration, parameter estimation and model diagnostics as appeared in several literature [21, 27]. Consider a GWPR model of the form

$$Y_i \sim \text{Poisson} \left[\exp\left(\sum_k \beta_k(\mathbf{u}_i)x_{ik}\right) \right],$$

Table 1. Different kernel functions

Kernel	Functional form
Box-car	$\omega_{ij} = \begin{cases} 1 & \text{if } d_{ij} < b, \\ 0 & \text{otherwise.} \end{cases}$
Bi-square	$\omega_{ij} = \begin{cases} \left(1 - \left(\frac{d_{ij}}{b}\right)^2\right)^2 & \text{if } d_{ij} < b, \\ 0 & \text{otherwise.} \end{cases}$
Tri-cube	$\omega_{ij} = \begin{cases} \left(1 - \left(\frac{ d_{ij} }{b}\right)^3\right)^3 & \text{if } d_{ij} < b, \\ 0 & \text{otherwise.} \end{cases}$
Exponential	$\omega_{ij} = \exp\left(-\frac{ d_{ij} }{b}\right).$
Gaussian	$\omega_{ij} = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right).$

where ω_{ij} is the j th element of the diagonal matrix when a model is calibrated for i calibration point, d_{ij} is the distance between i and j th points and b is the chosen bandwidth.

with the likelihood function

$$L(\mathbf{u}_i) = \sum_{j=1}^N \left(-\widehat{Y}_j(\beta(\mathbf{u}_i)) + Y_j \log \widehat{Y}_j(\beta(\mathbf{u}_i)) \right) \cdot \omega_{ij}(\|\mathbf{u}_i - \mathbf{u}_j\|).$$

The parameter vector β is estimated by solving

$$\frac{\partial L(\mathbf{u}_i)}{\partial \widehat{\beta}(\mathbf{u}_i)} = 0. \quad (1)$$

Equation (1) gives

$$\frac{\partial}{\partial \widehat{\beta}} \left[\sum_{j=1}^N \left(-\widehat{Y}_j(\beta(\mathbf{u}_i)) + Y_j \log \widehat{Y}_j(\beta(\mathbf{u}_i)) \right) \cdot \omega_{ij}(\|\mathbf{u}_i - \mathbf{u}_j\|) \right] = 0. \quad (2)$$

Maximum likelihood estimates are obtained by using the iterative re-weighted least square method using (2) which can also be referred as local Fisher's scoring method and gives as

$$\beta^{(l+1)}(\mathbf{u}_i) = \left(\mathbf{X}^T \Psi(\mathbf{u}_i) \Gamma^{(l)}(\mathbf{u}_i) \mathbf{X} \right)^{-1} \mathbf{X}^T \Psi(\mathbf{u}_i) \Gamma^{(l)}(\mathbf{u}_i) \mathbf{y}^{*(l)}(\mathbf{u}_i), \quad (3)$$

where $\beta^{(l+1)}(\mathbf{u}_i)$ is a vector of local parameter estimates specific to location i while superscript $(l+1)$ gives the number of iterations. At l th stage, the parameter vector

$$\beta^{(l)}(\mathbf{u}_i) = (\beta_0^{(l)}(\mathbf{u}_i), \beta_1^{(l)}(\mathbf{u}_i), \dots, \beta_K^{(l)}(\mathbf{u}_i))^T,$$

where X^T denotes the transpose of the design matrix X , $\Psi(\mathbf{u}_i)$ is the spatial weight matrix and $\Gamma^{(l)}(\mathbf{u}_i)$ is the variance weights matrix for the location i . The vector of adjusted dependent variables $\mathbf{y}^{*(l)}(\mathbf{u}_i)$ is given by

$$\mathbf{y}^{*(l)}(\mathbf{u}_i) = \left(y_1^{*(l)}(\mathbf{u}_i), y_2^{*(l)}(\mathbf{u}_i), \dots, y_N^{*(l)}(\mathbf{u}_i) \right)^T,$$

