EXPLANATORY ANALYSES OF WORK TRIP GENERATION USING MIXED GEOGRAPHICALLY WEIGHTED REGRESSION (MGWR)

M. Shahri *[1], M. A. Ghannadi [1]

[1] Dept. of Geoscience Engineering, Arak, Iran – (shahri, m.ghannadi@arakut.ac.ir)

Commission IV, WG IV/3

KEY WORDS: Trip production, Trip attraction, Spatial non-stationarity, Mixed geographically weighted regression, Ordinary Least Square, Moran's I

ABSTRACT:

In transportation planning, forecasts have commonly followed the sequential four-step model in which, trip generation (production and attraction) plays a critical role. Among the methods applied to model trip generation, regression with Gaussian distribution of errors are recognized as the most prevailing techniques to describe the relationships between production/attraction and explanatory variables by estimating the global, fixed coefficients. Considering that, trip generation is almost impressed by spatial factors which vary over the study area; the main objective of this research is to apply Mixed Geographically Weighted Regression (MGWR) on 253 traffic analysis zones (TAZs) in Mashhad, Iran, by applying travel demand data and relating factors in 2018 to investigate the spatial non-stationarity which are not revealed when global specifications are applied. The influence of certain explanatory variables on response variables may be global, whereas others are local, accordingly, MGWR performs better compared with geographically weighted regression. The results of Moran's I as spatial autocorrelation index performing on residuals of global, mixed models proved the reliability of the proposed model over the traditional one. The spatial model indicated an improvement in model performance using goodness-of-fit criteria with the coefficient of determination varying from 0.84-0.95 compared with 0.76 and 0.6 in the conventional model. The results of such analysis can provide descriptive and predictive tools at the planning-level for decision-makers.

1. INTRODUCTION

Travel demand modelling is one of the most important steps of transportation planning applied to forecast travel characteristics and application of transportation under different socioeconomic states alternative transport service and land-use configurations. Traditional four-step models consist of four main stages, composing trip generation, trip distribution, modal split and assignment. The model starts by defining a study area, dividing it into multiple Traffic Analysis Zones (TAZ), and considering the entire transport network in the system. Then, the trip generation model is evolved in which, land use, socio-economic and demographic data are used to estimate the total number of trips generated by each zone.

Considering the fact that land use often are divided into two broad categories including residential and non-residential, trips can be home-based (HB) or non-home-based (NHB). Home-based trips are those in which one end of the trip is the traveller's home while in non-home-based trips neither ends of the trip are the traveller's home. Trip ends are modelled as productions or attractions. The home-end of a trip is always the production and the non-home-end is the attraction (for NHB trips, the origin is the production and the destination is the attraction). Trips can also be categorized by trip purpose, for example trips for work, education, shopping, etc. Among these, home-based work trips usually contribute a major portion of total trips in a metropolitan area. Work trips are regular in terms of frequency and time of departure/destination, highly tied to morning and evening peaks. Such concentration extensively affects the design of transportation infrastructures. Additionally, a large number of non-work trips such as shopping, personal

business and childcare are planned according to work trips. The critical role of work trips in daily activities and the related issues have widely been investigated by many researchers from different points of view (Hendrickson and Plank, 1984, Palma and Rochat, 2000, Sohn, 2005, Nurul Habib et al., 2009).

In terms of spatial units, trip generation models are constructed either as aggregate (based on TAZs), or disaggregate. Among techniques suggested to predict trip generation, regressions are most attracted because of simplicity and easy implementation. In current ordinary least square (OLS) regression techniques applied in trip generation (with assuming Gaussian distribution for errors), fixed coefficients are estimated to describe the relationships between number of trips with a specific purpose and every one of explanatory variables. In other words, we assume that the same relationships are held throughout the entire study area (2014). Such assumption in conventional OLS models hides some substantial spatial aspects influencing trip making; therefore, the accuracy of such models might be doubted. Indeed, some explanatory factors might have strong predictive power to estimate number of trips in one location, but might act weaker elsewhere (spatial non-stationarity). Furthermore, when dealing with spatial data, it is expected that the measurements made at nearby locations maybe closer in value than measurements made at locations further apart. This phenomenon which is regarded as spatial autocorrelation (Anselin, 1995) indicates the correlation of one variable with itself in space. In global OLS, the error terms of the model are assumed to be independent, so in case of any spatial autocorrelation, applying common OLS models might to be not true. To handle such restrictions, some advanced statistical

