Transportation Research Part A 78 (2015) 161–177

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

Factors affecting public transportation usage rate: Geographically weighted regression

Yu-Chiun Chiou [a,*], Rong-Chang Jou [b], Cheng-Han Yang [a]

[a] Department of Transportation and Logistics Management, National Chiao Tung University, 4F, 118 Sec. 1, Chung-Hsiao W. Rd., Taipei 10012, Taiwan, ROC
[b] Department of Civil Engineering, National Chi Nan University, No. 1, University Rd, Puli, Nantou County 54561, Taiwan, ROC

ARTICLE INFO

Article history:
Received 22 December 2014
Received in revised form 24 April 2015
Accepted 21 May 2015
Available online 7 June 2015

Keywords:
Public transportation usage rate
Tobit regression
Geographically weighted regression
Spatial autocorrelation

ABSTRACT

As the number of private vehicles grows worldwide, so does air pollution and traffic congestion, which typically constrain economic development. To achieve transportation sustainability and continued economic development, the dependency on private vehicles must be decreased by increasing public transportation usage. However, without knowing the key factors that affect public transportation usage, developing strategies that effectively improve public transportation usage is impossible. Therefore, this study respectively applies global and local regression models to identify the key factors of usage rates for 348 regions (township or districts) in Taiwan. The global regression model, the Tobit regression model (TRM), is used to estimate one set of parameters that are associated with explanatory variables and explain regional differences in usage rates, while the local regression model, geographically weighted regression (GWR), estimates parameters differently depending on spatial correlations among neighbouring regions. By referencing related studies, 32 potential explanatory variables in four categories, social-economic, land use, public transportation, and private transportation, are chosen. Model performance is compared in terms of mean absolute percentage error (MAPE) and spatial autocorrelation coefficient (Moran' *I*). Estimation results show that the GWR model has better prediction accuracy and better accommodation of spatial autocorrelation. Seven variables are significantly tested, and most have parameters that differ across regions in Taiwan. Based on these findings, strategies are proposed that improve public transportation usage.

1. Introduction

Although travellers typically differ in their transportation choices, private transportation has the advantages of better accessibility and convenience than public transportation, which have engendered rapid growth in the number of private vehicles, increasing both traffic congestion and environmental pollution. Many governments worldwide have focused on improving their public transportation services and restricting ownership and usage of private vehicles. However, public transportation patronage in many developed countries with adequate public transportation systems remains low. Thus, whether governments should develop their public transportation services has not been thoroughly explored. In literature, two main approaches are applied to conduct this exploration: individual and collective approaches. The individual approach

* Corresponding author. Tel.: +886 2 23494940.
E-mail address: ycchiou@mail.nctu.edu.tw (Y.-C. Chiou).

http://dx.doi.org/10.1016/j.tra.2015.05.016
0965-8564/

developed mode choice models based on data from a questionnaire survey of travellers. Covariates of travel time (e.g. walking time, waiting time, and on-board time) and travel cost (e.g. parking cost, fuel cost, and bus fare) of different transportation modes, along with demographics and trip characteristics of travellers, are commonly considered. Estimation results by the individual approach provide insights into the effects of covariates on travellers' mode choice behaviours, which is helpful when designing public transportation systems. However, aggregation of individual mode choice behaviours to represent regional usage rates usually leads to biased predictions due to small sample sizes or excessive sampling errors. In contrast, the collective approach directly correlates ridership or usage rate of public transportation to potential covariates, which is helpful for decisions about whether to invest in public transportation systems in specific regions in order to increase the market share of public transportation, but individual preferences for public transportation systems are hard to be predicted based on the collective models. Therefore, to support decisions about investing in and designing public transportation systems, these two approaches are necessary. The collective approach can investigate the necessity of public transportation investment in various regions, while the individual approach can support decisions regarding the design of public transportation systems. Although the individual approach has several advantages over the collective approach, especially at the level of details in associating the mode choice changes with the improved service quality of public transportation, the commonly adopted explanatory variables of the individual approach, such as travel time, waiting time, walking distance of various transportation systems and social-demographics and trip characteristics of individual travellers, are rather expensive to be collected for model estimation and prediction for a large-scale area. To assess governmental investment in public transportation in various regions nationwide, this study adopts the collective approach to identify the key factors contributing to usage rate of public transportation systems.

To objectively review the benefit of investment in public transportation, a demand-side response should be evaluated scientifically and systematically to assist decision-makers in fully utilizing limited resources towards policy objectives. This information can also be used as feedback to amend short- and long-term strategies as well as avoid ineffective investments. Previous studies on aggregate public transportation usage rate modelling were limited to large unit (e.g., countries or cities) and thereby did not examine differences in inter-regional influential factors. Moreover, since the service level of transportation systems in neighbouring regions may be closely related (i.e. spatial autocorrelation), this study applies the geographically weighted regression (GWR) model to investigate the effects of key factors on public transportation usage rates locally rather than globally. The GWR model allows estimated parameters to vary across regions to accommodate potential spatial dependencies. Comparisons of model performance and estimation results with traditional global regression models are also conducted. As the dependent variable in this study is a percentage, which is truncated and does not have a normal distribution, the Tobit regression model (TRM) is used.

The rest of this paper is organized as follows. Section 2 presents a brief review of prior research. Section 3 briefly introduces the TRM and GWR models used in this paper along with performance indices for model comparisons. Section 4 addresses empirical data and their descriptive statistics. Section 5 presents estimation results along with implications of the two models. Model comparisons are given in Section 6. Section 7 presents conclusions and suggestions for further research.

2. Prior research

The models used to identify key factors contributing transit ridership and usage rate along with contributing factors considered are in previous studies are briefly reviewed.

2.1. Global and local regression models

Numerous studies have been conducted to examine the effects of key factors to transit ridership or usage rate globally or locally. The global regression models generally treat each observation independently and estimate one set of parameters to represent all observations. The commonly adopted models depend upon the dependent variables: transit ridership (i.e. number of boarding) and public transportation usage rate. For transit ridership, multiple regression analysis (Cervero, 1996; Kuby et al., 2004; Taylor et al., 2009; Mulley and Tanner, 2009; Swimmer and Klein, 2010; Blainey, 2010; Cervero et al., 2010; Souche, 2010) and simultaneous regression (Taylor et al., 2009; Swimmer and Klein, 2010) were commonly adopted. For public transportation usage rate, various models were constructed, including multiple regression analysis (Chow et al., 2006; Messenger and Ewing, 2007), simultaneous regression (Messenger and Ewing, 2007), the aggregate logit model (Buehler, 2011; Coldren et al., 2003; Chen et al., 2009), and the Tobit regression model (Boame, 2004) and so on.

Fotheringham et al. (2000) noted that the global regression model only estimates a set of parameters for relationships between an independent and dependent variables, and the estimated parameters do not vary with space therefore reflect spatial characteristics. That is, the major drawback of the models is that any geographical variation in the relations between variables is masked (Lloyd and Shuttleworth, 2005; Cardozo et al., 2012) by ignoring the existence of local variations due to the spatial autocorrelation.

In contrast, the local models, such as spatial proximity regression, GWR, distance-decay weighted regression, hierarchical linear model, are recommended when spatial autocorrelation is present. Several studies have been conducted for

modelling transit ridership and usage rate. For example, Blainey (2010), Blainey and Preston (2010), Gutiérrez et al. (2011), Pulugurtha and Agurla (2012), Cardozo et al. (2012) used of above local models to examine the key factors affecting number of passengers of public transportation at station level or at traffic analysis zone (TAZ) level locally. While Kobayashi and Lane (2007) and Chow et al. (2006, 2010) investigated key factors to public transportation usage rate locally. All these studies consistently show that the local models are preferred under the study area exhibiting spatial autocorrelation.

2.2. Key factors to transit ridership or usage rate

A review of above studies show that the explanatory variables considered can be broadly divided into demographic and socio-economic (e.g., population, employment, age group, and income level), land use (e.g. residential, industrial, commercial, or mixed use), availability of private transport (e.g., car ownership, road length, fuel price, parking fee, parking space), and accessibility to public transport (e.g., fare rate, station accessibility, service frequency, park and ride space, feeder service).