where

$$\begin{aligned} y_j^{*(l)}(\mathbf{u}_i) &= \left(\beta_0^{(l)}(\mathbf{u}_i) + \sum_k^K \beta_k^{(l)}(\mathbf{u}_i) x_{jk} \right) + \frac{Y_j - \widehat{Y}_j(\beta^{(l)}(\mathbf{u}_i))}{\widehat{Y}_j(\beta^{(l)}(\mathbf{u}_i))} \\ &= \eta_j(\beta^{(l)}(\mathbf{u}_i)) + \frac{Y_j - \widehat{Y}_j(\beta^{(l)}(\mathbf{u}_i))}{\widehat{Y}_j(\beta^{(l)}(\mathbf{u}_i))} \end{aligned}$$

and η_j is a linear predictor of j th observation. The sets of the local parameter can be estimated by repeating the iterative procedure for each location i . The final estimate can be written as

$$\beta(\mathbf{u}_i) = (\mathbf{X}^T \Psi(\mathbf{u}_i) \Gamma(\mathbf{u}_i) \mathbf{X})^{-1} \mathbf{X}^T \Psi(\mathbf{u}_i) \Gamma(\mathbf{u}_i) \mathbf{y}^*(\mathbf{u}_i).$$

The prediction of each observation j at each regression points i would be

$$\widehat{Y}_{ij} = \mathbf{C} \mathbf{y}^*(\mathbf{u}),$$

where the i th row of the matrix \mathbf{C} is

$$\mathbf{c}_i = \mathbf{x}_i (\mathbf{X}^T \Psi(\mathbf{u}_i) \Gamma(\mathbf{u}_i) \mathbf{X})^{-1} \mathbf{X}^T \Psi(\mathbf{u}_i) \Gamma(\mathbf{u}_i).$$

To predict the number of events at a location, the diagonal elements of the matrix on the lefthand side in equation (3) are utilised. The mapping from adjusted dependent variable to linear predictor can be given below

$$\widehat{\boldsymbol{\eta}} = \mathbf{S} \mathbf{y}^*,$$

where

$$\widehat{\boldsymbol{\eta}} = (\widehat{\eta}_1(u_1), \widehat{\eta}_2(u_2), \dots, \widehat{\eta}_N(u_N))^T \text{ and } \mathbf{y}^* = (y_1^*(\mathbf{u}_1), y_2^*(\mathbf{u}_2), \dots, y_N^*(\mathbf{u}_N))^T.$$

Analogous to the hat matrix, \mathbf{S} is the matrix with the (ij) th element S_{ij} being

$$S_{ij} = C_{ij} \frac{y_i^*(\mathbf{u}_j)}{y_j^*(\mathbf{u}_j)},$$

where C_{ij} is the (ij) th element of the matrix \mathbf{C} .

2.4. Model diagnostic tool

To compare and evaluate the GWPR model performance with five aforementioned kernel weighting functions, we have used three most common methods of model evaluation: Akaike information criterion (AIC) [33], corrected Akaike information criterion (AICc) [9] and geographically weighted deviance (GW) [27]. The AIC function for a model with bandwidth b is defined as

$$\text{AIC}(b) = D(b) + 2K(b),$$

where D is the deviance of the model, K is the effective number of parameters in the model and b is the bandwidth used in the kernel weighting function for the model. A model with minimum AIC value has the higher goodness of fit [27]. However, in case of a local spatial regression model with a small degrees of freedom, a bias adjustment in the AIC is introduced [29]. The Akaike information criterion with bias correction (AICc) can then be defined as

$$\text{AICc}(b) = D(b) + 2 \left(K(b) + \frac{K(b)(K(b) + 1)}{N - K(b) - 1} \right),$$

where N is the number of spatial observations. For a geographically weighted Poisson regression model the deviance statistic, referred to as geographically weighted deviance (GWD) [27], can be given by

$$\text{GWD}(b) = \left[\sum_{i=1}^N Y_i \log \left(Y_i / \hat{Y}_i(\boldsymbol{\beta}(\mathbf{u}_i), b) \right) - \sum_{i=1}^N \left(Y_i - \hat{Y}_i(\boldsymbol{\beta}(\mathbf{u}_i), b) \right) \right].$$

where Y_i is observed number of events in the i th location, \hat{Y}_i is the predicted value for the i th location and $\boldsymbol{\beta}(\mathbf{u}_i)$ is the parameter vector for the set of K explanatory variables. The most suitable would be the one with the lowest values of the Akaike information criterion (AIC), corrected Akaike information criterion (AICc) and geographically weighted deviance (GWD).

