International Journal of Geographical Information Science
Vol. 24, No. 3, March 2010, 383–401

Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices

Bo Huang[a]*, Bo Wu[b] and Michael Barry[c]

[a]*Department of Geography and Resource Management, The Chinese University of Hong Kong, Shatin, NT, Hong Kong;* [b]*Spatial Information Research Center, Fuzhou University, Fuzhou, Peoples Republic of China;* [c]*Department of Geomatics Engineering, University of Calgary, Calgary, AB, Canada*

(*Received 10 April 2008; final version received 30 November 2008*)

By incorporating temporal effects into the geographically weighted regression (GWR) model, an extended GWR model, geographically and temporally weighted regression (GTWR), has been developed to deal with both spatial and temporal nonstationarity simultaneously in real estate market data. Unlike the standard GWR model, GTWR integrates both temporal and spatial information in the weighting matrices to capture spatial and temporal heterogeneity. The GTWR design embodies a local weighting scheme wherein GWR and temporally weighted regression (TWR) become special cases of GTWR. In order to test its improved performance, GTWR was compared with global ordinary least squares, TWR, and GWR in terms of goodness-of-fit and other statistical measures using a case study of residential housing sales in the city of Calgary, Canada, from 2002 to 2004. The results showed that there were substantial benefits in modeling both spatial and temporal nonstationarity simultaneously. In the test sample, the TWR, GWR, and GTWR models, respectively, reduced absolute errors by 3.5%, 31.5%, and 46.4% relative to a global ordinary least squares model. More impressively, the GTWR model demonstrated a better goodness-of-fit (0.9282) than the TWR model (0.7794) and the GWR model (0.8897). McNamara's test supported the hypothesis that the improvements made by GTWR over the TWR and GWR models are statistically significant for the sample data.

Keywords: geographically and temporally weighted regression; geographically weighted regression; spatial nonstationarity; temporal nonstationarity; housing price; Calgary

1. Introduction

Location and time are important determinants of real estate prices. Location is an important factor because of the spatial dependence between real estate prices, even though there tends to be spatial heterogeneity across prices over a large area. Prices of proximate houses tend to be similar because they share common local neighborhood factors, such as similar physical characteristics (age, size, and exterior and interior features) and similar neighborhood amenities (socioeconomic status, access to employment opportunities, shopping, public service facilities, schools, etc). Differences between houses in the same neighborhood tend to be determined by the size of the lot and the size and quality of the top structure.

*Corresponding author. Email: bohuang@cuhk.edu.hk

ISSN 1365-8816 print/ISSN 1362-3087 online

DOI: 10.1080/13658810802672469
http://www.informaworld.com

A large-size lot, for example, tends to have a large house, a garage, and more bedrooms. In older neighborhoods, the date and quality of construction and the level of renovation tend to be major discriminating factors.

Housing price observations in many studies tend to be geo-referenced to account for spatial autocorrelation and general neighborhood characteristics. The time of an observation also matters in the determination of real estate prices. From a modeling perspective, it is generally accepted that real estate prices depend not only on recent market events but also on their lagged prices. Temporal effects include, for example, market trends, inflationary factors, and differential rates of obsolescence related to the age distribution of houses. Depreciation of housing amenities might occur at different rates related to housing characteristics at the beginning of the study period, the original value of the house, the specific amenities included in the housing package, and other factors omitted in models that do not specifically account for intertemporal heterogeneity (Dombrow *et al.* 1997).

Heterogeneous spatial and/or temporal effects may violate the basic assumption of statistical independence of observations which is typically required for unbiased and efficient estimation (Huang *et al.* 2009). Several studies have, therefore, tended to incorporate spatial or temporal characteristics in house price equations to eliminate dependencies, or nonrandom effects, in the residuals.

Prediction methods using regression can be grouped into two general categories: global and local regression models. Global spatial models are usually an improved form of the traditional hedonic model (Can 1992, Dubin 1992, Anselin 1998). Spatial or temporal effects are addressed by modeling the residual variance–covariance matrix directly or by inversing the residual variance–covariance matrix to eliminate dependency in the residuals. In Gelfand *et al.* (2004), a rich class of temporal hedonic models using a Bayesian framework is formulated by extending the different processes of the error term. An extended model of this study with spatially varying coefficient process is also developed in Gelfand *et al.* (2003). In fact, this model can be expanded to a spatio-temporal setting. Can and Megbolugbe (1997) constructed a distance-weighted average variable that captures both spatial and temporal information, which proves to be a significant explanatory variable in the hedonic model. Alternatively, Pace *et al.* (1998) introduced a filtering process, and their results greatly enhanced the accuracy of model estimations.

However, a major problem with global methods when applied to spatial or temporal data is that the processes being examined are assumed to be constant over space. For a specific model (e.g., the price of real estate), the assumption of stationarity or structural stability over time and space is generally unrealistic, as parameters tend to vary over the study area.

In order to capture the spatial variation, various localized modeling techniques have been proposed to capture spatial heterogeneity in housing markets. Eckert (1990) suggested that, based on the assumption that subsets are characterized by a lower variance, models generated for housing submarkets should yield greater explanatory power (and predictive accuracy) than those computed at the overall market level. Goodman and Thibodeau (1998) introduced the concept of hierarchical linear modeling, whereby dwelling characteristics, neighborhood characteristics, and submarkets interact to influence house prices. In a similar vein, McMillen (1996) and McMillen and McDonald (1997) introduced nonparametric local linear regression in nonmonocentric city models. Notably, Brunsdon *et al.* (1996), Fotheringham *et al.* (1996), and Fotheringham *et al.* (2002) proposed geographically weighted regression (GWR) as a local variation modeling technique.

GWR allows the exploration of the variation of the parameters as well as the testing of the significance of this variation, and this methodology has, therefore, received considerable attention in recent years. Pavlov (2000), Fotheringham *et al.* (2002), and Yu (2006) have all applied the GWR or GWR-similar methodology to housing markets. Brunsdon *et al.* (1999), in a study of house prices in the town of Deal in south-eastern England, examined the determinants of house price with GWR and found that the relationship between house price and size varied significantly through space. Despite the strength of GWR as opposed to global models and the success of GWR in capturing spatial variations, applying a GWR model to house price analysis that incorporates temporal effects remains a relatively unexplored area.

The objective of this article is to extend the traditional GWR model to problems involving both spatial and temporal nonstationarity in real estate data. This study seeks to contribute to the literature on the topic in the following three ways. First, we extend the traditional GWR model with temporality into a geographically and temporally weighted regression (GTWR) model and apply it to spatio-temporal real estate data analysis. Second, we propose the use of McNamara's test for comparing the statistically significant difference between the estimation methods according to their accuracies (Foody 2004). Third, we examine and compare the hedonic model, temporally weighted regression (TWR), GWR, and GTWR for modeling housing prices by means of a case study in the Canadian city of Calgary.