Table 1 summarizes the key factors significantly tested in selected studies. As shown in Table 1, population and route (public transportation) are two key factors confirmed by the largest number of the selected studies (9 out of 16), followed by income and accessibility (public transportation stations) (7 out of 16), and next followed by accessibility (traffic analysis zone, TAZ), car ownership, and frequency (public transportation) (6 out of 16), suggesting that a TAZ with a highly populated density, low income, low accessibility, low car ownership, high-quality public transportation service (high route density, high service frequency, high station accessibility) exhibits high public transportation patronage.

3. Method

This paper aims to develop a model of public transportation usage rate and to examine the effect of potential covariates. Two methods, TRM and GWR, are developed for comparisons and briefly introduced below.

3.1. Tobit regression model (TRM)

TRM allows us to incorporate the value bounded dependent variable while usage rate of public transportation is constrained to fall between zero and one. The structure model of the TRM is (Greene, 2003):

$$y_i^* = X_i\beta + \varepsilon_i \tag{1}$$

where $\varepsilon_i \sim N(0, \sigma^2)$. $N(0, \sigma^2)$ is a normal distribution with mean of 0 and variance of σ^2. y^* is a latent variable that is observed for values greater than 0 and censored at 0. The observed y is defined by the following measurement equation:

$$y_i = \begin{cases} y^* & \text{if } y^* > 0 \\ 0 & \text{if } y^* \leqslant 0 \end{cases} \tag{2}$$

The maximum likelihood estimation method is used to estimate the TRM. The log-likelihood function for the TRM is:

$$\ln L = \sum_{i=1}^{N} \left\{ d_i \left(-\ln\sigma + \ln\phi\left(\frac{y_i - X_i\beta}{\sigma}\right)\right)\right\} + (1-d_i)\ln\left(1 - \Phi\left(\frac{X_i\beta}{\sigma}\right)\right) \tag{3}$$

where N is the sample size. d_i is an indicator variable that equals 1 if $y_i > 0$ and 0, otherwise. $\phi(\cdot)$ is a probability density function of the standard normal distribution. $\Phi(\cdot)$ is a cumulated density function of the standard normal distribution. Eq. (3) is made up of two parts. The first part corresponds to the classical regression for uncensored observations, while the second part corresponds to the relevant probabilities that an observation is censored.

3.2. Geographically weighted regression (GWR)

The GWR estimates the parameter for each location (u_i, v_i) using a weighted least squares method, as shown in the following equation (Fotheringham and Charlton, 1998):

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_{ik}(u_i, v_i)x_{ik} + \varepsilon_i. \tag{4}$$

The weight matrix is a diagonal matrix, and each element on the primary diagonal line represents a function of the location i. The weight matrix is expressed as $W(i)$, and according to the principle of weighted least squares method, the original parameter at location i is estimated by $\beta(i) = (\beta_{i0}, \beta_{i1}, \ldots, \beta_{ip})^T$ as:

$$\widehat{\beta}(i) = [X^T W(i) X]^{-1} X^T W(i) Y \tag{5}$$

Table 1
Significant factors to public transportation patronage of selected studies.

Studies	Target variable		Significant factors																												
			Socio-economic									Land use						Private transportation							Public transportation						
	Ridership	Usage rate	Population	High-educated population	Households with kinds	Low-income households	Employment	Unemployment rate	Income	Residential area	Commercial area	Industrial area	Land use mix	Urban area/ structure	Accessibility	Car ownership	Fuel/ Parking cost	Parking space	Road length	One-way street	Speed limit	Travel permit	Routes	Frequency	Accessibility	Feeder service	Fare rate	Park&ride space	Terminal/ transfer stations		
Cervero (1996)	√						√			√					√									√	√		√	√		
Kuby et al. (2004)	√		√			√	√																	√	√		√	√		
Chow et al. (2006)		√	√		√			√	√						√	√							√							
Kobayashi and Lane (2007)		√	√			√	√			√					√				√				√		√					
Messenger and Ewing (2007)		√	√				√		√							√	√						√	√	√					
Chen et al. (2009)		√	√				√		√	√	√				√								√	√			√		√	
Taylor et al. (2009)	√		√			√		√	√	√				√		√	√		√											
Sun et al. (2009)	√		√						√										√								√			
Blainey (2010)	√						√									√		√				√		√	√					√	
Blainey and Preston (2010)	√		√												√	√							√	√	√						
Cervero et al. (2010)	√		√																					√	√	√		√			
Souche (2010)	√								√					√			√	√				√	√	√			√				
Swimmer and Klein (2010)	√		√	√		√	√							√			√						√			√	√				
Gutiérrez et al. (2011)	√		√				√						√		√								√								
Cardozo et al. (2012)	√						√																√								
Pulugurtha and Agurla (2012)	√		√						√	√	√	√				√				√	√										
Total	12	4	12	1	1	4	9	2	7	5	2	1	1	3	6	6	4	2	3	1	1	2	9	6	7	4	4	3	4		

where

$$W(i) = \begin{bmatrix} w_1(i) & & & \\ & w_2(i) & & \\ & & \ddots & \\ & & & w_n(i) \end{bmatrix} = diag[w_1(i) \quad w_2(i) \quad \cdots \quad w_n(i)\,].X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{p1} \\ 1 & x_{12} & \cdots & x_{p2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \cdots & x_{pn} \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

The weight function $w_j(i)$ using weighted least square method is the spatial analysis function for different conditions at each data point. i is the regression point for the parameters at any location. When the data point j is closer to i, the impact is enhanced. For example, in the global regression model, if the distance between data point j and regression point i is within a given radius d, the weight function $w_j(i) = 1$ (i.e., in this model, the data point weights are identical). Otherwise, $w_j(i) = 0$, but this will generate a non-continuous spatial weights problem. Based on the above principles, there are several weight functions options. Two functions, Gaussian and bi-square, are commonly adopted, which can be respectively expressed as:

$$w_j(i) = \exp\left[-(d_{ij}/b)^2\right],\ j = 1, 2, \ldots, n \tag{6}$$

$$w_j(i) = \begin{cases} \left[1 - \frac{d_{ij}^2}{b^2}\right]^2, \text{ if } d_{ij} \leqslant b \\ 0 \text{ if } d_{ij} > b \end{cases}, j = 1, 2, \ldots, n \tag{7}$$

where d_{ij} is the distance between regression point i and data point j and b is bandwidth. When b is fixed and the data point j is distal to regression point i, the weight is closer to 0. For the bi-square function, if the distance between two locations is larger than the preset bandwidth, the weight is 0. In this study, for parameter estimation of large or coastal areas, there are less than 30 weighted samples (i.e. neighbouring regions); therefore, this weight function was excluded from this study.

Fotheringham et al. (2000) noted that GWR parameters can be estimated with different bandwidths. b is generated through cross-validation (CV), including the least squares method followed by minimisation, as shown in the following equation:

$$\mathrm{CV} = \sum_{i=1}^{n} [y_i - \widehat{y}_{\neq i}(b)]^2 \tag{8}$$

where $\widehat{y}_{\neq i}(b)$ is the optimal value for y_i and represents the observed value when the regression point i is excluded during calibration; therefore, when b is very small, the model only calibrates the proximal samples other than i. Thus, when the CV value is minimal, b indicates the required bandwidth.

To estimate GWR parameters requires selection of a spatial weight function and bandwidth. The results from GWR are sensitive to the bandwidth of the given weight function (Fotheringham et al., 2002). Based on the data point, bandwidth may be constant (fixed kernel) or variable (adaptive kernel). Fixed-kernel bandwidth is used for a large sample size to demonstrate that the impact is greater for proximal points but the distance value is smaller (or near 0), which is of the same as the global regression model. However, if the sample size is small, then the adaptive-kernel is adopted, and thus, bandwidth is a variable. When the regression points are densely populated, the ever-changing kernel can be used to discern the optimal spatial bandwidth and produce a small bandwidth. Therefore, the primary difference between the two kernels is sample size (Fotheringham et al., 2002). In this study, there were 348 samples, and the distance matrix between the samples was large; therefore, a fixed-kernel bandwidth was adopted.