---

* Corresponding author

methods have been introduced in last decades. Geographically Weighted Regression (GWR), a local regression technique, is one of these methods that provides calibration of multiple regression models. (Brunsdon et al., 2002). A major highlight of GWR is that regression coefficients can be estimated at arbitrary spatial locations. Additionally, residuals of GWR have more spatial randomness than the errors resulting from OLS. GWR coefficients are estimated using a spatial weight matrix in which the data around the sample points are weighted using a distance decay function, i.e. the closer observations have a stronger effect on local regression coefficients. The main output from GWR is a set of location-based parameters that can be indicated on a map and are analysed to derive information on spatial non-stationarity. It is also recognized as a useful tool for regional analysis and policymaking. Over the past years, GWR technique has been applied to many scientific fields such as social sciences (Powers et al., 2021) ,health (Wang and Wu, 2020) urban development and planning (Yu et al., 2021), climatology (Xu and Zhang, 2021) and transportation (Pagliara and Mauriello, 2020, Yu and Peng, 2019) . Although many researchers agree that spatial characteristics often influence travel behaviour (Stead, 2001, Lloyd and Shuttleworth, 2005, Cardozo et al., 2012), very few studies consider the relative importance and significance of spatial circumstances specifically on travel demand steps.

Although GWR appears to be suitable method to investigate spatial non-stationarity, in practical cases, the effect of certain independent variables upon response variable may be global, whereas others are local. Therefore, in 1999 Brunsdon et al. proposed a new version of the model called Mixed Geographically Weighted Regression (MGWR). Some coefficients in this model are assumed to be constant, while others are allowed to vary spatially across the study area. Among studies which employed MGWR (Zeng et al., 2016), no research is found in travel demand studies particularly in work trip analysis. Taking into account the importance role of spatial properties of data related to work trips and its explanatory variables, and filling the gap in the literature, the main purpose of this study is to reveal some aspects of spatial patterns which are hidden when employing global OLS in travel demand analysis and to explore the relationship between work trips and its related factors through presenting a mixed model (MGWR).

In this research, Mashhad is studied as one of the most populated city in Iran. As in other metropolitan areas, the socio-economic growth over the past three decades and the land-use diversity affecting the generation of travel through the TAZ have contributed to a comprehensive study and robust study of travel demand in this city. It emphasizes the need to develop predictive models. To this end, primarily, results of conventional OLS for work trip production and attraction will be presented. Next, GWR will be developed and variables are tested for the spatial non-stationarity state and examining whether MGWR could represent a clear enhancement over pure GWR. Models will also be compared in terms of goodness-of-fit criteria such as AICc, MSE, ANOVA, and adjusted coefficient of determination. In following, we discuss the descriptive power of mixed models for investigating local variations in candidate variables and spatial non-stationarity in the study area. To explore spatial dependencies among residuals of OLS and mixed models, Univariate Moran is I as one of the most common methods of detecting spatial dependencies will be employed. The method presented in this paper not only has never been adopted in the literature on the topic, but allows for a more precise comprehension of spatial relationship between trip making and related factors at the local scale. Such

descriptive analysis provides useful information for decision-makers along with predictive models.

2. DATA PREPRATION

In this paper, several types of databases are applied; the first includes total number of work trips generated by 253 TAZs in Mashhad collected based on household interview in year 2018 at the residential-end of trips to update the comprehensive transportation studies and to construct and validate the travel demand models. Table 1 indicates the portion of trips based on different purposes in study area in 2018. As evident, work trips constitute major portion of trips in the region.

Trip Purpose	Portion
Work	26%
Non-Home Based (NHB)	10%
Pilgrimage	6%
Shopping	13%
Personal	10%
Recreation	14%
Education	21%

Table 1. Portion of trips based on purposes in Mashhad in 2018

Category	Variable	Description	Min	Max
General	TAZ	Traffic analysis zone	1	253
	Area m2	TAZ area in km2	1.65	19.04
Trip generation Production/Attraction	Pw	Work trip production	0	23206
	Aw	Work trip Attraction	0	1335898
Independent variables	EL_i	Number of employees living in ith TAZ	0	14120
	CAR_i	Average car ownership in ith TAZ	0	0.31
	Pi	Population per TAZ	0	42564
	EW_i	Number of employees working in ith TAZ	0	3850
	Se_i	Number of salesmen and businessmen	6	8758
	BU_i	Number of business units (commonly shopping centers and retail stores) in th TAZ	1	1829

Table 2. Descriptive analysis of data

Other types of data including a broad range of socio-economic characteristics were obtained from Statistical Census of Iran for the corresponding year. Table 2 also indicates a descriptive analysis of data available for this study.