This article is structured as follows. In Section 2, we present a basic framework for GWR and then extend it to include temporal data. Section 3 offers some key technical implementation details of the GTWR model, including spatial and parameter optimal selection, and model comparison criterion. In Section 4, a case study of housing prices in Calgary is reported. In Section 5, the results for different models are compared and analyzed. Finally, we summarize and draw conclusions.

2. Geographically weighted regression model

2.1. The model and the parameter estimation

The GWR model extends the traditional regression framework by allowing parameters to be estimated locally so that the model can be expressed as

$$Y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)X_{ik} + \varepsilon_i \quad i = 1, \ldots, n \tag{1}$$

where (u_i,v_i) denotes the coordinates of the point i in space, $\beta_0(u_i,v_i)$ represents the intercept value, and $\beta_k(u_i,v_i)$ is a set of values of parameters at point i. Unlike the 'fixed' coefficient estimates over space in the global model, this model allows the parameter estimates to vary across space and is therefore likely to capture local effects.

To calibrate the model, it is assumed that the observed data close to point i have a greater influence in the estimation of the $\beta_k(u_i,v_i)$ parameters than the data located farther from observation i. The estimation of parameters $\beta_k(u_i,v_i)$ is given by Equation (2)

$$\hat{\beta}(u_i, v_i) = \left[X^T W(u_i, v_i)X\right]^{-1} X^T W(u_i, v_i)Y \tag{2}$$

where $W(u_i,v_i)$ is an $n \times n$ matrix whose diagonal elements denote the geographical weighting of observation data for observation i, and the off-diagonal elements are zero. The weight matrix is computed for each point i at which parameters are estimated.

2.2. Weighting matrix specification

The weight matrix in GWR represents the different importance of each individual observation in the data set used to estimate the parameters at location i. In general, the closer an observation is to i, the greater the weight. Thus, each point estimate i has a unique weight matrix.

In essence, there are two weighting regimes that can be used: fixed kernel and adaptive kernel. For the fixed kernel, distance is constant but the number of nearest neighbors varies. For the adaptive kernel, distance varies but the number of neighbors remains constant. The most commonly used kernels are Gaussian distance decay-based functions (Fotheringham *et al.* 2002):

$$W_{ij} = \exp\left(-\frac{d_{ij}^2}{h^2}\right) \tag{3}$$

where h is a non-negative parameter known as bandwidth, which produces a decay of influence with distance and d_{ij} is the measure of distance between location i and j. Using point coordinates (x_i,y_i) and (x_j,y_j), the distance is usually defined as a Euclidean distance

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{4}$$

According to Equations (3) and (4), if i and j coincide, the weight of that observation will be unity, and the weight of other data will decrease according to the Gaussian curve when the distance between i and j increases. Other commonly used weighting functions include the bi-square function (Fotheringham *et al.* 2002) and the tri-cube kernel function (McMillen 1996).

To avoid (1) exaggerating the degree of nonstationarity present in the areas where data are sparse or (2) mask subtle spatial nonstationarity where the data are dense (Paez *et al.* 2002), adaptive weighting functions are used to change the kernel size to suit localized observation patterns. Kernels have larger bandwidths where the data points are sparsely distributed and smaller ones where the data are plentiful. By adapting the bandwidth, the same number of nonzero weights is used for each regression point i in the analysis. For example, the adaptive bi-square weighting function is the following:

$$W_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{h_i}\right)^2\right]^2, & \text{if } d_{ij} < h_i \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where h_i stands for the bandwidth particular to location i.

2.3. Choosing an appropriate bandwidth

In the process of calibrating a GWR model, the weighting model should first be decided. This can be done by cross-validation. Suppose that the predicted value of y_i from GWR is denoted as a function of h by $\hat{y}_i$ (h), the sum of the squared error may then be written as

$$\mathrm{CVRSS}(h) = \sum_i (y_i - \hat{y}_{\neq 1}(h))^2 \tag{6}$$

In practice, plotting CVRSS(h) against the parameter h can provide guidance on selecting an appropriate value of the parameter or it can be obtained automatically with an optimization technique by minimizing Equation (6) in terms of goodness-of-fit statistics or the corrected Akaike information criterion (AIC) (Hurvich *et al.* 1998, Fotheringham *et al.* 2002).

3. Extending GWR with temporal variations

Since complex temporal effects can also lead to nonstationarity in real estate prices, this article demonstrates how to incorporate temporal information in the GWR model to develop a GTWR model that captures both spatial and temporal heterogeneity and improves its goodness-of-fit.

3.1. Accounting for spatio-temporal nonstationarity

In practice, the GWR model accounts for spatial nonstationarity in parameter estimates by constructing a weight matrix based on distances between estimation point i and all other observations. Conventionally, the time variable is accommodated separately by adjusting the sale price observations to a common date, often using some adapted form of present value or future value calculation (Wang 2006).

As an alternative (or perhaps complementary) approach, we accounted for spatio-temporal nonstationary in parameter estimates by constructing the weight matrix based on distances determined from (x, y, t) coordinates between observation i and all other observations in line with the GWR technique. Thus, the GTWR model can be expressed as

$$Y_i = \beta_0(u_i, v_i, t_i) + \sum_k \beta_k(u_i, v_i, t_t)X_{ik} + \varepsilon_i \tag{7}$$

Essentially, the problem here is to provide estimates of $\beta_k(u_i,v_i,t_i)$, for each variable k and each space–time location i. Similarly, the estimation of $\beta_k(u_i,v_i,t_i)$ can be expressed as follows:

$$\hat{\beta}(u_i, v_i, t_i) = \left[X^T W(u_i, v_i, t_i)X\right]^{-1} X^T W(u_i, v_i, t_i)Y \tag{8}$$

where $W(u_i,v_i,t_i) = \mathrm{diag}(\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in})$ and n is the number of observations. Here the diagonal elements $\alpha_{ij}(1 \leq j \leq n)$ are space–time distance functions of (u,v,t) corresponding to the weights when calibrating a weighted regression adjacent to observation point i. Thus, the spatio-temporal GTWR model relies on the appropriate specification of the space–time distance decay function α_{ij}. To calibrate the model, it is still assumed that the observed data points 'close' to point i in the space–time coordinate system have a greater influence in the estimation of the $\beta_k(u_i,v_i,t_i)$ parameters than the data located farther from observation i. In this sense, the definition of 'close' incorporates two variables: temporal closeness and spatial closeness. Hence, defining and measuring the so-called closeness in a space–time coordinate system is a key problem in the GTWR model.