3.3. Model performance

Mean absolute percentage error (MAPE) and corrected MAPE were used to compare the accuracy between global and local regression models and discern the optimal model. MAPE is a relative value and is not influenced by actual and estimated values; thus, it can objectively reflect the difference between actual and estimated values. A smaller MAPE entails that the model's estimation is more accurate, as shown in the following equation:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \frac{|y_i - \widehat{y}_i|}{y_i} \times 100\% \tag{9}$$

where y_i: the actual value of dependent variables, that is, the usage rate of public transportation. $\widehat{y}_i$: the estimated value of y_i ($i = 1, 2, \ldots, n$). n: the number of samples.

The corrected MAPE generates more accurate differences between the actual and estimated values when it can represent the sample group (i.e., the average actual values for the samples in the group). It represents the accuracy for this group more effectively and can solve the equation when the actual value of y_i is 0 using the original MAPE, as shown in the following equation:

$$\text{Corrected} \quad \text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \widehat{y}_i|}{\bar{y}} \times 100\% \tag{10}$$

where $\bar{y}$: the average of all y_i, that is, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$.To test the spatial autocorrelation of the original data and residuals of the estimated models, Moran's I is used which can be expressed as:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{11}$$

where w_{ij}: the weight between observations i and j.

$$E(I) = \frac{-1}{(n-1)} \tag{12}$$

$$Var(I) = \frac{n[(n^2 - 3n + 3)W_1 - nW_2 + 3W_0]\} - K[(n^2 - n)W_1 - 2nW_2 + 6W_0^2]}{(n-1)(n-2)(n-3)W_0^2} - E(I)^2 \tag{13}$$

where $W_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}$. $W_1 = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{(w_{ij}+w_{ji})^2}{2}$. $W_3 = \sum_{i=1}^{n}(w_{\cdot i} + w_{i\cdot})^2$. $K = \frac{\left[\sum_{i=1}^{n}(y_i - \bar{y})^4/n\right]}{\left[\sum_{j=1}^{n}(y_i - \bar{y})^2/n\right]^2}$.

Thus, the standardized I is $Z(I) = \frac{[I - E(I)]}{\sqrt{Var(I)}}$.

4. Data

4.1. Data collection

Taiwan has 22 counties/cities, each of which has various townships/districts, such that Taiwan has a total of 368 townships/districts. To ensure that the transportation conditions in each study area are homogeneous, detailed zones, i.e., townships/districts, were used. Additionally, since collecting complete data is difficult for townships on offshore islands, only 348 inland townships/districts were the samples.

Corresponding variable data were primarily from Statistics Department and Directorate General of Highways of the Ministry of Transportation and Communications (MOTC), and the Statistics Department of the Ministry of the Interior (MOI). Dependent variables used in the models, usage rates for public transportation, cars, and motorcycles, were obtained *via* a telephone survey by the Statistics Department of the MOTC during October 12 to December 30, 2010. This survey used a systematic sampling technique based on the population in each area. In total, 12,120 questionnaires were valid. Demographic data were from the census conducted by the Department of Household Registration of the MOI on January 5, 2011. Census data included household and family data; personal data; economic, cultural, educational, and fertility data; as well as family and marital status. The population was stratified based on gender and age group as well as township/district. Land use data were collected from the Construction and Planning Agency of MOI. Moreover, operational data for intercity and city buses were obtained from geographic information system (GIS) graphics and tables provided by the Directorate General of Highways of MOTC and the transportation bureaus of local governments.

The public primarily used motorised transportation, including cars, motorcycles, and public transportation. Therefore, the usage rate for public transportation was adversely affected by that for cars and motorcycles, and the usage rate for each transportation mode was also impacted by internal and external factors. Public transportation usage was the dependent variable. Table 2 defines the explanatory variables and their expected effects on public transportation usage.

4.2. Descriptive statistics

Many studies have demonstrated that population density influences transportation mode choice. Messenger and Ewing (2007) noted that the relationship between population density and choice of public transportation is not direct, but it is an important factor and often used as an indicator when determining whether public transportation expansion is needed. A remote area is defined by the MOI as "a township where population density is lower than one fifth of the national average population density." Therefore, this classifies the townships into five groups based on population density *vs.* national population density: remote, rural, suburban, urban, and downtown.

Table 3 shows the distribution of the public transportation usage rate for each group, including the number of townships, maximum, minimum, average, and standard deviation. The highest average public transportation usage rate was downtown (0.150). The lowest average usage rate was in the suburban (0.081), suggesting that as urbanisation increases, public transportation usage rate increases. However, the public transportation usage rate was generally low.

Fig. 1 lists the average usage rate for each transportation mode in each area. Average motorcycle usage rate was highest for all areas, followed by average car usage rate. Average public transportation usage rate was lowest. As mention, the usage rate of public transportation was highest in the downtown areas, followed by that in the remote area, likely because

Table 2
Definitions of explanatory variables and their expected effects to public transportation usage rate.

Variables	Expected sign	Unit
Socio-economic		
Population density	+	Person/km^2
Number of households	+	Household
Percentage of minors	+	Percentage
Percentage of elders	+	Percentage
Percentage of handicappers	+	Percentage
Number of low-income households	+	Household
Number of employed people	+	Person
Percentage of employed people in primary and secondary industrial sectors	Undecided	Percentage
Percentage of employed people in tertiary industrial sectors	Undecided	Percentage
Number of collegiate students	Undecided	Person
Average income	–	NT $1000
Land use		
Residential area	+	ha
Commercial area	Undecided	ha
Industrial area	Undecided	ha
Private transportation		
Car ownership rate	–	Rate
Motorcycle ownership rate	–	Rate
Road length	–	km
Public transportation		
Number of intercity bus routes	+	Route
Total length of intercity bus routes	+	km
Average daily frequency of intercity bus routes	+	Frequency
Average bus age of intercity bus routes	+	Year
Number of city bus routes	+	Route
Total length of city bus routes	+	km
Average daily frequency of city bus routes	+	Frequency
Average bus age of city bus routes	+	Year
Number of mass rapid transit (MRT) routes	+	Route
Total length of MRT routes	+	km
Average daily frequency of MRT routes	+	Frequency
Number of MRT stations	+	Station
Distance to the nearest interchange	+	km
Distance to the nearest rail station	–	km
Distance to the nearest high-speed rail station	–	km
Distance to the nearest domestic airport	–	km

Note: "+" indicates an expected positive effect; "–" indicates an expected negative effect.

Table 3
Public transportation usage rates of five groups.

Group	Number	Min	Max	Ave	SD
Group 1: Remote	62	0.000	0.302	0.095	0.074
Group 2: Rural	72	0.000	0.293	0.084	0.059
Group 3: Suburban	72	0.000	0.375	0.081	0.066
Group 4: Urban	71	0.000	0.485	0.082	0.091
Group 5: Downtown	71	0.012	0.480	0.150	0.134

Note: Number: number of townships/districts. Max: maximum usage rate. Min: minimum usage rate. Ave: average usage rate. SD: standard deviation of usage rate.

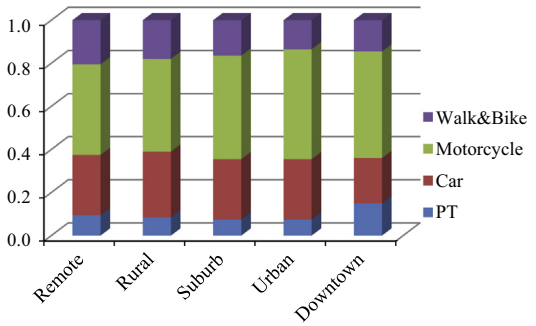


Fig. 1. Average usage rates of transportation modes in each group.

downtown areas have better public transportation services and stricter regulations on the usage of private vehicles than other areas. However, those in remote areas may be "captive" riders due to their social, economic or physical disadvantages, and thereby rely on public transportation.

5. Results

The parameters estimated by the global model (i.e., TRM) are identical for each area; while the parameters estimated by the local model (i.e., GWR) varied among regions. Linear and logarithmic forms of the three models were also developed, but only the forms with the better performance are reported for brevity. The R software was used to estimate two models.