3. Application: road traffic accident

Statisticians have prompted to find ways to gain a better understanding of road traffic accidents (RTA) due to the loss of lives and high cost to the societies. More than 1.2 million people die every year in roadway-related accidents and over fifty million faces different form of injuries throughout the world. By 2030, road traffic accidents are forecasted to be the fifth main cause of deaths globally [25]. Moreover, it has been considered as one of the significant causes of health problems in terms of death and

disability [6, 19]. Therefore, statistical analysis of RTA data has become one of the major concern for the development and implementation of the standards and diverse road safety policies [26]. Statistical modelling, analysis and prediction of road traffic accidents would provide a greater depth of understanding of the causes, associated factors and implement or revise policies and countermeasures to reduce the number of RTA [22]. An extensive amount of research on road traffic accident has been performed in many different contexts to achieve different goals [26]. Many researchers tried to quantify the effect of significant determinants (explanatory variables)[1, 6, 17], or to use the statistical models to forecast the number of accidents [2, 12, 13], or to evaluate the efficiency of identified safety countermeasures [21, 29, 32] .

Majority of the studies in RTA show that the number of road traffic accidents significantly varied over space or regions [1, 6, 17]. Therefore, statistical modelling and analysis should be contemplated with the spatial correlation and geographical factors for road traffic accident research. Recently, a considerable amount of literature has been found on the theme of spatial modelling and analysis of RTA. Several studies applied the geographically weighted Poisson regression (GWPR) model in accident analysis [21, 29]. As mentioned in the introduction, selecting the best kernel weighting function with a GWPR model is a complex task for any data as it depends on how the neighbouring units affect an observed outcome. Furthermore, we found that only bi-square and Gaussian kernel functions have been used to evaluate the GWPR models for road traffic accident data [21, 29]. To the best of our knowledge, there is no study in the literature, which evaluates all the kernel weighting functions for the traffic accidents data. Therefore, we attempt to evaluate geographically weighted Poisson regression (GWPR) models for different kernel weighting functions to choose the most suitable one for the road traffic accident data in this study.

3.1. Data and software

The road traffic accidents (RTA) data of 2017 have been collected for the Sultanate of Oman for this study. The data includes the number of road traffic accidents by eleven governorates in Oman, which have been considered as spatial units in this study. The RTA database in Oman is maintained by the Royal Omani Police (ROP). ‘Statistical Summary Bulletins’ are published annually by the Directorate of Road Traffic as part of the ROP [30]. In this study, we also use explanatory variables such as population size, population density, number of registered vehicles, number of unemployed persons, speed driving, season (February-June-July or otherwise). Population size and population density data are collected from the National Centre for Statistics & Information (NCSI) monthly reports, called ‘Monthly Statistical Bulletin’ [28]. The number of jobs seekers (unemployment) data in each governorate of Oman were collected from the Public Authority of Manpower register. Number of registered vehicles in each governorate, speed driving, season are collected from the RTA database.

There are several tools and softwares that can be implemented to model and analyse road traffic accidents. The open-source software R has been extensively used as a sophisticated tool to fit models

and analyse data in many areas [34]. We have implemented the R software in our model framework to perform the spatial analysis and fit the geographically weighted Poisson regression models.

4. Results and discussion

In this study, five geographically weighted Poisson regression (GWPR) models are evaluated and compared with different kernel weighting functions including box-car, bi-square, tri-cube, exponential and Gaussian kernel weighting function. We have considered GWPR models with different kernel weighting functions to evaluate the model performance and find out the most suitable one. There are eleven spatial units in this study, namely, governorates (regions) in Oman. The bandwidth size also plays a key role during model calibration to give weights for the nearest spatial units. The adaptive bandwidth was employed for each kernel weighting functions in our work. Three common diagnostic tools have been used to evaluate the performance of the GWPR models: geographically weighted deviance (GWD), Akaike information criterion (AIC) and the corrected Akaike Information Criterion (AICc).

