

Wang, Lu; Wang, Guangxing; Yu, Huan; Wang, Fei

## Article

# Prediction and analysis of residential house price using a flexible spatiotemporal model

Journal of Applied Economics

## Provided in Cooperation with:

University of CEMA, Buenos Aires

*Suggested Citation:* Wang, Lu; Wang, Guangxing; Yu, Huan; Wang, Fei (2022) : Prediction and analysis of residential house price using a flexible spatiotemporal model, Journal of Applied Economics, ISSN 1667-6726, Taylor & Francis, Abingdon, Vol. 25, Iss. 1, pp. 503-522, <https://doi.org/10.1080/15140326.2022.2045466>

This Version is available at:

<https://hdl.handle.net/10419/314172>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>



## Prediction and analysis of residential house price using a flexible spatiotemporal model

Lu Wang, Guangxing Wang, Huan Yu & Fei Wang

To cite this article: Lu Wang, Guangxing Wang, Huan Yu & Fei Wang (2022) Prediction and analysis of residential house price using a flexible spatiotemporal model, Journal of Applied Economics, 25:1, 503-522, DOI: [10.1080/15140326.2022.2045466](https://doi.org/10.1080/15140326.2022.2045466)

To link to this article: <https://doi.org/10.1080/15140326.2022.2045466>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 03 Apr 2022.



Submit your article to this journal [↗](#)



Article views: 3911



View related articles [↗](#)




View Crossmark data [↗](#)



Citing articles: 10 View citing articles [↗](#)

# Prediction and analysis of residential house price using a flexible spatiotemporal model

Lu Wang<sup>1</sup>, Guangxing Wang<sup>2</sup>, Huan Yu<sup>3</sup>  and Fei Wang<sup>4</sup>

<sup>1</sup>College of Natural Resources and Environment, South China Agricultural University, Guangzhou;

<sup>2</sup>Department of Geography and Environmental Resources, Southern Illinois University, Carbondale, United States; <sup>3</sup>College of Earth Sciences, Chengdu University of Technology, China Chengdu, China; <sup>4</sup>Data science team, CraGurus, Cambridge, United States United States

## ABSTRACT

House price prediction has traditionally been approached using linear or spatial linear hedonic models and focused on big cities. In this study, we developed a flexible spatiotemporal model (FSTM) to explore the spatiotemporal characteristics of the residential house price and the impact factors in middle-small cities. The FSTM integrated both spatial and temporal components of the residential house price, accounted for its spatiotemporal characteristics, and reproduced its spatial variability and temporal trends. The results showed that the governmental policy had a significant influence on the house price and led to the characteristics being different from those in big cities. The significant factors also included the density of roads, the density of banks, density of supermarkets, the area used by public and user shared area within a building. This study implied that FSTM provided the potential for spatiotemporal prediction of the residential house price in the middle-small cities.

## ARTICLE HISTORY

Received 11 August 2020

Accepted 18 February 2022



## KEYWORDS

Driving factors; middle-small city; residential house price; spatiotemporal modeling; Shunde district; urbanization

## 1. Introduction

During the past several decades, the increase of the global economy has led to rapid urban development and city sprawling (Miller, 1973). Urban planning and land use efficiency are the key forces of urban development stage, especially in the rapidly urbanizing eastern region of China. On the contrary, China's urban development projects have many effects on the society, which lead to the house price increasing all around the country. On the other hand, urban house price is an important driving force and constraint for urban upgrading and development in China. Therefore, investigating the factors that affect house price and spatiotemporally quantifying their effects are an urgent need of social-economic healthy development and sustainability in the era of new urbanization in China.

Moreover, the spatial variability and temporal dynamics of house price are subject to many factors including the process of city sprawling, the level of economic development, infrastructure condition, locations, road density, bank density, supermarket density,

**CONTACT** Huan Yu  [yuhuan0622@126.com](mailto:yuhuan0622@126.com)  College of Earth Sciences, Chengdu University of Technology, No. 1 Erxianqiao East 3rd Road, Chengdu, 610059, China

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

landscape aesthetics, environmental quality, one's desirability, etc. The factors can be summarized as socioeconomic variables, location variables and environmental variables (Brueckner, 1990; Davis & Heathcote, 2007; Scott, 1983). The factors vary spatially and temporally, and may be autocorrelated spatially and temporally, which determines the spatiotemporal variability of city house price.

Characterizing the spatial distribution and temporal trend of city house price is thus challenging. For this purpose, various methods have been developed, including real estate price forecast models, housing price indexes, valuation models or AVMS, and spatial prediction model.

Real estate price forecast models of predicting city house price are mainly based on segmentation functions (Colwell & Munneke, 2009) and regression models (Majumdar, Munneke, Gelfand, Banerjee, & Sirmans, 2006) to explain gradient changing characteristics of house price; and one-directional adjacency matrices in spatial autoregressive model to explore the determined factors of land price by spatial autoregressive model and Monte Carlo simulation results Yokoi and Ando (2012).

Housing price indexes are often integrated with hedonic house price models to quantify the hedonic house price indices within a short-run equilibrium model (Diewert, Haan, & Hendriks, 2015; Goodman, 1978), and integrated with spatial models to capture additional quality from observed and unobserved land characteristics (Glumac, Herrera, & Licheron, 2019).

The valuation model or AVMS such as hedonic models often integrated with theoretical models, such as: Geoghegan (2002) used the integrated hedonic model to test how different types of open spaces affect house price, Leggett and Bockstael (2000) considered the possibility that pollution sources to analyze the impacts of water pollution on property values, Rosen (1974) used the integrated model of Hedonic Price Method (HPM) to estimate the effect of changing demands on shadow or Willingness to Accept (WTA) land prices among smallholder farmers, Jiao and Liu (2010) put forward a geographic field-based spatial hedonic models to evaluate environmental amenities by formulating the relationships between house price and various factors, Gong and de Haan (2018) propose a hedonic house price models using geospatial data and a semi-parametric method to measure the impact on aggregate price change and get almost identical the house price indices with different hedonic models.

Geographically Weighted Regression (GWR) is a technique that extends the traditional regression framework by allowing spatial parameters to be explicitly estimated (Oshan & Fotheringham, 2018). Hu, Yang, Li, Zhang, and Xu (2016) utilized geographically weighted regression (GWR) locally models the relationship of house price with the driving factors or variables having significant contributions to the spatiotemporal variability of the house price. The methods can realize multi-scale weighted interpolation of house price and be used to spatially explore the relationship between house price as the dependent variable and one or more explanatory variables or factors given a study area. Zhang & Yen (2020) identified the relationship between estimates of land value and property value changes due to BRT (dependent variable) based variables for analysis methods. Yang et al., (2020) found the factors of closely associated with residential land expansion, and indicated that the relationships of the impact factors with land price are spatially non-stationary (Hu et al., 2016).

Diverse interpolation methods of spatial statistical models, such as Kriging methods (Tsutsumi & Seya, 2008), Inverse Distance Weighted (IDW), Multifractal Inverse Distance Weighted (MIDW) (Hu, Cheng, Wang, & Xu, 2013), self-adaptive weighted gliding averages, has been used in land price analysis, and other interpolation methods are often integrated with fractal theory and GIS technologies, such as grid discretization to evaluate the spatial distribution characteristics of house price (Hu, Cheng, Wang, and Xie, 2012), and so forth. Hu et al. (2013) pointed out that the MIDW interpolation method was a more efficient alternative for characterizing the spatial distribution of house prices compared with the IDW and Kriging method. Furthermore, a geographically and temporally weighted regression (GTWR) model can consider a spatial non-stationary by calculating local statistics and spatial relationships between variables, which is established to estimate the spatial and temporal impacts of land price variation and environmental regulation using panel data (Wang et al., 2020). The above spatial statistical relevant methods take into account the spatial autocorrelation of house price and its impact factors, but can only be utilized to analyze and model the spatial characteristics of house price, and usually require a relatively large number of house price samples. The methods may not work well to model the spatial patterns of house price only for cities where spatially well-distributed house price samples are available. More importantly, the methods lack the ability of modeling temporal trends of house price and their variability over space.