3. METHOGOLOGY

3.1 Mixed Geographically Weighted Regression (MGWR)

It is assumed that the coefficients of spatially stationary models such as general linear models (GLMs) are constant across the study region. When spatial non-stationary exist, the estimated coefficients will be the function of (u_i, v_i), which denoted the spatial coordinate of ith point. GWR is described by the following equation:

$$\eta_i = \sum_k \beta_k(u_i, v_i) x_{k,i} = x_i^t \beta(U_i) \qquad (1)$$

where
$x_{i,k}$ = kth independent variable
β_k = the corresponding local coefficients
(u_i, v_i) = coordination of the ith location
$\beta_k(u_i, v_i)$ = varying coefficients based on the location

The expected value of response of the ith observation, $E(y_i)$ is related to the linear predictor through a link function, such as f:

$$f(E[y_i]) = \eta \qquad (2)$$

In order to define the log-likelihood of the observation, a distributional function of the exponential family is applied:

$$\log g(y_i | \eta_i, \varphi_i) = \frac{y_i \eta_i - b(\eta_i)}{a(\varphi_i)} + c(y_i, \varphi_i) \qquad (3)$$

where
η_i = canonical
φ_i = dispersion parameter
a, b and c = functional components of the exponential family

Such formulation covers many commonly applied regression models such as Gaussian, Poisson and logistic variants of GWR, where in GWR (Gaussian): $y_i \sim [\eta_i, \sigma 2]$. In geographically weighted generalized linear models (GWGLM), a vector of local coefficients is estimated by concentrating on the ith regression point through solving the following maximization problem of the geographically weighted log-likelihood of the model:

$$\hat{\beta}(U_i) = \arg\max \sum_j^k \left\{ \log g(y_i | \hat{\eta}_{(i)j}, \hat{\varphi}_{(i)j}) . w_{ij}(\|U_i - U_j\|) \right\} \qquad (4)$$

where hat symbol = estimated value

$$\hat{\eta}_{(i)j} = \eta_j(\hat{\beta}(U_i)), \hat{\varphi}_{(i)j} = \varphi(\hat{\beta}(U_i)) \qquad (5)$$

These two working variables are fitted canonical and dispersion parameters for estimating the response at the jth location with coefficients at the ith regression point. The spatial weight of the jth observation at the ith regression point, w_{ij} is recognized as a non-negative and monotonously decreasing function of the distance between the regression point i and the jth observation location, such as a Gaussian kernel function:

$$w_{ij} = \exp\left(-\frac{1}{2} \frac{\|U_i - U_j\|}{G}\right) \qquad (6)$$

where G = kernel bandwidth

Bandwidth controls the distance decay and amount of locality of the weighting function. Increasing the bandwidth softens the local fluctuations and brings the coefficients closer to the global value. If the kernel is considered fixed, the bandwidth is assumed to be constant at each regression point in the study area. Alternatively, adaptive spatial cores adapt to data densities at different locations (TAZ). When the adaptive kernel is applied, the optimal number of neighbouring TAZs is chosen in order to find the given number of closest TAZs and ensure that they contain a given number of local samples. The next step is to estimate the weights using the given kernel and set the value of each TAZ according to Bi-square function. In practical cases, the results of GWR are not seriously sensitive to the weighting function type, but very sensitive to bandwidth. If the bandwidth is adaptive, AICc is applied to optimize the bandwidth:

$$AIC_c(G)D(G) + 2K(G) + 2\frac{K(G)(K(G)+1)}{n - K(G) - 1} \qquad (7)$$

where
G = bandwidth
D = deviance of parameters
K = effective number of parameters
n = number of samples (TAZs)

In situations where some independent variables are global in nature, and some others are local, a MGWR is suggested in which some coefficients in the equation are assumed to be constant and the others are allowed to vary across the studied area. MGWR includes linear terms of explanatory effects on the response in canonical parameter:

$$\eta_i = \sum_k \beta_i(u_i, v_i) x_{ki} + \sum_l \gamma_{li} z_{li} + \varepsilon_i \qquad (8)$$

where
z_{li} = independent variable
γ_{li} = fixed coefficient

Combines geographically local scoring and backfitting algorithms to compute coefficient and index estimates for model diagnostics, including information criteria such as coefficient standard errors, degrees of freedom, and AICc (corrected AIC) can be used to determine the optimal bandwidth size and model comparison. AICc estimation has the advantage of being generally applied, meanwhile it can be employed to compare whether or not the results from MGWR better fits compared with global model by considering model's degree of freedom.

3.2 Model Calibration and Assessment

A correlation matrix was constructed to investigate if the selected variables for both trip production and attraction models were highly correlated or not. If two explanatory variables are highly correlated, applying them simultaneously into the same model must be avoided to minimize the effects of collinearity. In addition, to test the correlation of explanatory variables, the variance inflation factor (VIF) is calculated. VIF is defined as VIF=$1/(1 - R_k^2)$, in which, R_k^2 is the coefficient of determination of a regression of kth variable on all the other variables and its range from 5 to 10 or higher indicates a serious collinearity. to decide whether or not to keep one variable in the model is based on logical and statistical significance. Eventually, the model with the minimum correlated variables and the best goodness-of-fit criteria are selected. The goodness-of-fit in this study is evaluated in several ways. First, t-statistics significance at the 5% level to decide on whether or not to keep a variable in the model or not. Akaike Information Criterion (AIC) generated for

OLS and corrected Akaike Information Criterion (AICc) calculated for MGWR are also employed for model comparison. Three other goodness-of-fit measures are defined to examine the local and mixed model, improvement over a global model: 1) The adjusted coefficient of determination (Adjusted R2) in the range of +1 and −1, (2) the mean square error (MSE), and 3) ANOVA (Analysis of Variance). The higher adjusted correlation coefficient and lower MSE and AICc, the better model fits the data. These criteria are used to determine which model could interpret data better.