Before accounting for the spatio-temporal distance function, it might be helpful to discuss some underlying notions on the measurement of 'closeness'. Suppose that the observed data are located in a three-dimensional space–time coordinate system and consider those points close to location i. For instance, if the space–time coordinate system has the same scale effect on distance, we can draw a sphere of certain radius, say r, around a

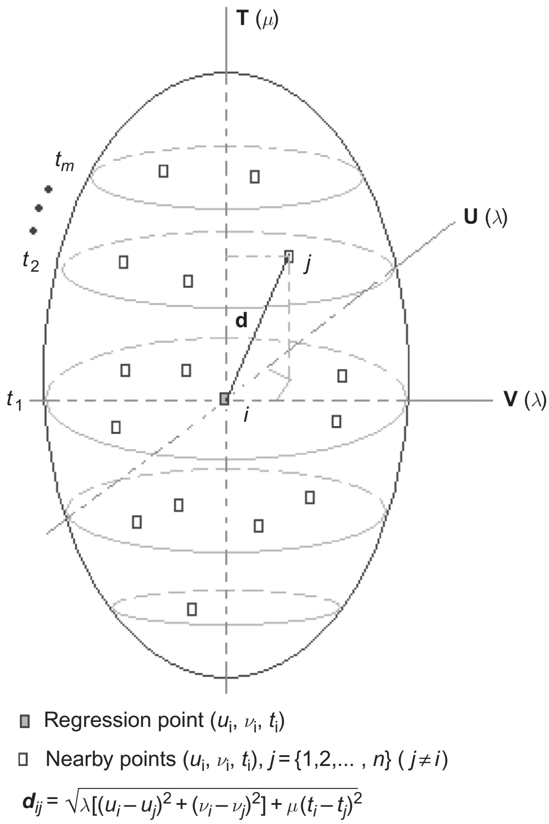

Figure 1. An illustration of spatio-temporal distance.

particular regression point i and calibrate a regression model using ordinary least squares (OLS) only on the observations within this sphere. The $\beta_k(u_i,v_i,t_i)$ obtained can then be considered as an estimate of the associations between the variables in and around i.

However, location and time are usually measured in different units (in our case, location in meters and time in days), thus they have different scale effects. It seems to be more appropriate to use an ellipsoidal coordinate system to measure the 'closeness' between a regression point and its surrounding observed points. Figure 1 shows an example of the proposed spatio-temporal distance. It suggests that a simple and straightforward way of modeling a distance of time is to integrate it directly with spatial distance into the spatio-temporal distance function.

Given a spatial distance d^S and a temporal distance d^T, we can combine them to form a spatio-temporal distance d^{ST} such that

$$d^{ST} = d^S \otimes d^T \tag{9}$$

where $\otimes$ can represent different operators. If the '+' operator is adopted to measure the total spatio-temporal distance d^{ST}, then it is expressed as a linear combination between d^S and d^T.

$$d^{ST} = \lambda d^S + \mu d^T \tag{10}$$

where λ and μ are scale factors to balance the different effects used to measure the spatial and temporal distance in their respective metric systems. Therefore, if the parameters λ and μ are adjusted appropriately, d^{ST} can be used to measure the extent of 'closeness' in a spatio-temporal space. It should be noted that the scale factors are necessary; otherwise, if d^S is much larger than d^T, d^{ST} will be dominated by d^S. This may degrade the temporal effect, and vice versa.

3.2. *Weighting specification for spatio-temporal variations*

Following Equation (10), specifically, if the Euclidean distance and Gaussian distance–decay-based functions are used to construct a spatial–temporal weight matrix, we will have

$$\left(d_{ij}^{ST}\right)^2 = \lambda\left[\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2\right] + \mu\left(t_i - t_j\right)^2 \tag{11}$$

where t_i and t_j are observed times at locations i and j.

$$\begin{aligned}
\alpha_{ij} &= \exp\left\{-\left(\frac{\lambda\left[\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2\right] + \mu\left(t_i - t_j\right)^2}{h_{ST}^2}\right)\right\} \\
&= \exp\left\{-\left(\frac{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2}{h_S^2} + \frac{\left(t_i - t_j\right)^2}{h_T^2}\right)\right\} \\
&= \exp\left\{-\left(\frac{\left(d_{ij}^S\right)^2}{h_S^2} + \frac{\left(d_{ij}^T\right)^2}{h_T^2}\right)\right\} \\
&= \exp\left\{-\frac{\left(d_{ij}^S\right)^2}{h_S^2}\right\} \times \exp\left\{-\frac{\left(d_{ij}^T\right)^2}{h_S^T}\right\} \\
&= \alpha_{ij}^S \times \alpha_{ij}^T
\end{aligned} \tag{12}$$

where $\alpha_{ij}^S = \exp\{-(d_{ij}^S)^2/h_S^2\}$, $\alpha_{ij}^T = \exp\{-(d_{ij}^T)^2/h_S^T\}$, $(d_{ij}^S)^2 = (u_i - u_j)^2 + (v_i - v_j)^2$, $(d_{ij}^T)^2 = (t_i - t_j)^2$, h_{ST} is a parameter of spatio-temporal bandwidth, and $h_S^2 = h_{ST}^2/\lambda$ and $h_T^2 = h_{ST}^2/\mu$ are parameters of the spatial and temporal bandwidths, respectively. As such, the weighting construct of the GTWR model retains a diagonal matrix, whose diagonal elements are multiplied by $\alpha_{ij}^S \cdot \alpha_{ij}^T$ ($1 \leq j \leq n$). Thus, it follows that we can build a spatially weighted matrix W^S and a temporally weighted matrix W^T and then combine them to form a spatio-temporal weight matrix $W^{ST} = W^S \times W^T$. Other weighted matrix combination methods, such as Kronecker products, proposed by Langville and Stewart (2004) for spatio-temporal covariance can also be examined, but these fall outside the scope of this discourse.

Based on Equation (11), after the distances between location i and all observations are computed, the weighting functions can then be constructed. In theory, if there is no temporal variation in the observation data, then the parameter μ can be set to 0 (i.e., $\mu = 0$), which, in turn, degrades the distance calculation to the traditional GWR distance. If, on the other hand, the parameter λ is set to 0 (i.e., $\lambda = 0$), only the temporal distances and the temporal nonstationarity are considered. This will lead to a TWR. In most real cases, however, neither μ nor λ equals zero, and both spatial and temporal distances will be modeled using Equation (11).