5.1. TRM

Downtown (Group 5) is the basis for comparisons in this study; four dummy variables (Table 4) are included in the TRM model to explore differences in influential factors among the groups. Significance level (α) was set at 0.05. The dummy variables used to indicate the group of regions is useful to examine the spatial heterogeneity among regions by introducing interaction terms. If the absolute t-value was <1.645, the variable did not reach significance and was excluded. Linear and logarithmic functions were used and compared in the models. The estimation results are as follows.

This study used the three major explanatory variables categories and the dependent variable (public transportation usage rate) for estimation with the TRM model. Estimation results show that for the linear and logarithmic forms, respectively, R_{adj}^2 was 0.566 and 0.289, respectively, and Akaike's information criterion (AIC) values were −1929.94 and −56.46, respectively. Both indicators show that the linear performs better than the logarithmic form.

Table 5 shows estimation results by the linear TRM. The F value was 27.645; thus, the null hypothesis (H_0) was rejected, indicating that at least one coefficient was significantly different from 0. The estimation results also suggest that the number of low-income households, average daily service frequency of intercity bus routes, and number of city bus routes were positively correlated, while percentage of minors, average city bus age, motorcycle ownership rate, and road length were negatively correlated. When data from downtown areas were compared, significant differences were identified in percentage of minors, number of city bus routes and average city bus age in a remote area, as were the number of low-income households in a rural area, number of city bus routes in a suburban area, and number of low-income households in an urban area.

5.2. GWR

In contrast to estimations using global models, local models represent each local area. Notably, GWR may explain the impact of characteristics from each dataset on the use rate for various transportation modes after geographic coordinates from each dataset are converted into X/Y coordinates. Moreover, the GWR can clearly discern the level of impact of the characteristic variables in each area on the public transportation usage rate. For subsequent analyses, GWR was used to analyse the relationship between characteristic variables and public transportation usage rate.

First, the spatial bandwidth for GWR must be determined to identify the number of spatial units necessary in the surrounding weighted kernel. In this study, the geometric centre of each township (county and city) was the kernel, and samples in each space were geographically weighted for GWR descriptive statistics and regression analysis. In total, 348 data points (samples) were used. The public transportation usage rate was a dependent variable. To compare the GWR results with those by global linear regression model, significant variables from global linear regression were adopted as explanatory variables for GWR.

Given that the sample size was large (N = 348), this study adopted a fixed kernel bandwidth to reflect the impact of distance between the regression point and a data point. A minimal coefficient of variation (CV) value was applied to determine the required bandwidth for the minimum CV value. The minimum CV value was estimated at 1.089, and the corresponding bandwidth was 25.238 km as shown in Fig. 2; the weight value for significant impact was discerned from samples within a distance of 25.238 km from the kernel of each dataset.

Using descriptive statistics from R statistical software, original data were input and each regression point was a kernel to generate different weights according to the distance between data points and the kernel. Using GWR, each sample was weighted to estimate the parameters for each area. Therefore, each estimated dataset from each sample and each variable

Table 4
Dummy variables.

Group	D_1	D_2	D_3	D_4
Group 1: Remote	1	0	0	0
Group 2: Rural	0	1	0	0
Group 3: Suburban	0	0	1	0
Group 4: Urban	0	0	0	1
Group 5: Downtown	0	0	0	0

Table 5
Estimated TRM parameters.

Variable	Coefficient	*t*-Value	*p*-Value
Constant	0.282	5.968	0.000
Socio-economic			
Percentage of minors	−0.445	−2.669	0.008
Number of low-income households	0.005	3.450	0.001
Public transportation			
Average daily frequency of intercity bus routes	0.010	5.272	0.000
Number of city bus routes	0.015	7.925	0.000
Average age of city buses	−0.004	−2.669	0.007
Private transportation			
Motorcycle ownership rate	−0.149	−5.970	0.000
Road length	−0.032	−5.927	0.000
Dummy			
D_1	0.072	1.087	0.278
D_2	0.059	2.692	0.007
D_3	0.029	1.832	0.068
D_4	0.003	0.193	0.847
Interaction terms			
Percentage of minors × D_1	0.472	1.942	0.053
Number of city bus routes × D_1	−0.110	−2.212	0.028
Average age of city buses × D_1	−0.009	−2.047	0.041
Number of low-income households × D_2	−0.015	−1.917	0.056
Number of city bus routes × D_3	0.041	2.254	0.025
Number of low-income households × D_4	0.008	2.351	0.019
F-value	27.645		
R_{adj}^2	0.566		
AIC	−1929.94		

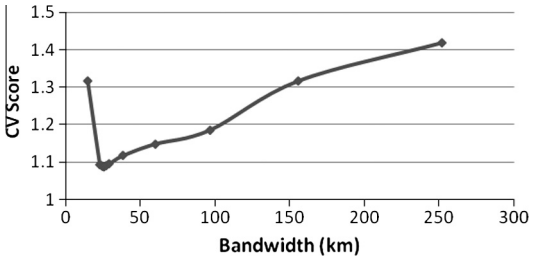


Fig. 2. Schematic diagram of spatial bandwidth.

had a weighted average and weighted standard deviation (SD). Because the sample size was too large, Tables 6 and 7 show values as min, max, median, and mean.

For the estimation results from the parameters, each dataset had coefficient values and *t*-values. In total, 348 estimation results were obtained. The tables would be excessively large if the estimation results for each dataset were shown. Therefore, Tables 8 and 9 show the coefficient and min value, first quartile value, median, mean, third quartile value and maximum *t*-value for each explanatory variable.

Based on the estimated GWR model, the effect of explanatory variables on public transportation usage rate can be spatially investigated as shown in Figs. 3–9. The estimated parameters of the GWR model vary across regions, indicating that each parameter included local characteristics, and therefore, an analysis of model performance through estimation could be a reference for government and public transportation operators, and GWR could be a better choice to analyse the key factors influencing public transportation use in the future. For percentage of minors, Fig. 3(a) and (b) show the distributions of estimated parameters and *t*-values across regions in Taiwan, respectively. The effect of percentage of minors is most significant in northern (especially Taipei Metropolitan City), mid-eastern and southern areas (Fig. 3(b)). The significant positive correlations that existed between percentage of minors and public transportation usage rate were primarily distributed in the eastern area, suggesting that minors in this area had weak transportation autonomy and thus relied on public transportation. Bus frequency could be increased during rush hour or a dynamic information system could be created to reduce uncertainties of this population group when waiting for public transportation. Significant negative correlations for public transportation were primarily distributed in the north central area and were increasingly strong moving northward. We speculate that this is due to the daily life in the north, where parents prefer transporting children on their own rather than

Table 6
Descriptive statistics for the GWR averages.

Variable	Min	Median	Ave	Max
Public transportation usage rate	0.016	0.068	0.099	0.338
Percentage of minors	0.157	0.199	0.202	0.274
Number of low-income households	25.210	225.660	324.590	1133.310
Average daily frequency of intercity bus routes	3.300	18.960	22.810	57.830
Number of city bus routes	0.000	1.710	11.911	78.825
Average age of city buses	4.725	11.701	10.384	13.000
Motorcycle ownership rate	0.438	0.662	0.661	0.846
Road length	26.510	116.310	119.260	225.240

Table 7
Descriptive statistics for the GWR SDs.

Variable	Min	Median	Ave	Max
Public transportation usage rate	0.004	0.038	0.048	0.137
Percentage of minors	0.011	0.026	0.026	0.057
Number of low-income households	2.828	172.008	264.914	847.114
Average daily frequency of intercity bus routes	0.326	10.511	13.423	46.802
Number of city bus routes	0.000	3.596	11.822	57.253
Average age of city buses	0.000	1.598	1.578	4.199
Motorcycle ownership rate	0.017	0.058	0.092	0.329
Road length	9.570	65.690	67.100	156.800

Table 8
Estimated GWR coefficients.

