International Journal of Geographical Information Science, 2014
Vol. 28, No. 5, 1186–1204, http://dx.doi.org/10.1080/13658816.2013.878463

Taylor & Francis
Taylor & Francis Group

A geographically and temporally weighted autoregressive model with application to housing prices

Bo Wu[a], Rongrong Li[b] and Bo Huang[b,c,d,e]*

[a]*Key Lab of Spatial Data Mining and Information Sharing of Ministry of Education, Fuzhou University, Fuzhou, People's Republic of China;* [b]*Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong;* [c]*Department of Geography and Resource Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong;* [d]*Yuen Yuen Research Centre for Satellite Remote Sensing, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong;* [e]*Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, People's Republic of China*

(*Received 18 May 2013; accepted 11 December 2013*)

Spatiotemporal autocorrelation and nonstationarity are two important issues in the modeling of geographical data. Built upon the geographically weighted regression (GWR) model and the geographically and temporally weighted regression (GTWR) model, this article develops a geographically and temporally weighted autoregressive model (GTWAR) to account for both nonstationary and auto-correlated effects simultaneously and formulates a two-stage least squares framework to estimate this model. Compared with the maximum likelihood estimation method, the proposed algorithm that does not require a prespecified distribution can effectively reduce the computation complexity. To demonstrate the efficacy of our model and algorithm, a case study on housing prices in the city of Shenzhen, China, from year 2004 to 2008 is carried out. The results demonstrate that there are substantial benefits in modeling both spatiotemporal nonstationarity and autocorrelation effects simultaneously on housing prices in terms of R^2 and Akaike Information Criterion (AIC). The proposed model reduces the absolute errors by 31.8% and 67.7% relative to the GTWR and GWR models, respectively, in the Shenzhen data set. Moreover, the GTWAR model improves the goodness-of-fit of the ordinary least squares model and the GTWR model from 0.617 and 0.875 to 0.914 in terms of R^2. The AIC test corroborates that the improvements made by GTWAR over the GWR and the GTWR models are statistically significant.

Keywords: GTWAR; two-stage least squares estimation; spatiotemporal nonstationarity; spatiotemporal autocorrelation; housing price

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

1. Introduction

Statistical modeling of quantitative relationships between a response variable and explanatory variables in geographical information science involves two important problems when the sample data has a spatial component (Anselin 1988, Fotheringham *et al.* 1996, LeSage 1999). One is that spatial or temporal autocorrelation exists between the observations; the other is that spatial or temporal nonstationarity occurs in the relationships we are modeling. Traditional models with multivariable regression and least square algorithms have by and large ignored these two issues that violate the traditional Gauss–Markov assumptions used in regression modeling. Spatial autocorrelation violates the Gauss–Markov assumption that

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

© 2014 Taylor & Francis

explanatory variables are fixed in repeated sampling, while spatial heterogeneity violates the Gauss–Markov assumption that a single linear relationship exists across the sample data observations. Therefore, alternative estimation procedures are attempted to successfully model these problems and draw appropriate inferences in the past decades (Anselin 1988).

A broad range of techniques have been developed to partially or fully model the aforementioned problems in the literature. An early effort was by Ord (1975) who proposed the use of autoregressive and moving average terms in regression models to account for spatial correspondences in the response variable and the residuals, respectively. Similarly, Cressie (1993) addressed spatial effects by modeling the residual variance–covariance matrix directly. To eliminate dependency in the residuals, Basu and Thibodeau (1998) argued that the remaining spatial effects were introduced into the error structure, and hence, the inverse of the covariance matrix can be modeled using simultaneous autoregressive SAR (Pace and Gilley 1998) or conditionally autoregressive CAR (Gelfand *et al.* 1998). Moreover, Pace *et al.* (1998) proposed a spatiotemporal autoregressive model to integrate the temporal correlations between observations and found it powerful in a residential real estate context. This methodology has been further improved by incorporating second-order spatial dependence effects within a Bayesian framework (Tu *et al.* 2004). Although these techniques have made significant contributions to the consideration of spatial and temporal process in regression modeling, they all assumed that there exists a universal relationship across the study region. This assumption of stationary or structural stability over time and space is generally unrealistic, as parameters tend to vary over the study area. As a result, there has been a recent shift in emphasis in spatial statistics away from globally fitted models to local ones where the aim is to identify spatial variations in relationship. One of the well-known models is nonparametric local linear regression in non-monocentric city models introduced by McMillen (1996) and McMillen and McDonald (1997). Notably, Brunsdon *et al.* (1996), Fotheringham *et al.* (1996), and Fotheringham *et al.* (2002) proposed geographically weighted regression (GWR), which allows the exploration of the variation of the parameters and the testing of the significance of space variation, and therefore received considerable attention. To model temporal heterogeneous effects, Huang *et al.* (2010) further developed a geographically and temporally weighted regression (GTWR) to deal with both spatial and temporal nonstationarity simultaneously by incorporating temporal effects into the standard GWR model.

Although spatial heterogeneity and dependence are often related in the context of modeling (LeSage 1999), few studies have attempted to jointly construct and estimate spatiotemporal autocorrelation and heterogeneity simultaneously, although there is a clear need to do so. On the one hand, much evidence shows that spatial heterogeneity exists in the presence of spatial autocorrelation; hence, omitting spatial heterogeneity can produce heteroskedastic disturbances. On the other hand, an inadequate model that fails to capture spatial heterogeneity will result in residuals that exhibit spatial dependence. Therefore, it is not a trivial matter to build an integrated model with spatiotemporal auto-correlated and heterogeneous effects (LeSage 1999). To model both spatial nonstationarity and spatial autocorrelation in a complicated process, Brunsdon *et al.* (1998) proposed a geographically weighted regressor with spatially lagged objective variable mode (GWRSL) to exploit the spatial variations. But their model neglects temporal information, though temporal information between observations might play an important role. Another important practical problem is that the maximum likelihood estimation suggested for these models is often computationally very challenging when the sample size is large. Furthermore, the maximum likelihood procedure requires distributional assumptions that the researcher may not know in advance.

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

This study attempts to extend the GTWR model by integrating GTWR with a local autoregressive model and provides an efficient parameter estimation technique for the proposed model with a computationally simple two-stage least squares procedure on the basis of the algorithm proposed by Kelejian and Prucha (1998). Three contributions to the literature are made possible. First, we improve the GTWR model which the authors developed previously. Second, we propose a novel geographically and temporally weighted autoregressive model (GTWAR) to fully capture spatiotemporal variations. Third, we formulate a computationally simple two-stage least squares estimate and inference framework of the GTWAR model for consistent parameter estimation.

The remainder of this article is organized as follows. In Section 2, we briefly summarize the widely used GWR model, followed by an improved GTWR model based on our previous work, and then we integrate the improved GTWR (IGTWR) with the autoregressive (AR) model to produce the GTWAR model. Section 3 offers the detailed estimation technique for the GTWAR model with a two-stage least squares procedure. In Section 4, a case study of housing prices in the city of Shenzhen, China, is reported. Finally, the study is summarized, and conclusions are drawn.