Let τ denote the parameter ratio μ/λ and $\lambda \neq 0$. We can rewrite Equation (11) by normalizing the coefficient of d^S,

$$\frac{\left(d_{ij}^{ST}\right)^2}{\lambda} = \left[\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2\right] + \tau\left(t_i - t_j\right)^2 \tag{13}$$

and we have $\tilde{W}^{ST} = (1/\lambda)\,\alpha_{ij}^S \cdot \alpha_{ij}^T = W^{ST}/\lambda$. Because $W(u_i,v_i,t_i)$ multiplied by a constant in Equation (8) will not change the estimation of $\beta_k(u_i,v_i,t_i)$, we can see that only the parameter ratio $\tau = \mu/\lambda$ plays an important role in constructing weights. In fact, the essential effect of τ is to enlarge/reduce the temporal distance effect to match with spatial distance. Without loss of generality, we set $\lambda = 1$ to reduce the number of parameters in practice, and so only μ has to be determined. μ can also be optimized using cross-validation in terms of R^2 or AIC if no a priori knowledge is available.

4. Implementation

To examine the applicability of GTWR, a case study was implemented using the housing prices observed between 2002 and 2004 in the city of Calgary, Canada. The spatial and temporal heterogeneities were first tested using statistical hypotheses, and we then developed various price models using four different approaches and examined their goodness-of-fit. First, the housing data was analyzed using a global OLS without any spatial or temporal considerations. Following this, we performed regressions with the traditional GWR and the proposed TWR and GTWR, respectively, for the housing data. For comparison purposes, we applied McNamara's test to assess the statistical significance of differences between different models.

4.1. Study data

Calgary is located in southern Alberta on the eastern edge of the Rocky Mountain Foothills at the merging of the Bow and Elbow rivers. It is the largest city in Alberta and the fifth largest in Canada.

As portrayed in Figure 2, the study area was in a large part of the northwest and southeast market analysis areas. The northwest area is near the central part of the city, north of the Bow River from the downtown area. The sample contains several established neighborhoods, including Parkdale, St. Andrew Heights, Hounsfield Heights, West Hillhurst, Hillhurst, Rosedale, Sunnyside, and crescent Heights. Most of them are the relatively old residential communities, with some homes dating back to around 1910. Some of the southeast sample area is still under development. The neighborhoods are Cougar Ridge, West Springs, Patterson Heights, Couch Hill, Aspen Woods, Strathcona Park, Christie Park, Springbank Hill, Signal Hill, and Discovery Ridge, which are all fairly new residential communities. Most of these were developed after 1980.

A set of 5000 observations were available: the data (1) included full information on age, living area, land area, garage type, condition, and other variables; (2) were of the most common construction types; (3) were of the most common occupancy type; and (4) were of the most common zoning types. According to Sirmans *et al.* (2005), using the observed price is generally considered to be more suitable for minimizing the bias than other measures such

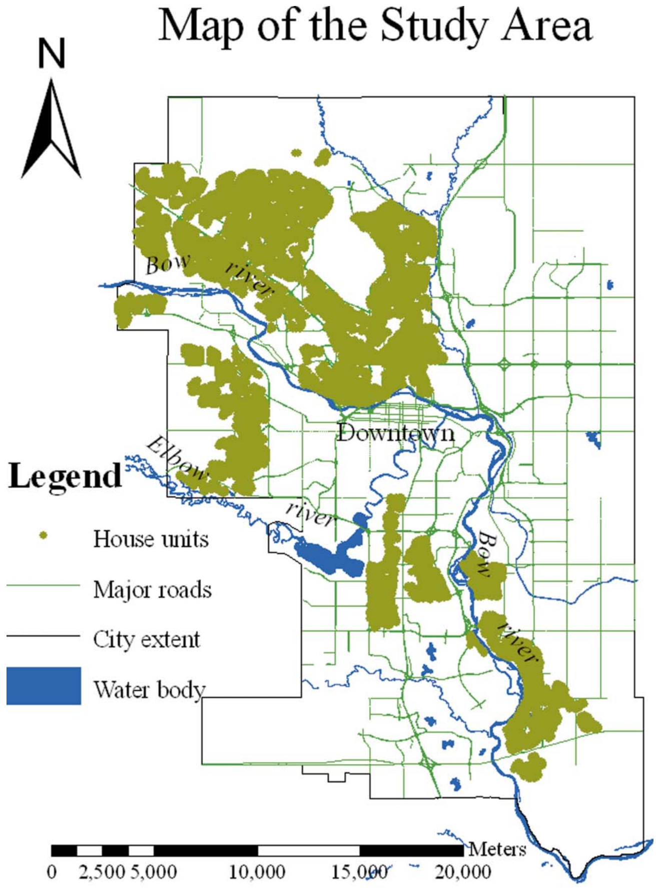

Figure 2. A summary map of the study area.

as an owner's self-assessment. A recent selling price was taken as the dependent variable, standing as a proxy for the market value of the house. The explanatory variables comprised three groups, which included a total of 33 variables. These were conflated into 11 variables: living area, land area, quality, structure type, renovation, garage, condition, green space, traffic condition, view, and age. Averages for the age of the houses, land area, and living area at each house unit (an entire house) were calculated using ArcGIS. Discrete variables, such as the number of houses without a garage and the number of houses with a view in each unit, were also counted using location-related joins, and then the percentage of properties without garages and the percentage of properties with a view in each unit were calculated. The percentage of green and open space area in each unit was also computed.

4.2. Spatial and temporal nonstationarity diagnosis

For the type of data in the sample, an analyst should first consider whether the GWR-based models (TWR, GWR, and GTWR) can describe the data set significantly better than an OLS model. In other words, we need to assess whether there is significant spatial and/or temporal nonstationarity over the study area before applying GTWR models. In earlier research, Fotheringham *et al.* (2002) assessed the degree of nonstationarity visually and constructed experimental distributions using Monte Carlo procedures. However, this technique is computationally demanding. Brunsdon *et al.* (1999) suggest a test of the null hypothesis that the functions $\beta_k(u,v)$ are constants for all points (u,v) in the study area. If there is no evidence to reject this hypothesis, it suggests that an ordinary global regression model is an adequate descriptor of the data, i.e.

$$H_0 : \frac{\partial \beta_i}{\partial u} \equiv \frac{\partial \beta_i}{\partial v} \equiv 0 \tag{14}$$

against

$$H_1 : \frac{\partial \beta_i}{\partial u} \neq \frac{\partial \beta_i}{\partial v} \neq 0 \tag{15}$$

or more specifically, H_0 will be compared against a subset of H_1 corresponding to a GWR type of estimation of the $\beta(u,v)$'s. This basically states that if both models are expressed in the hat matrix form, for normally distributed y, then the expression

$$F = \left[\frac{(y^T R_0 y) - (y^T R_1 y)}{v}\right]\left[\frac{(y^T R_1 y)}{\delta}\right]^{-1} \tag{16}$$

where $R_z = (I - S_z)^T (I - S_z)$, $z \in \{0,1\}$, $S_1 = X(X^T W X)^{-1} X^T W$, $S_0 = X(X^T X)^{-1} X^T$, $v = Tr(R_0 - R_1)$, and $\delta = Tr(R_1)$ have an approximate F distribution with degrees of freedom given by $(v^2/v', \delta^2/\delta')$, where $v' = Tr[(R_0 - R_1)^2]$ and $\delta' = Tr\left(R_1^2\right)$. This F-test is based on an analysis of variance and uses generalized degrees of freedom to compare with the improved sum of squares accounted for by the GWR estimates as compared with the global OLS estimates. This suggests that GWR/OLS comparisons can be expressed in the form of an analysis of variance (ANOVA) table, with the residual mean squares (MS) for both GWR and OLS being compared.