In addition to the characteristics of spatial variability, house price usually shows obviously temporal change and autocorrelation. In most of the research, the temporal changes of house price are usually explained separately with the spatial characters (Chen & Fik, 2017; Du & Huang, 2018). The spatiotemporal characteristics of house price which requires a spatiotemporal modeling approach that can characterize non-separable spatiotemporal co-variances (Gamerman, 2010; Gneiting & Guttorp, 2010; Montero & Mínguez, 2018).

This is especially true in medium-small cities of developing countries such as China, in which people are increasingly moving from countryside areas to cities. First of all, in large cities when house price reaches a high level, it tends to increase slowly and then to be stable, while in medium-small cities house price increases slowly at the beginning of the city expansion and then quickly, often showing an exponential rise. The reason is mainly that at the beginning the populations of middle-small cities are relatively small and not large enough to cause the obvious change of the house price, before the populations of middle-small cities reach a certain level, the increasing populations will lead to the dramatically increasing demand of houses, or the special location of the small town around the big cities, with spillover real estate purchasing power that will result in the quick increase of house price; and the process will continue until the populations or reach to the scale of large cities or closing to the real estate development capacity of small cities. Moreover, the change rates of house price in middle-small cities usually are larger than those in large cities. In addition, the driving factors that affect the change of house prices in middle-small cities are different from those in large cities. For example, large cities may have more and larger city centers than middle-small cities, and the distance from a city center on house price in large cities may have a more obvious effect than that in

middle-small cities. A flexible spatiotemporal modeling method should be developed and used to accurately model and analyze the spatiotemporal variability of house prices and make its predictions at middle-small cities (Lindström et al., 2013)

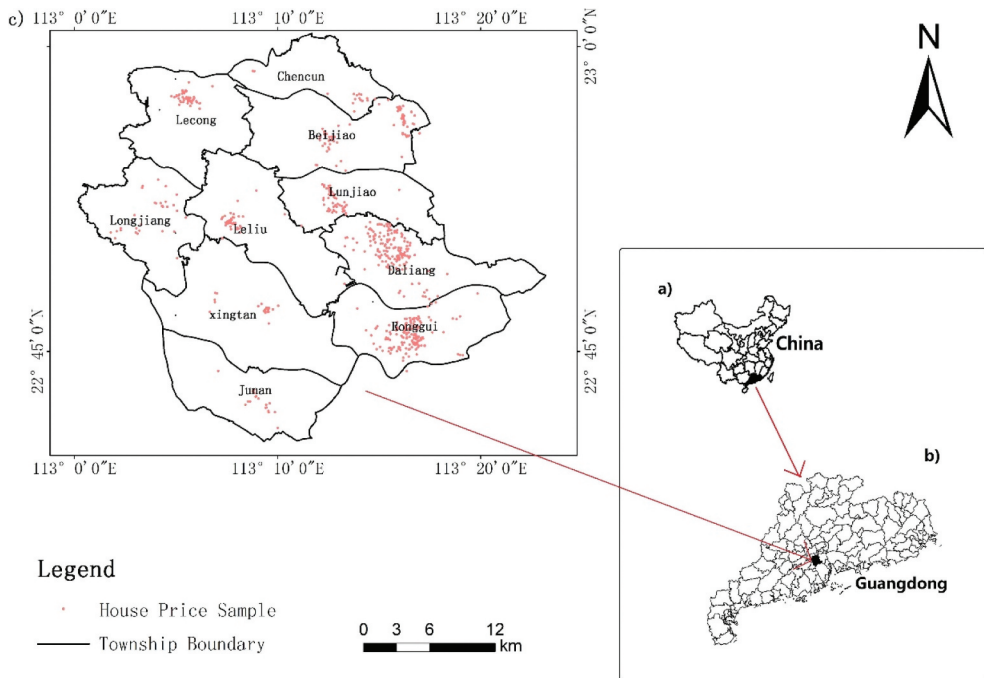
Most of the existing studies related to city house price deal with spatial distribution of house price at a single time and ignore its dynamic changes. The dynamic change of house price is influenced by location condition, macroeconomic policy, environmental factors, etc. The factors affecting the change of house price are constantly changing. Modeling the spatiotemporal variability of house price should take into account the spatiotemporal characteristics of house price and the driving factors.

The general objective of this paper was to develop a flexible spatiotemporal model (FSTM) as a general method to model the spatiotemporal variability of house price and to fill the knowledge gap for prediction of house price for middle-small cities. The flexible FSTM model means different time variations and spatial variations of housing price can be considered in the model in an easy manner, as an approximation of the real complex relationships in housing price from the temporal and spatial domains. The specific purpose was to validate the method by investigating and analyzing the spatiotemporal variability of residential house price, and developing and examining the FSTM in Shunde District of Foshan City located in Pearl River, Southern Guangdong of China. Moreover, this study explored the major factors that affect the residential house price and quantify their effects. In addition, this study could also enhance understanding the relationship of the residential house price with the driving factors and provide the technologic basis for the scientific management and planning of land resources in the urbanization of middle-small cities for China and other developing country in the world.

## 2. Materials and Methods

### 2.1. Study area

This study was conducted in a middle-small city – Shunde District, Foshan, Guangdong of China. Shunde District is located at the latitude and longitude of 23°00'N and 113°12'E (Figure 1), adjacent to the south of Guangzhou, the west of Shenzhen, the southeast of Hong Kong and the north of Macao in the Pearl River Delta. Shunde District has an area of 806.08 square km and a population of 2.54 million. It is a typical representative of middle-small cities with the fastest economic development in southern China during the past 30 years. The urbanized area occupied 89% of the district in 2013. The gross domestic product of Shunde District in 2016 was about 291 billion RenMingBi (RMB) (about \$45 billion). The public transportation system in Shunde District is extremely efficient with railways or metros, highway, and other means of transportations. Shunde District also has a long history of industrial development with significant effect of land use since the 1990s, which is called “China’s home appliances” and “the world’s food capital”.



**Figure 1.** Location of the study area – Shunde District in a) China and b) Guangdong; and c) spatial distribution of residential house price samples.

## 2.2. Residential house price

The residential house price data sets of Shunde district were obtained from the residential house price database for years of 1999 to 2010; and the land price updating database for years of 2003, 2006 and 2010 obtained from Land Bureau of Shunde District. Other data sets included land use data for years from 2006 to 2010, Greenland map and water map, and socioeconomic data from the Statistical Yearbook of Shunde for the years of 1992, 2000, 2003, 2008 and 2012. A total of 1498 residential house price samples were collected from 653 sample areas. In this study, more than one house prices were often collected in the sample areas and thus we used the average values of the house price in each of the sample areas to analyze its characteristics of spatial distribution. The coordinates of each of the sample areas were created from a topographic map in ArcGIS software. The residential house price data were used to conduct spatial analysis and modeling.