Our secondary data consist of seven variables: number of RTA as the dependent variable and six variables: population density, population size, number of registered vehicles, number of unemployed persons, speed driving and seasons as explanatory variables for the Sultanate of Oman in 2017. There are eleven governorates considered as the eleven spatial units for the GWPR model. A primary analysis has been conducted by fitting a Poisson generalised linear model to learn about the significant factors. Table 2 provides the fitting of the global Poisson regression model in generalised linear modelling (GLM) framework. However, the Poisson GLM model shows that population size, number of registered vehicles, number of unemployed persons and season are significant.

Table 2. Poisson GLM fitting

Coefficients	Estimate	Standard Error	Z-value	p-value
Intercept	4.40e+00	5.19e-02	84.82	< 2e-16***
Population size	8.42e-07	4.15e-07	2.03	0.042*
Population density	8.28e-04	1.46e-03	0.57	0.570
No. of registered vehicles	-3.16e-06	7.75e-07	-4.08	4.54e-05***
No. of unemployed persons	2.42e-05	1.46e-05	1.66	0.097
Speed driving	1.48e-03	1.31e-03	1.13	0.260
Season	7.08e-03	2.96e-03	2.39	0.017*

Level of significance: * $p < 0.05$, ** $p < 0.001$, *** $p < 0.0001$

The mean and variance of the number of accidents in Oman are found to be 350 and 110,577, respectively per governorate per year. This clearly indicates the over-dispersion in the data and the inadequacy of the fitted global Poisson GLM model in Table 2. Therefore, we fitted a generalised Poisson GLM [10] and a negative binomial GLM [18] and the evaluation of model fitting is shown in Table 3 in the next

section. However, the high variation of the number of accidents at different governorate level provides an indication that different factors may have impacted the RTA differently at each governorate level and local level models may explain the variation more clearly. This motivates applying the geographically weighted Poisson regression.

4.1. GWPR model evaluation

One of the main aims of this study is to evaluate different kernels for the GWPR models with Oman road traffic accident data and suggest a model with the most suitable one. The results, shown in Table 3, are obtained from the evaluation of the five GWPR models with five different kernel weighting functions. As mentioned earlier, the most suitable model is the one which has the lowest values of GWD, AIC and AICc. As expected, the Poisson GLM is found to be the worst model as the model gives the highest residual deviance, AIC and AICc values in comparison to any GWPR model.

Although bi-square and Gaussian kernel functions have been found in the literature, in particular, for modelling and analysing road traffic accident data, however, the other kernel methods are also found to be equally potential from the analytical point of view. Box-car kernel weighting function is the simplest form of the kernel functions and GWPR model with box-car function (GWPR₁) gives the highest values for all three indicators: GWD equals to 107.76, AIC 121.76 and AICc 159.09. Gaussian kernel function seems to be the most popular kernel method for spatial analysis and also for road traffic accident data. Interestingly, the GWPR model with Gaussian kernel (GWPR₅) found to be the second least preferred method for our data as the model gives second largest values of GWD equals to 98.95, AIC 113.22 and AICc 153.74. The GWPR model with a bi-square kernel function gives a substantially better result than the model with Gaussian kernel function (Table 3). Comparison between the GWPR models with bi-square and tri-cube kernel function is difficult as the GWD and AIC values are lower for the model with the tri-cube kernel while the AICc values are higher (GWPR₂ vs GWPR₃).

Table 3. Evaluation of different GWPR models for different kernel weighting functions

Model	Kernel weighting function	Deviance* or GWD	AIC	AICc
GLM (Poisson)	—	107.76*	202.29	239.62
GLM (Generalised Poisson)	—	107.76*	108.14	145.48
GLM (Negative binomial)	—	11.60*	137.18	209.18
GWPR ₁	Box-car	107.76	121.76	159.09
GWPR ₂	Bi-square	84.14	99.00	147.75
GWPR ₃	Tri-cube	80.56	95.79	150.61
GWPR ₄	Exponential	75.37	90.46	143.09
GWPR ₅	Gaussian	98.95	113.22	153.74

Deviance* is used for GLMs (Poisson, generalised Poisson and negative binomial models)

Clearly, the GWPR model with exponential kernel weighting function (GWPR₄) has the lowest values

of GWD (75.36), AIC (90.46) and AICc (143.09) in comparison to those obtained from the other models. Here all three indicators give the lowest value for the model and can be considered as the most suitable one. Note that we used the adaptive kernel method for bandwidth selection as it has been suggested as the optimal bandwidth for spatial analysis.