The results of ANOVA tests on the observations in Calgary are shown in Table 1. In this table, the first column lists the residual sum of squares (RSS) (Brunsdon *et al.* 1999) of OLS, GWR, TWR, GTWR, and the difference between OLS and GWR-based models. The second column gives the degrees of freedom for each of these models. The third column, MS, gives the results of dividing the sums of squares by their respective degrees of freedom. The last two columns show the pseudo-F statistic and the p-value. We can note the reduction in RSS when GWR-based approaches were used. It can be seen from the F-test values in Table 1 that the statistics indicate that there is significant spatial and temporal nonstationarity over the study

Table 1. ANOVA comparison between GWR and OLS models.

Source of variation	RSS	DF	MS	F-test	p-value
OLS residuals	95.09	12	7.923		
TWR residuals	88.57	4947.9	0.018	9.08	0.00
GWR residuals	44.27	4367.2	0.010	8.10	0.00
GTWR residuals	36.99	4080.2	0.009	2.77	0.00
TWR/OLS improvement	6.52	40.1	0.163		
GWR/OLS improvement	50.81	620.8	0.082		
GTWR/OLS improvement	58.09	907.8	0.064		
GTWR/GWR Improvement	7.27	287.0	0.025		

RSS, residual sum of squares; DF, degree of freedom; MS, mean square.

area. Therefore, it is more appropriate to model the specified data set with GWR-based models (TWR, GWR, and GTWR). Moreover, it can be found that modeling spatial nonstationarity with the traditional GWR is inadequate for our data set, and we posit that a more accurate model can be established if temporal variation information were added to GWR.

4.3. Optimal parameter selection

As pointed out earlier in Section 3.2, the measurement units for location and time are usually different. In our case, Euclidean distance was quoted in meters and time in days. These units need to be harmonized in calculating the space–time distance before constructing spatio-temporal weighting matrices. We introduced a parameter τ to balance or harmonize the different spatial and temporal units. Therefore, one important issue is to optimize τ before implementing the GTWR model. In this article, we have used a validation procedure to obtain an appropriate parameter value in terms of goodness-of-fit.

Figure 3 provides the details of parameter selection. It is evident that the explanatory ability of the GTWR model with an inefficient τ parameter could be worse than that of GWR. For instance, if $\tau < 10$, the R^2 of GTWR is less than that of GWR, where $R^2 = 0.8856$. In the data set used to generate Figure 3, the optimal parameter of τ was found to be 35.

Another issue is the choice of spatio-temporal bandwidth. The bandwidth in the GTWR model determines the rate at which the regression weights decay around a given point (u,v,t). It is important to choose a suitable bandwidth to obtain reliable estimates of the spatio-temporal variations in the coefficients. If the bandwidth is small, weights decay quickly with distance, the values of the regression coefficients change rapidly over space, and the standard error of $\hat{\beta}(u,v,t)$ increases. On the other hand, larger bandwidths produce smoother results, but bias increases. An effective way to trade off standard error and bias is to verify by cross-validation. This argument has been supported by a large body of research (see, e.g., Brunsdon *et al.* 1999, Fotheringham *et al.* 2002).

4.4. Model comparison criterion

The main criteria adopted for comparing the three different local models (TWR, GWR, and GTWR) are the conventional goodness-of-fit measures using R^2. However, to judge whether there is statistical significance in the differences between the accuracies achieved by the three models, McNamara's test was performed, which is based on the standardized normal test statistic (Foody 2004):

$$Z_{12} = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}} \tag{17}$$

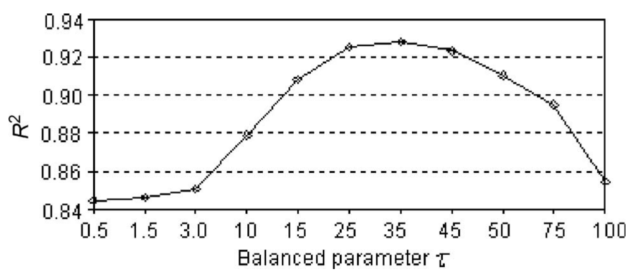

Figure 3. The parameter selection for the GTWR model.

where Z_{12} measures the pair-wise statistical significance of the difference between the accuracies of the first and second prediction models. f_{12} stands for the number of samples classified correctly and wrongly by the first and second models, respectively. Accordingly, f_{12} and f_{21} are the counts of classified samples on which the first and second models disagree. A lower prediction error (higher accuracy) is identified by the sign on Z_{12}. A negative sign indicates that the results from f_{12} are more accurate than those from model f_{21}. At the commonly used 5% level of significance, the difference of accuracies between the first and the second models is considered statistically significant if $|Z_{12}| > 1.96$.

5. Results comparison and analysis

5.1. Results of the global model

Using the Calgary house price data set, an OLS regression was first carried out and the results are reported in Table 2. The Durbin–Watson test indicated that the model is statistically significant and 76.31% of the variation in the house values can be explained by the model according to R^2.

These results also indicate that the assessed house values in Calgary can be modeled by the selected housing structural attributes and the neighborhood environment conditions. Therefore, the hypothesized relationships between the structural neighborhood attributes and the house values are supported by the data. Indeed, all determinants except one (i.e., garage) are statistically significant at the 95% confidence level according to their *t*-probabilities. In particular, living area, land area, quality, renovation, and house age are positively correlated with house values, whereas the traffic condition index is negatively correlated with house values. The larger the living area, the higher the price, and the presence of high quality can add value to the house significantly. Therefore, living area and quality are the most significant variables, with the *t*-values of 33.9 and 40.0, respectively, which show these variables have the strongest relationship with housing price. Interestingly, it is also found that age is a positive factor, which means the older the house, the more valuable it is. A possible reason is that in the samples, older houses are closer to the city centre and hence

Table 2. Hedonic model (OLS) parameter estimate summaries.