## 2.3. Factors affecting residential house price

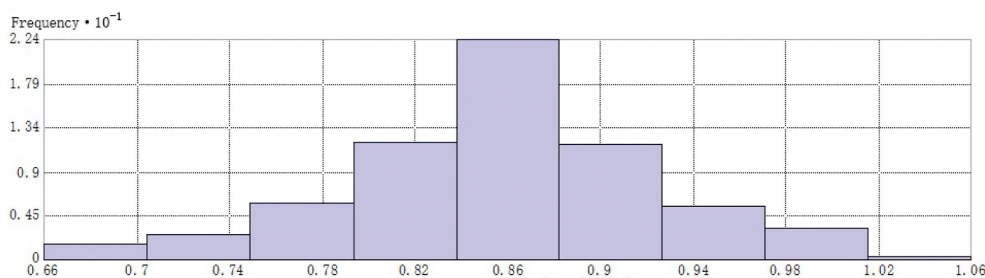
All the residential house price data had the detailed information of land use type, house type, ownership, land use term, garage type, and house structure, lot size, sale price, location, public area and so forth. The other attributes included school, hospital, distance to city center, distance to bus station, road density, population density, distance to metros or rail transits, market density, post office density, supermarket density, and other factors that affected residential house price.



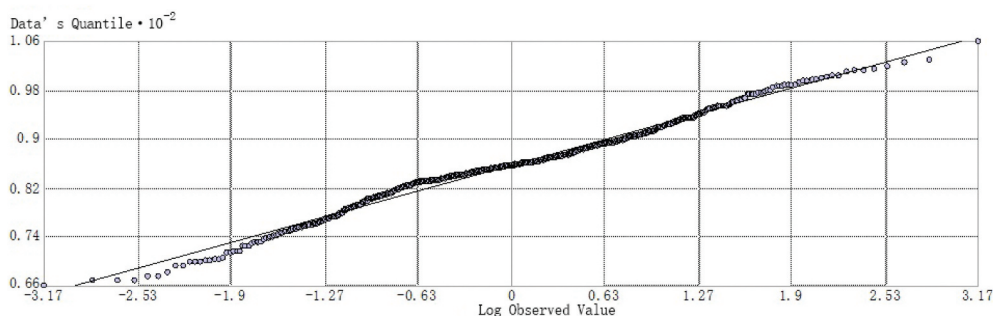
The housing price data were collected from Real estate transaction system of Shunde District. All the data were transformed to the same spatial coordinate system and spatial resolution. The descriptive statistic analysis of the house price data is in the following:

The data set of the residential house price from different regions and times in Shunde district had a mean and standard deviation. The original data set was not normally distributed for each year, but the logarithmic transformation values of the house price data had a normal distribution for each of the years. As examples, the frequency distribution for year 2007's house price data (Figure 2) and quantile–quantile (Q–Q) graph (Figure 3) for year 2010's house price data verified the normal distribution of the logarithmically transformed house price data.

A 3-dimensional perspective of the residential house price was generated by an exploratory spatial data analysis in Geographic information system (GIS), showing the overall trend of the residential house price values (Figure 4). In both the X and Y directions, there was a U-shaped trend of the residential house price values. That is, the house prices gradually decreased and after reaching the lowest values and then gradually increased. The U-shaped curves were consistent with the overall trends of the residential house prices in Shunde District, that is, the house prices were higher at the east, south, west and north parts and lower in the central area.

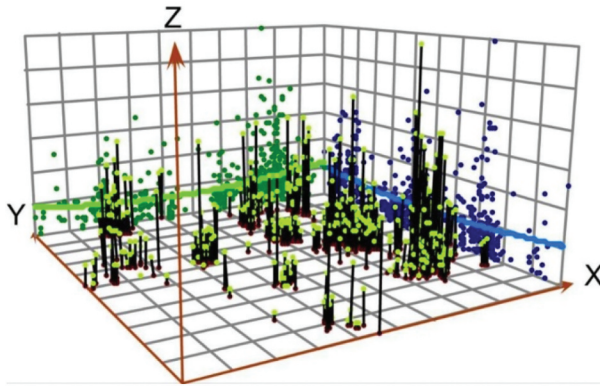


**Figure 2.** Frequency distribution of logarithmic transformation residential house price values for year 2007.



**Figure 3.** The quantile – quantile (Q-Q) graph of the logarithmically transformed residential house price values for 2010.





**Figure 4.** The 3-dimensional distribution of Shunde District residential house price data.

Urban land prices consisted of residential house price, commercial building price and industrial land price. In this study, only the residential house price data were used because of the data completion in the Shunde land price system. Moreover, the factors or variables that affected the changes of the residential house price and were taken into account included the distance of a house to the center of the district (DisCenter), the distance of a house to a metro or subway (DisSubTrain), density of post offices (DensityPost), density of supermarkets (DensitySupmarket), density of roads (DensityRoad), density of markets (DensityMarket), density of banks (DensityBank), public area (PublicArea), density of bus stations (DensityBus), density of telecommunication halls (DensityTeleCom), area used by public (AreaUsedP), user shared area within a building (UserShareAreaB), and density of population for each year, such as Density\_2003Pop meaning the density of population for 2003. The area used by public was defined as the built-up areas, such as roads, playgrounds and squares that are shared by the citizens who live in the same area or village. Moreover, the city population change and gross domestic product (GDP) of Shunde District greatly affect the city house price. As the city population and GDP increase, generally the house price increases. The increase of the urbanized or built-up area also affect the house price. In addition, in China the governmental policy has great impact on city house price. For example, in China banks are often state-owned, when the city economic development slows down, the government often let the banks lower down the interest rates and make loans easier to obtain to promote the economic development. However, the factors such as GDP and governmental policy have the temporal information only for the whole city at different times and lack the spatial data of the sample areas.

#### **2.4. Statistical analysis of data**

First, we tested whether or not the statistical distribution of the residential house price in Shunde district was normal based on the frequency distributions derived using the sample data and their logarithm transformation. Then, we analyzed the global and local spatial autocorrelation of the residential house price using Moran's I index and Moran's I scattergram (Anselin, 1995; Sokal, Oden, & Thomson, 1998). When local

spatial autocorrelation analysis was carried out, the point data of the house price were transferred into surface data using a Thiessen polygon method and the spatial weight matrix was obtained based on a first-order ROOK weight matrix of the transformed surface. That is, when the polygon  $i$  and the polygon  $j$  were adjacent,  $w_{ij} = 1$  and otherwise  $w_{ij} = 0$ . Moreover, significance analysis of the factor effects on the residential house price was conducted using Monte Carlo test.

## 2.5. FSTM for residential house price

To account for the spatial and temporal variability of residential house price, in this study we developed the FSTM by considering together the spatial and temporal effects of the factors or independent variables on the residential house price, and taking into account the spatial and temporal correlations in the residuals of the model predictions for Shunde District as follows (Lindström et al., 2013; Sampson, Szpiro, Sheppard, Lindström, & Kaufman, 2011):

$$y(s, t) = \mu(s, t) + v(s, t) \quad (1)$$

where  $y(s, t)$  represented the observed house price value at the location  $s$  and time  $t$ ;  $\mu(s, t)$  indicated the expected house price at the location  $s$  and time  $t$ ;  $v(s, t)$  meant the spatiotemporal error term. The expected house prices were estimated by following model:

$$\mu(s, t) = \sum_{i=1}^m \beta_i(s) + f(t) \quad (2)$$

where  $f(t)$  represented the time component and was estimated by a piecewise linear model,  $\beta_i(s)$  accounted for the space component and was estimated by a random effect model:

$$\beta_i(s) \sim N(X_i \alpha_i, \Sigma_{\beta_i}(\theta_i)) \quad (3)$$

where  $X_i$  represented the designed matrix consisting of the geospatial data collected;  $\alpha_i$  were the regression coefficients;  $\Sigma_{\beta_i}(\theta_i)$  was a  $n \times n$  covariance matrix. In this model, the regression coefficients for each location were considered as random variables and could explain the variability of the independent variables, that is in the random coefficients model, we assume coefficients corresponding to different locations are sampled from the same normal distribution, which is a standard assumption in the random coefficient model. It was assumed that the spatiotemporal error term  $v(s, t)$  in the model obeyed a normal distribution in which the spatial variance could be expressed as:

$$\mathbf{S}_v^1(\theta_v) \mathbf{0} \mathbf{0}^T \cdot \mathbf{0} \mathbf{0} \mathbf{S}_v^T(\theta_v) \quad (4)$$

where the size of each matrix  $\mathbf{S}_v^t(\theta_v)$ ,  $t = 1, \dots, T$ , was  $n_t \times n_t$  ( $n_t$  was the number of the house price samples collected at time  $t$ ). The FSTM was solved using the maximum likelihood estimation and specified as follows

$$y(s, t) = \alpha_0 + X_1(s)\alpha_1 + \dots + X_{12}(s)\alpha_{12} + \beta_{0s} + Z(s)\beta_{1s} + \hat{f}(t) + v(s, t) \quad (5)$$

where  $X_1(s), \dots, X_{12}$  represented the independent variables or factors that significantly contributed to improving the predictions of the house price;  $\alpha_1, \alpha_2, \dots, \alpha_{12}$  were the regression coefficients;  $Z(s)$  was the road density;  $\beta_{1s}$  represented the random slope of location  $s$ ;  $\beta_{0s}$  was the random intercept of location  $s$ ;  $\beta_{1s}$  and  $\beta_{0s}$  obeyed the independent normal distribution;  $\hat{f}(t)$  was the piecewise linear nonparametric function as the time factor. The parameters of the FSTM were estimated using R statistical software package (Lindström et al., 2013).

The FSTM is a multi-level mixed effects model that is comprised of a spatiotemporal trend model and spatiotemporal residuals. In order to improve computational efficiency in the modelling, Gaussian Markov Random Fields (Lindgren et al., 2011), predictive process (Banerjee, Ghaoui, & d'Aspremont, 2008) and fixed rank kriging (Cressie & Johannesson, 2008) were extended to the spatiotemporal data analysis. In FSTM, the maximum likelihood method was utilized to get standard errors (Szpiro et al., 2010; Lindström et al., 2013). It was also assumed that the spatiotemporal residuals  $v(s, t)$  be temporally independent but spatially correlated with each other with a common covariance for all time periods.

A 10-fold cross validation was used to estimate the model parameters, make the spatiotemporal predictions of the residential house price and assess the prediction accuracy. That is, the sample data were divided into a total of 10 sets and each data set was to estimate the model parameters and make the spatiotemporal predictions of the residential house price. The predictions were compared with the observed values of the residential house price and the prediction accuracies including the root mean square error (RMSE), the coefficient of determination ( $R^2$ ), the root mean squared prediction error (RMSPE) in Eq. (6) and the mean absolute prediction error (MAPE) in Eq. (7) were calculated (Seya, Tsutsumi, Yoshida, & Kawaguchi, 2011). Both the RMSPE and MAPE should tend to zero.

$$RMSPE = \sqrt{\frac{1}{n} \sum_{i=1}^n ((y_i - \hat{y}_i)/y_i)^2} \quad (6)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (7)$$

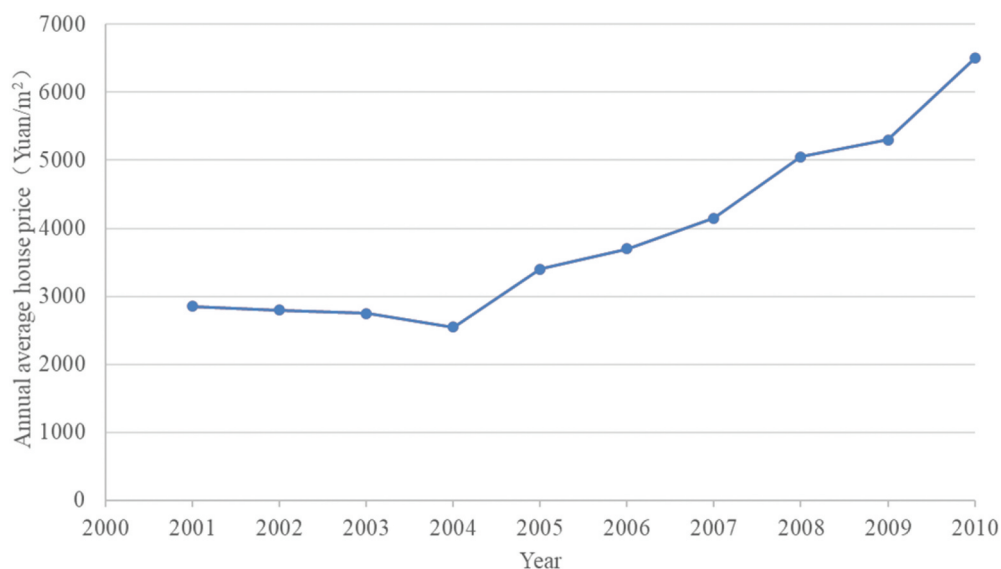
where  $y_i$  and  $\hat{y}_i$  are the observed and predicted house prices at the location  $i$ .

### 3. Results and Discussion

#### 3.1. Time Variation and impact factors of Residential House Price

Based on the original data, we calculated the annual average values of the residual house price for Shunde District and obtained the temporal trend characteristic  $f(t)$ . In Figure 5, the horizontal axis is the year and the vertical axis represents the annual average values of the residential house price (Chinese yuan/m<sup>2</sup>).

Overall, the main temporal change of residential land price in Shunde District slightly decreased from 2824 Chinese Yuan/m<sup>2</sup> to 2468 Chinese Yuan/m<sup>2</sup> with an overall decrease rate of 4.4% during the years of 2001 to 2004 and then continuously increased from 2468 Chinese Yuan/m<sup>2</sup> to 6545 Yuan/m<sup>2</sup> with an overall increase rate of 18.0% during the years



**Figure 5.** The temporal variability of the residential house price in Shunde District for the years of 2001 to 2010.

of 2004 to 2010. There were two quick increases for the time periods from the year 2004 to the year 2005 and the year 2009 to the year 2010, respectively. The temporal change characteristics of house price could be accounted for by the dynamics of economic development quantified by GDP, population, and urbanization (built-up area). It was found that the residential house price was statistically significantly correlated with the city GDP, population and built-up area at the significant level of 0.05, with the Pearson's correlation coefficients of 0.914, 0.86 and 0.725, respectively. This indicated that the factors had a significant influence on the residential house price during the years of 2001 to 2010.

Moreover, the governmental real estate market macro-control policy for city planning also obviously affected the dynamics of the residential house price of Shunde. For example, from the year of 2001 to 2004, the residential house price in Shunde slightly decreased. However, in the year of 2004, the government began to regulate the housing market by increasing the interest rate to control loans from bank in big cities, which led to the increase of urbanization cost, the slowing down of housing investment and the control of residential house price. The financial policy was not applicable in middle-small cities. This led to moving of money from big cities to middle-small cities and thus resulted in the great increase of the residential house price in Shunde. A similar change of the residential house price in Shunde occurred from the year of 2009 to 2010 due to the governmental house market policy that limited the residents to own two houses. Overall, GDP, population, urbanization and governmental policy for city house markets had significant effects on the residential house price of Shunde.

### 3.2. Spatial variability and statistical analysis of data

The data of residential house price in Shunde were collected in different years. In order to reduce the computation burden and simplify the spatial analysis, the data of different years were updated and calibrated to the level of the year 2010 (Figure 6) based on the base land prices, the real estate transaction data and the exchange of currency. The calibrated residential house price data of 2010 had a mean and standard deviation of 6506 Chinese Yuan (\$1001) per square meter and 4683 Chinese Yuan (\$702) per square meter. There was a relatively large coefficient of variation, 72%, which led to a potential difficulty for generating accurate predictions of the residential house price. Based on the sample values, the calibrated house price of Shunde was higher at the southeast, east and northeast parts and lower in the southwest, west and central areas (Figure 6). Some extremely large house price values were found in the southeast part and also scattered in the east and northeast parts.