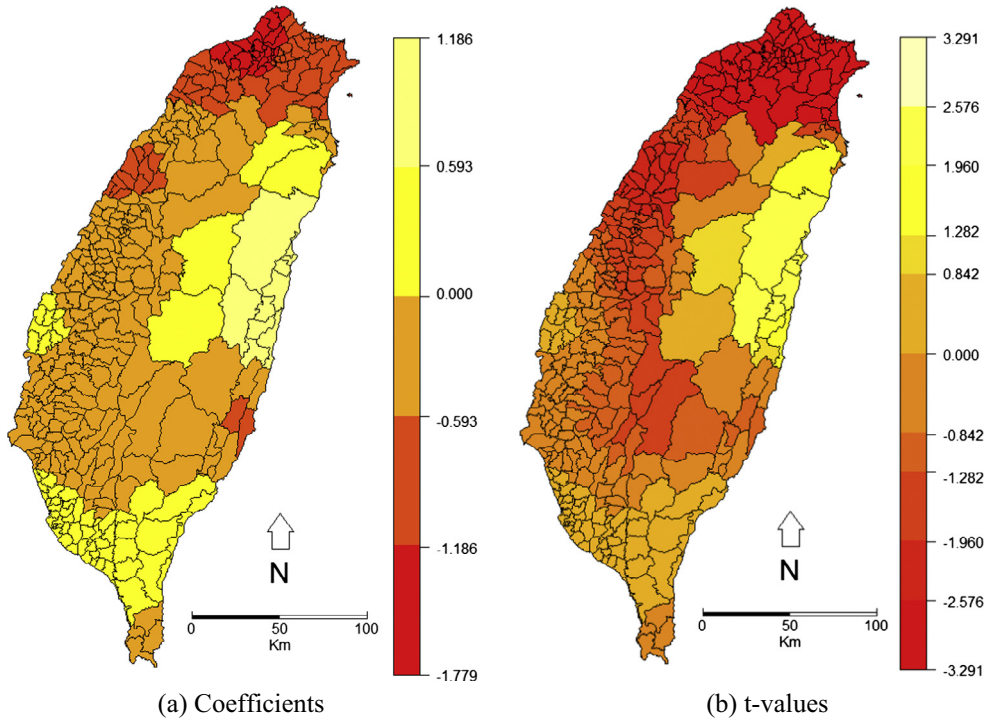
Variable	Min	First quartile	Median	Ave	Third quartile	Max
Constant	−0.336	0.063	0.158	0.173	0.273	0.549
Percentage of minors	−1.475	−0.578	−0.254	−0.332	−0.030	1.185
Number of low-income households	−0.024	−0.002	0.001	−0.001	0.002	0.011
Average daily frequency of intercity bus routes	−0.013	−0.001	0.003	0.004	0.009	0.034
Number of city bus routes	−0.005	0.002	0.011	0.009	0.013	0.099
Average age of city buses	−0.007	−0.002	0.0002	−0.001	0.001	0.013
Motorcycle ownership rate	−0.345	−0.057	−0.021	−0.010	0.023	0.465
Road length	−0.009	−0.003	−0.001	−0.002	−0.001	0.002

Table 9
Estimated GWR *t*-values.

Variable	Min	First quartile	Median	Ave	Third quartile	Max
Constant	−1.348	0.877	1.467	2.693	2.937	10.290
Percentage of minors	−6.586	−2.105	−1.061	−1.475	−0.133	2.333
Number of low-income households	−1.770	−0.333	0.220	0.213	1.086	2.710
Average daily frequency of intercity bus routes	−0.818	−0.373	0.448	1.037	1.895	6.374
Number of city bus routes	−0.401	0.256	1.284	2.301	4.181	8.337
Average age of city buses	−3.967	−0.678	0.038	−0.281	0.357	1.714
Motorcycle ownership rate	−5.168	−0.720	−0.325	−0.380	0.332	2.234
Road length	−6.913	−1.931	−1.023	−1.495	−0.546	0.461

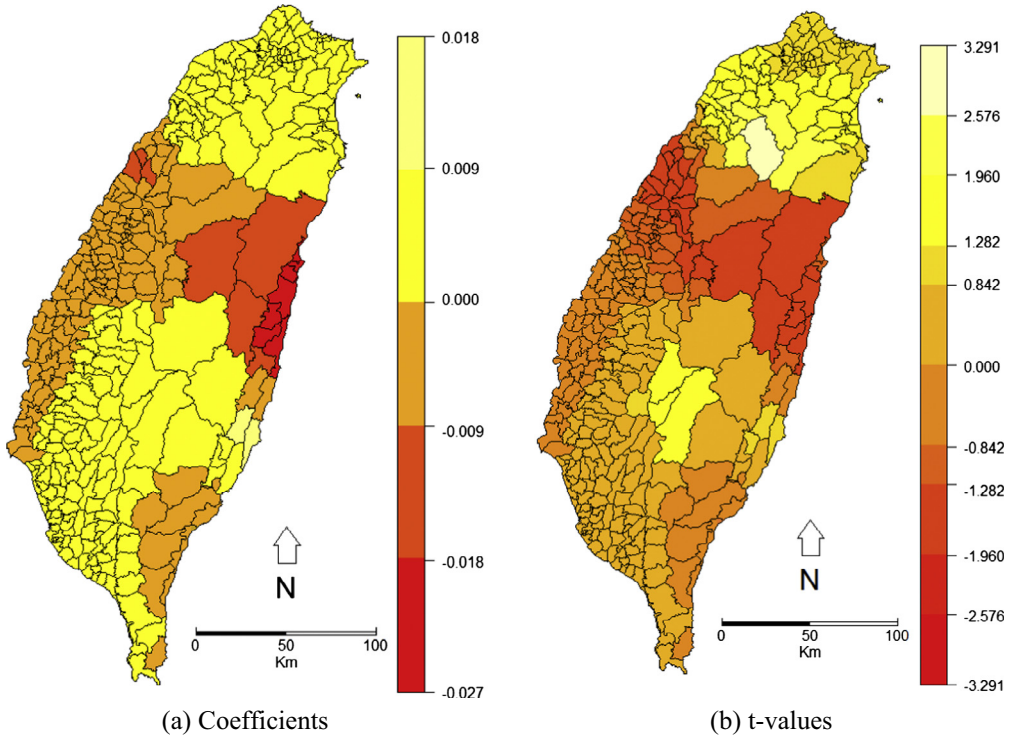
let children walk to school in heavy traffic, resulting in a negative correlation between percentage of minors and public transportation usage rate. The safety of taking local public transportation should be enhanced to ensure parents that their children are safe in urban environments.

Significant positive correlations that exist for the number of low-income households and public transportation usage rate are primarily in the south and centre of the northern region. Low-income households in these areas cannot likely afford private transportation, and thus use public transportation. Significant negative correlations between public transportation usage and number of low-income households were primarily in the central mountainous area. Given the convenience and overall lower travel cost (time and fuel price) of motorcycles when compared with that of public transportation (ticket price) and the insufficient public transportation service in the area, people from lower-income households prefer motorcycles to public transportation, resulting in low public transportation usage rates. Regardless of whether low-income households use public transportation, ticket subsidies for this group could increase their use of public transportation.