Variable	Coefficient	*t*-statistic	*t*-probability	95% confidence interval	
Intercept	11.746	436.58	0.000000	11.693	11.799
Living area	1.2082	33.997	0.000000	1.1385	1.2779
Land area	0.3202	11.251	0.000000	0.2644	0.3759
Quality	0.7350	40.052	0.000000	0.6990	0.7710
Structure type	–0.0982	–13.044	0.000000	–0.1130	–0.0835
Renovation	0.1969	16.648	0.000000	0.1737	0.2200
Garage	–0.0268	–0.5169	0.605235*	–0.1287	0.0750
Condition	0.0609	9.0784	0.000000	0.0478	0.0741
Green space	0.0143	2.0858	0.037045	0.0009	0.0278
Traffic condition	–0.0844	–9.8716	0.000000	–0.1017	–0.0677
View	0.1626	10.131	0.000000	0.1311	0.1940
Age	0.1878	12.068	0.000000	0.1573	0.2183
Diagnostic information					
R^2	0.7631				
Residual standard error	0.1379				
Residual sum of squares	95.09				
AIC	–5595.6				

have higher prices. Intuitively, one should expect a newer house to be more valuable than an older house, and this points to a general weakness in global models mentioned earlier.

5.2. Results of the TWR, GWR, and GTWR models

Using the same data set, GWR-based models (i.e., TWR, GWR, and GTWR) were also tested, and the results are reported in Tables 3 and 4. Because the output of local parameter estimates from TWR, GWR, and GTWR would be voluminous, Tables 3 and 4 only provide a five-column summary of the distribution of each parameter to indicate the extent of its variability. The signs of all the parameters between the lower quartile (LQ) and the upper quartile (UQ) in GWR are the same as GTWR, and the magnitude of all the parameters in the global models are between the minimum and the maximum values of those in TWR, GWR, or GTWR.

Tables 3 and 4 provide a detailed statistical comparison. It should be noted that the percentage of explanation of variance has increased from 76.31% in the global OLS model to 77.94% in TWR, 88.97% in GWR, and 92.82% in GTWR. Tables 3 and 4 reveal that the GTWR model is the best, even if the differences in degrees of freedom with the reduction in AIC (from −5595.6 for the global model to −5886.9 for TWR, −8693.9 for GWR, and −8850.4 for GTWR) are taken into account. By comparing the residual standard error and RSS, the decreased value further indicates that GTWR gives a better fit of data than the TWR, GWR, and global models. We posit that this is because GTWR can handle both spatial and temporal heterogeneities. Moreover, Table 3 shows that the GWR model achieved a better goodness-of-fit than that of TWR model in terms of R^2. A possible reason is that the experimental data only covered a short period (3 years), which indicates that the temporal nonstationary effect is less significant than that of spatial nonstationarity.

Given that the GWR-based models (TWR, GWR, and GTWR) are statistically valid, another issue is to decide which parameters vary significantly across the study area. The Monte Carlo significance test is usually used to check whether the parameter estimates

Table 3. TWR and GWR parameter estimate summaries.

	TWR (bandwidth = 0.7711)					GWR (bandwidth = 0.3881)				
Parameter	Min	LQ	Med	UQ	Max	Min	LQ	Med	UQ	Max
Intercept	11.64	11.69	11.74	11.79	11.81	9.02	9.47	9.53	11.62	12.70
Living area	1.13	1.17	1.20	1.23	1.24	0.48	1.08	1.23	1.36	2.37
Land area	0.21	0.28	0.34	0.39	0.44	0.14	0.35	0.46	0.65	1.57
Quality	0.65	0.69	0.72	0.76	0.77	−0.22	0.49	0.59	0.76	2.81
Structure type	−0.12	−0.11	−0.09	−0.08	−0.08	−0.44	−0.18	−0.13	−0.05	0.09
Renovation	0.15	0.17	0.19	0.21	0.21	−0.11	0.12	0.16	0.21	0.57
Garage	−0.07	−0.04	−0.02	0.01	0.06	−0.75	0.14	4.71	4.76	4.91
Condition	0.05	0.05	0.06	0.07	0.07	−0.29	0.04	0.05	0.07	0.16
Green space	0.01	0.01	0.02	0.02	0.03	−0.15	0.01	0.02	0.04	0.34
Traffic condition	−0.10	−0.09	−0.09	−0.08	−0.07	−0.31	−0.07	−0.05	−0.03	0.16
View	0.11	0.15	0.16	0.19	0.20	−0.15	0.09	0.18	0.24	0.44
Age	0.17	0.18	0.20	0.22	0.25	−0.21	0.11	0.22	0.48	0.94
Diagnostic information										
R^2			0.7794				0.8897			
Residual standard error			0.1873				0.0941			
Residual sum of squares			88.57				44.27			
AIC			−5886.9				−8693.9			

Table 4. GTWR parameter estimate summaries (bandwidth = 0.4502).

Parameter	Minimum	Lower quartile	Median	Upper quartile	Maximum
Intercept	8.62	9.48	9.60	11.81	12.43
Living area	−0.41	1.06	1.23	1.33	2.83
Land area	−0.18	0.37	0.49	0.67	1.39
Quality	−1.63	0.49	0.57	0.75	3.62
Structure type	−0.04	−0.18	−0.13	−0.06	0.65
Renovation	−0.20	0.11	0.15	0.19	0.78
Garage	−0.76	−0.01	4.68	4.74	4.84
Condition	−0.35	0.03	0.05	0.08	0.24
Green space	−0.39	0.01	0.02	0.04	0.62
Traffic condition	−0.33	−0.08	−0.06	−0.03	0.21
View	−0.19	0.08	0.20	0.25	1.54
Age	−0.37	0.13	0.21	0.51	1.03
Diagnostic information					
R^2	0.9282				
Residual standard error	0.0860				
Residual sum of squares	36.99				
AIC	−8850.4				

exhibit significant spatial variation (Fotheringham *et al.* 1998). However, this approach has the disadvantage of being computer-intensive and computation demanding. Hence, we adopted the alternative statistical test proposed by Leung *et al.* (2000) due to its simplicity and efficiency. They test the following hypothesis using the F statistic.

$$H_0 : \beta_{1k} = \beta_{2k} = \cdots = \beta_{nk}, \text{ for a given } k$$
$$H_1 : \text{not all } \beta_{ik}(i = 1, 2, \ldots, n) \text{ are equal}$$

By constructing the statistical value $V_k^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\beta}_{ik} - \frac{1}{n}\sum_{i=1}^{n}\hat{\beta}_{ik}\right)^2$, which reflects the spatial variation of the given set of the parameters with the sample variance of the estimated values of $\beta_{ik}(i = 1,2, \ldots, n)$, an approximated F distribution value is used to decide which hypothesis is appropriate (Leung *et al.* 2000). The large value of F supports the alternative hypothesis H_1. It should be noted that the p-value is now widely accepted in applied statistics because of its ease of use. For the proposed test, the p-value of the test statistic is the probability that the statistic could have been more extreme than its observed value under the null hypothesis. A large p-value supports the null hypothesis, whereas a small p-value supports the alternative hypothesis. A test can be carried out by comparing the p-value with a given significance level, for example, 0.05. If the p-value is less than 0.05, the null hypothesis is rejected; otherwise, it is accepted.