Most of the residential house price samples concentrated around the administrative center and the intersection of traffic lines, showing the spatial distribution patterns of clustering, including two old city centers of Daliang street and Ronggui street located in the east and southeast parts of Shunde, and the residential house groups: Lunjiao, Beijiao Town and Chencun Town in the northeast part, Lecong in the northwest, Longjiang in the west, Leliu and Xingtian in the central and Junan in the southwest of Shunde. The residential house price samples were also distributed along the main roads, including Guangzhou high-speed road, Bigui Road, Guangzhou intercity railway, Shunde Avenue, and Guangzhou highway line.

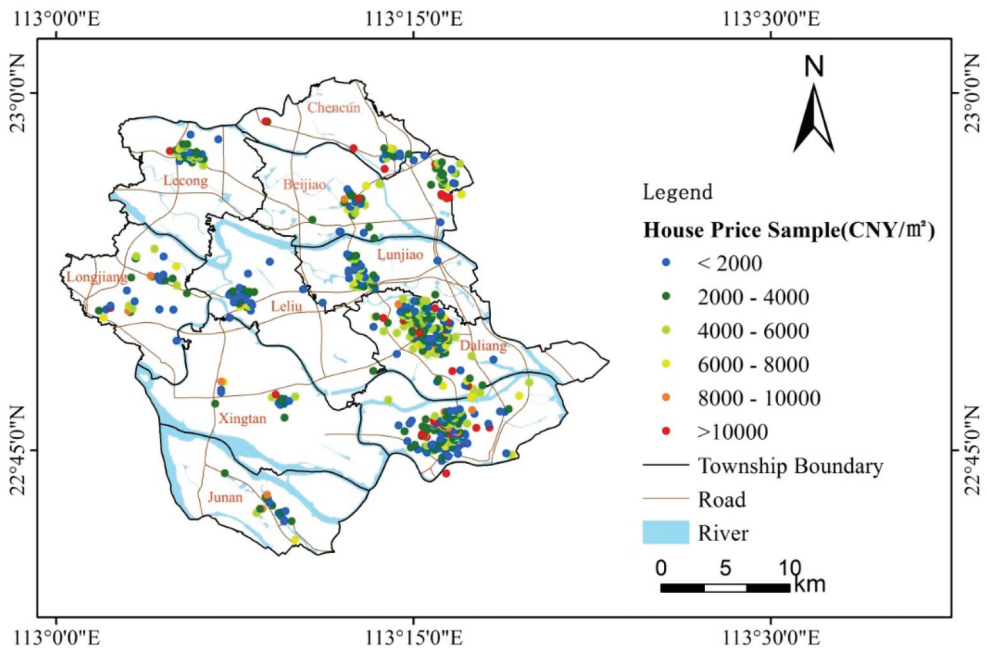
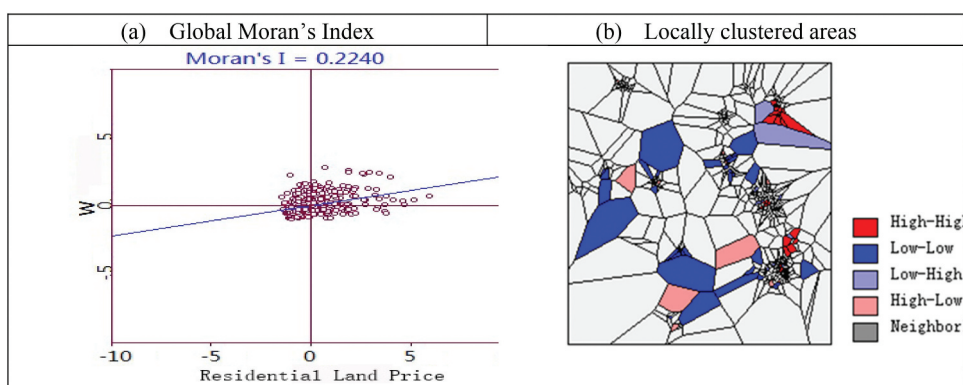


Figure 6. Spatial distribution of residential house price with township boundaries, rivers and roads.



**Figure 7.** Spatial autocorrelation of the residential house price in Shunde District: (a) global Moran's Index and (b) locally clustered areas.

The original values of the house price were not normally distributed, but the logarithmic transformation values of the house price had a normal distribution and quantile–quantile (Q–Q) graph of the house price data verified the normal distribution of the logarithmically transformed values. The spatial analysis led to a Moran's I index value of 0.224 with a z statistic of 10.71 and a p-value of 0.001, which indicated that overall the residential house price had a significant spatial autocorrelation (Figure 7a). In the Moran's I scatter plot, the points showed a trend from the first to the third Quadrant, implying positively spatial autocorrelation, the clustering of high house price values, and the clustering of low house price values. There were also some points falling in the second and fourth quadrant, showing negatively spatial autocorrelation and indicating high values surrounded with low and low values surrounded with high. At the significant level of 0.05, the significantly clustered areas of the residential house price in Shunde mainly concentrated in Ronggui Street, Daliang Street, Leliu Street and Beijiao Town (Figure 7b). The clustered areas of high values were mainly found in the central area of Daliang Street, Ronggui Street, Beijiao Town, and Chencun Town of the eastern part of Shunde District, being highly economic development areas. The clustered areas of low values mainly concentrated in the central and west of Shunde Lunjiao streets and Xingtian town, which indicated the cold spots of the residential house prices in Shunde District.

### 3.3. FSTM for Residential House Price

The FSTM of the residential house price was developed using the sample data. The FSM was then used to make predictions of the house price at the sample locations. The predicted values were compared with the observed prices to analyze the influence of each driving factor on the residential house price at each of the sample locations. The analysis was conducted based on the coefficients of FSTM and the p-values (Table 1).

The overall intercept was 2974.66 Chinese Yuan/m<sup>2</sup> with a p-value less than 0.05, indicating a statistical significance. The coefficients of the years 2002, 2003 and 2004 were negative and their absolute values increased, implying a temporal trend of decrease and corresponding to the decreased change of the residential house price from the year 2001

**Table 1.** The coefficients of flexible spatiotemporal model based on the sample data of residential house price.

	Coefficient	Standard error	p-value
(Intercept)	2974.66	222.83	0.000
as.factor(Time)2002	-120.30	277.37	0.665
as.factor(Time)2003	-164.37	260.14	0.528
as.factor(Time)2004	-456.48	290.14	0.116
as.factor(Time)2005	417.38	296.57	0.160
as.factor(Time)2006	575.53	298.60	0.054
as.factor(Time)2007	792.12	316.11	0.012
as.factor(Time)2008	1723.92	336.21	0.000
as.factor(Time)2009	2172.30	321.50	0.000
as.factor(Time)2010	3341.63	406.18	0.000
DisCenter	-130.66	135.56	0.335
DensityRoad	-411.78	135.61	0.002
DensityMarket	88.65	140.72	0.529
DensityBusS	-108.46	111.14	0.329
DensityTeleCom	-107.58	153.38	0.483
DensitySupmarket	-219.79	111.02	0.048
DensityPost	-57.82	99.85	0.563
DensityBank	423.41	197.04	0.032
AreaUsedP	234.20	78.76	0.003
UserShareAreaB	267.51	79.71	0.001
Density_2003Pop	79.39	112.92	0.482

to the year 2004, although their p-values were larger than the significant level of 0.05, indicating a statistical insignificance. This implied that the decrease of the residential house price was slight and not significant. Starting in 2005, the coefficients became positive, increased over years, and showed statistically significant effect on the residential house price at the significant level of 0.05 except the year 2006 in which the p-value was 0.054, slightly smaller than 0.05. This indicated that overall, the increase of the residential house price over the years was statistically significant.