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

2. GTWAR model

2.1. GWR model and parameter estimation

GWR is a relatively simple, but effective, technique that extends the traditional regression framework for exploring spatial nonstationarity. It allows different relationships to exist at different points in space, such that local rather than global parameters can be estimated. 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 $\beta_k(u_i, v_i)$ represents a set of values of parameters at point i. The coefficients in the GWR model are specific to location i rather than assumed to be constant. This model allows the parameter estimates to vary across space, and therefore, spatial nonstationarity can be captured. The parameters in the GWR model Equation (1) are usually calibrated using the local weighted least squares approach. The estimation of parameters $\beta_k(u_i, v_i)$ is given by Equation (2)

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

where the weighting matrix $W(u_i, v_i)$ is an $n \times n$ matrix whose off-diagonal elements are zero and whose diagonal elements denote the geographical weighting of observation data for observation i, that is, $W(u_i, v_i) = \mathrm{diag}(W_{i1}, W_{i2}, \ldots, W_{in})$, and the weight matrix has to be computed for each point i at which the parameters are estimated.

To estimate parameters in the GWR Equation (1), it is important to choose the criterion to decide on the weighting matrix. The weights are usually obtained through a spatial kernel function. Two types of spatial kernels are often used, that is, fixed and adaptive kernels (Paez *et al.* 2002). In a fixed kernel function, an optimum spatial kernel (bandwidth) will be calculated and applied over the study area. The most commonly used weighting function is the Gaussian function

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

where h is a nonnegative parameter known as bandwidth, which produces a decay of influence with distance d_{ij} between locations i and j.

On the other hand, the adaptive kernel function seeks a certain number of nearest neighbors to adapt the spatial kernel to ensure a constant size of local samples. In such a way, the weighting functions adapt themselves in size to variations, such that the kernels have larger bandwidths where the data points are sparsely distributed and have smaller bandwidths where the data are plentiful. The commonly used adaptive weighting is the bi-square function:

$$W_{ij} = \begin{cases} \left[1 - (d_{ij}/h_i)^2\right]^2, & \text{if } d_{ij} < h_i \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where h_i stands for different bandwidths, which express the number or proportion of observations to consider in the estimation of regression at location i.

To calibrate a GWR model, a cross-validation (CV) approach by minimizing residual sum of squares between predicted values and actual observations is suggested. 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

$$\mathrm{CV}(h) = \sum_i (y_i - \hat{y}_i(h))^2 \tag{5}$$

In practice, plotting CV against the parameter h will provide guidance on selecting an appropriate value of the parameter or it can be obtained automatically with an optimization technique by minimizing Equation (5) in terms of the goodness-of-fit statistics or the corrected Akaike Information Criterion (AIC) (Hurvich *et al.* 1998, Fotheringham *et al.* 2002).

2.2. *Improved geographically and temporally weighted regression*

Although the traditional GWR technique seems to be appealing in addressing spatial nonstationarity and usually provides better model fit for certain data (Fotheringham *et al.* 2002), the temporal nonstationarity has not been accounted for in such models. Huang *et al.* (2010) proposed a GTWR model to deal with both spatial and temporal heteroskedasticity problems simultaneously by constructing the weight matrix based on distances determined from (u, v, t) coordinates between observation i and all other observations in line with the GWR technique. The GTWR model can be expressed as

$$Y_i = \beta_0(u_i, v_i, t_i) + \sum_k \beta_k(u_i, v_i, t_i) X_{ik} + \varepsilon_i \tag{6}$$

To specify the weighting matrix in the GTWR model, the spatial distance d_{S}^2 and the temporal distance d_{T}^2 were measured separately and then were combined with specified operators '$\otimes$' to form a spatiotemporal distance $d_{\mathrm{ST}}^2 = d_{\mathrm{S}}^2 \otimes d_{\mathrm{T}}^2$, such that the weights can be computed with appropriate kernel function according to the spatiotemporal distance.

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

In the previous research, the authors adopt only the simple '+' operator to measure the spatiotemporal distance with a linear combination between spatial distance and temporal distance, such that

$$d_{\mathrm{ST}}^2 = d_{\mathrm{S}}^2 \otimes d_{\mathrm{T}}^2 = \lambda d_{\mathrm{S}}^2 + \mu d_{\mathrm{T}}^2 \tag{7}$$

where λ and μ are adjustment parameters to balance the different scale effects used to measure spatial and temporal distance in the respective coordinate systems. Using Gaussian distance–decay-based kernel functions to construct a spatialtemporal weight matrix, it equals to build a spatially weighted matrix W^{S} and a temporally weighted matrix W^{T}, respectively, and then combine these to form a spatiotemporal weight matrix $W^{\mathrm{ST}} = W^{\mathrm{S}} \times W^{\mathrm{T}}$ (Huang *et al.* 2010). The proposed '+' operator provides a simple and straightforward way for modeling spatiotemporal distance and satisfies our intuition to measure the 'close' between regression point and observed points with an ellipsoidal coordination system. This specification assumes an orthogonal spatiotemporal coordination system. Thus, the distance measured in space dimensions has no effect on temporal distance and is inadequate for modeling their interactions. In this article, we define the operator '$\otimes$' in a relative complex way and form an IGTWR model.

$$\begin{cases} d_{ij}^{\mathrm{ST}} = d_{ij}^{\mathrm{S}} \otimes d_{ij}^{\mathrm{T}} = \lambda d_{ij}^{\mathrm{S}} + \mu d_{ij}^{\mathrm{T}} + 2\sqrt{\lambda\mu d_{ij}^{\mathrm{S}} d_{ij}^{\mathrm{T}}}\cos(\xi), & t_j < t_i \\ d_{ij}^{\mathrm{ST}} = \infty, & t_j > t_i \end{cases} \tag{8}$$

where t_i and t_j are observed times at locations i and j, and λ, μ and $\xi \in [0, \pi]$ are adjustment parameters, which can also be optimized with CV procedures in terms of R^2 or AIC if no *a priori* knowledge can be obtained. Figure 1 shows an example of the improved spatiotemporal distance.

It can be observed from Figure 1 that there are two distinct improvement of IGTWR compared with the original GTWR on modeling spatiotemporal distance. One is that the parameter ξ is introduced to control the interaction of space and time effects. Note that if $\xi = 0$, we have $d_{\mathrm{ST}}^2 = d_{\mathrm{S}}^2 \otimes d_{\mathrm{T}}^2 = \left(\sqrt{\lambda} d_{\mathrm{S}} + \sqrt{\mu} d_{\mathrm{T}}\right)^2$, indicating that space and time happen to have the maximal effects. If $\xi = \pi/2$, the additive interacted item is equal to zero and Equation (8) returns to the previous version. Another specification is that the spatiotemporal neighborhood impact is estimated only within prior sales. Accordingly, we only take into account previous 'time neighbors' sale transactions and estimate their impact on the current transaction.