3.3 Exploring the Spatial Autocorrelation

Spatial autocorrelation analyses the amount of correlation of observations in space. When the level of dependency is more than expected degree, the nearby observations show positive autocorrelation. If the dependency is negative, high observations are surrounded by low values and the spatial autocorrelation is negative. There is no spatial autocorrelation if the data are distributed such that relationships between nearby samples cannot be explored. As mentioned earlier, it is assumed that errors in one observation in a regression model cannot be related to errors in other observations (Fotheringham et al., 2002). When the significant spatial autocorrelation exists in residuals, employing conventional OLS might be questionable. The most common technique for evaluating the spatial dependencies is global Moran's I ranging from -1 to 1 introduced by (Anselin, 1995) and is defined as:

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\left(\sum_{i \neq j}^{n} \sum_{j=1}^{n} w_{ij} \right) \left(\sum_{i=1}^{n} (y_i - \bar{y}) \right)} \qquad (9)$$

where   n = number of TAZs

$\bar{y}$ = global mean value for residuals of a regression

y_i and y_j = residuals of regression models at ith and jth TAZ

w_{ij} = spatial weight matrix

The larger the absolute value of Moran's I, the more significant the spatial autocorrelation and a value of zero implies the perfect spatial randomness (Anselin, 1995).

4. RESULTS AND DISCUSSIONS

Two separate models were built for trip production and attractions. In the first step, a correlation matrix is constructed to see if the variables of interest are highly correlated. An initial model is then built based on the selected explanatory variables and a stepwise approach with 5% confidence intervals. A global OLS regression model is constructed to examine the relationship between dependent and independent variables. Next, GWR in which all explanatory variables are local-dependent is developed to investigate whether variables are spatially non-stationary or not. In case, the models include both global and local variables, the MGWR model is built. The results of models, corresponding statistics, and some discussion around the results are presented in following.

4.1 Results of OLS for Trip Production and Trip Attraction

The results of constructing the global model to identify the relationship between work trips and related factors resulting from correlation matrix for both trip production and trip attraction have been summarized in Table 3 and Table 4 respectively. Based on tables, the magnitude of term "constant" for trip production and attraction models are rather high. This represents a perturbation in the model formulation. Therefore, it should be read as the collective effect of independent variables not included in the specification. Since models were constructed through stepwise procedure, the effects of other variables and their significance had already been evaluated and suggested models were the most appropriate based on available variables. The estimated coefficients relating to trip production indicates the direct relationship between the number of work trip produced and EL_i (Table 3). It is common knowledge that the more commuters, the more commutes. According to the t-statistic, the null hypothesis that the true value is zero can be rejected at the 5% significance level.

The direct and positive relationship between number of business units (BU_i) and employees working in a TAZ (EW_i) with trip attraction has been tabulated in Table 4. The corresponding t-statistics reject the null hypothesis stating that the values are zero. The results show that collinearity is not a matter of concern in the global model, as indicated by the low VIF index for the variables of trip attraction. As evident, the estimated coefficients for all variables appeared in trip production and attraction models are the global values which are fixed throughout the study area. It is likely that the effects of each independent variable on dependent variable might be spatially varying. In case, where significant spatial dependencies prevail, one solution is that such effect can be developing local and semi-parametric models.

Variable	Coeff	Std err	t-stat	Sig.
Constant	457.25	217.13	2.11	0.036
EL_i	1.40	0.50	28.01	0.000
Goodness of fit				
AICc: 4593.58				
BIC= 4604.09				
R square: 0.76				
Adjusted R square: 0.75				

Table 3. Results of Global OLS for trip production

Variable	Coeff	Std err	t-stat	Sig.	VIF
Constant	1288.30	286.68	4.49	0.926	
BU_i	4.04	0.46	8.873	0.000	1.11
EW_i	4.79	0.30	16.14	0.000	1.20
Goodness of fit					
AICc: 4726.99					
BIC: 4740.96					
R square: 0.60					
Adjusted R square: 0.59					

Table 4. Results of Global OLS for trip attraction

4.2 Results of MGWR

Two models are compared to examine the geographic variability of the kth variance coefficient. A fitted GWR model and a model where only the kth coefficient is fixed and the other coefficients are left in the fitted GWR model. Assume that these

two models are the original model and the switched model respectively. It is concluded that if the original GWR performs better than the compared switched model based on the defined criterion (as was used for selecting the optimum kernel bandwidth), the non-stationary state of the kth coefficient is confirmed. In another approach, and by applying the Gaussian model, the difference of deviance (Diff of deviance) between the two models should follow the F-statistic under the null hypothesis that there is no difference in the performance of the two models(Nakaya.T, 2014). If the null hypothesis is rejected, it can be inferred that the term shows significant spatial variation at the significance level employed for the test.