For the factors that drove the spatial variability of the residential house price, there were only five factors that had the p-values smaller than 0.05, including DensityRoad, DensitySupmarket, DensityBank, AreaUsedP and AreaShareAreaB. DensityRoad with a negative coefficient significantly affected the residential house price, meaning that the higher the density of roads, the lower the residential house price. Within a city, the higher road density often implies an old downtown where there is much noise and the houses were old and had lower quality, which would reduce one's desirability of buying houses. DensitySupmarket also had a significant negative effect on the residential house price. In addition to noise, super markets often cause safe problems. Moreover, the residential houses are usually far away from super markets. Banks (DensityBank) provided convenience for residents and thus had significantly positive an effect on the residential house price. The areas used by public (AreaUsedP) provided facilities such as recreation and sport areas for the residents and thus enhanced the extraction for the residents to buy the houses. The shared areas within buildings had a positive coefficient, seemed not reasonable because the residents paid more for the residents shared areas. In fact, the residents liked to live in the buildings that had good facilities such as elevators and sport rooms.



Although other spatially driving factors did not significantly affect the residential house price, their coefficients were meaningful. DisCenter had a negative coefficient, suggesting that the larger the distances to the city center, the smaller the residential house price values. DensityBusS, DensityTeleCom and DensityPost all had negative coefficients because of noise and safety reasons. Moreover, the population densities of all the years were involved in the FSTM but removed except the population density of the year 2003 with a positive coefficient.

### 3.4. The spatiotemporal predictions and accuracy assessment using FSTM

The obtained FSTM was used to make the spatiotemporal predictions of the residential house price that consisted of the spatial distribution maps of the residential house price for the years from 2001 to 2010. The 10-fold cross validation led to a RMSE value of 2087 Chinese Yuan/m<sup>2</sup>, a relative RMSE of 58.8% (Table 2), a RMSPE value of 0.568 and a MAPE value of 0.41 (Table 3). The RMSPE and MAPE were statistically slightly significantly larger than zero at the significant level of 0.05 (Table 3).

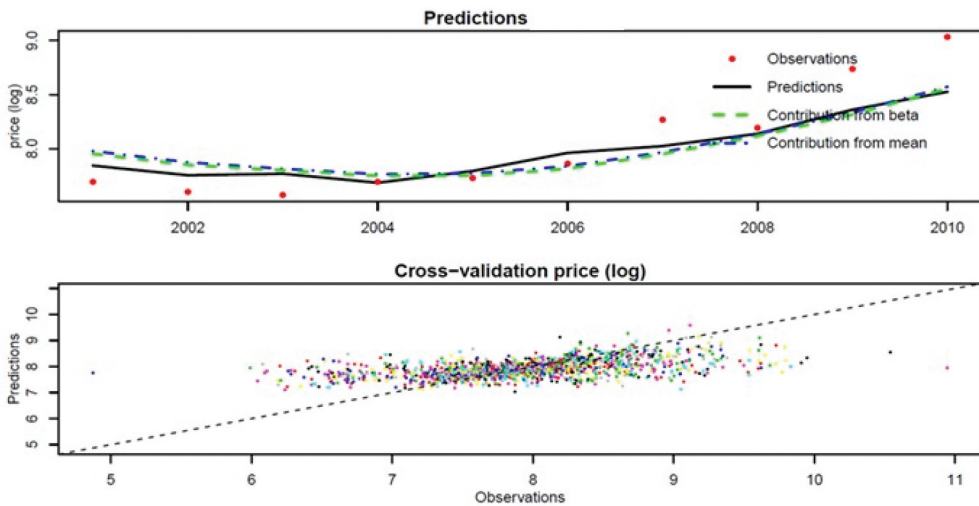
In Table 2, the large RMSE and relative RMSE values were caused mainly by the large overall coefficient of variation, 86.4%. The annual relative RMSE ranged from 31.9% to 114.2%. The extremely large value of the annual relative RMSE, 114.2%, occurred in the year of 2005 mainly because in 2005 there was one extremely large house price, 56,616 Chinese Yuan/m<sup>2</sup>, collected from a business center, while the annual sample mean of the year 2005 was 3377 Chinese Yuan/m<sup>2</sup>. When developing the FSTM, we did not

**Table 2.** The accuracy assessment of predicted residential house price values from FSTM based on root mean square error (RMSE) and relative RMSE (=RMSE×100/sample mean).

Year	Sample mean	Average estimate	RMSE	RRMSE
2001	2825	2825	1726	61.1
2002	2763	2763	1784	64.6
2003	2695	2695	1472	54.6
2004	2468	2468	1288	52.2
2005	3377	3377	3857	114.2
2006	3710	3710	1546	41.7
2007	4170	4170	1488	35.7
2008	5055	5055	2950	58.3
2009	5442	5442	2010	36.9
2010	6727	6598	2147	31.9
Overall	3551	3551	2087	58.8

**Table 3.** Comparison of performance results for models as a spatial predictor (RMSPE and MAPE).

RMSPE	MAPE	R Square	Cross Validation Method	Independent Variable Set
FSM	0.568	0.410	0.275	10 fold cross validation PricePer ~ List((Time)+ (DisCenter+DensityRoad+DensityMarket +DensityCar+DensityCom+DensitySupmarket+DensityPost +DensityBank+LandComArea+LandShareArea +Density_2003Pop)