2.3. *Geographically and temporally weighted autoregressive model*

LeSage (1999) argued that it is better to cover both spatial and temporal heterogeneity and spatial autocorrelation effects in a mixed model because spatial heterogeneity and autocorrelation are often related in the context of modeling even though the two problems are theoretically distinct. One of the advantages of the combined model is that it can enlarge GWR-based models by including a spatial lag to the dependent variable to reduce unstable estimates. In this article, we combine the aforementioned IGTWR model with the autocorrelation regression model to set up a GTWAR.

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

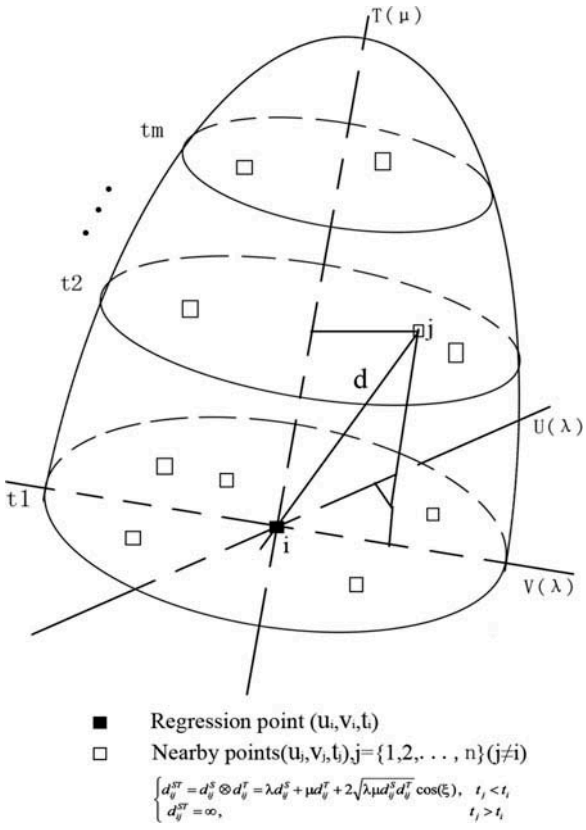

Figure 1. An illustration of improved spatiotemporal distance.

Spatial autocorrelation is constructed using an explanatory variable $\overline{W}Y$ in a linear regression relationship to explain variation Y in across the spatial sample of observations (Can and Megbolugbe 1997).

$$Y_i = \rho_i \overline{W} Y + \varepsilon_i \tag{9}$$

where ρ_i represents a regression parameter to be estimated at the point i and ε_i denotes the stochastic disturbance in the relationship; $\overline{W}$is the $N \times N$ spatial matrix of known constants. $\overline{W}$ contains nonnegative elements of neighboring properties whose diagonal elements are zeros to prevent each observation from predicting itself. A standardized transformation is often used to convert the matrix $\overline{W}$ to have row-sums of unity in applied work. Clearly, Equation (9) is a locally based autoregressive model. The parameter ρ_i is varying from geographical locations, which encapsulates a spatial diffusion process to affect the Y-variable. This implies that the level of spatial association between all adjacent zones is varied. Integrating Equations (6) and (9), we have

$$Y_i = \beta_0(u_i, v_i, t_i) + \rho(u_i, v_i, t_i)\overline{W}Y_i + \sum_k \beta_k(u_i, v_i, t_t)X_{ik} + \varepsilon_i \tag{10}$$

where $\rho(u_i, v_i, t_i)$ is a scalar autoregressive parameter at point i.

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

3. Two-stage least square estimation technique

Since the proposed GTWAR model contains spatially lagged dependent variables, which is typically correlated with the disturbance, the ordinary least squares (OLS) estimator is no longer appropriate in that it can lead to inconsistent estimation. On the other hand, the widely used maximum likelihood estimation procedures proposed by Brunsdon *et al.* (1998) are often computationally very challenging when the sample size is large, and it usually assumes the data to be normally distributed or requires specifying other distributional assumptions that we are not likely to know *a priori* in actual situations. To avoid these problems, we formulate a two-stage least squares estimation and inference technique for the GTWAR model based on the algorithm by Kelejian and Prucha (1998).

Let $Z_i = (X_i, (\overline{W}Y)_i) = W_i^{-1/2}(h)(X, \overline{W}Y), \delta_i = (\beta_i^{\mathrm{T}}, \rho_i)^{\mathrm{T}}$, applying a Cochrane–Orcutt type transformation to Equation (10) yields

$$Y_i = Z_i\delta_i + \phi_i \tag{11}$$

As proposed by Kelejian and Prucha (1998), we design an instrument matrix H, such that the spatially lagged model can be estimated with ordinary least square. In this way, the parameters can be written as

$$\hat{\delta}_i = (\hat{Z}_i^{\mathrm{T}}\hat{Z}_i)^{-1}\hat{Z}_i^{\mathrm{T}}Y = (Z_i^{\mathrm{T}}H_i(H_i^{\mathrm{T}}H_i)^{-1}H_i^{\mathrm{T}}Z_i)^{-1}Z_i^{\mathrm{T}}H_i(H_i^{\mathrm{T}}H_i)^{-1}H_i^{\mathrm{T}}Y_i \tag{12}$$

where $\hat{Z}_i = P_{H_i}Z_i = (X_i, (\overline{W}\hat{Y})_i)$, $\bar{W}_i\hat{Y}_i = P_{H_i}(\overline{W}\hat{Y})_i = H_i(H_i^{\mathrm{T}}H_i)^{-1}H_i^{\mathrm{T}}$.

Generally, the ideal instrument matrices H are $E(Y_i) = E[W_i^{-1/2}(h)Y] = W_i^{-1/2}(h)E(Y)$, and according to Equation (10), the expected operator can be formulated as

$$\begin{aligned} E(Y_i) &= W_i^{-1/2}(h)E(Y) = W_i^{-1/2}(h)[(I-\rho_i\overline{W})^{-1}X\beta_i + (I-\rho_i\overline{W})^{-1}E(\varepsilon_i)] \\ &= W_i^{-1/2}(h)(I-\rho_i\overline{W})^{-1}X\beta_i = W_i^{-1/2}(h)\sum_{j=0}^{\infty}\rho_i^j\overline{W}^jX\beta_i, \qquad \text{when } |\rho_i| < 1 \end{aligned} \tag{12}$$

Consequently, in this case, $W_i^{-1/2}(h)E(Y)$ is seen to be formed as a linear combination of the columns of the matrices $W_i^{-1/2}(h)[X, \overline{W}X, \overline{W}^2X, \cdots]$. It is for this reason that we postulate H is composed of a subset of the linearly independent columns of $W_i^{-1/2}(h)[X, \overline{W}X, \overline{W}^2X, \ldots]$. If the number of determinates in the regression model is large, we approximate the H just with the first three columns, i.e. $W_i^{-1/2}(h)[X, \overline{W}X, \overline{W}^2X, \ldots]$. Based on the weighting function of GTWR in the Gaussian form (Huang *et al.* 2010), for the ith regressive point, the weighting matrix is thus $W_i^{-1/2}(h) = \mathrm{diag}([\exp(d_{i1}/h)^2, \ldots, \exp(d_{iq}/h)^2])$. We also note that the matrices H have full column rank, which could be relaxed at the expense of working with generalized inverses, since the orthogonal projection of any vector onto the space spanned by the columns of H is unique even if H does not have full column rank. To make the estimators remain well asymptotically and work well, the instruments H and ρ_i satisfy furthermore the following constraints (Kelejian and Prucha 1998, 2006).