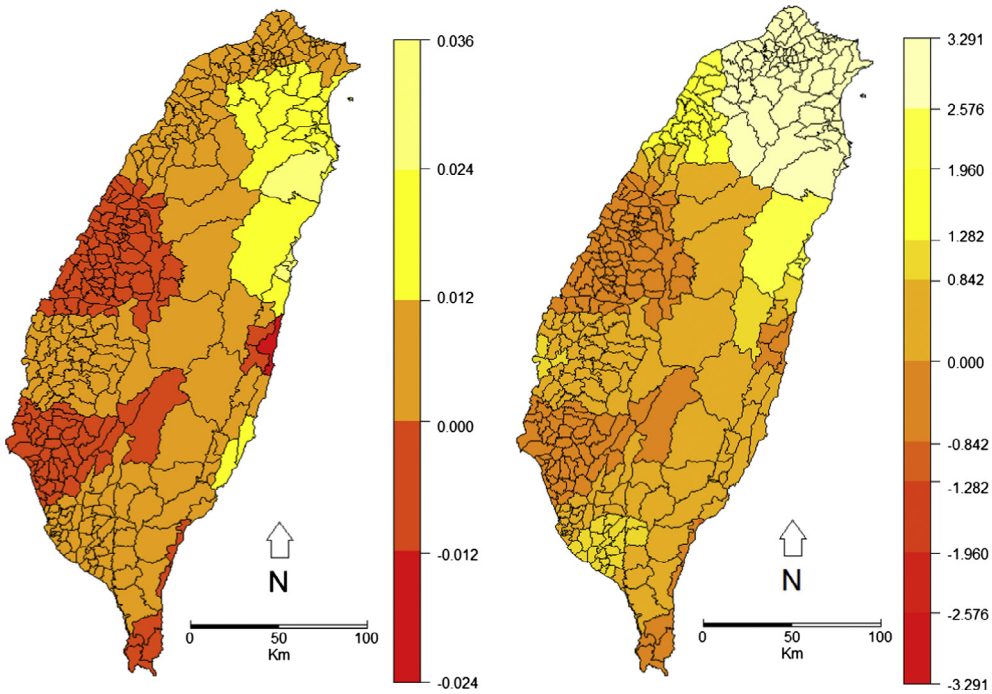
(a) Coefficients (b) t-values

Fig. 3. Spatial distribution of estimated coefficients and *t*-values of percentage of minors.



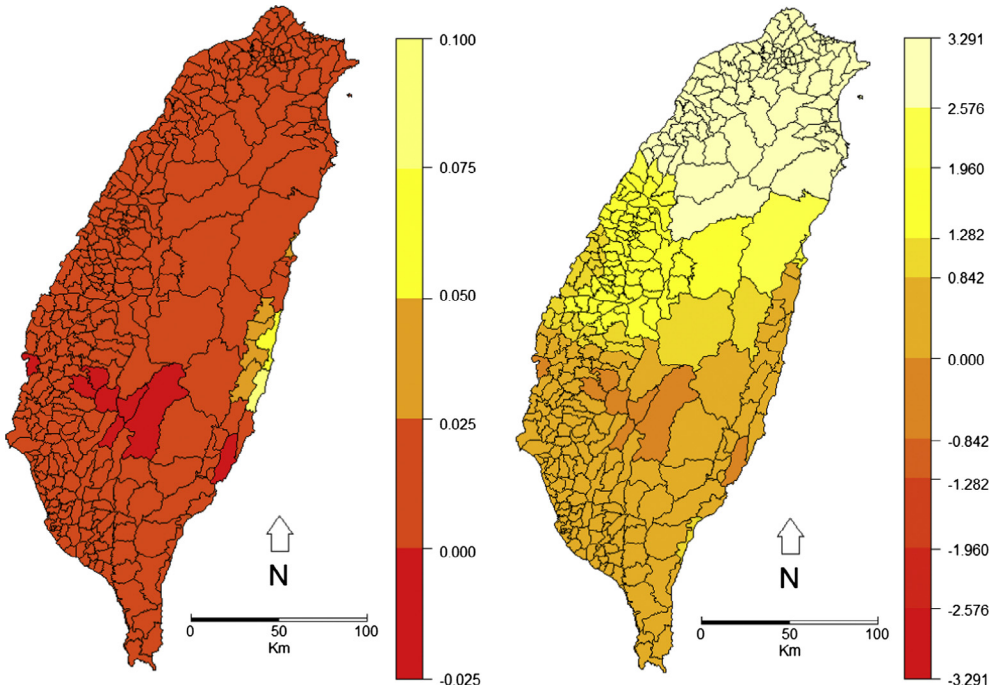
(a) Coefficients (b) t-values

Fig. 4. Spatial distribution of estimated coefficients and *t*-values of number of low-income households.



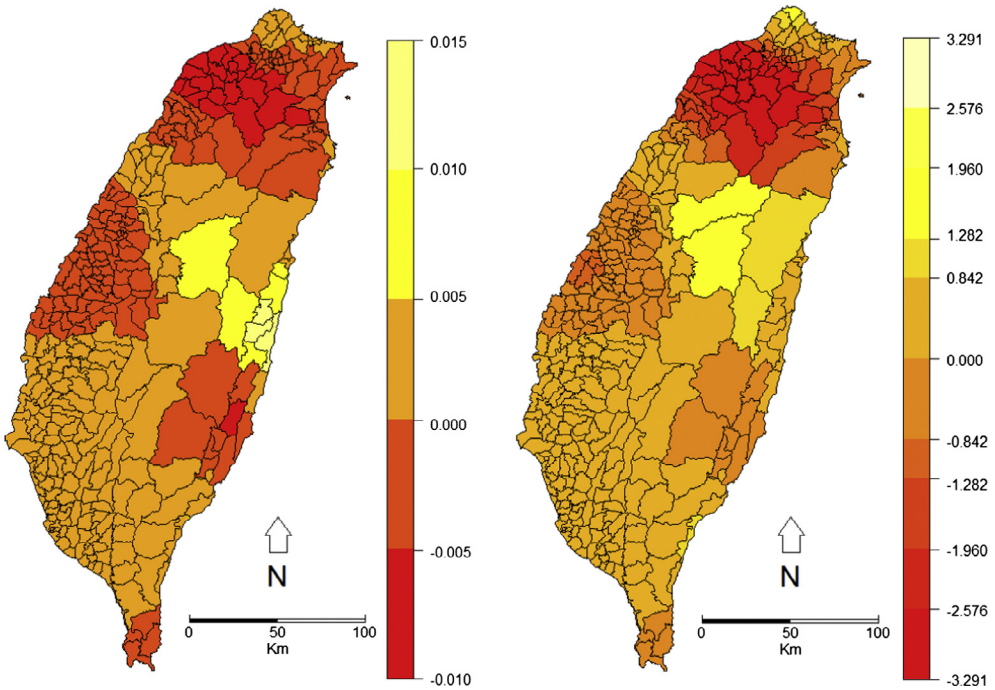
(a) Coefficients (b) t-values

Fig. 5. Spatial distribution of estimated coefficients and *t*-values of average daily frequency of intercity bus routes.



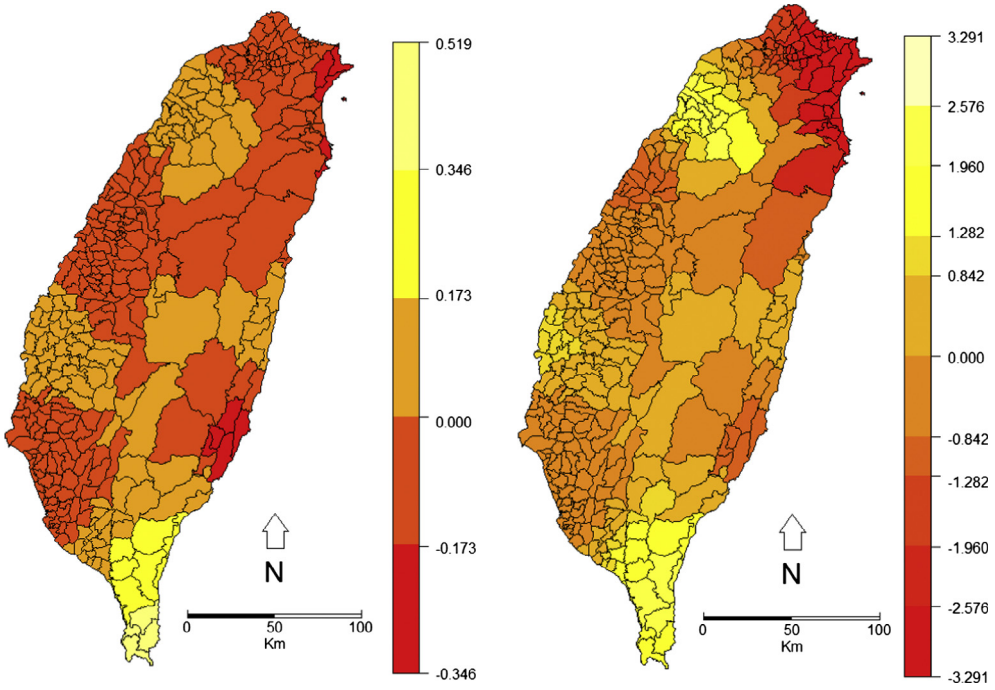
(a) Coefficients (b) t-values

Fig. 6. Spatial distribution of estimated coefficients and *t*-values of number of city bus routes.