Table 5 indicates the results of non-stationarity test for the variables appeared in trip production and trip attraction primary models. It includes rows of local terms with a "Diff of Criterion" column, which shows the difference in the model comparison indicator between the original and switched GWR models. A positive "Diff of Criterion", especially greater than or equal to two, suggests that spatial variability does not exists (Nakaya.T, 2014). Accordingly, the "Diff of Criterion" for term "constant" in trip production model is positive (9.65) which suggests that the variables would better be considered as fixed. The negative values of "Diff of Criterion" relating to all variables of trip attraction model, except the "constant" term are indicators for spatially varying state and therefore, it is suggested that the model is better to be reconstructed by considering "constant" as global fixed variable and the rest as local in the process of modelling. The five factors representing the local estimated coefficients resulting from MGWR in addition to the global estimated coefficient of "constant" for trip production have been summarized in Table 6.

Accordingly, it is evident that although the spatially varying F test resulting from previous step suggested that the constant is better to be considered as global, the variable is not statistically significant when a mixed model is developed. Although the estimated global coefficient of EL_i which is the average over the study area and the mean values resulting from MGWR are almost the same (1.40 vs. 1.63), the local coefficients of EL_i vary in the range of (1.23-2.09). Comparing the AICc of global OLS (4726.99) and the corresponding value obtained from MGWR (4556.14) indicates the improvement of the mixed model over the full global one.

Similar analyses for variables of trip attraction model reveals interesting points. As mentioned earlier, according to the results of spatial non-stationary test in Table 5, the "constant" has been considered as fixed in MGWR model for trip attraction. Comparing the estimated coefficients of mentioned variables in global OLS and MGWR indicates that both models yield approximately the same values for explanatory variables (4.79 vs. 4.89 for EW_i and 4.04 vs. 4.82 for BU_i). However, as can be seen in Table 7, the values of local coefficients of EW_i and BU_i vary in a broad range of (2.11-12.57) and (0.40-15.01) respectively indicating that considering these variables as fixed might be controversial. Additionally, the AICc values of MGWR is lower as compared with the corresponding values for global model, another indication for the model improvement.

Category	Variable	F	Degree of Freedom for F test	DIff of Criterion[*]
Trip production	Constant	1.01	7.84	9.65
	EL_i	4.13	8.70	-16.26

	variables			
Trip attraction variables	Constant	0.74	11.52	20.28
	BU_i	6.77	10.87	-45.28
	EW_i	4.56	8.05	-18.17

*Difference in model comparison indicator between the original and switched GWR models

Table 5. Results of testing spatial non-stationarity for explanatory variables of both models

Local Variables	Min	Lower Quartile	Mean	Upper Quartile	Max
EL_i	1.23	1.49	1.63	1.80	2.09
Global Variable	Coeff	Str err	t-stat	Sig.	
Constant	-142.65	238.25	-0.60	<0.05	
Goodness-of-fit					
AICc: 4556.14					
BIC: 4620.32					
R square: 0.81					
Adjusted R square: 0.79					

Table 6. Results of MGWR for trip production

Local Variables	Min	Lower Quartile	Mean	Upper Quartile	Max
EW_i	2.11	3.77	4.89	5.95	12.57
BU_i	0.40	3.00	4.82	5.79	15.01
Global Variable	Coeff	Str err	t-stat	Sig.	
Constant	1093.16	265.34	4.12	<0.05	
Goodness-of-fit					
AICc: 4647.51					
BIC: 4733.51					
R square: 0.76					
Adjusted R square: 0.72					

Table 7. Results of MGWR for trip attraction

The results are local coefficients, which can be mapped for visual inspection. It's true that local map generation is one of the main advantages of GWR, but it's worth noting that finding the best way to do it is not easy. (Shoff et al., 2012) because of coherency nature of variations and therefore the results must be interpreted with due caution.

As mentioned, Number of employees living in ith TAZ (EL_i) is positively significant in the global trip production OLS (Table 3), but the local coefficients represent a varying range from 1.23 to 2.09. The relationship between EL_i with trip production is consistent with the expectations, exhibiting a positive correlation. Results of local coefficients of EL_i have been mapped in (Figure 1). Accordingly, the most intense effects of number of employees living in TAZs upon work trip production can be found in the southern parts, and a portion located in centre and southwest of study area. The positive but weaker effects are indicated in northwest, north, northeast, east, and southeast.