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

$$\begin{cases} Q_{HH} = \lim\limits_{n->\infty} n^{-1} H^{\mathrm{T}} H \\ Q_{HZ_i} = p \lim\limits_{n->\infty} n^{-1} H^{\mathrm{T}} Z_i \\ \det\left|I - \rho_i \overline{W}\right| \neq 0 \end{cases} \tag{13}$$

where Q_{HH} and the matrix $I - \rho_i \bar{W}$ are finite and nonsingular and Q_{HZ} is finite and full column rank. In practice, we should be checking whether these constraints satisfy the necessary premise before estimating the parameters with the two-stage least squares framework. In our experiments, we observed that the absolute value of estimated ρ_i is likely to be larger than one, which invalidates the limitation of the autocorrelation coefficient and results in the matrix $I - \rho_i \bar{W}$ being singular. For instance, if the ith observations invalidate the constraints, then we solve it using quadric optimization programming technique:

$$\begin{cases} \min \quad \frac{1}{2} \delta_i^{\mathrm{T}} \hat{Z}_i^{\mathrm{T}} \hat{Z}_i \delta_i - \hat{Z}_i^{\mathrm{T}} Y_i \\ \mathrm{St.} \quad \delta_i = (\beta_i^{\mathrm{T}}, \rho_i)^{\mathrm{T}}, |\rho_i| < 1 \end{cases} \tag{14}$$

Let $U_i^{\mathrm{T}} = (1, x_{i1}, x_{i2}, \ldots, x_{in}, (\bar{W}\hat{Y})_{i1})$ be the ith row of Z_i. Then the fitted value of Y_i at point i is obtained by

$$\hat{Y}_i = U_i^{\mathrm{T}} [Z^{\mathrm{T}} W^{-1} H (H^{\mathrm{T}} W^{-1} H)^{-1} H^{\mathrm{T}} W^{-1} Z)^{-1} Z^{\mathrm{T}} W^{-1} H (H^{\mathrm{T}} W^{-1} H)^{-1} H^{\mathrm{T}} W^{-1} Y_i \tag{15}$$

Denote the vector of the fitted values of Y_i and the vector of the residuals at the N locations by $\hat{Y} = (\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_n)$ and $\hat{\Phi} = (\hat{\varphi}_1, \hat{\varphi}_2, \ldots, \hat{\varphi}_n)$, respectively, and L is the hat matrices of Y, then $\hat{Y} = LY$, where

$$L = \begin{pmatrix} U_1^{\mathrm{T}} [Z^{\mathrm{T}} W^{-1} H (H^{\mathrm{T}} W^{-1} H)^{-1} H^{\mathrm{T}} W^{-1} Z)^{-1} Z^{\mathrm{T}} W^{-1} H (H^{\mathrm{T}} W^{-1} H)^{-1} H^{\mathrm{T}} W^{-1} \\ \vdots \\ U_n^{\mathrm{T}} [Z^{\mathrm{T}} W^{-1} H (H^{\mathrm{T}} W^{-1} H)^{-1} H^{\mathrm{T}} W^{-1} Z)^{-1} Z^{\mathrm{T}} W^{-1} H (H^{\mathrm{T}} W^{-1^{-1}} H)^{-1} H^{\mathrm{T}} W^{-1} \end{pmatrix} \tag{16}$$

is a matrix and I is an identity matrix of order N.

The global fitted residuals are $\hat{\hat{\Phi}} = (Y - \hat{Y}) = (I - L)Y$ and the sum-square-error is $\mathrm{RSS} = \hat{\hat{\Phi}}^{\mathrm{T}} \hat{\hat{\Phi}} = Y^{\mathrm{T}} (I - L)^{\mathrm{T}} (I - L) Y$.

4. Application to house prices

To examine the applicability of GTWAR, a case study was implemented using housing prices observed between 2001 and 2008 in the city of Shenzhen, China. First, the global OLS approach was used to analyze the housing data without any spatial or temporal consideration (taking space coordinates and time as exogenous determinants). Then, the proposed GTWAR model was employed to attempt to analyze the housing data including spatial and temporal considerations. Subsequently, we tested parameters representing significant spatiotemporal variations across the study area. For comparison purpose, three different GWR-based approaches namely GWR, GTWR, and GWRSL were also implemented using the same data set, and we examined their goodness-of-fit in terms of R^2 and AIC criteria.

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

4.1. Study data

Shenzhen, the special economic zone of China, covering an area about 1984 km^2, is located in southern Guangdong Province on the eastern edge of the Daya and Dapeng Bay, east of the Pearl River, southeast of Dongguan, and north of Hong Kong. As portrayed in Figure 2, Shenzhen consists of six administrative districts: Luohun, Futian, Nanshan, Yantian, Bao'an, and Longgang, of which the last two are not located in the Special Zone. Since China opened its doors to the global market, Shenzhen has experienced some drastic changes owing to its special status accorded by the Central Government of China. The real estate prices of Shenzhen are continuously increasing at an alarming rate associated with rapid industrialization and urbanization and consequent demands for various categories of real estate.

The experimental data on residential estate prices come from the Shenzhen municipal bureau of land resources and housing management. 406 observations were available from the study area, which (1) provided full information on housing age, land area, green rate, traffic condition, and other variables; (2) were of the most common construction types and occupancy type; and (3) provided the exact geo-coded coordinates for each transaction to enable us to conduct the spatiotemporal procedures. According to Sirmans *et al.* (2005), using the observed price is generally thought to minimize the bias inherent in other measures such 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 13 variables: whether or not located in the Special Zone (SZYN), land area (LANDA), property management spend (MAGT), green space (GREEN), proximity to bus (TRAFF), floor area ratio (FLOOR), land price (LPRICE), structure type (STRU), quality (QUAL), distance from the nearest major road (DROAD), distance from the nearest school (DSCHL), distance from the nearest hospital (DHOSL), and distance from the downtown area (DCENT). Averages for the land area, floor area ratio and land price at each residential estate (observation point) were calculated using ArcGIS. Discrete variables such as whether located in the Special Zone, estate developer, proximity to bus, and the structure type in each unit were also counted using location-related joins, and then the proximity variables measuring the minimum Euclidean distances to the nearest road, school, downtown area,