(a) Coefficients (b) t-values

Fig. 7. Spatial distribution of estimated coefficients and *t*-values of average age of city buses.



(a) Coefficients (b) t-values

Fig. 8. Spatial distribution of estimated coefficients and *t*-values of motorcycle ownership rate.

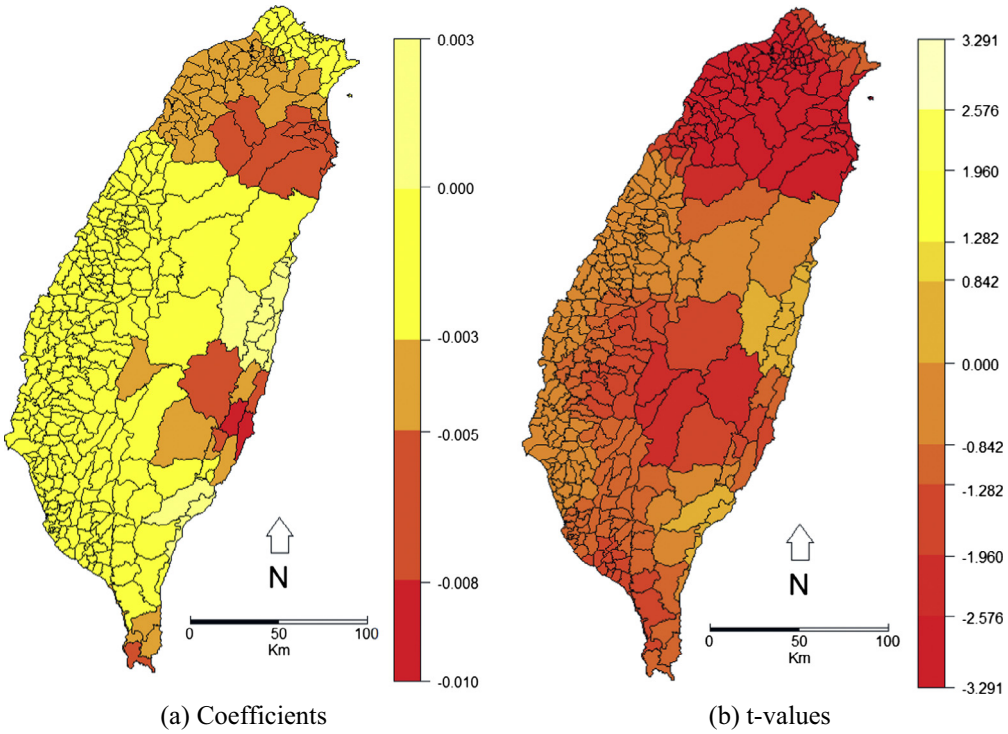


Fig. 9. Spatial distribution of estimated coefficients and *t*-values of road length.

Significant positive correlations between average daily frequency of intercity bus routes and public transportation usage rate were primarily in the north. Public bus routes cover counties and cities and are important to residents who live near administrative boundaries or commute across counties and cities. Daily frequency affects passenger waiting time and is especially important to commuting. Therefore, daily frequency could be increased to increase the public transportation usage rate.

Significant positive correlations between the number of city bus routes and the public transportation usage rate were primarily in the north central area, which has many city buses. As expected, city bus routes primarily serve city residents. The number of routes is an effective indicator of coverage density for a road network. By increasing the number of routes, the access time/distance to bus stations or stops could be minimized. Therefore, the number of city bus routes could be increased and thereby increase the public transportation usage rate.

Significant negative correlations between average age of city buses and public transportation usage rate were primarily in the north, suggesting that in the north, which has a sufficient supply of city buses, requirements for public transportation services are also high. Therefore, replacing buses in urban areas could improve service quality.

These improvement strategies for public transportation service-related factors may not be appropriate for suburbans with low population densities. Because the number of residents who use public transportation in these areas is low, the public transportation network has few routes, and some rural areas may lack any public transportation service. However, because para-transit has the advantages of low volume and route flexibility, transportation-related departments should set para-transit-related regulations to supplement inadequate public transportation in remote areas.

Significant negative correlations between motorcycle ownership rate and public transportation usage rate were primarily in the north, demonstrating that the motorcycle ownership rate is high in this area. Moreover, motorcycles are very convenient and cost little to operate; thus, they are the preferred mode of transportation. Therefore, the cost of motorcycle ownership could be increased or the number of motorcycles in each household could be restricted to affect a shift towards public transportation usage.

Significant negative correlations between road length and public transportation usage rate were primarily in remote mountainous areas. Because public transportation is less accessible than private transportation, longer road length suggests greater accessibility to roads in this area, which would increase the public preference for private transportation, further decreasing the public transportation usage rate. Therefore, to simultaneously meet local transportation demands and build roads, public transportation service routes could be increased or para-transit could be provided to enhance public transportation (the-last-mile) service, increasing accessibility to public transportation and thereby attracting users.

6. Comparisons

The overall explanatory power of each model illustrated the extent to which the model could explain dependent variables, primarily public transportation usage rate. Table 10 compares the explanatory power and Moran's *I* test of the global and local regression models. Comparison results show that the explanatory power of the local regression model was significantly higher than that of the global regression model, which is consistent with findings in previous studies (e.g. Chow et al., 2006, 2010; Gutiérrez et al., 2011; Cardozo et al., 2012; Pulugurtha and Agurla, 2012). The R^2 and $R_{adj}{}^2$ are largely improved from 0.588 and 0.566 to 0.771 and 0.767. Additionally, analysis of the residuals shows better results (closer to the expected value) in the GWR model than in the TRM model. The Moran's *I* test also supports that the GWR model successfully accommodates spatial autocorrelation, but obviously the TRM model fails to do so.

The R^2 explanatory power of GWR for each area ranged from 12.54% (minimum) to 86.43% (maximum) as shown in Fig. 10. Explanatory power increased from southwest to northeast Taiwan and increased slightly from southwest to southeast Taiwan. It should be noted that although the overall explanatory power of GWR has been largely improved, there are still more than half of the regions with R^2 less than 0.37. How to further improve the goodness of fit of GWR in these regions remains a further study.

Table 10
Comparisons of explanatory power and spatial autocorrelation.

Goodness of fit	TRM	GWR
R^2	0.588	0.771
$R_{adj}{}^2$	0.566	0.767
Spatial autocorrelation		
Moran's *I*	0.2919934	−0.0111799
Expectation	−0.0028818	−0.0028818
Variance	0.0010593	0.0010570
Z-score	9.060	−0.255
p-Value	<2.2e−16	0.799

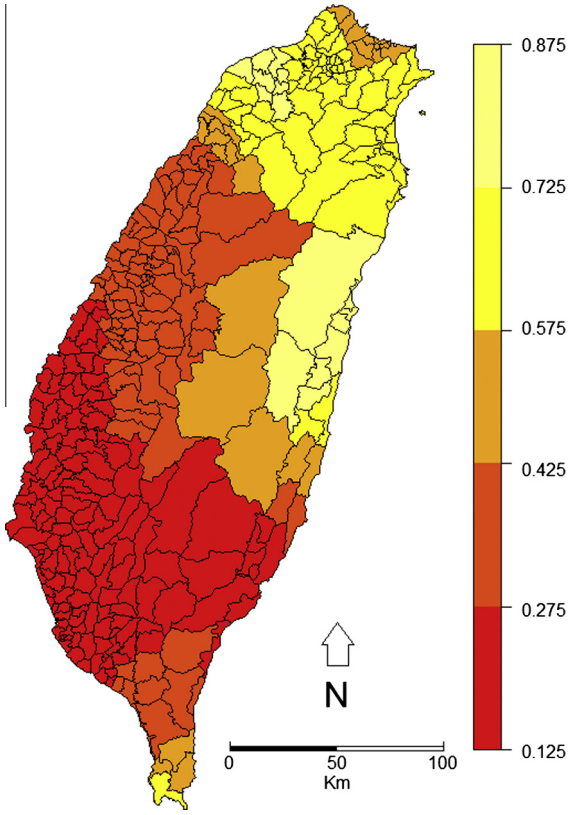


Fig. 10. Spatial distribution of the determination coefficient (R^2) in each region.