These maps allow us to visually identify TAZs where specific explanatory factors have a strong statistical impact on the model, as well as important local variations not captured by the global OLS model. The local coefficients related to the EW_i have the same sign as in the global OLS but with different intensities delineating specific spatial patterns; this indicates the importance of the spatial dimension and, therefore, the necessity

of dealing with it (Figure 2). The strong and positive effects are explored in the east, northeast and western parts of the study area. In such areas, the effect of number of employees working in ith TAZs is relatively higher than other TAZs. The positive but weaker influences are found in the broad extend located in the central part of Mashhad.

Local maps can be regarded as a useful tool in policy making process. For instance, special consideration can be given to the areas where EW_i and EL_i have strong influence on work trip generation during certain time of the day particularly morning and evening peak hours. It is not surprising if such regions experience higher traffic congestion rather than other areas during this time. Due to the strong influence of employees in such regions, different types of travel demand strategies with the focus of managing the demand of this group of community could be helpful.

This cannot be claimed but interventions such as flexible work times, teleworking, and compressed workweeks can be applied in areas with highly positive significant EL_i and EW_i coefficients. Flexible scheduling allows employees to shift the work trips to non-peak hours of the day and telecommuting allows them to work from home or non-office locations can help to reduce the number of work trips as well. Vanoutrive et al. state that employers can encourage a more sustainable commuting by promotion of alternative modes (Vanoutrive et al., 2010). Providing facilities and services to make non-SOV commute options more appealing and viable, for example, securing work place parking for bikes as well as shower and locker facilities, provision of free vanpool vehicles, shuttle services, and car sharing programs for the employees could be considered. Financial incentives such as instituting parking charges, unbundling free or subsidized parking from employee benefits, proving a few days of free parking each month for employees using non-SOV modes could be other helpful policies. These strategies are in line with the findings of (Curtis, 1981). Additionally, solutions relating to urban structure and development management which affect employees travel behaviour have been studied by previous researchers. According to (Lee, 2016) for example, many workers combine commuting and commuting, and local work-life balance is expected to encourage people to commute by public transport to some extent. going to and from work.

In spite of positively estimated coefficient of BU_i, resulting from both global OLS and the MGWR, the local coefficients of MGWR show markedly regional differentiation in terms of magnitude (Figure 3). The local coefficients of business units vary in the range of 0.40 to 15.01 indicating how the effects of this variable spatially vary upon work trip attraction. The strong influences are explored in the distinct area located in the central part of the city, which is known as Mashhad CBD. This could be explained trough locating non-residential units and concentrating shopping centers, malls, and retail stores.
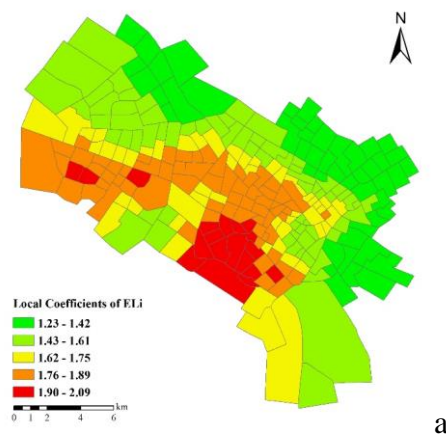

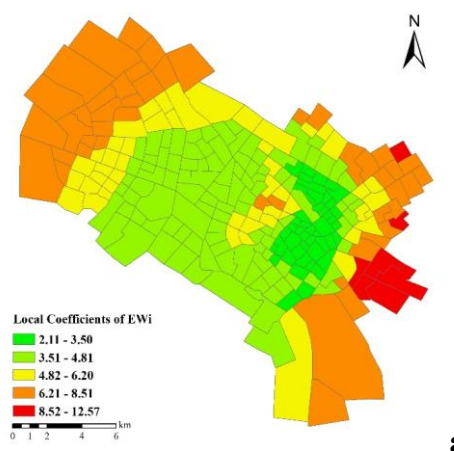**Figure 1.** Local coefficients of EL_i resulting from trip production models

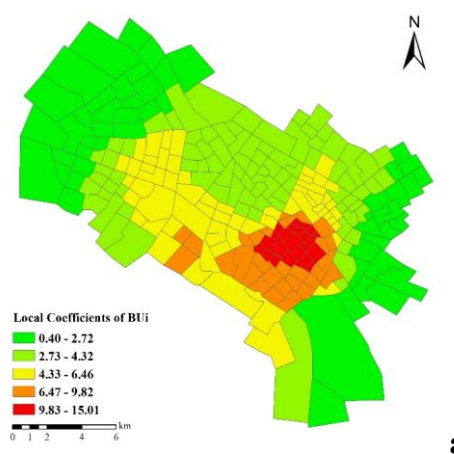**Figure 2.** Local coefficients of EW_i resulting from trip attraction models