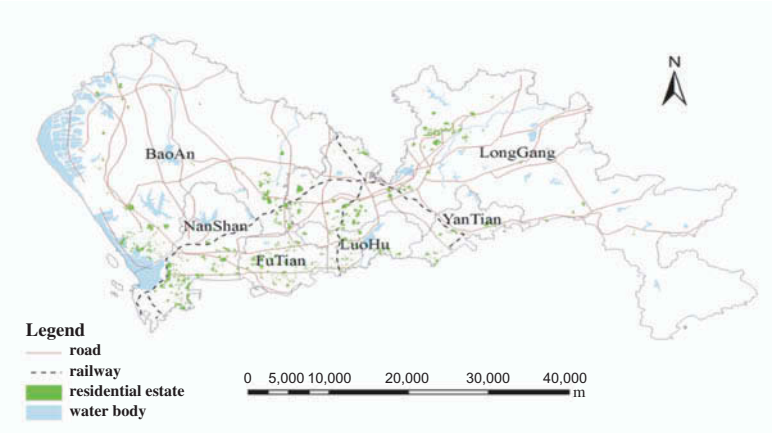

Figure 2. A summary map of the study area.

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

and so on were calculated using the ArcGIS spatial analyst tools. The percentage of green and open space area in each unit was also computed.

4.2. Global OLS model

Before investigating possible spatiotemporal variations in the determinants of housing prices across Shenzhen city, the global regression equation representing the average relationships of the spatial units between the level of housing price and various factors is presented in Table 1.

The F-test indicates that the OLS model is statistically significant. Moreover, the R^2 of the model is 0.617, which means that the equation explains 61.7% of the variance of the level of house values in Shenzhen. This result indicates that the assessed house values can be modeled by the selected housing structural attributes and the neighborhood environment variables. Therefore, the hypothesized relationships between the structural, neighborhood attributes and the house values are supported by the data. Indeed, LANDA, DROAD, QUAL, LPRICE, MAGT, and TRAFF determinants are statistically significant at 90% confidence level according to their *t*-probabilities. In particular, land area, quality, land price, traffic condition, and property management spend (MAGT) are positively associated with house values, whereas the *Y* coordinate is negatively associated with the house values. A high land price will push up the price of houses and high quality of build will significantly add to the value of the house. Since the first-class residential estates usually charge more property management spend (MAGT) in China, the MAGT also has a

Table 1. Hedonic model (OLS) parameter estimate summaries.

Variable	Coefficient	Std. Error	*t*-Statistic	*t*-Probability
Intercept**	7.9322	0.1263	62.8144	0.0000
SZYN	−0.0114	0.0756	−0.1508	0.8801
LANDA*	0.3962	0.2211	1.7921	0.0738
DROAD*	0.3706	0.2089	1.7745	0.0767
DSCHL	−0.1211	0.1281	−0.9453	0.3450
DHOSL	0.1154	0.1326	0.8705	0.3845
FLOOR	0.0866	0.0777	1.1154	0.2653
STRU	−0.0720	0.0696	−1.0355	0.3010
QUAL**	0.2301	0.0988	2.3290	0.0203
DCENT	−0.0955	0.1591	−0.6005	0.5485
LPRICE**	1.5082	0.2763	5.4588	0.0000
MAGT**	1.6585	0.2603	6.3717	0.0000
TRAFF**	0.9607	0.2854	3.3663	0.0008
GREEN	0.1348	0.1287	1.0476	0.2954
X	0.0266	0.1116	0.2384	0.8117
*Y***	−0.4925	0.1914	−2.5737	0.0104
Time**	0.8546	0.0753	11.3509	0.0000
Diagnostic information				
R^2	0.617			
AIC	1848.7			
Residual standard error	0.1389			
F-Statistic	36.5934			
P-Value	0.0000			

Note: * and **Denote 10% and 5% statistical significance, respectively.

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

positive effect on the housing price accordingly. It is also found that time is a positive factor, which means the house prices are continuously rising in Shenzhen city with its recent rapid industrialization and urbanization. Interestingly, the Y coordinate is a negative factor, whereas the X coordinate is not a statistically significant factor. The gradient price from south to north across the Y coordinate is possibly due to the fact that Bao'an and Longgang located in the north of Shenzhen city are two less developed districts; thus, the house value is relatively lower. Indeed, the results demonstrate that space and time are two important determinates associated with house values in Shenzhen city. Therefore, we further examine the spatial and temporal variations in the following section.

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

4.3. *Results of the GTWAR model*

To consider the spatial and temporal variation of relationships between the level of house price and statistically significant determinants, the aforementioned GTWAR model is implemented. Several problems must be solved before we apply the GTWAR model to house prices. The prime one is that we must test whether time and space autocorrelation exist among these data. If there is no autocorrelation, GTWR is adequate for spatiotemporal variations; otherwise, the GTWAR model is more appropriate. In this article, Global Moran'I is used for this purpose because of its simplicity and effectiveness (Anselin 1988).

It can be seen from Table 2 that the value of Moran'I for spatial, temporal, and spatiotemporal data is 0.1013, 0.1422, and 0.2055, respectively, and all of their t-statistic values are larger than 1.96. This indicates that there exist spatiotemporal autocorrelations in the dependent variable. Therefore, the GTWAR model can achieve better results by accounting for spatiotemporal autocorrelation. The second critical problem is to determine the spatiotemporal weighted matrix W^{ST}. In this case, the observations are unevenly distributed across the study area. Therefore, the adaptive kernel function is employed. This adaptive kernel might present a more reasonable means of representing the degree of spatial nonstationarity in the study area.

To obtain an optimum size of nearest neighbors for the adaptive kernel, a common approach is to maximize the R^2 of the GTWAR model. According to Huang *et al.* (2010), only the parameter ratio $\tau = \mu/\lambda$ plays an important role in constructing weights. Hence, we set $\lambda = 1$ to reduce the number of parameters, and so only three parameters, q, μ, and ξ, need to be determined, where q is the number of the nearest observed points. Note that the minimization of the CV function is subject to the constraint that the number of nearest neighbors should satisfy $q \geq 30$. This imposed constraint is to prevent explanatory variables tending towards multicollinearity and small sample bias. Figure 3 presents the parameters μ and ξ against R^2 with CV procedures. We can see that the maximal R^2 is achieved when $\mu = 290$ and $\xi = 0$.

Using the optimal parameters, GWR and GTWAR models were attempted, and the results are reported in Table 3. It can be observed that the optimized bandwidth of GWR

Table 2. Tests of space and time autocorrelation of the dependent variable with Moran' I statistics.

Models for weights	Moran' I	t-Stat	P-value
Spatial weights	0.1031	3.9611	0.0008
Temporal weights	0.1422	4.9788	0.0000
Spatiotemporal weights	0.2055	10.0009	0.0000

Table 3. GWR and GTWAR parameter estimate summaries.