Table 11
Comparisons of model prediction accuracy.

Group	TRM		GWR	
	MAPE	Corrected MAPE	MAPE	Corrected MAPE
1	126.61	49.93	79.01	35.97
2	72.39	45.12	59.11	40.01
3	70.64	45.51	69.83	37.41
4	106.86	53.41	72.00	39.20
5	78.12	36.85	38.30	20.49
All	89.89	44.84	63.26	32.64

The MAPE and Corrected MAPE were applied to compare the accuracy of the TRM and GWR in predicting the public transportation usage rates of various groups as shown in Table 11. As shown in Table 11, GWR performs better than the TRM model with lower MAPE and Corrected MAPE values of all groups. The order of accuracy for each group varied with the different models, but the Corrected MAPE was significantly lower than the MAPE. The denominator of the corrected MAPE (Eq. (9)) was the average of all actual values, and therefore the difference between $|y_i - \widehat{y}_i|$ and $\bar{y}$ was smaller than $|y_i - \widehat{y}_i|$ and y_i from the original MAPE. The Corrected MAPE shows that the accuracy for downtown is the highest, suggesting that isomorphism downtown is great. Additionally, analysis of the sample and groups show that the GWR has the highest accuracy.

7. Conclusions

To better understand the key factors to public transportation use, this study respectively applies global (TRM) and local (GWR) regression models to identify the key factors to the public transportation usage rates of a total of 348 regions (township or districts) in Taiwan. In terms of corrected MAPE and Moran's *I*, GWR model performs better than TRM model, suggesting the superior of the local model under the existence of spatial autocorrelation. According to the estimation results of GWR, seven variables are significantly tested, and most have parameters that differ across regions in Taiwan. For example, the percentage of minors, the most significantly tested variable, has a negative effect on public transportation usage rate on the north-western part of Taiwan where is highly populated, but show a positive effect on public transportation usage rate on southern and eastern parts of Taiwan which are considered as remote areas, suggesting the tendency of spatial clustering effect. With the introduction of interaction terms, TRM can also show similar results with GWR model that the percentage of minors has a positive effect on public transportation usage rate in remote areas (with the lowest population density), but shows a negative effect in other areas. Anyhow, the introduction of interaction terms requires of subjective judgement and trial-and-error efforts, the local models having higher flexibility in parameters estimation are recommended for the cases with existence of spatial autocorrelation. Based on these findings, strategies that increase public transportation use are then proposed.

In this study, variables of the original linear model were maintained in GWR for comparison purposes. Future studies can alter the variables or increase the variety of variables. Additionally, it should be noted that although the overall explanatory power of GWR has been largely improved, there are still more than half of the regions with R^2 less than 0.37. How to further improve the goodness of fit of GWR in these regions deserves a further study. For data collection, difficulty in gathering information on private transportation-related variables and personal data, such as travel time and cost, may decrease the explanatory power of models. To periodically collect data of these variables for policy analysis is essential. To account for the competition among transportation modes, the GWR-based simultaneous equation model and aggregate logit model for simultaneously modelling usage rates of all modes can be developed. Meanwhile, the mixed GWR model, a semi-local approach, which is able to estimate both global and local parameters can be adopted, so some coefficients with small variation over space are kept constant over all study areas. Last but not least, GWR, a geospatial statistical analysis, differs from general statistical analysis wherein panel data are used to estimate weights. The variables used herein were cross-sectional data, which are appropriate for GWR. In future studies, data for related variables can be tracked continuously, and other weighting methods can be incorporated into the model for comparison by GWR.

Acknowledgments

The authors would like to thank the insightful comments and constructive suggestions from three anonymous reviewers, which have helped clarify several weak points of the original version of this paper. This study was sponsored by the Institute of Transportation, Ministry of Transportation and Communications of the Republic of China, under contract MOTC-IOT-101-MEB012.

References

Blainey, S., 2010. Trip end models of local rail demand in England and Wales. J. Transp. Geogr. 18 (1), 153–165.
Blainey, S.P., Preston, J.M., 2010. A geographically weighted regression based analysis of rail commuting around Cardiff, South Wales. In: Paper Presented at the 12th World Conference on Transportation Research, Lisbon, Portugal.

Boame, A.K., 2004. The technical efficiency of Canadian urban transit systems. Transp. Res. Part E 40 (5), 401–416.
Buehler, R., 2011. Determinants of transport mode choice: a comparison of Germany and the USA. J. Transp. Geogr. 19 (4), 644–657.
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. Appl. Geogr. 34, 548–558.
Cervero, R., 1996. Commuter and Light-rail Transit Corridors: The land Use Connection, TCRP Report 16, Transportation Research Board, National Research Council, Washington, D.C.
Cervero, R., Murakami, J., Miller, M., 2010. Direct ridership model of bus rapid transit in Los Angeles County, California. Transp. Res. Rec. 2145, 1–7.
Chow, L.F., Chi, H., Zhao, F., 2010. Subregional transit ridership models based on geographically weighted regression. In: Paper Presented At 89th Annual Meeting of Transportation Research Board, Paper #10e3810.
Chow, L.F., Zhao, F., Li, M.T., Liu, X., Ubaka, I., 2006. Transit ridership model based on geographically weighted regression. J. Transp. Res. Board 1972, 105–114.
Chen, C., Chen, J., Barry, J., 2009. Diurnal pattern of transit ridership: a case study of the New York City subway system. J. Transp. Geogr. 17 (3), 176–186.
Coldren, G.M., Koppelman, F.S., Kasturirangan, K., Mukherjee, A., 2003. Modeling aggregate air-travel itinerary shares: logit model development at a major US airline. J. Air Transp. Manage. 9 (6), 361–369.
Fotheringham, A.S., Brunsdon, C., Charlton, M.E., 2002. Geographically Weighted Regression. Wiley, Chichester.
Fotheringham, A.S., Charlton, M.E., 1998. Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. Environ. Plan. A 30 (11), 1905–1927.
Fotheringham, A.S., Charlton, M.E., Brunsdon, C., 2000. Quantitative Geography. Sage, London.
Greene, W., 2003. Econometric Analysis. Prentice Hall, New Jersey.
Gutiérrez, J., Cardozo, O., García-Palomares, J.C., 2011. Transit ridership forecasting at station level: an approach based on distance-decay weighted regression. J. Transp. Geogr. 19, 1081–1092.
Kobayashi, T., Lane, B., 2007. Spatial heterogeneity and transit use. In: Paper Presented at the 11th World Conference on Transportation Research. USA: Berkeley.
Kuby, M., Barranda, A., Upchurch, C., 2004. Factors influencing light-rail station boardings in the United States. Transp. Res. A 38 (3), 223–247.
Lloyd, C., Shuttleworth, I., 2005. Analysing commuting using local regression techniques: scale, sensitivity, and geographical patterning. Environ. Plan. A 37 (1), 81–103.
Messenger, T., Ewing, R., 2007. Transit-oriented development in the sun belt. J. Transp. Res. Board 1996, 145–153.
Mulley, C., Tanner, M., 2009. The vehicle kilometres travelled by private car: a spatial analysis using geographically weighted regression. In: Proceedings of the 32nd Australasian Transport Research Forum (ATRF). Auckland: ATRF.
Pulugurtha, S.S., Agurla, M., 2012. Assessment of models to estimate bus-stop level transit ridership using spatial modeling methods. J. Public Transport. 15 (1), 2012.
Souche, S., 2010. Measuring the structural determinants of urban travel demand. Transp. Policy 17, 127–134.
Swimmer, C.R., Klein, C.C., 2010. Public transportation ridership levels. J. Econ. Educat. 10 (1), 40–46.
Taylor, B.D., Miller, D., Iseki, H., Fink, C., 2009. Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas. Transp. Res. Part A 43 (1), 60–77.