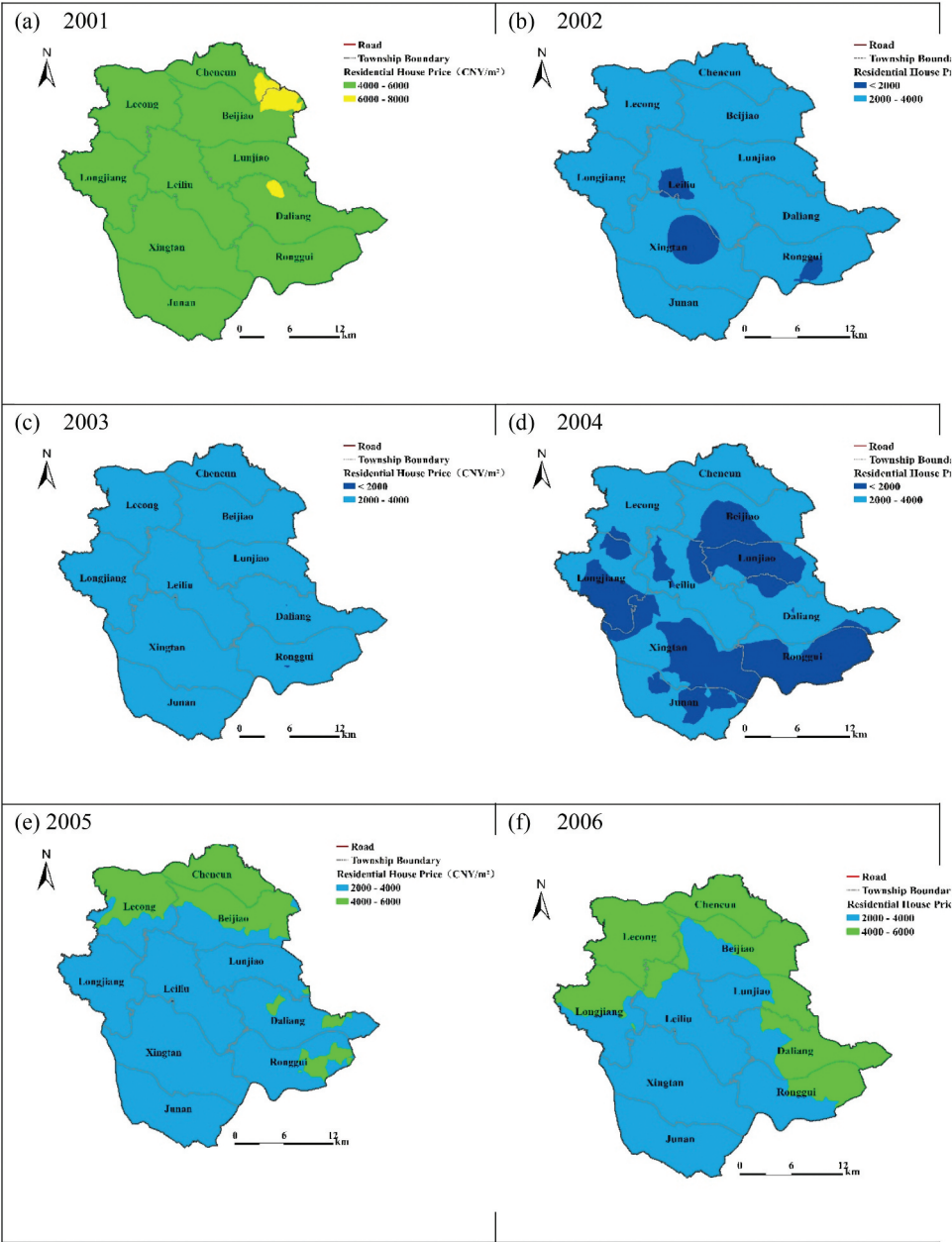


**Figure 8.** The relationship of the predictions from the FSTM with the observations: top) the log transformations of the observations and estimates over the years from 2001 to 2010; and bottom) Q-Q plot of the log transformations of the predictions against the log transformations of the observations.

remove this extremely large value because this happened in the practice. Moreover, the used house price samples were clustered in several business centers and not evenly distributed in the space, which might also have led to the large errors.

In [Figure 8](#), the obvious over-predictions took place before the year 2004, that is, when the overall house price values were smaller, while the large under-predictions existed after the year 2007, that is, when the house price values were larger. The reason was mainly because, although FSTM integrated the spatial and temporal variability of the residential house price into a model, given a time (year) the model took into account the relationships of the house price with the predictors globally and ignored their locally spatial variability. This indicated the potential improvement of modeling the spatiotemporal variability of the house price by introducing the local modeling into the FSTM.

The maps clearly showed the spatiotemporal variability of the residential house price ([Figure 9](#)). Overall, the residential house price decreased from the year 2001 to the year 2004 and then increased until the year 2010. The residential house price was higher in the north, northeast and east parts of Shunde and lower in the southeast, south, southwest and west parts. The residential house price obviously decreased in the east central, southeast and west parts from 2004 to 2005 and increased in the southeast, south, southwest and west parts over the years from 2005 to 2010. The clustering of higher residential house price also occurred in the fast development areas of business such as Town Daliang and Town Ronggui.



**Figure 9.** The spatiotemporal prediction maps of the residential house price for years 2001 to 2010 for Shunde District.

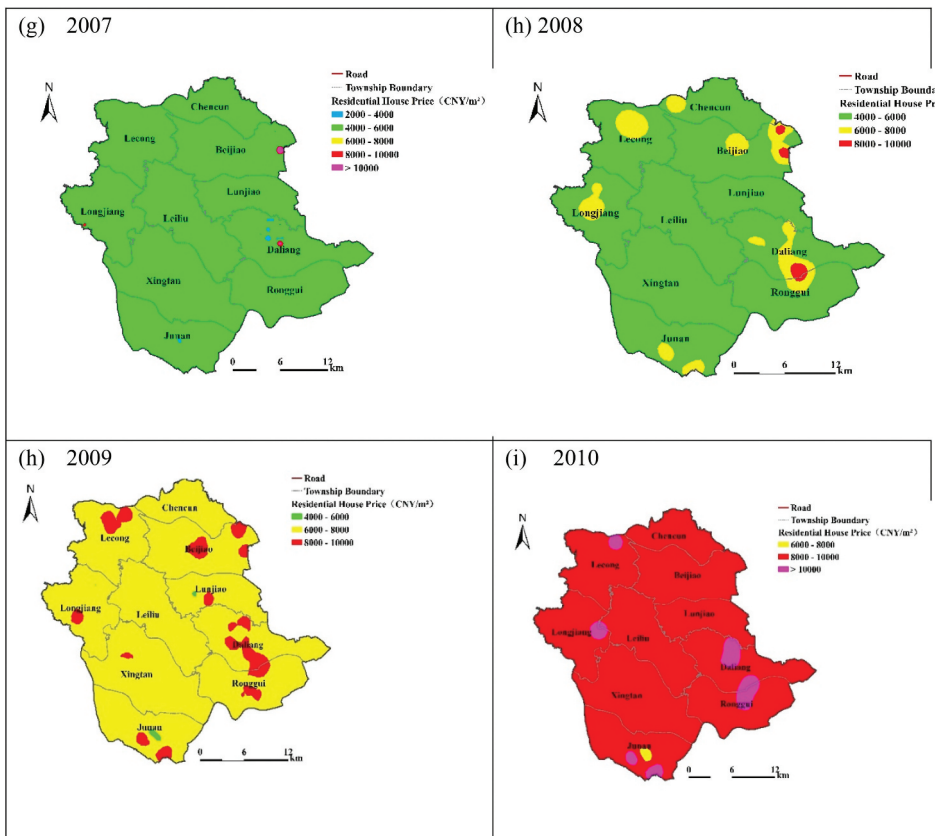


Figure 9 (Continued).

#### 4. Conclusions

Quantitatively understanding the variable spatiotemporal characteristics of residential house price, especially in middle-small cities in the fast developing regions, is crucial to the development and planning of cities in China. Shunde District has experienced a rapid urbanization, which has ranked the top 100th national economic zone for five consecutive years since the year of 2005. Its residential house price changed with the economic development process, which showed a unique characteristic different from those in big cities.

Based on a total of 1498 residential house price samples, we developed the FSTM to explore the spatiotemporal characteristics of the residential house price in Shunde as an example of the middle-small cities in China. The obtained FSTM integrated both spatial and temporal components, accounted for the spatiotemporal characteristics and reproduced the spatial variability and temporal trends of the residential house price in Shunde. The model had a flexible correlation structure that allowed for the non-separability of space and time by modeling a long-term trend using an empirical orthogonal basis function with spatially correlated random fields of coefficients, which led to the

spatiotemporal results of the residential house price at the same time. The governmental policy had significant influence on the house price, which led to the slight decrease of the house price of Shunde from 2001 to 2004 and fast increased from 2005 to 2010. The factors that had significant effects on the residential land price included DensityRoad, DensitySupmarket, DensityBank, AreaUsedP and UserShareAreaB. The higher house prices concentrated in the north and northeast parts and the business central areas of Shunde. The lower house prices were distributed in the west, south and central parts.

However, the errors of the predicted house price from the FSTM are relatively large mainly because of the large coefficients of variation from the annual sample values of the residential house price. On the other hand, the house price values were spatially and temporally autocorrelated with each other. However, given a year and a set of values for the predictors, the FSTM provides a global prediction. The integration of FSTM with geographically weighted regression could offer great potential for the improvement of spatially and temporally predicting the residential house price.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work was supported by the National Natural Science Foundation of China [U1901601]; National Natural Science Foundation of China [41871357]; National Natural Science Foundation of China [41971226].