Figure 3. Local coefficients of BU_i resulting from trip attraction models

4.3 Results Testing Spatial Autocorrelation

Moran's I results for all model residuals are shown in Table 8 to compare the reliability of the conventional OLS and MGWR spatial autocorrelation treatments. As can be seen, significant positive spatial autocorrelation is found for OLS in trip production model (Moran's I=0.030, z-score=2.04, p-value=0.010); while no spatial dependency is explored for the

MGWR of corresponding model specified by Moran's I and z-score equal to -0.038, -2.05 and p-value equals to 0.040. Similar analysis for trip attraction model indicates the positive and significant spatial dependency among the residuals of global OLS (Moran's I=0.056, z-score = 3.56, p-value= 0.000); while the local model has successfully removed spatial dependencies of residuals (Moran's I=-0.037, z-score = -1.95, p-value= 0.050). It reflects that with the same dependent and independent variables, the OLSs violate the assumption of residual independence. In return, MGWR has improved the reliability of the relationships by eliminating the spatial autocorrelations of residuals.

Unlike the global OLS, the spatial patterns of local (r_i^2) in MGWR represent a marked regional differentiation. In trip production model, local (r_i^2) is characterized by higher values (0.84 –0.95) in the southeast and some central parts of the study area (Figure 4a). Accordingly, it can be seen that the model did not fit very well to data for TAZs located in northeast part of Mashhad. The similar analysis of (r_i^2) resulting from rip attraction model indicates that how well the model has fit to data in the west part of the study area (Figure 4b). In areas located in a major northern part, trip attraction is not adequately explained by the selected explanatory variables with the local (r_i^2) falling below the OLS threshold and this could imply additional covariates were needed to explain the work trip attraction. Figure 4 helps to realize additional explanatory factors were required and where could those factors be applied. It is worth to note that we indicated that MGWR is potentially non-stationary; therefore, A model calibrated at one location cannot be expected to replicate data particularly well elsewhere unless the process being modelled is relatively stable. So local (r_i^2) reflects mixture of two issues: how well the model replicates the data and how stationary are the processes being modelled.

5. CONCLUSION

OLS is the most widely known method for calibrating trip generation models and assumes that the relationship between the dependent and independent variables is stationary. The key question in this study is whether explicit spatial non-stationary relationships between the trip generation and potential explanatory variables exist or not. This can efficiently be done through employing GWR, which brings about location-specific parameter estimates that can be mapped to explore the variations. The MGWR model is suitable for situations where certain explanatory variables affect the dependent variable globally and others locally. Local estimation coefficients obtained from MGWR show significant differences across the study area in terms of magnitude, indicating that the descriptive power of such models is stronger than that of OLS. Moreover, our results show that MGWR represents a significant improvement in model performance over the global model, as indicated by lower AICc, higher values of the coefficient of determination, and reduced spatial autocorrelation of the residuals.

Although strategic policies for the areas which are affected more by a particular predictor have been suggested, the detailed analyses of such areas still need further research. Such analyses are beneficial to urban planners, transport engineers and other analysts who deal with issues related to zoning and development of neighbourhoods.

The advantage of local models is highlighting the areas demanding more attention in terms of factors generating trips and consequently request more travel demand managements, once they are included in detailed engineering study sites. Detecting such areas could be beneficial in allocating the limited budget and resources in future urban development plans.
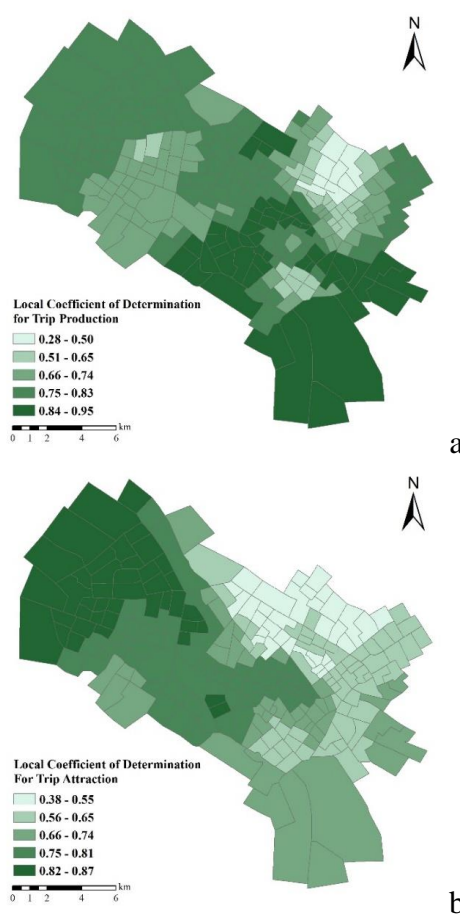
Category	Model	Moran's I	z-score	p-value
Trip production	OLS	0.030	2.04	0.010
	S-GWR	-0.038	-2.05	0.040
Trip attraction	OLS	0.056	3.56	0.000
	S-GWR	-0.037	-1.95	0.050

Table 8. Results of Moran's I for residuals of global OLS and MGWR

Figure 4. Local coefficients of determination for a) trip production and b) trip attraction models

REFERENCES

2014. The SAGE Handbook of Social Network Analysis. London: SAGE Publications Ltd.

Anselin, L. 1995. Local Indicators of Spatial Association—LISA. *Geographical Analysis*, 27, 93-115.