4.2. GWPR model estimate

In our analysis, GWPR model (with exponential kernel weighting function and adaptive bandwidth) (GWPR₄) found to be the most suitable one for road traffic accident data in Oman. Therefore, we produce the local parameter estimates from the fitted model and also the maps for the parameter estimates for the variables: population density, population size, number of registered vehicles, number of unemployed persons, speed driving and season. Table 4 provides the summary statistics of the local parameter estimates obtained from the GWPR₄ model.

Table 4. Summary of local parameter estimates by GWPR model with exponential kernel weighting function

Variable	Min.	1st Quartile	Median	3rd Quartile	Max.
Intercept	$4.37e+00$	$4.41e+00$	$4.44e+00$	$4.45e+00$	4.450
Population density	$-6.90e-04$	$-2.97e-04$	$9.08e-04$	$1.51e-03$	0.002
Population size	$6.78e-07$	$7.62e-07$	$8.11e-07$	$8.87e-07$	0.000
No. of registered vehicles	$-3.45e-06$	$-3.20e-06$	$-2.94e-06$	$-2.74e-06$	0.000
No. of unemployed persons	$1.39e-05$	$1.75e-05$	$2.43e-05$	$2.95e-05$	0.000
Speed driving	$8.38e-04$	$1.82e-03$	$2.49e-03$	$2.81e-03$	0.003
Season	$2.96e-03$	$4.11e-03$	$4.65e-03$	$5.46e-03$	0.009

The parameter estimates of the GWPR model are necessarily the estimated values of local coefficients, which are plotted by regions (governorates) in Figure 1. The maps in Figure 1 demonstrate the spread of local coefficients values over the geographical regions (eleven governorates) for Oman. It is clear from Figure 1 that the effects of different explanatory variables are varying over the regions in Oman. The estimated coefficients show that population size, density, number of unemployed persons, speed driving and season have a positive impact on the number of accidents in the governorates in Oman. Interestingly, the number of registered vehicles found to have a negative impact on accidents in Oman. The parameter estimates of population density range from 0.0005 to 0.0025 and the number of accidents in east governorates (both North and South Sharqiyah) are affected the most by population density in comparison to other regions while the governorates in the north-west are less affected as displayed in Figure 1(a). The estimates of the parameter for Population size range between $7.00e-7$ and $1.05e-06$ and Musandam (in the most north of the country) is the highest affected governorate by population size, see Figure 1(b). From Figure 1(d) and 1(f), we see that Dhofar is the governorate in the south, where the number of RTA affected by the unemployment and season. Moreover, the region is less affected by