	GWR (q = 51)			GTWAR (q = 44)		
Parameter	Min	Med	Max	Min	Med	Max
Intercept	3.5191	7.7126	9.663	0.8045	8.0195	16.663
LANDA	−4.75	0.23641	5.011	−6.6985	0.57093	6.192
DROAD	−2.3997	−0.30953	2.9283	−7.2365	−0.0164	8.8843
QUAL	−15.535	−0.3040	19.686	−10.28	−0.9405	5.4904
LPRICE	−3.7342	2.1249	12.279	−6.3766	1.3003	10.584
MAGT	−3.0212	2.6433	6.1173	−9.7942	1.7758	8.753
TRAFF	−3.292	0.9541	4.8424	−5.5886	1.1023	9.9913
ρ				0.0203	0.1101	0.9244
Diagnostic information						
R^2	0.736			0.914		
RSS	37.2			12.0		
AIC	832.7			816.6		
F-statistic	1.63			3.84		
P-value	0.0000			0.0000		

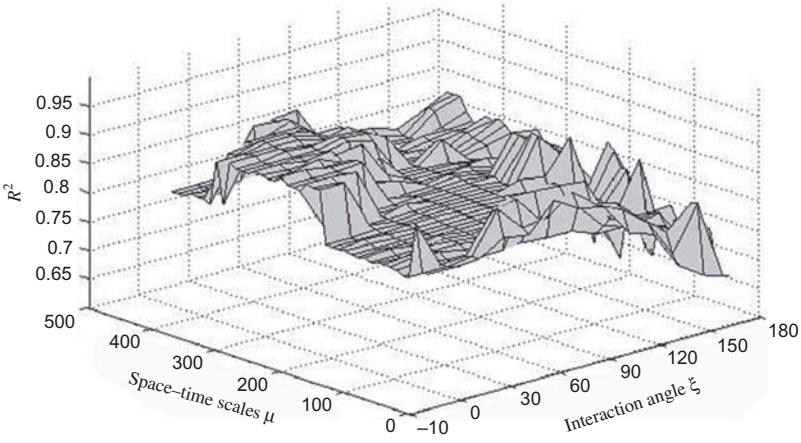

Figure 3. The R^2 values against the parameters μ and ξ.

and GTWAR is 51 and 44, respectively. Intuitively, the house price data may exhibit more nonstationary effects if the spatial and temporal nonstationarity are considered simultaneously. Thus, the GTWAR model has a smaller bandwidth than the GWR model. Because the output of the local parameter estimates from GWR and GTWAR would be voluminous, Table 3 provides a three-number summary of the distribution of each parameter to indicate the extent of the variability. A detailed statistical comparison is also included. Note that the percentage of explanation of variance has increased from 61.7% in OLS model to 73.6% in GWR model, and GTWAR explains a considerably larger proportion of variation of 91.4%. This is not surprising, because an increase was to be expected given the difference in degrees of freedom. It can also be seen from Table 3 that the GTWAR model is the best, even if differences in degrees of freedom with the reduction in the AIC from the global model (from 1848.7 for linear model to 832.7 for

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

GWR and 816.6 for GTWR, respectively) are taken into account. By comparing the residual sum of squares, the decreased value further indicates that GTWAR gives a better fit of data than the OLS and GWR models. GTWAR can handle the issues of both spatial and temporal heterogeneities, as well as autocorrelations. In this case, the impact of those variables varies over the space, and the statistical significance further indicates that neither the OLS model nor the GWR model is valid alternatives.

The two remaining problems are further considered. One is whether the IGTWR really counts compared with GTWR by adding an additional parameter; the other is whether it is necessary to add the lag item to the model.

To validate the enhancement of the IGTWR, the fitted capabilities of the IGTWR and GTWR models with various parameters are compared. It can be seen from Figure 4 that the IGTWR with additional parameter ξ usually outperforms GTWR in terms of R^2. To determine whether the spatial lag is significant to the model, we plot the ρ values and their corresponding t-statistical test with GTWAR model in Figure 5.

It can be calculated that only about half of the observations do exhibit significant spatial autocorrelations, whereas the other half have weak spatial autocorrelations with ρ values less than 0.05 in this study. In Table 3, the median value of ρ is about 0.1,

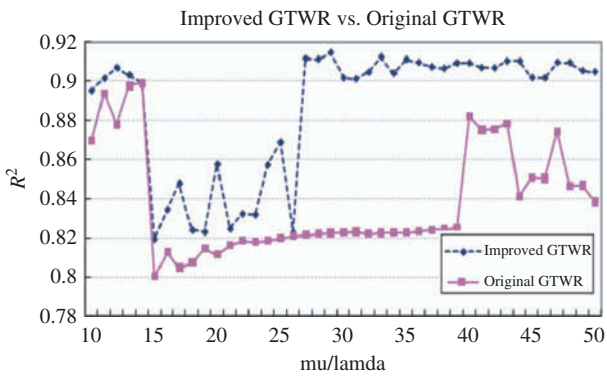

Figure 4. Comparison of the performances of improved GTWR and GTWR.

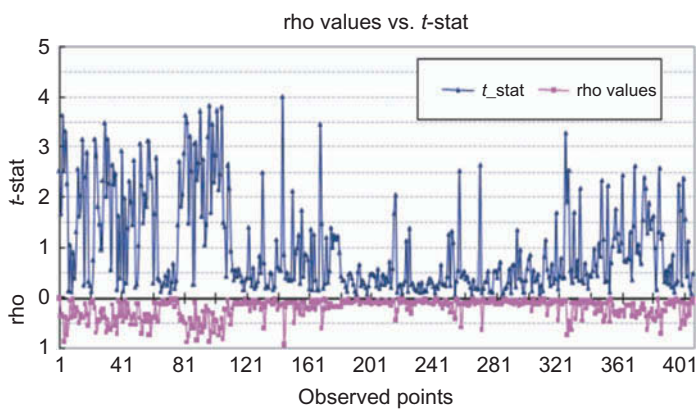

Figure 5. The estimated ρ and its t-test value.

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

suggesting that spatial autocorrelation of the tested data is rather weak. One important reason is that we have used only about 400 observations,[1] sparsely scattered across the study area. However, it is still reasonable to account for spatial autocorrelation in some regions where the observations are relatively dense.

4.4 Significant test and model comparison

From the statistical viewpoint, two critical questions are necessary to discuss. One is whether each set of parameter estimates exhibits significant variation across the study area (Brunsdon *et al.* 1998, Leung *et al.* 2000); the other is whether the proposed GTWAR model describes the relationship significantly better compared with other local models, such as GWR, GTWR, and GWRSL. The first one is a parameter nonstationarity test question, for which we use the statistical test proposed by Leung *et al.* (2000) because of its simplicity and efficiency. They proposed to test the following hypothesis with F statistic:

$$H_0 : \beta_{1k} = \beta_{2k} = \cdots = \beta_{nk}\text{, for a given } k$$

$$H_1 \text{ : not all } \beta_{ik}(i = 1, 2, \cdots, n) \text{ are equal.}$$

By constructing 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 $\beta_{ik}(i = 1, 2, \cdots, n)$, an approximated F distribution statistical 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 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, we reject the null hypothesis. Otherwise, we accept it.