Table 5 lists the F-statistic value of each variable and its corresponding p-value. Those statistically significant values at the 5% level are marked with a asterisk '*'. It can be found that intercept, living area, and land area have significant spatial and temporal variation in the local parameter estimates for all three models. However, the temporal variation of quality, traffic condition, and age variables are not significant in the TWR model, whereas they show significant spatial variation in the GWR and GTWR models. We can also infer from Table 5 that the GWR model fits the data set better than the TWR model, because the data exhibits more spatial variation in the variables than those of the temporal dimension. Quantitative support for this deduction is demonstrated by the R^2 statistic in Table 3. Moreover, the GTWR model integrates both potential temporal and spatial variation of variables, thus the

Table 5. Nonstationarity of parameters in the TWR, GWR, and GTWR models.

Parameter	TWR		GWR		GTWR	
	F value	*p*-value	*F* value	*p*-value	*F* value	*p*-value
Intercept	4.01	0.0000*	8.28	0.0000*	6.59	0.0000*
Living area	2.49	0.0365*	25.45	0.0000*	34.70	0.0000*
Land area	3.11	0.0140*	21.28	0.0000*	14.29	0.0000*
Quality	1.92	0.1044	3.95	0.0003*	8.21	0.0000*
Structure type	0.03	0.9978	1.04	0.3757	1.17	0.1483
Renovation	0.73	0.5695	2.03	0.0755	0.94	0.5045
Garage	0.41	0.8379	0.83	0.7315	0.84	0.6457
Condition	0.32	0.8606	0.71	0.8621	0.35	0.9998
Green space	0.04	0.9977	0.95	0.4570	1.29	0.0655
Traffic condition	0.03	0.9979	2.60	0.0006*	1.55	0.0271*
View	0.32	0.8603	1.25	0.2316	1.19	0.1048
Age	0.52	0.7196	134.37	0.0000*	43.56	0.0000*

number of nonstationary variables in the GTWR model should not be fewer than those of the TWR and GWR models.

One important characteristic of the GWR-based technique is that the local parameter estimates that denote local relationships are mappable and thus allow for visual analysis. Taking the coefficients of 'living area' as an example, we can group them into several intervals and color each interval to visualize the spatial variation patterns of this variable.

The spatial distributions of the parameter estimates for 'living area' of TWR, GWR, and GTWR are shown in Figure 3. It can be seen that for the TWR model, there is no significant spatial variation (i.e., [1.13, 1.24]) over time (temporal distance leads to the spatial variation in this case). This is somewhat trite, because TWR only models temporal heterogeneity, which indicates that the spatial variation of this coefficient is not obvious. Moreover, the spatial variation of 'living area' in GWR and GTWR share analogous distributions, except that the spatial variation in the GTWR model portrays heterogeneity in more detail. It can be inferred that the spatio-temporal nonstationarity of the GTWR model is dominated by the spatial effect for the test data set.

It can also be seen from Figure 4 that spatial variation of the parameter 'living area' in GTWR shows two major trends: living area varies from high in the east to low in the west and from low in the outer zones of the study area to high in the inner zone. This suggests that the living area of a house had the most important influence on housing prices in the eastern part of the city.

5.3. *Significant difference comparison of GWR-based models*

As demonstrated above, all GWR-based models (TWR, GWR, and GTWR) show significant improvements over the OLS model in terms of R^2 and AIC measures. However, it is still necessary to investigate whether the GTWR model performs significantly better than the GWR and TWR models from a statistical viewpoint.

The McNamara's test was implemented to test the significant difference between the TWR, GWR, and GTWR models. It was assumed that if the difference between the predicted price and the actual price were no more than a predefined percentage threshold, the model could be considered 'correct'. The predicted accuracies of each model can then be commutated and Z values obtained by Equation (17). Two percentage thresholds $\varepsilon = 0.1\%$ and

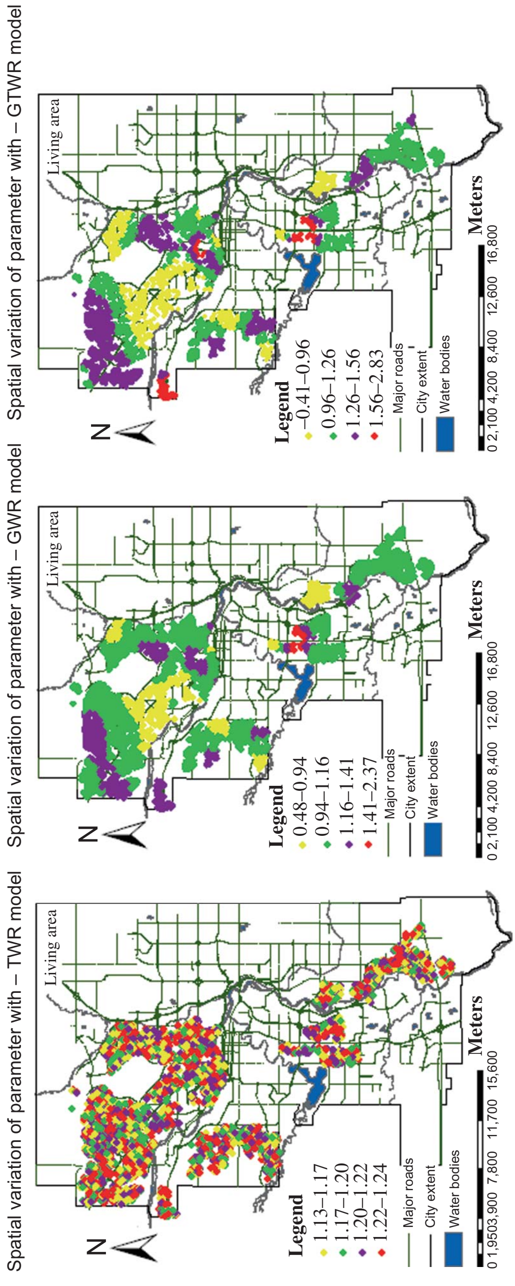

Figure 4. Spatial variation of the living area coefficient.

Table 6. Significance comparison for GWR-based models.

	Error tolerance = 0.1%			Error tolerance = 0.5%		
Models comparison	TWR	GWR	GTWR	TWR	GWR	GTWR
OLS	NA	−5.33	−7.21	NA	−13.67	−19.80
GWR	–	NA	−2.30	–	NA	−9.35
GTWR	–	–	NA	–	–	NA

$\varepsilon = 0.5\%$ were specified, and the results are listed in Table 6. They clearly indicate (negative value) that the GTWR model performed better than GWR. The Z values between TWR and GWR within 0.1 and 0.5% error bounds are −5.33 and −13.67, respectively, indicating that GWR substantially outperformed TWR. Also, the Z values between GTWR and GWR within 0.1 and 0.5% error bounds are −2.30 and −9.35, respectively, both of them less than −1.96. These results demonstrate a significant difference between the TWR, GWR, and TWGR models at the 95% confidence level. It is clear from these comparisons that GTWR outperforms both GWR and TWR in the model accuracy for the sample data.