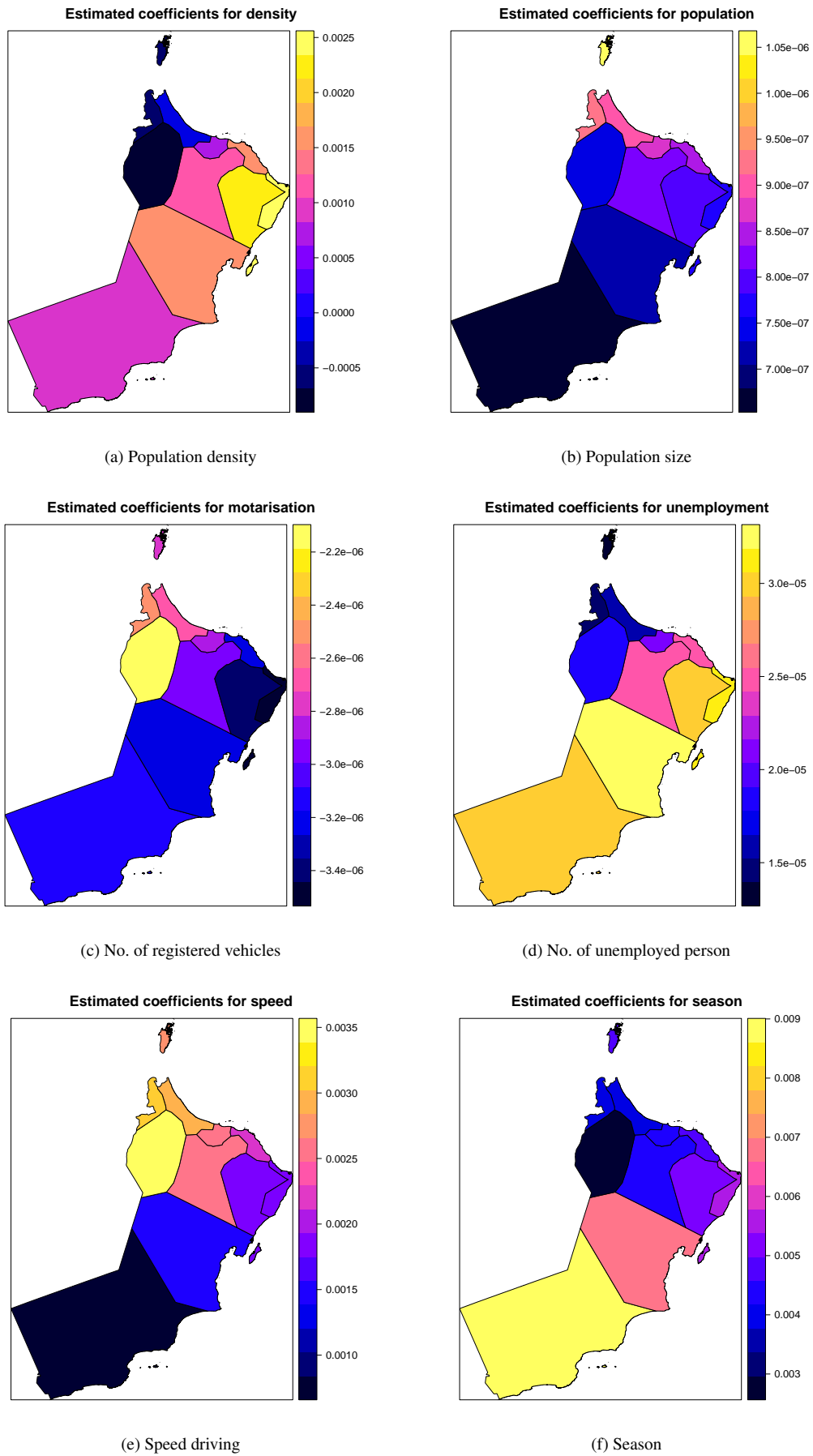


Figure 1. The estimates of local coefficients from the GWPR model with the exponential kernel.

population size, the number of registered vehicles and speed driving than other governorates in Oman as showing in Figure 1(b), 1(c) and 1(e). The variable season is found to have the highest impact on the region Dhofar because Dhofar celebrates the ‘Salalah Tourism Festival’ in the June-August period and the number of accident is significantly higher during that period.

Our model showed the spatial variation of the estimates of the parameters for the data and considered the effect of neighbouring regions. The map gives a clear indication to the policy-makers, relevant government departments and researchers to consider the associated factors related to the number of accidents in Oman and take appropriate actions and measures to prevent accidents in Oman.

5. Conclusion

This study aimed to fit geographically weighted Poisson regression (GWPR) models with different kernel weighting functions and evaluate the models to find out the most suitable one in the context of road traffic accident data. We compared the performance of five GWPR models with five kernel weighting functions including box-car, bi-square, tri-cube, exponential and Gaussian weighting function by implementing the model formulation and evaluation to the road traffic accident data in Oman. The study would help the decision-makers to assess the factors associated with the number of accidents so that necessary and appropriate measures can be taken to reduce the number of accidents. The model framework has been implemented through the open-source software R to calibrate the GWPR models, fit the most suitable and estimates the parameters. We used three popular information criteria– geographically weighted deviance (GWD), Akaike information criterion (AIC) and the corrected Akaike Information Criterion (AICc) to evaluate the performance of models.