Table 4 presents the F-statistic value of each variable and corresponding P-value. Those statistically significant values at 5% level are marked with star '*'. We see that all variables except ρ exhibit significant spatial and temporal variation in the local parameter

Table 4. Statistics results of GTWAR model testing.

Parameter	Statistics	Value	NDF	DDF	P-value
Intercept	F(1)	11.5047	27.7494	201.4599	0.0000*
LANDA	F(2)	5.0511	29.1683	201.4599	0.0000*
DROAD	F(3)	1.7036	22.3773	201.4599	0.0293*
QUAL	F(4)	7.1720	27.5176	201.4599	0.0000*
LPRICE	F(5)	8.0314	29.7324	201.4599	0.0000*
MAGT	F(6)	4.3612	26.3485	201.4599	0.0000*
TRAFF	F(7)	3.6348	28.8777	201.4599	0.0000*
ρ	F(8)	0.0011	28.2389	201.4599	0.9999

Note: NDF and DDF the degrees of freedom of the numerator and denominator of the corresponding F-distributions, respectively.

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

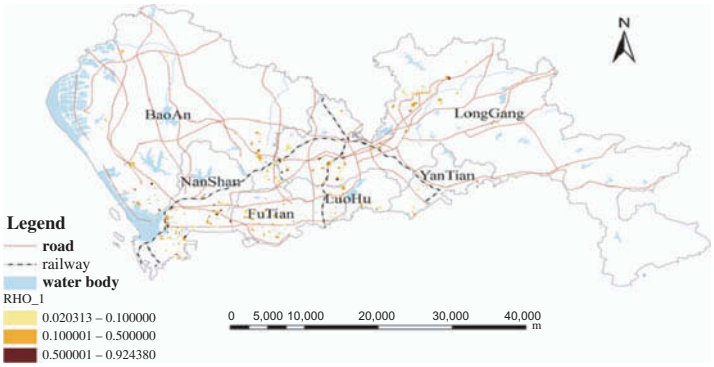

Figure 6. Spatial variation of the spatiotemporal lag coefficient ρ.

estimates. Though the spatiotemporal distribution of the ρ parameter does not display statistically significant spatiotemporal nonstationarity, it is found from Figure 6 that there exists significant spatial variation in Nanshan, Bao'an, and LongGang districts, whereas the ρ values are relatively smooth in FuTian and LuoHu districts. One possible interpretation of the result is that Bao'an and LongGang, as two fringe districts of Shenzhen, were undergoing a rapid development during our considered period. Therefore, various categories of housing estates, including many newly developed high-grade residences and older apartments, co-existed. On the other hand, as a primary residential zoning, luxurious apartments and common residences spread over the Nanshan district. As a result, all the three districts show significant spatial variations.

One important characteristic of the GWR-based technique is that the local parameter estimates which denote local relationships can be mapped out and thus allow for visual analysis. Taking the coefficients of 'traffic condition' and 'apartment quality' as examples, we can group them into several classes and color each class to view the spatial variation patterns of these variables.

It can also be seen from Figure 7 that the spatial variation of the variable 'traffic condition' in GTWAR shows two major trends: traffic condition varies from high in the north to low in the south and from high in the outer zones of the study area to low in the inner zone. This suggests that the 'traffic condition' of a residential zone had the most important influence on housing prices in the outer part of the city. It is reasonable that Futian and Luohu, two central districts, have a convenient access to road networks, which reduces the reliance on traffic conditions. On the other hand, 'Apartment quality' does not show apparent spatial variation, indicating that the 'Apartment quality' of a residential zone had the most important influence on all housing prices.

The second question is a goodness-of-fit test for the GTWAR model. Usually, a complex model will fit a given data set better than a simpler one. However, from the practical point of view, if a complex model does not perform significantly better than the simple one, it means that the simple model can be applied and interpreted. It is hence natural to ask whether the increase in modeling complexity can be significantly traded off with the gain in the GWR-based model. Here we adopt another F-test proposed by Brunsdon *et al.* (1998) for this problem. This F-test is based on analysis of variance and uses generalized degrees of freedom to compare with the improved sum of squares accounted for by the GWR-based estimates as compared with the global OLS estimates. This suggests that these comparisons can be expressed in the form of an ANOVA table, with residual mean squares for these

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

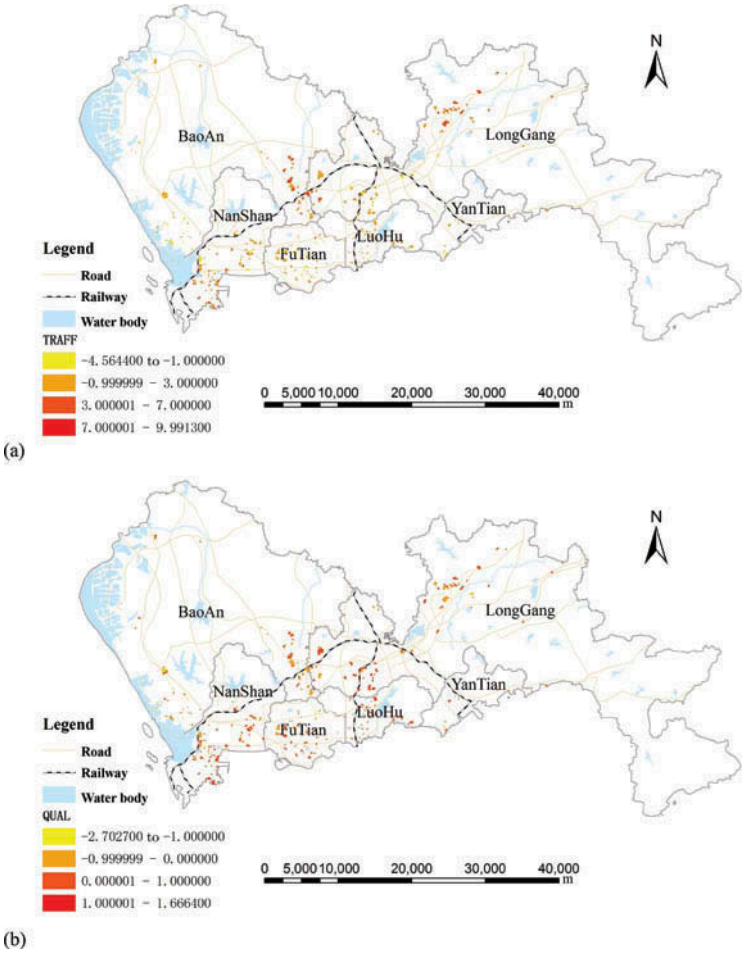

Figure 7. Spatial variation of (a) 'traffic condition' and (b) 'apartment quality' coefficients.

models being compared. To investigate whether the GTWAR model performs significantly better than the other GWR-based models, R^2 and AIC measurements are also listed in Table 5. According to Fotheringham *et al.* (2002), a 'serious' difference between two models is

Table 5. ANOVA comparisons of GTWAR with GWR, GWRSL, GTWR, and IGTWR models.