6. Conclusions

As most previous studies have demonstrated that the hypothesis of a stationary housing market is unlikely to be supported, this study took a nonstationary approach for analyzing housing prices in the city of Calgary. Our analysis reveals that spatio-temporal heterogeneity prevails in the real estate data that evolve over both time and space in the sample area and that traditional GWR for spatial nonstationarity only is, therefore, inadequate to model such data.

We extended the GWR model to incorporate time to deal with both spatial and temporal heteroscedasticity simultaneously. GTWR achieved a better modeling accuracy than both the global OLS model with no spatio-temporal nonstationarity incorporated and the GWR model, which deals with spatial nonstationarity only in our sample data. Compared with the global OLS model, TWR and GWR increased the R^2 values from 0.763 to 0.779 and 0.889, respectively, and GTWR yielded a considerably higher R^2 of 0.928. The RSS for the GTWR also yielded a 46.4% improvement over OLS and a 15.6% improvement over GWR. Statistical tests showed that there was a significant difference between GTWR, GWR, and TWR, and therefore we conclude that it is meaningful to incorporate temporal nonstationarity into a GWR model, and GTWR can provide an additional useful methodology for computer-assisted mass estimation of real property prices.

Some limitations still remain in our study and further work is required. For example, only three years of temporal information was available. The inadequacy in temporal heterogeneity can be expected to degrade the model performance of TWG and GTWR. How GTWR would perform if applied to the data covering a longer period merits further investigation. In addition, we have used a simple weighting system based on a linear combination of temporal and spatial distances. More efficient weighting schemes still need to be designed to yield better results. We have also not attempted to duplicate the procedures used by the City of Calgary's appraisal department, where the city is divided into a number of appraisal areas and an OLS model is developed for each area. Further tests of the GTWR methodology, using a number of different models, against the localized OLS models, which perhaps incorporate response surface analysis and fuzzy boundaries, also merit investigation.

Acknowledgments

This research is funded by the Hong Kong Research Grants Council (RGC) under CERG project no. CUHK 444107 and the Natural Sciences and Engineering Research Council (NSERC) of Canada under discovery grant no. 312166-05. Their support is gratefully acknowledged. We also thank the two anonymous reviewers for their insightful comments that have been very helpful in improving this article.

References

Anselin, L., 1998. GIS research infrastructure for spatial analysis of real estate markets. *Journal of Housing Research*, 9, 113–133.

Brunsdon, C., Fotheringham, A.S., and Charlton, M., 1996. Geographically weighted regression: a method for exploring spatial non-stationarity. *Geographical Analysis*, 28, 281–298.

Brunsdon, C., Fotheringham, A.S., and Charlton, M., 1999. Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Sciences*, 39, 3, 497–524.

Can, A., 1992. Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, 22, 453–474.

Can, A. and Megbolugbe, I., 1997. Spatial dependence and house price index construction. *Journal of Real Estate Finance and Economics*, 14, 203–222.

Dombrow, J., Knight, J.R., and Sirmans, C.F., 1997. Aggregation bias in repeat sales indices. *Journal of Real Estate Finance and Economics*, 14, 75–88.

Dubin, R.A., 1992. Spatial autocorrelation and neighborhood quality. *Regional Science and Urban Economics*, 22, 432–452.

Eckert, J.K., ed., 1990. *Property appraisal and assessment administration*. Chicago, IL: International Association of Assessing Officers.

Foody, G.M., 2004. Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. *Photogrammetric Engineering and Remote Sensing*, 70, 627–633.

Fotheringham, A.S., Brunsdon, C., and Charlton, M., 2002. *Geographically weighted regression*. Chichester, UK: John Wiley and Sons.

Fotheringham, A.S., Charlton, M.E., and Brunsdon, C., 1996. The geography of parameter space: an investigation of spatial non-stationarity. *International Journal of Geographical Information Science*, 10, 605–627.

Fotheringham, A.S., Charlton, M.E., and Brunsdon, C., 1998. Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, 30, 1905–1927.

Gelfand, A.E., *et al*., 2001. The dynamics of location in home prices. *Journal of Real Estate Finance and Economics*, 29, 149–166.

Gelfand, A.E., *et al*., 2003. Spatial modeling with spatially varying coefficients processes. *Journal of the American Statistical Association*, 98, 387–396.

Goodman, A.C. and Thibodeau, T.G., 1998. Housing market segmentation. *Journal of Housing Economics*, 7, 121–143.

Huang, B., Zhang, L., and Wu, B., 2009. Spatio-temporal analysis of rural-urban land conversion. *International Journal of Geographical Information Science*, 23, 379–398.

Hurvich, C.M., Simonoff, J.S., and Tsai, C.L., 1998. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society* (Series B), 60, 271–293.

Langville, A.N. and Steward, W.J., 2004. The Kronecker product and stochastic automata networks. *Journal of computational and applied mathematics*, 164, 429–447.

Leung, Y., Mei, C.L., and Zhang, W.X., 2000. Statistical tests for spatial nonstationarity based on the geographically weighted regression model. *Environment and Planning A*, 32, 9–32

McMillen, D.P., 1996. One hundred fifty years of land values in Chicago: a nonparametric approach. *Journal of Urban Economics*, 40, 100–124.

McMillen, D.P. and McDonald, J.F., 1997. A nonparametric analysis of employment density in a polycentric city. *Journal of Regional Science*, 37, 591–612.

Pace, R.K., Barry, R., and Sirmans, C.F. 1998. Spatial statistics and real estate. *Journal of Real Estate Finance and Economics*, 17, 5–13.

Paez, A., Uchida, T., and Miyamoto, K., 2002. A general framework for estimation and inference of geographically weighted regression models: location-specific kernel bandwidths and a test for local heterogeneity. *Environmental and Planning A*, 34, 733–754.

Pavlov, A., 2000. Space varying regression coefficients: a semi-parametric approach applied to real estate markets. *Real Estate Economics*, 28, 249–283.

Sirmans, G., Macpherson, D., and Zietz, A., 2005. The composition of hedonic pricing models. *Journal of real Estate Literature*, 13, 1, 3–41.

Wang, P., 2006. *Exploring spatial effects on housing price: the case study of the city of Calgary*. Master dissertation. University of Calgary, Canada.

Yu, D.-L., 2006. Spatially varying development mechanisms in the Greater Beijing area: a geographically weighted regression investigation. *The Annals of Regional Science*, 40, 1, 173–190.