The study found that GWPR models can substantially capture the heterogeneity of the spatial factors over the regions or spatial units. The crucial finding to emerge from this study is that the GWPR model with the exponential kernel function and the adaptive bandwidth is the most suitable for modelling, fitting and analysing road traffic accident data. Although several studies suggested that the GWPR model with bi-square kernel function with the adaptive bandwidth is more suitable [21, 27], our results suggest that the GWPR model with the exponential kernel weighting function outperformed over other weighting functions. Some studies used the GWPR model with the Gaussian weighting kernel function assuming that the Gaussian kernel function would give better result than others [29]. We found that the GWPR model with the tri-cube kernel function or even Gaussian kernel function is less adequate than the model with the exponential kernel. As we applied the model formulation to a specific dataset the GWPR model with exponential kernel weighting is not generalisable for other types of datasets. However, the proposed model formulation shows– how to apply and choose the best GWPR models for different kernels with a specific dataset.

References

- [1] G. Al-Hasani, M. Asaduzzaman, and A.H. Soliman, *Comparison of spatial regression models with road traffic accidents data*, in *ICSTA'19 International Conference*. AVESTIA, 2019, pp. 31:1–31:4.
- [2] H. Al-Reesi, S.S. Ganguly, S. Al-Adawi, L. Laflamme, M. Hasselberg, and A. Al-Maniri, *Economic growth, motorization, and road traffic injuries in the Sultanate of Oman, 1985–2009*, *Traffic Injury Prevention* 14 (2013), pp. 322–328.
- [3] L. Anselin, *Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity*, *Geographical Analysis* 20 (1988), pp. 1–17.
- [4] S. Barua, K. El-Basyouny, and M.T. Islam, *Effects of spatial correlation in random parameters collision count-data models*, *Analytic Methods in Accident Research* 5 (2015), pp. 28–42.
- [5] P.E. Bidanset and J.R. Lombard, *Optimal kernel and bandwidth specifications for geographically weighted regression: an evaluation using automated valuation models (avms) for mass real estate appraisal*, in *Applied Spatial Modelling and Planning*, Routledge, 2016, pp. 131–144.
- [6] A. Boulieri, S. Liverani, K. de Hoogh, and M. Blangiardo, *A space–time multivariate Bayesian model to analyse road traffic accidents by severity*, *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 180 (2017), pp. 119–139.
- [7] W.J. Browne, H. Goldstein, and J. Rasbash, *Multiple membership multiple classification (MMMC) models*, *Statistical Modelling* 1 (2001), pp. 103–124.
- [8] C. Brunson, A.S. Fotheringham, and M.E. Charlton, *Geographically weighted regression: a method for exploring spatial nonstationarity*, *Geographical Analysis* 28 (1996), pp. 281–298.
- [9] J.E. Cavanaugh, *Unifying the derivations for the akaike and corrected akaike information criteria*, *Statistics & Probability Letters* 33 (1997), pp. 201–208.
- [10] P. Consul and F. Famoye, *Generalized Poisson regression model*, *Communications in Statistics - Theory and Methods* 21 (1992), pp. 89–109.
- [11] T.M. Davies and A.B. Lawson, *An evaluation of likelihood-based bandwidth selectors for spatial and spatiotemporal kernel estimates*, *Journal of Statistical Computation and Simulation* (2019), pp. 1–22.
- [12] K. El-Basyouny and T. Sayed, *Urban arterial accident prediction models with spatial effects*, *Transportation Research Record* 2102 (2009), pp. 27–33.
- [13] B. Flahaut, M. Mouchart, E. San Martin, and I. Thomas, *The local spatial autocorrelation and the kernel method for identifying black zones: a comparative approach*, *Accident Analysis & Prevention* 35 (2003), pp. 991–1004.
- [14] A.S. Fotheringham, R. Crespo, and J. Yao, *Geographical and temporal weighted regression (GTWR)*, *Geographical Analysis* 47 (2015), pp. 431–452.
- [15] A. Fotheringham, C. Brunson, and M. Charlton, *Geographically weighted regression: the analysis of spatially varying relationships*. (2002).
- [16] I. Gollini, B. Lu, M. Charlton, C. Brunson, and P. Harris, *GWmodel: an R package for exploring spatial heterogeneity using geographically weighted models*, arXiv preprint arXiv:1306.0413 (2013).
- [17] A.M. Hezaveh, R. Arvin, and C.R. Cherry, *A geographically weighted regression to estimate the comprehensive cost of traffic crashes at a zonal level*, *Accident Analysis & Prevention* 131 (2019), pp. 15–24.