Brunsdon, C., Fotheringham, A. S. & Charlton, M. 2002. Geographically weighted summary statistics — a framework for localised exploratory data analysis. *Computers, Environment and Urban Systems*, 26, 501-524.

Cardozo, O. D., García-Palomares, J. C. & Gutiérrez, J. 2012. Application of geographically weighted regression to the direct forecasting of transit ridership at station-level. *Applied Geography*, 34, 548-558.

Curtis, F. 1981. Work trip travel demand modelling in Kingston Township, Ontario, Canada. *Transportation Planning and Technology*, 7, 11-19.

Fotheringham, S., Brunsdon, C. & Charlton, M. 2002. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships, Chichester, Wiley.

Hendrickson, C. & Plank, E. 1984. The flexibility of departure times for work trips. *Transportation Research Part A: General*, 8, 25-36.

Lee, J. 2016. Impact of Neighborhood Walkability on Trip Generation and Trip Chaining: Case of Los Angeles. *Journal of Urban Planning and Development-asce*, 142, 05015013.

Lloyd, C. & Shuttleworth, I. 2005. Analysing Commuting Using Local Regression Techniques: Scale, Sensitivity, and Geographical Patterning. *Environment and Planning A: Economy and Space*, 37, 81-103.

Nakaya.T 2014. GWR4 user manual.

Nurul Habib, K. M., Day, N. & Miller, E. J. 2009. An investigation of commuting trip timing and mode choice in the Greater Toronto Area: Application of a joint discrete-continuous model. *Transportation Research Part A: Policy and Practice*, 43, 639-653.

Pagliara, F. & Mauriello, F. 2020. Modelling the impact of High Speed Rail on tourists with Geographically Weighted Poisson Regression. *Transportation Research Part A: Policy and Practice*, 132, 780-790.

Palma, A. D. & Rochat, D. 2000. Mode choices for trips to work in Geneva: an empirical analysis. *Journal of Transport Geography*, 8, 43-51.

Powers, S. L., Matthews, S. A. & Mowen, A. J. 2021. Does the relationship between racial, ethnic, and income diversity and social capital vary across the United States? A county-level analysis using geographically weighted regression. *Applied Geography*, 130, 102446.

Shoff, C., Yang, T.-C. & Matthews, S. A. 2012. What has geography got to do with it? Using GWR to explore place-specific associations with prenatal care utilization. *GeoJournal*, 77, 331-341.

Sohn, J. 2005. Are commuting patterns a good indicator of urban spatial structure? Journal of *Transport Geography*, 13, 306-317.

Stead, D. 2001. Relationships between Land Use, Socioeconomic Factors, and Travel Patterns in Britain. *Environment and Planning B: Planning and Design*, 28, 499-528.

Vanoutrive, T. U., Van Malderen, L., Jourquin, B., Thomas, I., Verhetsel, A. & Witlox, F. D. S. W. E. 2010. Mobility management measures by employers: overview and exploratory analysis for Belgium. (2010) *EUROPEAN JOURNAL OFTRANSPORT AND INFRASTRUCTURE RESEARCH*.

Wang, S. & Wu, J. 2020. Spatial heterogeneity of the associations of economic and health care factors with infant mortality in China using geographically weighted regression and spatial clustering. *Social Science & Medicine*, 263, 113287.

Xu, H. & Zhang, C. 2021. Investigating spatially varying relationships between total organic carbon contents and pH values in European agricultural soil using geographically weighted regression. *Science of the Total Environment*, 752, 141977.

Yu, D., Zhang, Y., Wu, X., Li, D. & Li, G. 2021. The varying effects of accessing high-speed rail system on China's county development: A geographically weighted panel regression analysis. *Land Use Policy*, 100, 104935.

Yu, H. & Peng, Z.-R. 2019. Exploring the spatial variation of ridesourcing demand and its relationship to build environment and socioeconomic factors with the geographically weighted Poisson regression. *Journal of Transport Geography*, 75, 147-163.

Zeng, C., Yang, L., Zhu, A. X., Rossiter, D. G., Liu, J., Liu, J., Qin, C. & Wang, D. 2016. Mapping soil organic matter concentration at different scales using a mixed geographically weighted regression method. *Geoderma*, 281, 69-82.