## Notes on contributors

**Lu Wang** is an associate professor in the College of Natural Resources and Environment at South China Agricultural University. She received Ph.D. from the University of Chinese Academy of Sciences. Her research interests include land use policy, land resources management, and GIS application.

**Lu Guangxing Wang** is a professor in Department of Geography and Environmental Resources, Southern Illinois University. His research interests include remote sensing, GIS, spatial statistics and their applications to geography, natural and environmental resources.

**Huan Yu** is a professor in College of Earth Sciences, Chengdu University of Technology. He received a doctorate in geographic information science and mainly focused on the landscape geochemistry study using geospatial techniques, remote sensing for ecological environment, and intelligent space-time simulation of surface landscape evolution.

**Fei Wang** is a senior data scientist in CarGurus. He received PhD from University of Michigan. His research includes big data computation, causal inference and experiment, and neural networks.

## ORCID

Huan Yu  <http://orcid.org/0000-0001-6107-6373>

## References

- Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, 27(2), 93–115.
- Banerjee, O., Ghaoui, L. E., & d'Aspremont, A. (2008). Model selection through sparse maximum likelihood estimation for multivariate Gaussian or binary data. *Journal of Machine Learning Research*, 9(3), 485–516.
- Brueckner, J. K. (1990). Growth controls and land values in an open city. *Land Economics*, 66(3), 237–248.
- Chen, Y. H., & Fik, T. (2017). Housing-market bubble adjustment in coastal communities—A spatial and temporal analysis of housing prices in Midwest Pinellas County, Florida. *Applied Geography*, 80, 48–63.
- Colwell, P. F., & Munneke, H. J. (2009). Directional land value gradients. *The Journal of Real Estate Finance and Economics*, 39(1), 1–23.
- Cressie, N., & Johannesson, G. (2008). Fixed rank kriging for very large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(1), 209–226.
- Davis, M. A., & Heathcote, J. (2007). The price and quantity of residential land in the United States. *Journal of Monetary Economics*, 54(8), 2595–2620.
- Diewert, W. E., Haan, J. D., & Hendriks, R. (2015). Hedonic regressions and the decomposition of a house price index into land and structure components. *Econometric Reviews*, 34(1–2), 106–126.
- Du, X., & Huang, Z. (2018). Spatial and temporal effects of urban wetlands on housing prices: Evidence from Hangzhou, China. *Land Use Policy*, 73, 290–298.
- Gamerman, D. (2010). Dynamic spatial models including spatial time series. In Gelfand, A. E., Diggle, P. J., Fuentes, M., & Guttorp, P. (eds.), *Handbook of Spatial Statistics* (pp. 437–448). Chapman & Hall/ CRC, Boca Raton.
- Geoghegan, J. (2002). The value of open spaces in residential land use. *Land Use Policy*, 19(1), 91–98.
- Glumac, B., Herrera, M., & Licheron, J. (2019). A hedonic urban land price index. *Land Use Policy*, 81, 802–812.
- Gneiting, T., & Guttorp, P. (2010). Continuous parameter spatio-temporal processes. *Handbook of Spatial Statistics*, 97, 427–436.
- Gong, Y., & de Haan, J. (2018). Accounting for Spatial Variation of Land Prices in Hedonic Imputation House Price Indices: A Semi-Parametric Approach. *Journal of Official Statistics*, 34(3), 695–720.
- Goodman, A. C. (1978). Hedonic prices, price indices and housing markets. *Journal of Urban Economics*, 5(4), 471–484.
- Hu, S., Cheng, Q., Wang, L., & Xie, S. (2012). Multifractal characterization of urban residential land price in space and time. *Applied Geography*, 34, 161–170.
- Hu, S., Cheng, Q., Wang, L., & Xu, D. (2013). Modeling land price distribution using multifractal IDW interpolation and fractal filtering method. *Landscape and Urban Planning*, 110, 25–35.
- Hu, S., Yang, S., Li, W., Zhang, C., & Xu, F. (2016). Spatially non-stationary relationships between urban residential land price and impact factors in Wuhan city, China. *Applied Geography*, 68, 48–56.
- Jiao, L., & Liu, Y. (2010). Geographic field model based hedonic valuation of urban open spaces in Wuhan, China. *Landscape and Urban Planning*, 98(1), 47–55.
- Leggett, C. G., & Bockstael, N. E. (2000). Evidence of the effects of water quality on residential land prices. *Journal of Environmental Economics and Management*, 39(2), 121–144.
- Lindgren, F., Rue, H., & Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 73(4), 423–498.
- Lindström, J., Szpiro, A. A., Sampson, P. D., Oron, A. P., Richards, M., Larson, T. V., & Sheppard, L. (2013). A flexible spatio-temporal model for air pollution with spatial and spatio-temporal covariates. *Environmental and Ecological Statistics*, 21, 1–23.

- Majumdar, A., Munneke, H. J., Gelfand, A. E., Banerjee, S., & Sirmans, C. (2006). Gradients in spatial response surfaces with application to urban land values. *Journal of Business & Economic Statistics*, 24(1), 77–90.
- Miller, Z. R. (1973). *The urbanization of modern America: A brief history*. New York: Harcourt Brace Jovanovich.
- Montero, J., Mínguez, R., & Fernández-Avilés, G. (2018). Housing price prediction: Parametric versus semi-parametric spatial hedonic models. *Journal of Geographical Systems*, 20(1), 27–55.
- Oshan, T. M., & Fotheringham, A. S. (2018). A comparison of spatially varying regression coefficient estimates using geographically weighted and spatial-filter-based techniques. *Geographical Analysis*, 50(1), 53–75.
- Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34–55.
- Sampson, P., Szpiro, A., Sheppard, L., Lindström, J., & Kaufman, J. (2011). Pragmatic estimation of a spatio-temporal air quality model with irregular monitoring data. *Atmospheric Environment*, 45(36), 6593–6606.
- Scott, J. T. (1983). Factors affecting land price decline. *American Journal of Agricultural Economics*, 65(4), 796–800.
- Seya, H., Tsutsumi, M., Yoshida, Y., & Kawaguchi, Y. (2011). Empirical comparison of the various spatial prediction models: In spatial econometrics, spatial statistics, and semiparametric statistics. *Procedia-Social and Behavioral Sciences*, 21, 120–129.
- Sokal, R. R., Oden, N. L., & Thomson, B. A. (1998). Local spatial autocorrelation in a biological model. *Geographical Analysis*, 30(4), 331–354.
- Szpiro, A., Sampson, P., Sheppard, L., Lumley, T., Adar, S., & Kaufman, J. (2010). Predicting intra-urban variation in air pollution concentrations with complex spatio-temporal dependencies. *Environmetrics*, 21(6), 606–631.
- Tsutsumi, M., & Seya, H. (2008). Measuring the impact of large-scale transportation projects on land price using spatial statistical models. *Papers in Regional Science*, 87(3), 385–401.
- Wang, Q., Wang, Y., Chen, W., Zhou, X., Zhao, M., & Zhang, B. (2020). Do land price variation and environmental regulation improve chemical industrial agglomeration? A regional analysis in China. *Land Use Policy*, 94, 104568.
- Yang, S., Hu, S., Wang, S., & Zou, L. (2020). Effects of rapid urban land expansion on the spatial direction of residential land prices: Evidence from Wuhan, China. *Habitat International*, 101, 102186.
- Yokoi, T., & Ando, A. (2012). One-directional adjacency matrices in spatial autoregressive model: A land price example and Monte Carlo results. *Economic Modelling*, 29(1), 79–85.
- Zhang, M., & Yen, B. T. H. (2020). The impact of Bus Rapid Transit (BRT) on land and property values: A meta-analysis. *Land Use Policy*, 96: 104684, 1–14.