- [18] J.M. Hilbe, *Negative binomial regression*, Cambridge University Press, 2011.
- [19] M.M. Islam and A.Y. Al Hadhrami, *Increased motorization and road traffic accidents in Oman*, *Journal of Emerging Trends in Economics and Management Sciences* 3 (2012), p. 907.
- [20] W. John Braun and V. Rousson, *An autocorrelation criterion for bandwidth selection in nonparametric regression*, *Journal of Statistical Computation and Simulation* 68 (2000), pp. 89–101.
- [21] Z. Li, W. Wang, P. Liu, J.M. Bigham, and D.R. Ragland, *Using geographically weighted Poisson regression for county-level crash modeling in California*, *Safety Science* 58 (2013), pp. 89–97.
- [22] D. Lord and F. Mannering, *The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives*, *Transportation Research Part A: Policy and Practice* 44 (2010), pp. 291–305.
- [23] B. Lu, M. Charlton, C. Brunson, and P. Harris, *The minkowski approach for choosing the distance metric in geographically weighted regression*, *International Journal of Geographical Information Science* 30 (2016), pp. 351–368.
- [24] B. Lu, P. Harris, M. Charlton, and C. Brunson, *The GWmodel R package: further topics for exploring spatial heterogeneity using geographically weighted models*, *Geo-spatial Information Science* 17 (2014), pp. 85–101.
- [25] F. Mannering and C. Bhat, *Analytic methods in accident research: methodological frontier and future directions*, *Analytic Methods in Accident Research* 1 (2014), pp. 1–22.
- [26] F. Mannering, *Temporal instability and the analysis of highway accident data*, *Analytic Methods in Accident Research* 17 (2018), pp. 1–13.
- [27] T. Nakaya, A.S. Fotheringham, C. Brunson, and M. Charlton, *Geographically weighted Poisson regression for disease association mapping*, *Statistics in Medicine* 24 (2005), pp. 2695–2717.
- [28] NCSI, *Monthly Statistical Bulletin September 2017*, National Centre for Statistics & Information, Sultanate of Oman, 2017.
- [29] A. Pirdavani, T. Bellemans, T. Brijs, and G. Wets, *Application of geographically weighted regression technique in spatial analysis of fatal and injury crashes*, *Journal of Transportation Engineering* 140 (2014), p. 04014032.
- [30] R.O. Police, *Facts and Figures*, Director General of Traffic, 2017.
- [31] M.A. Quddus, *Modelling area-wide count outcomes with spatial correlation and heterogeneity: an analysis of London crash data*, *Accident Analysis & Prevention* 40 (2008), pp. 1486–1497.
- [32] K.A. Rhee, J.K. Kim, Y.I. Lee, and G.F. Ulfarsson, *Spatial regression analysis of traffic crashes in Seoul*, *Accident Analysis & Prevention* 91 (2016), pp. 190–199.
- [33] Y. Sakamoto, M. Ishiguro, and G. Kitagawa, *Akaike information criterion statistics*, Dordrecht, The Netherlands: D. Reidel 81 (1986).
- [34] R.C. Team, et al., *R: A language and environment for statistical computing*, Vienna, Austria, 2013.
- [35] C.H. Wang and N. Chen, *A geographically weighted regression approach to investigating the spatially varied built-environment effects on community opportunity*, *Journal of Transport Geography* 62 (2017), pp. 136–147.
- [36] W. Wang and D. Li, *Structure identification and variable selection in geographically weighted regression models*, *Journal of Statistical Computation and Simulation* 87 (2017), pp. 2050–2068.
- [37] P. Xu and H. Huang, *Modeling crash spatial heterogeneity: random parameter versus geographically weighting*, *Accident Analysis & Prevention* 75 (2015), pp. 16–25.