Source of Variation	RSS	DF	MS	F	p-value	R^2	AIC
OLS	72.1	16	4.51	24.8	0.00	0.617	1848.7
GWR	37.2	249.0	0.15	1.63	0.00	0.736	832.7
GWRLS	36.2	248.3	0.14	1.70	0.00	0.743	828.6
GTWR	17.6	198.0	0.09	3.15	0.00	0.875	826.7
IGTWR	14.7	193.2	0.07	3.77	0.00	0.895	821.4
GTWAR	12.0	169.1	0.06	3.84	0.00	0.914	816.6
GTWAR/GWR improvement	25.2	79.9	0.31	–	–	17.8%	16.1
GTWAR/GWRSL improvement	24.2	80.7	0.29	–	–	17.1%	12.0
GTWAR/GTWR improvement	5.6	28.9	0.19	–	–	3.9%	10.1
GTWAR/IGTWR improvement	2.7	24.1	0.11	–	–	1.9%	4.8

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

generally regarded as one in which the difference in AIC values between the models is larger than 3. In this way, we can judge whether GTWAR is significantly better than other models.

The results of ANOVA tests on the observations are shown in Table 5. In this table, the first column lists the residual sum of squares of OLS, GWR, GTWR, and GTWAR residuals. The second column gives the degrees of freedom for each of these models. The third column, mean square (MS), gives the results of dividing the sums of squares by their respective degrees of freedom. The subsequent two columns show the pseudo F-statistic and the P-value. The last two columns are R^2 and AIC values. It can be found that the reduction in the residual sum of squares (RSS) is reduced when using GWR-based approaches. The F-test in Table 5 shows that there is a significant spatial and temporal nonstationarity over the study area. Therefore, it is more appropriate to model the specified data set with local models. Moreover, we can see that modeling spatial nonstationarity with the traditional GWR and GTWR is inadequate for this data set, and a higher level of accuracy can be achieved if temporal variation and lag information are added. The ANOVA test shows that the GTWAR model offers a significant improvement over the global OLS and other GWR-based models. In addition, the AIC of the GTWAR model (816.6) is far less than the AIC of the global OLS model (1848.7), GWR model (832.7), and GTWR (819.9). This indicates that even taking into account the added complexity of the GTWAR model, it still performs better than the OLS, GWR, and GTWR models. Results from Table 5 justify the hypothesis that significant nonstationary relationships between the house value and each of the four structural and neighborhood attributes exist in the City of Shenzhen.

5. Conclusions

This article has proposed a GTWAR model to exploit spatiotemporal variations and autocorrelations and formulated a two-stage least squares framework. The proposed model can deal with heterogeneous and auto-correlated effects simultaneously arising from spatial and temporal variations. Moreover, the algorithm can effectively reduce computation complexity of the GTWAR model without prespecified error distribution compared with the maximum likelihood estimation method.

The experimental results of our case study of housing prices in Shenzhen suggest that GTWAR achieves a better modeling accuracy than both the GWR model, which deals with spatial nonstationarity only, and the GTWR model with both spatial and temporal nonstationarity. Compared with the local regressive models, GTWAR can improve the R^2 of OLS, GWR, and GTWR from 0.617, 0.736, and 0.875, respectively, to 0.914. The AIC criterion also shows a 16.1 reduction over GWR and a 4.8 reduction over GTWR. The statistical tests indicate that there is a significant difference among GTWAR, GWR, and GTWR, and we therefore conclude that it is meaningful to incorporate both spatiotemporal lagging and temporal nonstationary effects into the GWR model.

Some limitations also exist in our case study. For example, only 406 observations were considered because of a shortage of available data. Whether GTWAR would perform better if applied to data covering more observations needs to be further investigated. Also, the proposed model has to determine several parameters, which results in the optimal parameters selection with the CV procedure imposing a heavy computational load. A possible solution is to seek the benefit of intelligent algorithms or high performance by distributed parallel computing techniques, which we will study further.

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

Funding

This work was supported in part by the National High-Tech R&D Program (863 Program) through Grant 2009AA122004 and in part by the Hong Kong Research Grants Council through Grant CUHK 444612. Their support is gratefully acknowledged.

Note

1. According to Equation (8), we estimate only the impact of previous and current neighboring samples on a current point and the earliest several points are generally not considered. Hence, we use only 400 samples rather than 406 samples.

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015

References

Anselin, L., 1988. *Spatial econometrics: methods and models*. Dordrecht: Kluwer Academic.

Basu, A. and Thibodeau, T., 1998. Analysis of spatial autocorrelation in house price. *Journal of Real Estate Finance and Economics*, 17 (1), 61–85.

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., 1998. Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Sciences*, 39 (3), 497–524.

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

Cressie, N., 1993. *Statistics for spatial data*. Revised Ed. NewYork: Wiley-Interscience, 923.

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

Gelfand, A., *et al*., 1998. Spatio-temporal modeling of residential sales data. *Journal of Business & Economic Statistics, 16,* 312–321.

Huang, B., Wu, B., and Barry, T., 2010. Geographically and temporally weighted regression for spatiotemporal modeling of house prices. *International Journal of Geographical Information Science*, 24 (3), 383–401.

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.

Kelejian, H. and Prucha, R., 1998. Generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Finance and Economics*, 17 (1), 99–121.

Kelejian, H. and Prucha, R., 2006. Estimation problems in models with spatial weighting matrices: which have blocks of equal elements. *Journal of Regional Science*, 4, 507–515.

LeSage, J.P., 1999. *The theory and practice of spatial econometrics*. Available from: http://raptor1.bizlab.mtsu.edu/s-drive/TEFF/sbook.pdf. [Accessed 28 August 2010].

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.

Ord, J.K., 1975. Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70, 120–126.

Pace, R.K. and Gilley, O.W., 1998. Generalizing the OLS and grid estimators. *Real Estate Economics*, 26 (2), 331–347.

Pace, R.K., *et al*., 1998. Spatiotemporal autoregressive models of neighborhood effects. *Journal of Real Estate Finance and Economics*, 17, 15–33.

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*, A34, 733–754.

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

Tu, Y., Yu, S.M., and Sun, H., 2004. Transaction-based office price indexes: a spatiotemporal modeling approach. *Real Estate Economics*, 32, 297–328.

Downloaded by [Chinese University of Hong Kong] at 22:22 05 January 2015